SECTION 1
NEW VISION FOR FUTURE AEROSPACE VEHICLES AND SYSTEMS

Daniel S. Goldin
Samuel L. Venneri
Ahmed K. Noor

Since the middle of the last century, aerospace technology has become a unique, indispensable part of our world. Commercial aviation has made it possible for more people and cargo to travel faster than at any previous time in history.

But today we are reaching the capacity limits of the airspace system, while transportation demand—both passenger and freight—is expected to increase in the long term. In space, our assets have improved weather prediction, provided global communication and navigation facilities to link people and businesses more effectively, and enabled spectral imaging of Earth to enhance our environmental stewardship and land use. Space-based observations are also taking on unprecedented importance in national security strategy. However, the high cost of launching payloads to Earth orbit has impeded progress in the exploration of space, as well as in its commercial use and development.

To keep pace with the projected increase in demand for mobility, a bold integrated vision is needed for future aerospace transportation systems. The vision for the next century is based on new technologies, and today some revolutionary ideas, many developing under programs at NASA and other government agencies, are emerging.

1.1 LEARNING FROM LIVING SYSTEMS

In the past few years, a growing number of engineers have been seeking inspiration from living systems—birds, insects, and other biological models—in an effort to produce breakthroughs in vehicle concepts, reshape our frame of reference, and change the definition of what is possible. Knowledge of the form and function of biological systems is used not as a blueprint, but as an architectural and operational

analog (see Figure 1.1). In many cases, we are finding that nature's basic design concepts—and, in some cases the actual designs—can be applied to aerospace systems.

Over the eons, nature has developed design approaches that make biological systems far more reliable and efficient than human-made systems. Biological systems evolve, develop, learn, and adapt on their own. The goal is to make our aerospace systems equally adaptive—that is, capable of undergoing modifications according to changing environmental circumstances, thereby enhancing their safety and performance.

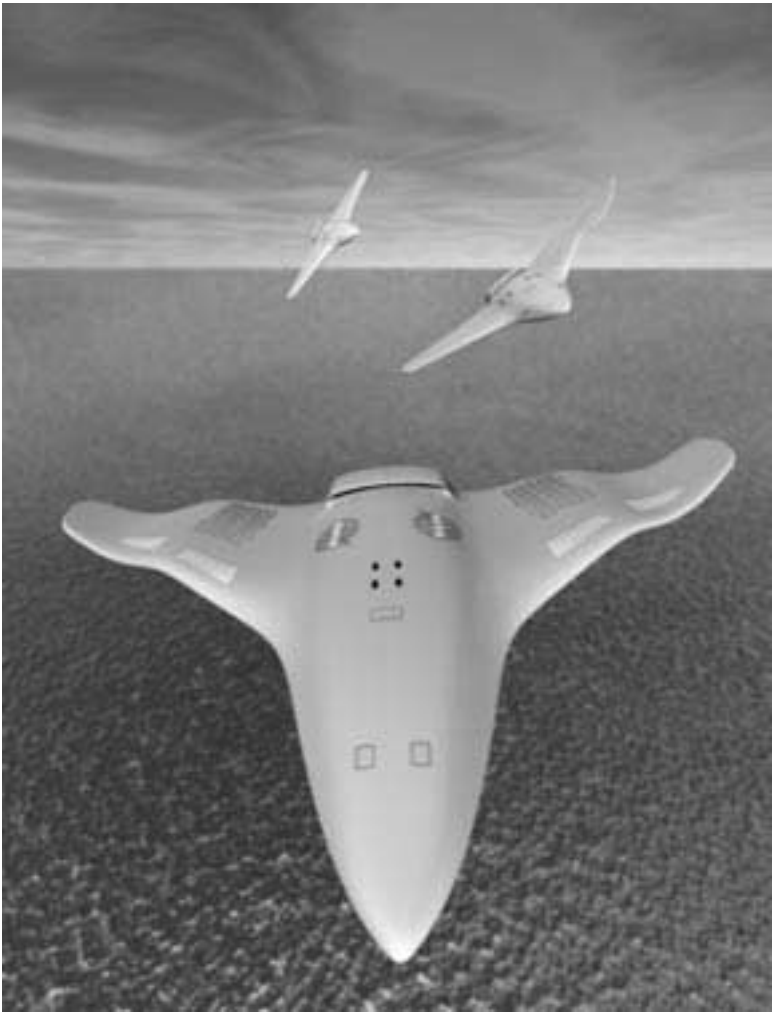


FIGURE 1.1 Smart vehicle.

NEW VISION FOR FUTURE AEROSPACE VEHICLES AND SYSTEMS

Flying birds display remarkable features of maneuverability that to date have no technological parallel. A bird can morph and rotate its wings in three dimensions. It has the ability to control the airflow over its wings by moving the feathers on its wingtips, which are more efficient than the flaps and rigid, pivoting tail surface of current aircraft.

A bird also is made from self-sensing and self-healing materials and has a fully integrated aerodynamic and propulsion system. Its skin, muscles, and organs have a nervous system that detects fatigue, injury, or damage and signals the brain. It is designed to survive.

Insects manage unsteady flow well. In insect flight, the flow is always separated. Understanding the mechanism by which insects dynamically manage unstable flow to generate lift could provide insight into the development of micro-air vehicles for a variety of missions. Also, insects have elaborate systems of sense organs. Organs in the cuticle, for example, detect bending strains and enable the insect to control its movements.

The application of biological concepts and principles to the development of technologies for engineering systems has led to the emergence of biomimetics, neuromimetics, and neuromorphic engineering. Biomimetics is the science of developing synthetic materials, devices, and structures that have the hierarchical organization, functionality, and strategies for optimization similar to biological systems. Neuromimetics deals with the application of neuroscience to building intelligent distributed knowledge networks. Neuromorphic engineering aims at the study of neuromechanical interactions, as well as the design and fabrication of systems whose architecture and design principles are based on biological nervous systems.

1.2 REVOLUTIONARY VEHICLES

Our national bird, the eagle, can serve as an operational model for a future aircraft. Instead of being built of mechanically connected parts and systems, eagles use fully integrated, embedded smart materials—nerves, sensory receptors, and muscles—that enable exceptional levels of efficiency and control. We are beginning to mimic this capability through the synergistic coupling of bio-, nano-, and infotechnologies.

Molecularly designed materials, or nanomaterials, for example, have the promise to be 100 times stronger than steel and only one-sixth the weight. Their maximum strain can be 10 to 30%, much higher than any current structural material. Wings made from such materials could morph and continuously deform for optimal control during takeoff, cruise, and landing. Also, like those of birds, the wings and body will be integrated for exceptional strength and light weight.

The entire surface of the wing will be covered with tiny embedded sensors and actuators. Sensors, like the nerves of a bird, will monitor and analyze temperatures, pressures, and vibrations. In response to sensor data, actuators embedded in the structure will function like a bird's wing muscles. The actuators will smoothly change the shape of the wing to adapt to flying conditions. The control surfaces will be integrated with the wing, instead of appended to it. And active flow control effectors will help mitigate adverse aircraft motion in turbulence.

Across the entire vehicle, intelligent systems composed of sensors, actuators, microprocessors, and adaptive controls will provide a full-system, distributed

knowledge network—a central nervous system to effect an adaptive physical response. The vehicles will be able to monitor their own performance, their environment, and even their operators. Eventually, vehicles may be able to anticipate and avoid impending failure, crashes, mishaps, and incidents through a distributed self-assessment and repair capability.

The application of high-temperature, molecularly designed materials to aircraft engines will be equally influential. Through successful application of advanced lightweight materials in combination with intelligent flow control and active cooling, thrust-to-weight ratio increases of up to 50% and fuel savings of 25% are possible for conventional engines. Further advances in integrating these technologies will result in novel engine concepts that may simplify highly complex rotating turbomachinery.

Among the novel aircraft propulsion concepts considered are adaptive propulsion system cycles, distributed mini-engines, an exoskeletal engine, and pulse detonation engines. Adaptive propulsion system cycles with variable inlets and exhaust nozzles could be configured to mix the flow from the fan and hot exhaust to minimize noise on takeoff and enhance efficiency during cruise.

In the distributed engine concept, the two or three large engines are replaced by many mini- or micro-engines. The use of several small engines allows thrust modulation to replace the control surfaces in the tail and rudder. Micro-engines can be more highly integrated with the vehicle airframe, potentially allowing a form of flow control that could reduce drag (see Figure 1.2).

The exoskeletal engine derives its name from its nonstandard design and is ideal for high Mach number applications. The turbomachinery blades are mounted inside

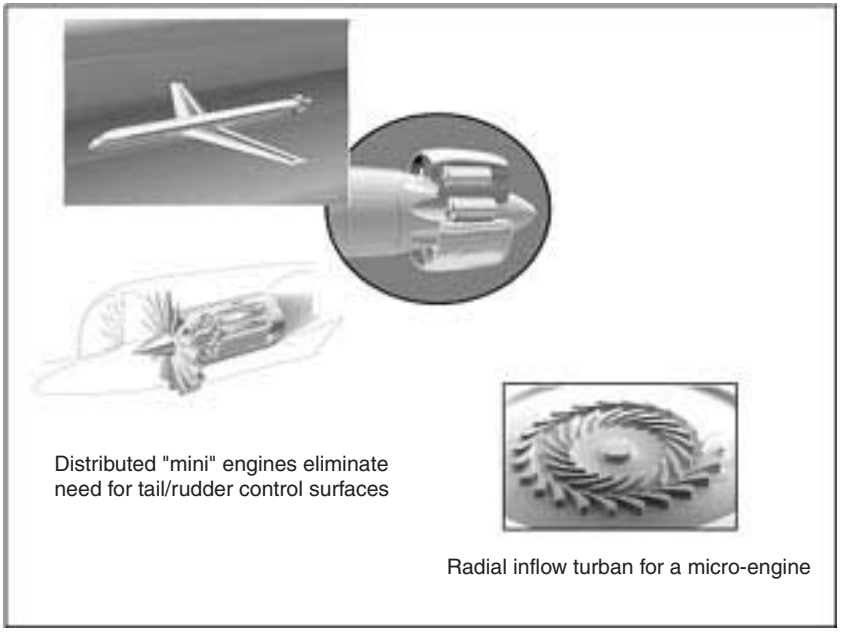


FIGURE 1.2 Distributed mini-engines.

a drum, rather than on the outer perimeter of a disk. This drum rotor places the blades in compression, allowing advanced composite and ceramic materials to be used, enabling lighter engines and higher operating temperatures (see Figure 1.3).

An open-cycle, pulse-detonation wave engine is an air-breathing, intermittent combustion jet engine that uses gas dynamics, instead of turbomachinery, to compress incoming air. It relies on a traveling detonation wave for the combustion and compression elements of the propulsive cycle. Combined with intelligent engine control capability, such an approach can lead to integrated internal flow management and combustion control. However, the problems associated with noise and the design of a complex set of valves and ducting in this concept need to be addressed.

Carried a step further, all of these concepts have the potential to integrate airflow over the airframe and through the propulsion systems for unprecedented efficiency, stability, and directional control. This will require new approaches to fully integrated airframe-propulsion design. Integrated propulsion and vehicle technology advancements could enable sustained supersonic flight with minimal impact from sonic booms or other environmental concerns.

In the very long term, comparable advances in electrical energy storage and generation technology, such as fuel cells, could lead to emissionless aircraft propulsion. Future aircraft might be powered entirely electrically. In one concept, thrust may be produced by a fan driven by highly efficient, compact electric motors powered by advanced hydrogen-oxygen fuel cells. However, several significant technological issues, such as efficient generation and storage, and an adequate infrastructure necessary for delivering the fuel to vehicles, must still be resolved in order to use hydrogen as a fuel.

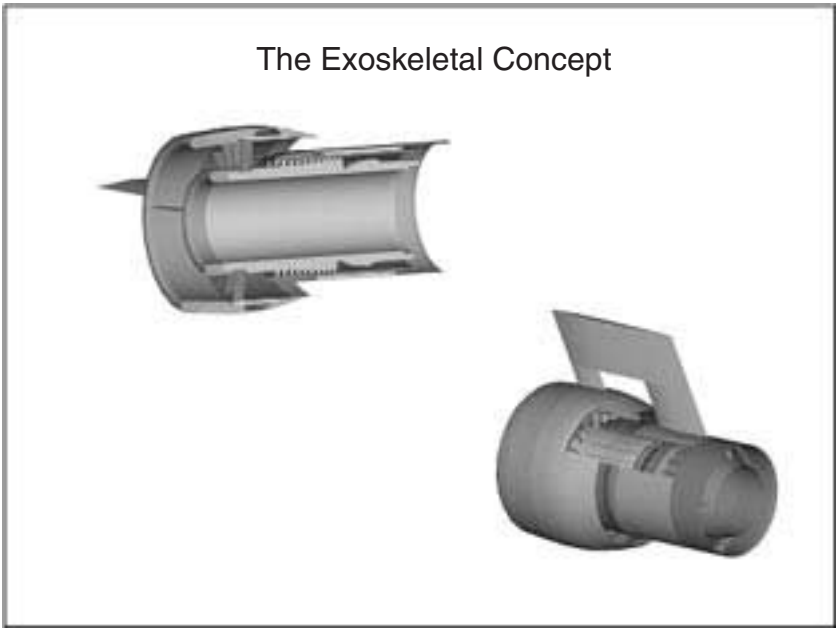


FIGURE 1.3 Exoskeletal engine design.

1.3 FUTURE SPACE TRANSPORTATION

Opening the space frontier to new levels of exploration and commercial endeavor requires a significant reduction in the cost of transportation systems and an increase in their reliability and safety. The long-term goal is to make aerospace vehicles capable of operating from the Earth's surface all the way to orbit, and to eliminate the distinction between a commercial airliner and a commercial launch vehicle.

In the near term, NASA will address crew safety by integrating sensor and information systems into the vehicle for improved health management and by providing in-flight crew-escape systems. The goal for Earth-to-orbit systems includes reducing the launch cost to $1,000 a pound and then to $100 a pound, while reducing the probability of failure currently from about 1 in 250 to 1 in 1,000 for the loss of the vehicle and 1 in 10,000 for the loss of crew, and eventually to 1 in 1,000,000 for both.

To address these goals, NASA has developed an integrated plan for space transportation which provides a phased strategy to ensure continued safe access to space. Investment in technical and programmatic risk-reduction activities will support full-scale development of commercially competitive, privately owned and operated Earth-to-orbit reusable launch vehicles, or RLVs (see Figure 1.4). The plan includes developing an integrated architecture with systems that build on evolutionary technologies for Earth-to-orbit launch vehicles, and developing revolutionary technologies for reusable hypersonic vehicles, and in-space transportation systems.

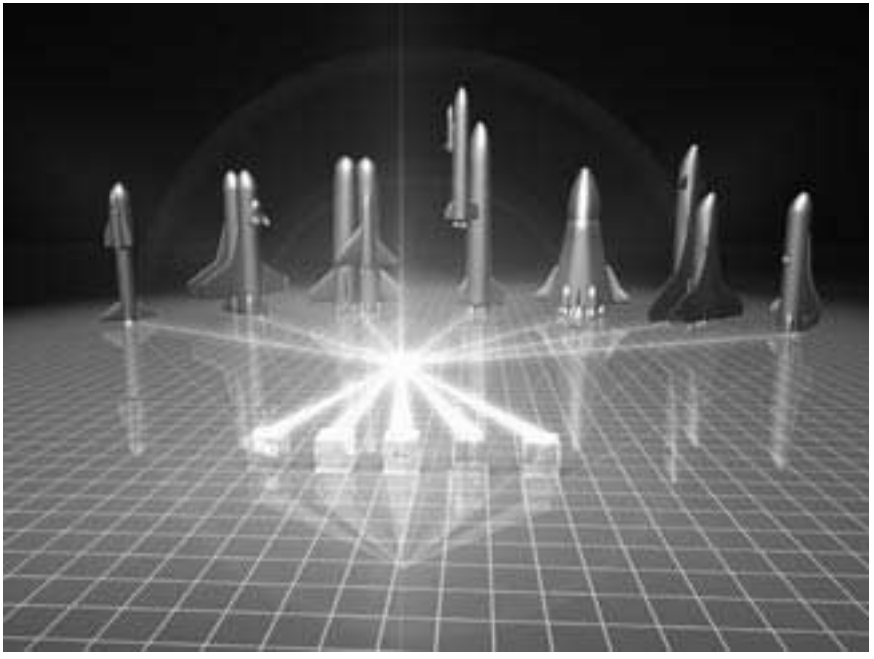


FIGURE 1.4 Variety of vehicle concepts for reusable launch vehicles.

NEW VISION FOR FUTURE AEROSPACE VEHICLES AND SYSTEMS

Setting the stage for aircraft-like operations for space access and reducing the life-cycle cost of Earth-to-orbit vehicles will require advances in a number of technologies, including propulsion, airframes and cryotanks, all-weather thermal protection, avionics and flight control, vehicle health management and monitoring, and automation technology to reduce ground operations costs.

Among the revolutionary propulsion concepts under consideration are supersonic combustion ramjet (scramjet), rocket-based combined cycle (RBCC), turbine-based combined cycle (TBCC), and magnetic launch assist.

The scramjet is a ramjet engine in which the airflow through the whole engine remains supersonic. Unlike rockets, the scramjet requires only onboard fuel, with the oxidizer obtained from the atmosphere. This offers significant weight and volume reductions, while the performance, which varies significantly with Mach number, exceeds that of a rocket over most of the flight trajectory.

The RBCC concept blends the performance of the rocket, ramjet, and scramjet into one system for seamless ground-to-orbit operation. It uses a rocket engine integrated with a ramjet/scramjet flow path. The rocket engine consumes atmospheric oxygen when flying in the atmosphere. When it leaves the atmosphere, the inlet ducts close and the rocket engine consumes stored liquid oxygen. The concept offers safety, reliability, and cost advantages by making vehicles smaller and more efficient.

The TBCC is similar to the RBCC in that it uses both jet propulsion and rocket propulsion. Initially, inlet ducts provide air to turbojets that have common exhaust ducting with a rocket propulsion system. At low to moderate Mach numbers, the turbojets provide thrust. The system then closes off inlet flow from the atmosphere, shuts down the turbojets, and converts to rocket propulsion.

In magnetic launch, the vehicle is accelerated horizontally along a track, using magnetic levitation to eliminate friction. Instead of the initial liftoff acceleration being provided by the onboard engines, linear electric motors are used to accelerate the launch vehicle to an initial speed of up to 1,000 km/hr. This initial boost can significantly reduce the amount of fuel needed to reach orbit and thus potentially increase payload for a given vehicle weight.

A number of other propulsion concepts are being explored for potential use in outer planet and interstellar space missions, including electric propulsion (see Figures 1.5 and 1.6), fission propulsion, laser-powered propulsion, antimatter propulsion, mini-magnetospheric plasma propulsion that couples to the solar wind for thrust, and space sails (both solar and laser driven). Space-based, electrically driven rotating tethers are also being evaluated for use to boost the velocity of near-Earth spacecraft.

1.4 FUTURE AIRSPACE SYSTEM

The current U.S. national airspace system is a complex collection of facilities, equipment, procedures, and airports that operate nonstop, 24 hours a day throughout the year. The FAA and NASA are currently making substantial efforts to modernize the equipment and procedures in order to enhance the efficiency of the airspace and the safety of its operations. These efforts include development of a suite of decision-support tools to improve gate-to-gate air traffic management. They also include deployment of technologies for transitioning from ground-based navigation

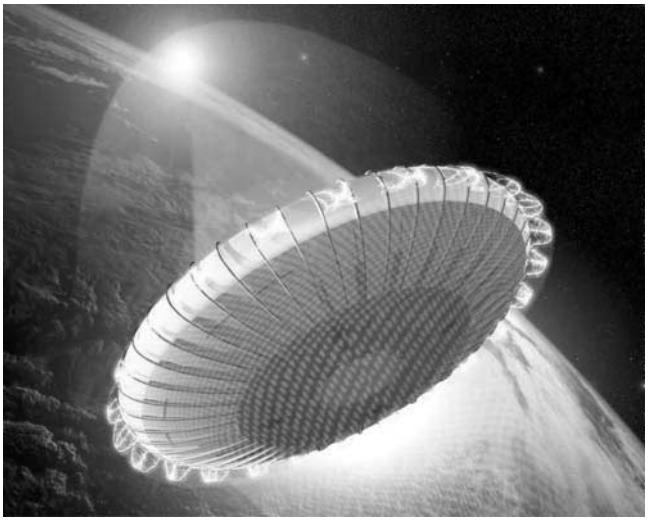


FIGURE 1.5 Beamed energy propulsion is one of several concepts under study for future space missions; topics of investigation also include laser power and fission propulsion, antimatter, and sails to catch the solar wind.

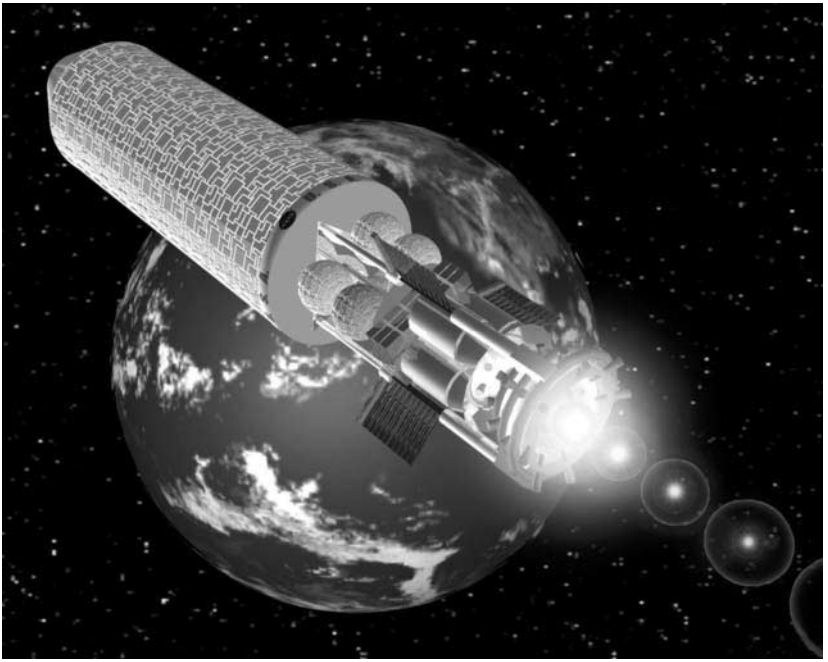


FIGURE 1.6 Today it is only an idea, but a magnetized target fusion propulsion system may some day enable rapid, affordable access to any point in the solar system.

NEW VISION FOR FUTURE AEROSPACE VEHICLES AND SYSTEMS

and surveillance systems to precise and reliable satellite-based systems in order to provide broader, more uniform coverage over the entire country for increased safety and expanded capacity (see Figures 1.7, 1.8 and 1.9).

The air traffic management system must be robust, able to tolerate equipment failures and interferences such as severe weather, and automatically adaptable to deal with increases in traffic. One approach for achieving scalability is to link air traffic management with an automatic flight management system on board each aircraft.

Such an air traffic management system will be built on global systems, such as GPS and future constellations of communication satellites. This will allow a precise approach to every runway without requiring installation of expensive, ground-based gear, such as instrument landing equipment, at every airport.

The system will enable the air traffic controller to have both regional and global views of the airspace. And while it will have a high degree of automation, both in the air and on the ground, human oversight will be retained.

Improved methods of weather data collection, processing, and transmission will be integrated into the airspace system. A significant challenge upon which the new architecture depends is the robustness of the global communication, navigation, and surveillance systems. The new system must allow graceful degradation—that is, be able to tolerate multiple failures and still be safe. Robustness in the system needs functional redundancy of communication, navigation, and surveillance systems, for example. This includes the ability to transition automatically and seamlessly from a space-based system to a ground-based system, such as LORAN.

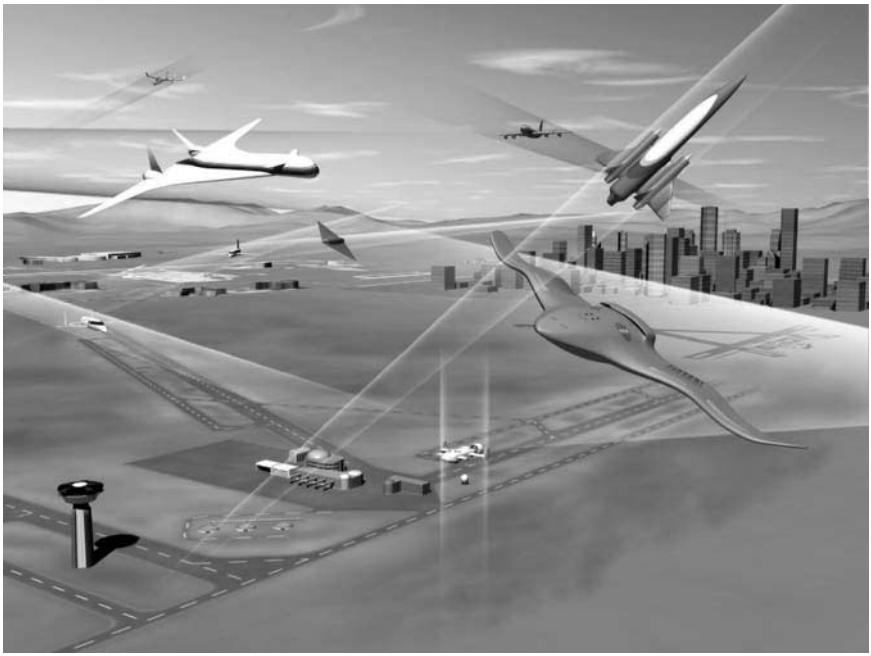


FIGURE 1.7 Highways in the sky: Sophisticated future technologies that will be able to monitor and, if necessary, take control of aerospace vehicles could one day raise the safety and security of flight to permit unprecedented ease of travel.

FIGURE 1.8 Futuristic spaceport.

FIGURE 1.9 Intermodal transportation systems engineering is key to optimal operations.

With a robust global system maintaining precise knowledge of position and trajectory for every aircraft, it will no longer be necessary to restrict flying to predetermined corridors. Optimal flight paths will be determined in advance and adjusted along the way for changes in weather and aircraft traffic.

Each aircraft will become capable of avoiding traffic. It will have an advanced flight management system that provides full knowledge of all aircraft in its area and enables it to coordinate, through an airborne internet, with other planes. The

FIGURE 1.10 NASA 507, a certified Lancair Columbia 300 general aviation aircraft, has state-of-the-art avionics, color moving map/GPS, and side-arm control stick.

pilot will be able to look at his flight path at different scales—from a strategic view of the entire route, showing other aircraft and weather systems, to a tactical view showing the immediate surroundings and flight path over the next few minutes.

Other technologies for improving safety and situational awareness may use advanced sensors, digital terrain databases, accurate geopositioning, and digital processing to provide three-dimensional moving displays showing aircraft, landing and approach patterns, runway surfaces, fixed and moving objects on the ground, and other relevant information (see Figure 1.10).

In the long term, full situational awareness will be achieved through interactive technologies (including voice recognition, eye tracking, and physiological monitoring). Cognitive neuroscience may allow observers to assess pilots' state of alertness to let them know when they are showing signs of fatigue and give them warning before they make mistakes.

1.5 NEEDED: INTELLECTUAL INFRASTRUCTURE

Although aerospace technologies have made significant advances in the last century, they have not reached a plateau. Technology advances in this century will be driven as much by the need for enhanced security, safety, and environmental compatibility as by the desire for improved efficiency and performance and reduced operating

costs. Combining biological, nanoscale, information, and other technologies can lead to revolutionary vehicle concepts and airspace management. Realizing the vision requires the creation of knowledge organizations, linking diverse interdisciplinary teams from NASA, other government agencies, universities, aerospace industry, and technology providers into hierarchical research and learning networks, configured as neural networks.

The learning networks will stimulate critical thinking and intellectual growth and promote intermingling among critical fields such as biological nanoscience, information science, and engineering disciplines. They will create a new generation of skilled scientists and engineers who can work across traditional disciplines and perform in rapidly changing environments. The research networks will link diverse, geographically dispersed teams and facilities.

The infrastructure enables collective intelligence, innovation, and creativity, through networking, to bear on the increasing complexity of future systems. Since a high percentage of experienced engineers in the aerospace industry will be eligible for retirement in a few years, the infrastructure could also offset the diminishing design team experience base in industry.

The airspace system can be integrated with the other two sets of transportation services—land and water—to form a comprehensive intermodal transportation system, functioning as one seamless whole, maximizing passengers' and shippers' options for convenience, efficiency, and reduced cost. It is a complex system of highly interconnected subsystems with multiple interfaces, shared information and infrastructure elements, and collaborative air-to-ground decision-making.

System integrity is absolutely essential to achieving this kind of far-reaching vision. A systems engineering approach must be used to define requirements, reinvent processes, formulate operational concepts, evaluate them, and then launch goal-oriented technology activities to transform the concepts into realities.

1.6 SMART VEHICLE, HEAL THYSELF

Intense work is currently being done at NASA, the Department of Defense, and other organizations to raise the bar for component vehicle technologies and to develop a suite of smart technologies and tools which in combination can lead to revolutionary vehicle concepts and aerospace systems. A smart vehicle can assess a situation, determine if action needs to be taken, and, if so, take it.

''Smartness'' can be characterized by self-adaptability, self-sensing, memory, and decision-making. Smart vehicles can assess their own health and perform self-repair. They will also know how to fly to a safe haven under emergencies.

Smart vehicle technologies are a blend of smart materials and structures, innovative actuators and sensors, and intelligent flow-control strategies—including sonic boom mitigation technologies, revolutionary propulsion ideas, and biologically inspired concepts. The field of smart materials and structures, for example, evolved over the past decades and increased its pace in the 1990s. It has inspired numerous innovative concepts in the United States and abroad.

Major demonstration programs have addressed structural health monitoring, vibration suppression, shape control, and multifunctional structural concepts for spacecraft and launch vehicles, aircraft and rotorcraft. The demonstrations have focused on showing potential system-level performance improvements using smart technologies in realistic aerospace systems.

NEW VISION FOR FUTURE AEROSPACE VEHICLES AND SYSTEMS

Some recent work related to smart vehicle technology has focused on the development of composite systems with active constituents, of distributed actuation systems, and of fiberoptic and compact integrated sensor systems (see Figure 1.11). Current trends aim at the atomic and molecular level to synthesize new materials that are functionally smart. Examples include molecularly imprinted polymers and other materials that contain inherent receptors for information. Other efforts are integrating diverse sensors on a single substrate and working on practical techniques to fabricate them.

Self-healing material concepts have received increasing attention in recent years. For example, self-healing plastics use material that has the ability to heal cracks when fracture occurs. Shape memory alloys in composites can stop propagating cracks by imposing compressive forces resulting from stress-induced phase transformation. Current research aims at developing adaptive, self-repairing materials and structures that can arrest dynamic crack propagation, heal cracks, restore structural integrity and stiffness, and reconfigure themselves to serve more functions (see Figure 1.12).

Controlling fundamental mechanisms in fluids has long been the focus of intense effort. Recent applications of airflow sensing and intelligent control to air vehicles include improving performance by increasing lift or reducing drag generated by a surface and maneuvering through the use of fluidic devices. Current activities aim at understanding the physics associated with shock wave formation in high-speed flight and developing designer fluid mechanics tools for all types of flow control—flow separation control, vortex control, laminar flow control, turbulent drag reduction, anti-noise, mixing enhancements, combustion control, circulation control, and favorable wave interference.

FIGURE 1.11 The Langley Macro-Fiber Composite (LaRC-MFC) is a high performance, low-cost, flexible piezoceramic actuator developed at NASA-Langley Research Center. The solid-state LaRC-MFC is being used to control deflections and vibrations on advanced composite helicopter rotor blades, fixed-wing aircraft control surfaces, and ultralightweight spacecraft components.

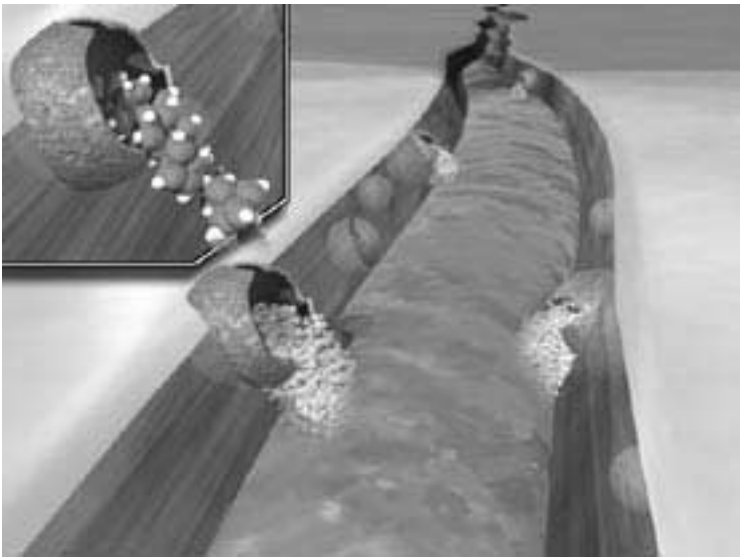


FIGURE 1.12 Micro-capsules releasing a healing agent into the crack through capillary action. The agent then interacts with the catalyst to form a solid polymer that fills the crack. (Concept developed by the researchers at the University of Illinois at Urbana-Champaign.)

A suite of high-payoff sonic boom mitigation technologies is being explored for reducing the sonic boom overpressure to an acceptable level to people on the ground (less than 0.3 psf). Techniques include airframe shaping, heat addition, particulate injection, leading-edge plasma generation, temporal and spatial variation of lift distribution, and adaptive flow control.

Indirect reduction of sonic boom amplitudes can also be achieved by decreasing vehicle gross weight or increasing vehicle lift-to-drag ratio by maintaining supersonic laminar flow. In addition, the use of intelligent propulsion control systems is being explored for efficient, reliable operation of the complex supersonic inlet/engine/nozzle system.

1.7 WORKING FOR MORE SECURE AIRSPACE

The tragic events of September 11, 2001, have had a tremendous impact on civil aviation. A number of technologies can be deployed, adapted, or developed to significantly enhance aviation security and restore public confidence in the system. These technologies can be grouped into three categories, ranging from direct aircraft protections to general security improvements.

Various overrides of the flight control system can prevent even a determined pilot from crashing an aircraft into a specific ground target. The technologies already in development for aviation safety purposes can be adopted for security applications as well.

Technology can be refined to create a ''refuse-to-crash aircraft.'' An automatic ground collision avoidance system, now an active project of the automobile industry, combined with aircraft control, can be applied to commercial airplanes for unusual attitude recovery and ground collision avoidance. The national and worldwide terrain/obstacle database, already under development by a number of organizations, can be extended to identify protective shells, or prohibited airspace, around selected areas such as specific high-risk facilities. The algorithms in the collision-avoidance system could interpret these protected areas as hard terrain and force the aircraft to avoid them.

Advanced autopilot, similar to the controls currently used in unmanned surveillance aircraft, can serve as a security backup. In the event of an emergency, or unauthorized deviation, manual flight controls could be disabled by a signal from the pilot or from ground controllers and allow automated safe flight to the nearest secure airport.

An aircraft damaged by a terrorist attack or system failure could, in many cases, be landed safely through the use of reconfigurable flight controls, a technology that has been demonstrated and is based on work by NASA, DOD, FAA, and the aerospace industry to compensate for damaged or failed systems.

Potentially threatening deviations from the flight paths can be automatically brought to the attention of ground controllers. Such anomaly detection requires the use of an advanced alerting system capable of assessing the probability that an aircraft is under malign control or the probability of a collision with some object or place. The system would build on existing efforts to develop airspace flow control automation, visualize air traffic patterns, and quantify aviation safety risk by monitoring and modeling air traffic patterns.

Linking passenger cabins and flight stations with ground communication networks into an integrated information environment can enable safety—or potential security—issues to be uncovered during a flight.

The use of biometric technologies (including fingerprint sensors, retinal scans, and facial recognition systems) for verification and identification can ensure that flight controls are used only by authorized pilots.

Enhanced airport and aircraft security can be achieved by using a multilayered suite of trace-detection and active-imaging technologies. Examples of instruments using these technologies are compact trace gas analyzers, biosensors, high-tech systems to detect molecular-level evidence of explosives and firearms, and high-speed, low-cost 3D imaging technologies integrated with better information technology databases and decision support systems, such as automated pattern-recognition devices.

These instruments could enable the implementation of a central security screening at a check-in location and a distributed, roving security detection system that encompasses the total airport environment and individual aircraft. Linkage of various distributed information databases could enable near-real-time identification of potential passenger threats.

SECTION 2

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

Section Editor: Mark Davies

PART 1

TRIGONOMETRIC FUNCTIONS AND GENERAL FORMULAE

John Barron

2.1 MATHEMATICAL SIGNS AND SYMBOLS

Sign, symbol	Quantity
$=$	equal to
$\neq$	not equal to
$\equiv$	identically equal to
Δ	corresponds to
$\approx$	approximately equal to
$\rightarrow$	approaches
$\simeq$	asymptotically equal to
$\sim$	proportional to
∞	infinity
$<$	smaller than
$>$	larger than
$\leq \leqslant \leqq$	smaller than or equal to
$\geq \geqslant \geqq$	larger than or equal to
$<<$	much smaller than
$>>$	much larger than
$+$	plus
$-$	minus
$\times$	multiplied by
$\frac{a}{b}$; a/b	a divided by b
$\lvert a \rvert$	magnitude of a
a^n	a raised to the power n
$a^{1/2}$; $\sqrt{a}$	square root of a
$a^{1/n}$; $\sqrt[n]{a}$	nth root of a
$\dot{a}$; $\langle a \rangle$	mean value of a
$p!$	factorial p, $1 \times 2 \times 3 \times \cdots \times p$
$\binom{n}{p}$	*binomial coefficient,* $\dfrac{n(n-1)\cdots(n-p+1)}{1 \times 2 \times 3 \times \cdots \times p}$
Σ	sum
Π	product
$f(x)$	function f of the variable x
$[f(x)]_a^b$	$f(b) - f(a)$
$\lim_{x \to a} f(x)$; $\lim_{x \to a} f(x)$	the limit to which $f(x)$ tends as x approaches a
Δx	delta x = finite increment of x
δx	delta x = variation of x
$\frac{df}{dx}$; df/dx; $f'(x)$	differential coefficient of $f(x)$ with respect to x

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

Sign, symbol	Quantity
$\frac{d^n f}{dx^n}$; $f^{(n)}(x)$	differential coefficient of order n of $f(x)$
$\frac{\partial f(x,y,...)}{\partial x}$; $\left(\frac{\partial f}{\partial x}\right)_{y...}$	partial differential coefficient of $f(x,y, \ldots)$ with respect to x, when $y \ldots$ are held constant
df	the total differential of f
$\int f(x)dx$	indefinite integral of $f(x)$ with respect to x
$\int_a^b f(x)dx$	definite integral of $f(x)$ from $x = a$ to $x = b$
e	base of natural logarithms
e^x, exp x	e raised to the power x
$\log_a x$	logarithm to the base a of x
lg x; log x; $\log_{10} x$	common (Briggsian) logarithm of x
lb x; $\log_2 x$	binary logarithm of x
sin x	sine of x
cos x	cosine of x
tan x; tg x	tangent of x
cot x; ctg x	cotangent of x
sec x	secant of x
cosec x	cosecant of x
arcsin x	arc sine of x
arcos x	arc cosine of x
arctan x, arctg x	arc tangent of x
arccot x, arcctg x	arc cotangent of x
arcsec x	arc secant of x
arcosec x	arc cosecant of x
sinh x	hyperbolic sine of x
cosh x	hyperbolic cosine of x
tanh x	hyperbolic tangent of x
coth x	hyperbolic cotangent of x
sech x	hyperbolic secant of x
cosech x	hyperbolic cosecant of x
arsinh x	inverse hyperbolic sine of x
arcosh x	inverse hyperbolic cosine of x
artanh x	inverse hyperbolic tangent of x
arcoth x	inverse hyperbolic cotangent of x
arsech x	inverse hyperbolic secant of x
arcosech x	inverse hyperbolic cosecant of x
i, j	imaginary unity, $i^2 = -1$
Re z	real part of z
Im z	imaginary part of z
$\lvert z \rvert$	modulus of z
arg z	argument of z
$\underline{z}^*$	conjugate of z, complex conjugate of z
$\underline{A}$, A', A^t	transpose of matrix A
A^*	complex conjugate matrix of matrix A
A^+	Hermitian conjugate matrix of matrix A
A, a	vector
$\lvert\mathbf{A}\rvert$, A	magnitude of vector
$\mathbf{A} \cdot \mathbf{B}$	scalar product

Sign, symbol	Quantity
$\mathbf{A} \times \mathbf{B}$, $\mathbf{A} \wedge \mathbf{B}$	vector product
∇	differential vector operator
$\nabla\varphi$, grad φ	gradient of φ
$\nabla \cdot A$, div $\mathbf{A}$	divergence of $\mathbf{V}$
$\nabla \times \mathbf{A}$, $\nabla \wedge \mathbf{A}$, curl $\mathbf{A}$, rot $\mathbf{A}$	curl of $\mathbf{A}$
$\nabla^2\varphi$, $\Delta\varphi$	Laplacian of φ

2.2 TRIGONOMETRIC FORMULAE

$$\sin^2 A + \cos^2 A = \sin A \operatorname{cosec} A = 1$$

$$\sin A = \frac{\cos A}{\cot A} = \frac{1}{\operatorname{sosec} A} = (1 - \cos^2 A)^{1/2}$$

$$\cos A = \frac{\sin A}{\tan A} = \frac{1}{\sec A} = (1 - \sin^2 A)^{1/2}$$

$$\tan A = \frac{\sin A}{\cos A} = \frac{1}{\cot A}$$

$$1 + \tan^2 A = \sec^2 A$$

$$1 + \cot^2 A = \operatorname{cosec}^2 A$$

$$1 - \sin A = \operatorname{coversin} A$$

$$1 - \cos A = \operatorname{versin} A$$

$$\tan \tfrac{1}{2}\theta = t; \quad \sin \theta = 2t/(1 + t^2); \quad \cos \theta = (1 - t^2)/(1 + t^2)$$

$$\cot A = 1/\tan A$$

$$\sec A = 1/\cos A$$

$$\operatorname{cosec} A = 1/\sin A$$

$$\cos (A \pm B) = \cos A \cos B \mp \sin A \sin B$$

$$\sin (A \pm B) = \sin A \cos B \pm \cos A \sin B$$

$$\tan (A \pm B) = \frac{\tan A \pm \tan B}{1 \mp \tan A \tan B}$$

$$\cot(A \pm B) = \frac{\cot A \cot B \mp 1}{\cot B \pm \cot A}$$

$$\sin A \pm \sin B = 2 \sin \tfrac{1}{2}(A \pm B)\cos \tfrac{1}{2}(A \mp B)$$

$$\cos A + \cos B = 2 \cos \tfrac{1}{2}(A + B) \cos \tfrac{1}{2}(A - B)$$

$$\cos A - \cos B = 2 \sin \tfrac{1}{2}(A + B) \sin \tfrac{1}{2}(B - A)$$

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

$$\tan A \pm \tan B = \frac{\sin (A \pm B)}{\cos A \cos B}$$

$$\cot A \pm \cot B = \frac{\sin (B \pm A)}{\sin A \sin B}$$

$$\sin 2A = 2 \sin A \cos A$$

$$\cos 2A = \cos^2 A - \sin^2 A = 2 \cos^2 A - 1 = 1 - 2 \sin^2 A$$

$$\cos^2 A - \sin^2 B = \cos (A + B) \cos (A - B)$$

$$\tan 2A = 2 \tan A/(1 - \tan^2 A)$$

$$\sin \tfrac{1}{2}A = \pm \left(\frac{1 - \cos A}{2}\right)^{1/2}$$

$$\cos \tfrac{1}{2}A = \pm \left(\frac{1 + \cos A}{2}\right)^{1/2}$$

$$\tan \tfrac{1}{2}A = \frac{\sin A}{1 + \cos A}$$

$$\sin^2 A = \tfrac{1}{2}(1 - \cos 2A)$$

$$\cos^2 A = \tfrac{1}{2}(1 + \cos 2A)$$

$$\tan^2 A = \frac{1 - \cos 2A}{1 + \cos 2A}$$

$$\tan \tfrac{1}{2}(A \pm B) = \frac{\sin A \pm \sin B}{\cos A + \cos B}$$

$$\cot \tfrac{1}{2}(A \pm B) = \frac{\sin A \pm \sin B}{\cos B - \cos A}$$

2.3 TRIGONOMETRIC VALUES

Angle	0°	30°	45°	60°	90°	180°	270°	360°
Radians	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$	π	$3\pi/2$	2π
Sine	0	$\frac{1}{2}$	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}\sqrt{3}$	1	0	-1	0
Cosine	1	$\frac{1}{2}\sqrt{3}$	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}$	0	-1	0	1
Tangent	0	$\frac{1}{3}\sqrt{3}$	1	$\sqrt{3}$	∞	0	∞	0

2.4 APPROXIMATIONS FOR SMALL ANGLES

$$\sin\theta = \theta - \theta^3/6; \quad \cos\theta = 1 - \theta^2/2; \quad \tan\theta = \theta + \theta^3/3; \ (\theta \text{ in radians})$$

2.5 SOLUTION OF TRIANGLES

$$\frac{\sin A}{a} = \frac{\sin B}{b} = \frac{\sin C}{c}; \quad \cos A = \frac{b^2 + c^2 - a^2}{2bc}$$

$$\cos B = \frac{c^2 + a^2 - b^2}{2ca}; \quad \cos C = \frac{a^2 + b^2 - c^2}{2ab}$$

where A, B, C and a, b, c are shown in Figure 2.1. If $s = \frac{1}{2}(a + b + c)$

$$\sin\frac{A}{2} = \sqrt{\frac{(s-b)(s-c)}{bc}}; \quad \sin\frac{B}{2} = \sqrt{\frac{(s-c)(s-a)}{ca}}$$

$$\sin\frac{C}{2} = \sqrt{\frac{(s-a)(s-b)}{ab}}$$

$$\cos\frac{A}{2} = \sqrt{\frac{s(s-a)}{bc}}; \quad \cos\frac{B}{2} = \sqrt{\frac{(s-b)}{ca}}$$

$$\cos\frac{C}{2} = \sqrt{\frac{s(s-c)}{ab}}$$

$$\tan\frac{A}{2} = \sqrt{\frac{(s-b)(s-c)}{s(s-a)}}; \quad \tan\frac{B}{2} = \sqrt{\frac{(s-c)(s)-a)}{s(s-b)}}$$

$$\tan\frac{C}{2} = \sqrt{\frac{(s-a)(s-b)}{s(s-\mathbf{c})}}$$

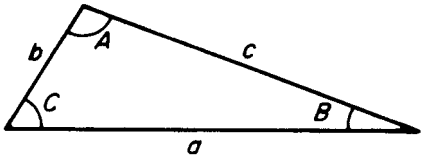


FIGURE 2.1 Triangle.

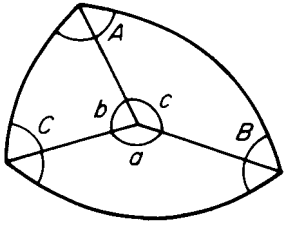


FIGURE 2.2 Spherical triangle.

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

2.6 SPHERICAL TRIANGLE

$$\frac{\sin A}{\sin a} = \frac{\sin B}{\sin b} = \frac{\sin C}{\sin c}$$

$$\cos a = \cos b \cos c + \sin b \sin c \cos A$$

$$\cos b = \cos c \cos a + \sin c \sin a \cos B$$

$$\cos c = \cos a \cos b + \sin a \sin b \cos C$$

where A, B, C and a, b, c are now as in Figure 2.2.

2.7 EXPONENTIAL FORM

$$\sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i} \quad \cos \theta = \frac{e^{i\theta} + e^{+i\theta}}{2}$$

$$c^{i\theta} = \cos \theta + i \sin \theta \; e^{-i\theta} = \cos \theta - i \sin \theta$$

2.8 DE MOIVRE'S THEOREM

$$(\cos A + i \sin A)(\cos B + i \sin B) = \cos(A + B) + i \sin (A + B)$$

2.9 EULER'S RELATION

$$(\cos \theta + i \sin \theta)^n = \cos n\theta + i \sin n\theta = e^{in\theta}$$

2.10 HYPERBOLIC FUNCTIONS

$$\sinh x = (e^x - e^{-x})/2 \quad \cosh x = (e^x + e^{-x})/2$$

$$\tanh x = \sinh x \cosh x$$

Relations between hyperbolic functions can be obtained from the corresponding relations between trigonometric functions by reversing the sign of any term containing the product or implied product of two sines, e.g.:

$$\cosh^2 A - \sinh^2 A = 1$$

$$\cosh 2A = 2 \cosh^2 A - 1 = 1 + 2 \sinh^2 A$$

$$= \cosh^2 A + \sinh^2 A$$

$$\cosh(A \pm B) = \cosh A \cosh B \pm \sinh A \sinh B$$

$$\sinh (A \pm B) = \sinh A \cosh B \pm \cosh A \sinh B$$

$$e^x = \cosh x + \sinh x \quad e^{-x} = \cosh x - \sinh x$$

2.11 COMPLEX VARIABLE

If $z = x + iy$, where x and y are real variables, z is a complex variable and is a function of x and y. z may be represented graphically in an Argand diagram (Figure 2.3).

Polar form:

$$z = x + iy = |z|(\cos \theta + i \sin \theta)$$

$$x = r \cos \theta \qquad y = r \sin \theta$$

where $r = |z|$.

Complex arithmetic:

$$z_1 = x_1 + iy_1; \qquad z_2 = x_2 + iy_2$$

$$z_1 \pm z_2 = (x_1 \pm x_2) + i(y_1 \pm y_2)$$

$$z_1 \cdot z_2 = (x_1 x_2 - y_1 y_2) + i(x_1, y_2 + x_2, y_1)$$

Conjugate:

$$z^* = x - iy \qquad z \cdot z^* = x^2 + y^2 = |z|^2$$

Function: another complex variable $w = u + iv$ may be related functionally to z by

$$w = u + iv = f(x + iy) = f(z)$$

which implies

$$u = u(x,y) \qquad v = v(x,y)$$

e.g.,

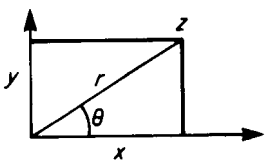


FIGURE 2.3 Argand diagram.

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

$$\cosh z = \cosh (x + iy) = \cosh x \cosh iy + \sinh x \sinh iy$$
$$= \cosh x \cos y + i \sinh x \sin y$$
$$u = \cosh x \cos y \qquad v = \sinh x \sin y$$

2.12 CAUCHY–RIEMANN EQUATIONS

If $u(x,y)$ and $v(x,y)$ are continuously differentiable with respect to x and y,

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

$w = f(z)$ is continuously differentiable with respect to z and its derivative is

$$f'(z) = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i \frac{\partial u}{\partial y} = \frac{1}{i} \left(\frac{\partial u}{\partial y} + i \frac{\partial v}{\partial y} \right)$$

It is also easy to show that $\nabla^2 u = \nabla^2 v = 0$. Since the transformation from z to w is conformal, the curves $u =$ constant and $v =$ constant intersect each other at right angles, so that one set may be used as equipotentials and the other as field lines in a vector field.

2.13 CAUCHY'S THEOREM

If $f(z)$ is analytic everywhere inside a region bounded by C and a is a point within C,

$$f(a) = \frac{1}{2\pi i} \int_C \frac{f(z)}{z - a} dz$$

This formula gives the value of a function at a point in the interior of a closed curve in terms of the values on that curve.

2.14 ZEROES, POLES, AND RESIDUES

If $f(z)$ vanishes at the point z_0 the Taylor series for z in the region of z_0 has its first two terms zero, and perhaps others also: $f(z)$ may then be written

$$f(z) = (z - z_0)^n g(z)$$

where $g(z_0) \neq 0$. Then $f(z)$ has a *zero* of order n at z_0. The reciprocal

$$q(z) = 1/f(z) = h(z)/(z - z_0)^n$$

where $h(z) = 1/g(z) \neq 0$ at z_0. $q(z)$ becomes infinite at $z = z_0$ and is said to have a *pole* of order n at z_0. $q(z)$ may be expanded in the form

$$q(z) = c_{-n}(z - z_0)^n + \cdots + c_{-1}(z - z_0)^{-1} + c_0 + \cdots$$

where c_{-1} is the *residue* of $q(z)$ at $z = z_0$. From Cauchy's theorem, it may be shown that if a function $f(z)$ is analytic throughout a region enclosed by a curve C except at a finite number of poles, the integral of the function around C has a value of $2\pi i$ times the sum of the residues of the function at its poles within C. This fact can be used to evaluate many definite integrals whose indefinite form cannot be found.

2.15 SOME STANDARD FORMS

$$\int_0^{2\pi} e^{\cos\theta} \cos(n\theta - \sin\theta)d\theta = 2\pi/n!$$

$$\int_0^{\pi} \frac{x^{a-1}}{1+x}\,dx = \pi \operatorname{cosec} a\pi$$

$$\int_0^{\pi} \frac{\sin\theta}{\theta}\,d\theta = \frac{\pi}{2}$$

$$\int_0^{\pi} x \exp(-h^2x^2)dx = \frac{1}{2h^2}$$

$$\int_0^{\pi} \frac{x^{a-1}}{1-x}\,dx = \pi \cot a\pi$$

$$\int_0^{\pi} \exp(-h^2x^2)dx = \frac{\sqrt{\pi}}{2h}$$

$$\int_0^{\pi} x^2 \exp(-h^2x^2)dx = \frac{\sqrt{\pi}}{4h^3}$$

2.16 COORDINATE SYSTEMS

The basic system is the rectangular Cartesian system (x,y,z), to which all other systems are referred. Two other commonly used systems are as follows.

Cylindrical Coordinates

Coordinates of point P are (x,y,z) or (r,θ,z) (see Figure 2.4), where

$$x = r\cos\theta \qquad y = r\sin\theta \qquad z = z$$

In these coordinates the volume element is $r\,dr\,d\theta\,dz$.

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

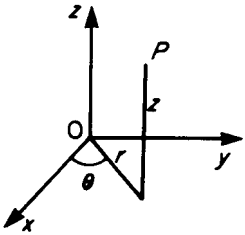


FIGURE 2.4 Cylindrical coordinates.

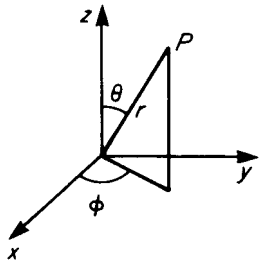


FIGURE 2.5 Spherical polar coordinates.

Spherical Polar Coordinates

Coordinates of point P are (x,y,z) or (r,θ,φ) (see Figure 2.5), where

$$x = r \sin \theta \cos \phi \qquad y = r \sin \phi \sin \phi \qquad z = r \cos \theta$$

In these coordinates the volume element is $r^2 \sin \theta \, dr \, d\theta \, d\phi$.

2.17 TRANSFORMATION OF INTEGRALS

$$\int\int\int f(x,y,z)dx\ dy\ dz = \int\int\int \varphi(u,v,w)J\ du\ dv\ dw$$

where

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} & \dfrac{\partial z}{\partial u} \\ \dfrac{\partial x}{\partial v} & \dfrac{\partial y}{\partial v} & \dfrac{\partial z}{\partial v} \\ \dfrac{\partial x}{\partial w} & \dfrac{\partial y}{\partial w} & \dfrac{\partial z}{\partial w} \end{vmatrix} = \frac{\partial(x,y,z)}{\partial(u,v,w}$$

is the Jacobian of the transformation of coordinates. For Cartesian to cylindrical coordinates, $J = r$, and for Cartesian to spherical polars, it is $r^2 \sin \theta$.

2.18 LAPLACE'S EQUATION

The equation satisfied by the scalar potential, from which a vector field may be derived by taking the gradient is Laplace's equation, written as:

$$\nabla^2\Phi = \frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2} = 0$$

In cylindrical coordinates:

$$\nabla^2\Phi = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\Phi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\Phi}{\partial\theta^2} + \frac{\partial^2\Phi}{\partial z^2}$$

In spherical polars:

$$\nabla^2\Phi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\Phi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial\Phi}{\partial\theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\Phi}{\partial\Phi^2}$$

The equation is often solved by setting

$$\Phi = U(u)V(u)W(w)$$

in the appropriate form of the equation, separating the variables and solving separately for the three functions, where (u,v,w) is the coordinate system in use.

In Cartesian coordinates, typically the functions are trigonometric, hyperbolic, and exponential; in cylindrical coordinates the function of z is exponential, that of θ trigonometric, and that of r a Bessel function. In spherical polars, typically the function of r is a power of r, that of φ is trigonometric, and that of θ is a Legendre function of $\cos\theta$.

2.19 SOLUTION OF EQUATIONS

Quadratic Equation

$$ax^2 - bx + c = 0$$

$$x = -\frac{b}{2a} \pm \frac{\sqrt{b^2 - 4ac}}{2a}$$

In practical calculations if $b^2 > 4ac$, so that the roots are real and unequal, calculate the root of larger modulus first, using the same sign for both terms in the formula, then use the fact that $x_1x_2 = c/a$ where x_1 and x_2 are the roots. This avoids the severe cancellation of significant digits which may otherwise occur in calculating the smaller root.

For polynomials other than quadratics, and for other functions, several methods of successive approximation are available.

Bisection Method

By trial find x_0 and x_1 such that $f(x_0)$ and $f(x_1)$ have opposite signs (see Figure 2.6). Set $x_2 = (x_0 + x_1)/2$ and calculate $f(x_2)$. If $f(x_0)f(x_2)$ is positive, the root lies in the interval (x_1,x_2); if negative in the interval (x_0,x_2); and if zero, x_2 is the root. Continue if necessary using the new interval.

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

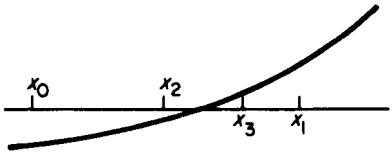


FIGURE 2.6 Bisection method.

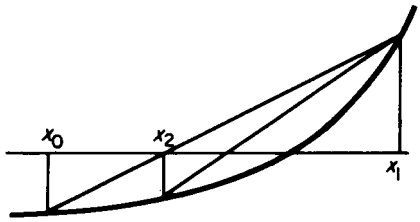


FIGURE 2.7 Regula falsi.

Regula Falsi

By trial, find x_0 and x_1 as for the bisection method; these two values define two points $(x_0, f(x_0))$ and $(x_1, f(x_1))$. The straight line joining these two points cuts the x-axis at the point (see Figure 2.7).

$$x_2 = \frac{x_0 f(x_1) - x_1 f(x_0)}{f(x_1) - f(x_0)}$$

Evaluate $f(x_2)$ and repeat the process for whichever of the intervals (x_0, x_2) or (x_1, x_2) contains the root. This method can be accelerated by halving at each step the function value at the retained end of the interval, as shown in Figure 2.8.

Fixed-Point Iteration

Arrange the equation in the form

$$x = f(x)$$

Choose an initial value of x by trial, and calculate repetitively

$$x_{k+1} = f(x_k)$$

This process will not always converge.

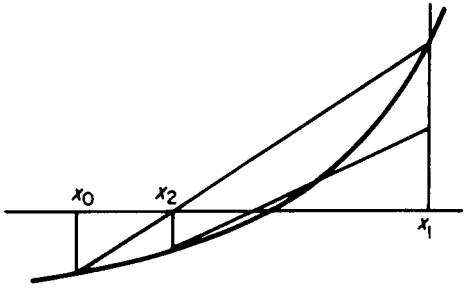


FIGURE 2.8 Accelerated method.

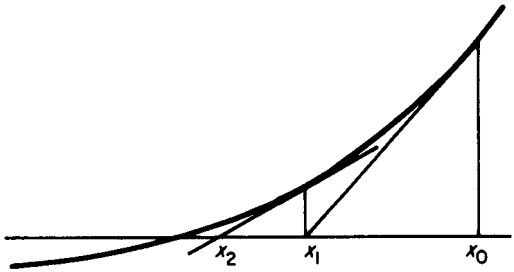


FIGURE 2.9 Newton's method.

Newton's Method

Calculate repetitively (Figure 2.9).

$$x_{k+1} = x_k - f(x_k)/f'(x_k)$$

This method will converge unless: (a) x_k is near a point of inflexion of this function; or (b) x_k is near a local minimum; or (c) the root is multiple. If one of these cases arises, most of the trouble can be overcome by checking at each stage that

$$f(x_{k+1}) < f(x_k)$$

and, if not, halving the preceding value of $|x_{k+1} - x_k|$.

2.20 METHOD OF LEAST SQUARES

To obtain the best fit between a straight line $ax + by = 1$ and several points (x_1,y_1), (x_2,y_2), . . . , (x_n,y_n) found by observation, the coefficients a and b are to be chosen so that the sum of the squares of the errors

$$e_i = ax_i + by_i - 1$$

is a minimum. To do this, first write the set of inconsistent equations

$$ax_1 + by_1 - 1 = 0$$

$$ax_2 + by_2 - 1 = 0$$

$$\vdots$$

$$ax_n + by_n - 1 = 0$$

Multiply each equation by the value of x it contains, and add, obtaining

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

$$a \sum_{i=1}^{n} x_i^2 + b \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i = 0$$

Similarly, multiply by y and add, obtaining

$$a \sum_{i=1}^{n} x_i y_i + b \sum_{i=1}^{n} y_i^2 - \sum_{i=1}^{n} y_i = 0$$

Lastly, solve these two equations for a and b, which will be the required values giving the least squares fit.

PART 2

CALCULUS

John Barron

2.21 DERIVATIVE

$$f'(x) = \lim_{\delta x \to 0} \frac{f(x + \delta x) - f(x)}{\delta x}$$

If u and v are functions of x,

$$(uv)' = u'v + uv'$$

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}$$

$$(uv)^n = u^{(n)}v + nu^{(n-1)}v^{(1)} + \cdots + {}^nC_p u^{(n-p)}v^{(p)} + \cdots + uv^{(n)}$$

where

$${}^nC_p = \frac{n!}{p!(n-p)!}$$

If $z = f(x)$ and $y = g(z)$, then

$$\frac{dy}{dx} = \frac{dy}{dz}\frac{dz}{dx}$$

2.22 MAXIMA AND MINIMA

$f(x)$ has a stationary point wherever $f'(x) = 0$: the point is a maximum, minimum, or point of inflexion according as $f''(x) <, >$ or $= 0$.

$f(x,y)$ has a stationary point wherever

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$$

Let (a,b) be such a point, and let

$$\frac{\partial^2 f}{\partial x^2} = A, \qquad \frac{\partial^2 f}{\partial x \partial y} = H, \qquad \frac{\partial^2 f}{\partial y^2} = B$$

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

all at that point, then:

If $H^2 - AB > 0$, $f(x,y)$ has a saddle point at (a,b).

If $H^2 - AB < 0$ and if $A < 0$, $f(x,y)$ has a maximum at (a,b),
but if $A > 0$, $f(x,y)$ has a minimum at (a,b)

If $H^2 = AB$, higher derivatives need to be considered.

2.23 INTEGRAL

$$\int_b^a f(x)dx = \lim_{N\to\infty} \sum_{n=0}^{N-1} f\left(a + \frac{n(b-a)}{N}\right)\left(\frac{b-a}{N}\right)$$

$$= \lim_{N\to\infty} \sum_{n=1}^{N} f(a + (n-1)\,\delta x)\delta x$$

where $\delta x = (b - a)/N$.

If u and v are functions of x, then

$$\int uv'dx = uv - \int u'vdx \quad \text{(integration by parts)}$$

y	$\int ydx$	
$\sin mx \sin nx$	$\frac{1}{2}\frac{\sin(m-n)x}{m-n} - \frac{1}{2}\frac{\sin(m+n)x}{m+n}$	$(m \neq n)$
	$\frac{1}{2}\left(x - \frac{\sin 2mx}{2m}\right)$	$(m = n)$
$\sin mx \cos nx$	$-\frac{1}{2}\frac{\cos(m+n)x}{m+n} - \frac{1}{2}\frac{\cos(m-n)x}{m-n}$	$(m \neq n)$
	$-\frac{1}{2}\frac{\cos 2mx}{2m}$	$(m = n)$
$\cos mx \cos nx$	$\frac{1}{2}\frac{\sin(m+n)x}{m+n} + \frac{1}{2}\frac{\sin(m-n)x}{m-n}$	$(m \neq n)$
	$\frac{1}{2}\left(x + \frac{\sin 2mx}{2m}\right)$	$(m = n)$

2.24 DERIVATIVES AND INTEGRALS

y	$\frac{dy}{dx}$	$\int y\,dx$
x^n	nx^{n-1}	$x^{n+1}/(n+1)$
$1/x$	$-1/x^2$	$\ln(x)$
e^{ax}	ae^{ax}	e^{ax}/a
$\ln(x)$	$1/x$	$x[\ln(x)-1]$
$\log_a x$	$\frac{1}{x}\log_a e$	$x\log_a\left(\frac{x}{e}\right)$
$\sin ax$	$a\cos ax$	$-\frac{1}{a}\ln(\cos ax)$
$\cos ax$	$-a\sin ax$	$-\frac{1}{a}\sin ax$
$\tan ax$	$a\sec^2 ax$	$-\frac{1}{a}\cos ax$
$\cot ax$	$-a\,\mathrm{cosec}^2 ax$	$\frac{1}{a}\ln(\sin ax)$
$\sec ax$	$a\tan ax\sec ax$	$\frac{1}{a}\ln(\sec ax+\tan ax)$
$\mathrm{cosec}\,ax$	$-a\cot ax\,\mathrm{cosec}\,ax$	$\frac{1}{a}\ln(\mathrm{cosec}\,ax-\cot ax)$
$\arcsin(x/a)$	$1/(a^2-x^2)^{1/2}$	$x\arcsin(x/a)+(a^2-x^2)^{1/2}$
$\arccos(x/a)$	$-1/(a^2-x^2)^{1/2}$	$x\arccos(x/a)-(a^2-x^2)^{1/2}$
$\arctan(x/a)$	$a/(a^2+x^2)$	$x\arctan(x/a)-\frac{1}{2}a\ln(a^2+x^2)$
$\mathrm{arccot}(x/a)$	$-a/(a^2+x^2)$	$x\,\mathrm{arccot}(x/a)+\frac{1}{2}a\ln(a^2+x^2)$
$\mathrm{arcsec}(x/a)$	$a(x^2-a^2)^{-1/2}/x$	$x\,\mathrm{arcsec}(x/a)-a\ln[x+(x^2-a^2)^{1/2}]$
$\mathrm{arccosec}(x/a)$	$-a(x^2-a^2)^{-1/2}/x$	$x\,\mathrm{arccosec}(x/a)+a\ln[x+(x^2-a^2)^{1/2}]$
$\sinh ax$	$a\cosh ax$	$\frac{1}{a}\cosh ax$
$\cosh ax$	$a\sinh ax$	$\frac{1}{a}\sinh ax$
$\tanh ax$	$a\,\mathrm{sech}^2 ax$	$\frac{1}{a}\ln\cosh ax$
$\coth ax$	$-a\,\mathrm{cosech}^2 ax$	$\frac{1}{a}\ln(\sinh ax)$
$\mathrm{sech}\,ax$	$-a\tanh ax\,\mathrm{sech}\,ax$	$\frac{2}{a}\arctan(e^{ax})$
$\mathrm{cosech}\,ax$	$-a\coth ax\,\mathrm{cosech}\,ax$	$\frac{1}{a}\ln\left(\tanh\frac{ax}{2}\right)$
$\mathrm{arsinh}(x/a)$	$(x^2+a^2)^{-1/2}$	$x\,\mathrm{arsinh}(x/a)-(x^2+a^2)^{1/2}$
$\mathrm{arcosh}(x/a)$	$(x^2-a^2)^{-1/2}$	$x\,\mathrm{arcosh}(x/a)-(x^2-a^2)^{1/2}$
$\mathrm{artanh}(x/a)$	$a(a^2+x^2)^{-1}$	$x\,\mathrm{artanh}(x/a)-\frac{1}{2}a\ln(a^2+x^2)$
$\mathrm{arcoth}(x/a)$	$-a(x^2-a^2)^{-1}$	$x\,\mathrm{arcoth}(x/a)+\frac{1}{2}a\ln(x^2-a^2)$
$\mathrm{arsech}(x/a)$	$-a(a^2-x^2)^{-1/2}/x$	$x\,\mathrm{arsech}(x/a)+a\arcsin(x/a)$
$\mathrm{arcosech}(x/a)$	$-a(x^2+a^2)^{-1/2}/x$	$x\,\mathrm{arcosech}(x/a)+a\,\mathrm{arsinh}(x/a)$
$(x^2+a^2)^{1/2}$		$\frac{1}{2}x(x^2+a^2)^{1/2}+\frac{1}{2}a^2\,\mathrm{arsinh}(x/a)$
$(x^2-a^2)^{1/2}$		$\frac{1}{2}x(x^2-a^2)^{1/2}-\frac{1}{2}a^2\,\mathrm{arcosh}(x/a)$

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

y	$\dfrac{dy}{dx}$	$\int y\,dx$
$(a^2 - x^2)^{1/2}$		$\begin{cases} \frac{1}{2}(x^2 \pm a^2)^{p+1}/(p+1) & (p \neq -1) \\ \frac{1}{2}\ln(x^2 \pm a^2) & (p = -1) \end{cases}$
$(x^2 \pm a^2)^p x$ $(a^2 - x^2)^p x$		$\begin{cases} -\frac{1}{2}(a^2 - x^2)^{p+1}/(p+1) & (p \neq -1) \\ -\frac{1}{2}\ln(a^2 - x^2) & (p = -1) \end{cases}$
$x(ax^2 + b)^p$		$(ax^2 + b)^{p+1}/2a(p+1 \quad (p \neq -1)$ $[\ln(ax^2 + b)]/2a \quad (p = -1)$
$(2ax - x^2)^{-1/2}$		$\arccos\left(\dfrac{a - x}{a}\right)$
$(a^2 \sin^2 x + b^2 \cos^2 x)^{-1}$		$\dfrac{1}{ab}\arctan\left(\dfrac{a}{b}\tan x\right)$
$(a^2 \sin^2 x - b^2 \cos^2 x)^{-1}$		$-\dfrac{1}{ab}\operatorname{artanh}\left(\dfrac{a}{b}\tan x\right)$
$e^{ax} \sin bx$		$e^{ax}\dfrac{a \sin bx - b \cos bx}{a^2 + b^2}$
$e^{ax} \cos bx$		$e^{ax}\dfrac{(a \cos bx + b \sin bx)}{a^2 + b^2}$

2.25 STANDARD SUBSTITUTIONS

Integral a function of	Substitute
$a^2 - x^2$	$x = a \sin\theta$ or $x = a \cos\theta$
$a^2 + x^2$	$x = a \tan\theta$ or $x = a \sinh\theta$
$x^2 - a^2$	$x = a \sec\theta$ or $x = a \cosh\theta$

2.26 REDUCTION FORMULAE

$$\int \sin^m x\,dx = -\frac{1}{m}\sin^{m-1} x \cos x + \frac{m-1}{m}\int \sin^{m-2} x\,dx$$

$$\int \cos^m x\,dx = -\frac{1}{m}\cos^{m-1} x \sin x + \frac{m-1}{m}\int \cos^{m-2} x\,dx$$

$$\int \sin^m x \cos^n x\,dx = \frac{\sin^{m+1} x \cos^{n-1} x}{m+n} + \frac{n-1}{m+n}\int \sin^m x \cos^{n-2} x\,dx$$

If the integrand is a rational function of $\sin x$ and/or $\cos x$, substitute $t = \tan \frac{1}{2}x$, then

$$\sin x = \frac{1}{1+t^2}, \qquad \cos x = \frac{1-t^2}{1+t^2}, \qquad dx = \frac{2dt}{1+t^2}$$

2.27 NUMERICAL INTEGRATION

Trapezoid Rule (Figure 2.10)

$$\int_{x_1}^{x_2} y\,dx = \tfrac{1}{2}h(y_1 + y_2) + O(h^3)$$

Simpson's Rule (Figure 2.10)

$$\int_{x_1}^{x_2} y\,dx = 2h(y_1 + 4y_2 + y_2)/6 + O(h^5)$$

Change of Variable in Double Integral

$$\iint f(x,y)dx\,dy = \iint F(u,v)|J|du\,dv$$

where

$$J = \frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} \\ \dfrac{\partial x}{\partial v} & \dfrac{\partial y}{\partial v} \end{vmatrix}$$

is the Jacobian of the transformation.

Differential Mean Value Theorem

$$\frac{f(x + h) - f(x)}{h} = f'(x + \theta h)\ 0 < \theta < 1$$

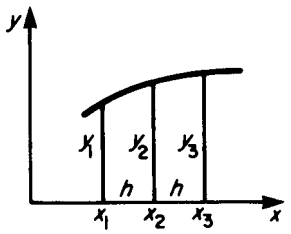


FIGURE 2.10 Numerical integration.

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

Integral Mean Value Theorem

$$\int_a^b f(x)g(x)dx = g(a + \theta h)\int_a^b f(x)dx \qquad h = b - a, \quad 0 < \theta < 1$$

2.28 VECTOR CALCULUS

Let $s(x,y,z)$ be a scalar function of position and let $\mathbf{v}(x,y,z) = \mathbf{i}v_x(x,y,z) + \mathbf{j}v_y(x,y,z) + \mathbf{k}v_2(x,y,z)$ be a vector function of position. Define

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}$$

so that

$$\nabla \cdot \nabla = \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

then

$$\text{grad } s = \nabla s = \mathbf{i}\frac{\partial s}{\partial x} + \mathbf{j}\frac{\partial s}{\partial y} + \mathbf{k}\frac{\partial s}{\partial z}$$

$$\text{div } \mathbf{v} = \nabla \cdot \mathbf{v} = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}$$

$$\text{curl } \mathbf{v} = \nabla \times \mathbf{v} = \mathbf{i}\left(\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z}\right) + \mathbf{j}\left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x}\right) + \mathbf{k}\left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y}\right)$$

The following identities are then true:

$$\text{div}^*(s\mathbf{v}) = s \text{ div } \mathbf{v} + (\text{grad } s) \cdot \mathbf{v}$$

$$\text{curl}(s\mathbf{v}) = s \text{ curl } \mathbf{v} + (\text{grad } s) \times \mathbf{v}$$

$$\text{div}(\mathbf{u} \times \mathbf{v}) = \mathbf{v} \cdot \text{cur } \mathbf{u} - \mathbf{u} \cdot \text{curl } \mathbf{v}$$

$$\text{curl}(\mathbf{u} \times \mathbf{v}) = \mathbf{u} \text{ div } \mathbf{v} - \mathbf{v} \text{ div } \mathbf{u} + (\mathbf{v} \cdot \nabla)\mathbf{u} - (\mathbf{u} \cdot \nabla)\mathbf{v}$$

$$\text{div grad } s = \nabla^2 s$$

$$\text{div curl } \mathbf{v} = 0$$

$$\text{curl grad } s = 0$$

$$\text{curl curl } \mathbf{v} = \text{grad}(\text{div } \mathbf{v}) - \nabla^2 \mathbf{v}$$

where ∇^2 operates on each component of $\mathbf{v}$.

$$\mathbf{v} \times \text{curl } \mathbf{v} + (\mathbf{v} \cdot \nabla)\mathbf{v} = \text{grad } \tfrac{1}{2} \mathbf{v}^2$$

potentials:

If curl $\mathbf{v} = 0$, $\mathbf{v} = \text{grad } \phi$ where ϕ is a scalar potential.

If div $\mathbf{v} = 0$, $\mathbf{v} = \text{curl } \mathbf{A}$ where $\mathbf{A}$ is a vector potential.

PART 3

SERIES AND TRANSFORMS

John Barron

2.29 ARITHMETIC SERIES

Sum of n terms,

$$\begin{aligned} S_n &= a + (a + d) + (a + 2d) + \cdots + [a + (n - 1)d] \\ &= n[2a + (n - 1)d]/2 \\ &= n(a + I)/2 \end{aligned}$$

2.30 GEOMETRIC SERIES

Sum of n terms,

$$S_n = a + ar + ar^2 + \cdots + ar^{n-1} = a(1 - r^n)/(1 - r)(|r| < 1)$$

$$S_\infty = a/(1 - r)$$

2.31 BINOMIAL SERIES

$$(1 + x)^p = 1 + px + \frac{p(p - 1)}{2!}x^2 + \frac{p(p - 1)(p - 2)}{3!}x^3 + \cdots$$

If p is a positive integer the series terminates with the term in x^p and is valid for all x; otherwise the series does not terminate, and is valid only for $-1 < x < 1$.

2.32 TAYLOR'S SERIES

Infinite form

$$f(x + h) = f(x) + hf'(x) + \frac{h^2}{2!}f''(x) + \cdots + \frac{h^n}{n!}f^{(n)}(x) + \cdots$$

Finite form

$$f(x + h) = f(x) + hf'(x) + \frac{h^2}{2!}f''(x) + \cdots + \frac{h^n}{n!}f^{(n)}(x) + \frac{h^{n+1}}{(n + 1)!}f^{(n+1)}(x + \lambda h)$$

where $0 \leqslant \lambda \leqslant 1$.

2.33 MACLAURIN'S SERIES

$$f(x) = f(0) + xf'(0) + \frac{x^2}{2!}f''(0) + \cdots + \frac{x^n}{n!}f^{(n)}(0) + \cdots$$

Neither of these series is necessarily convergent, but both usually are for appropriate ranges of values of h and x respectively.

2.34 LAURENT'S SERIES

If a function $f(z)$ of a complex variable is analytic on and everywhere between two concentric circles center a, then at any point in this region

$$f(z) = a_0 + a_1(z - a) + \cdots + b_1/(z - a) + b_2/(z - a)^2 + \cdots$$

This series is often applicable when Taylor's series is not.

2.35 POWER SERIES FOR REAL VARIABLES

	Math	*Comp*
$e^x = 1 + x + \frac{x^2}{2!} + \cdots$	all x	$\lvert x\rvert \leq 1$
$\ln(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots$		$-1 < x \leq 1$
$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$	all x	$\lvert x\rvert \leq 1$
$\cos x = 1 - \frac{x^2}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$	all x	$\lvert x\rvert \leq 1$
$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$	all x	$\lvert x\rvert \leq 1$
$\tan x = x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + \cdots$		$\lvert x\rvert < \frac{\pi}{2}$
$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots$		$\lvert x\rvert \leq 1$
$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \cdots$	all x	$\lvert x\rvert \leq 1$
$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \cdots$	all x	$\lvert x\rvert \leq 1$

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

The column headed "Math" contains the range of values of the variable x for which the series is convergent in the pure mathematical sense. In some cases a different range of values is given in the column headed "Comp," to reduce the rounding errors which arise when computers are used.

2.36 INTEGER SERIES

$$\sum_{n=1}^{N} n = 1 + 2 + 3 + 4 + \cdots + N = N(N+1)/2$$

$$\sum_{n=1}^{N} n^2 = 1^2 + 2^2 + 3^2 + 4^2 + \cdots + N^2 = N(N+1)(2N+1)/6$$

$$\sum_{n=1}^{N} n^3 = 1^3 + 2^3\ 3^3\ 4^3 + \cdots + N^3 = N^2(N+1)^2/4$$

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = \ln(2) \text{ (see } \ln(1+x))$$

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{2n-1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \frac{\pi}{4} \text{ (see } \arctan x)$$

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \cdots \frac{\pi^2}{6}$$

$$\sum_{n=1}^{N} n(n+1)(n+2)\cdots(n+r)$$

$$= 1 \cdot 2 \cdot 3 \cdots + 2 \cdot 3 \cdot 4 \cdots + 3 \cdot 4 \cdot 5 \cdots + \cdots$$

$$+ N(N+1)(N+2)\cdots(N+r)$$

$$= \frac{N(N+1)(N+2)\cdots(N+r+1)}{r+2}$$

2.37 FOURIER SERIES

$$f(\theta) = \frac{1}{2} a_0 + \sum_{n=1}^{\infty} (a_n \cos n\theta + b_n \sin n\theta)$$

with

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(\Theta) \cos n\Theta d\Theta$$

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(\Theta) \sin n\Theta d\Theta$$

or

$$f(\theta) = \sum_{n\pi-\infty}^{\infty} c_n \exp(jn\theta)$$

with

$$c_n = \frac{1}{2\pi}\int_0^{2\pi} f(\Theta)\exp(-jn\Theta)d\Theta = \begin{cases} \frac{1}{2}(a_n + jb_n) & n < 0 \\ \frac{1}{2}(a_n - jb_n) & n > 0 \end{cases}$$

The above expressions for Fourier series are valid for functions having at most a finite number of discontinuities within the period 0 to 2 of the variable of integration.

2.38 RECTIFIED SINE WAVE

$$f(\omega t) = \frac{1}{\pi} + \frac{1}{2}\cos\,\omega t + \frac{2}{\pi}\sum_{n=1}^{\infty}(-1)^{n+1}\frac{\cos\,2n\omega t}{4n^2 - 1}$$

$$f(\omega t) = \frac{\sin(\pi/p)}{\pi/p} + \frac{2p}{\pi}\sin\left(\frac{\pi}{p}\right)\sum_{n=1}^{\infty}(-1)^{n+1}\frac{\cos\,np\omega t}{p^2n^2 - 1}$$

See Figures 2.11 and 2.12.

2.39 SQUARE WAVE

$$f(\omega t) = \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\sin(2n - 1)\omega t}{(2n - 1)}$$

See Figure 2.13.

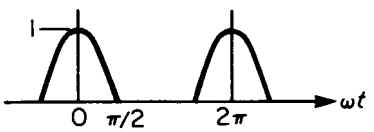


FIGURE 2.11 Half wave.

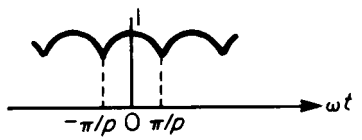


FIGURE 2.12 *p*-phase.

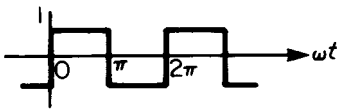


FIGURE 2.13 Square wave.

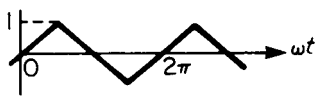


FIGURE 2.14 Triangular wave.

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

2.40 TRIANGULAR WAVE

$$f(\omega t) = \frac{8}{\pi^2} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{\sin(2n-1)\omega t}{(2n-1)^2}$$

See Figure 2.14.

2.41 SAWTOOTH WAVE

$$f(\omega t) = \frac{2}{\pi} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{\sin n\omega t}{n}$$

See Figure 2.15.

2.42 PULSE WAVE

$$f(t) = \frac{\tau}{T} + \frac{2\tau}{T} \sum_{n=1}^{\infty} \frac{\sin(n\omega\tau/T)}{n\pi\tau/T} \cos\left(\frac{2n\pi t}{T}\right)$$

See Figure 2.16.

2.43 FOURIER TRANSFORMS

Among other applications, these are used for converting from the time domain to the frequency domain.

Basic formulae:

$$\int_{-\infty}^{\infty} U(f)\exp(j2\pi ft)df = u(t) \leftrightarrows U(f) = \int_{\infty}^{\infty} u(t)\exp(-j2\pi ft)dt$$

Change of sign and complex conjugates:

$$u(-t) \leftrightarrows (U(-f),\ u^*(t) \leftrightarrows U^*(-f)$$

Time and frequency shifts (τ and ϕ constant):

$$u(t-\tau) \leftrightarrows U(f)\exp(-j2\pi f\tau)\exp(j2\pi\phi t)u(t) \leftrightarrows U(f-\phi)$$

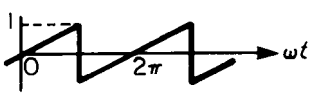


FIGURE 2.15 Sawtooth wave.

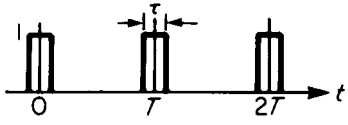


FIGURE 2.16 Pulse wave.

Scaling (T constant):

$$u(t/T) \leftrightarrows TU(fT)$$

Products and convolutions:

$$u(t)*v(t) \leftrightarrows U(f)V(f), \quad u(t)v(t) \leftrightarrows U(f)*V(f)$$

Differentiation:

$$u'(t) \leftrightarrows j2\pi f\,U(f), \quad -j2\pi tu(t) \leftrightarrows U'(f)$$

$$\partial u(t,\alpha)/\partial\alpha \leftrightarrows \partial(U - f,\alpha)/\partial\alpha$$

Integration ($U(0) = 0$, a and b real constants):

$$\int_{-\infty}^{t} u(\tau)d\tau \leftrightarrows U(f)/j2\pi f$$

$$\int_{a}^{b} v(t,\alpha)d\alpha \leftrightarrows \int_{a}^{b} V(f,\alpha)d\alpha$$

Interchange of functions:

$$U(t) \leftrightarrows u(-f)$$

Dirac delta functions:

$$\delta(t) \leftrightarrows 1 \quad \exp(j2\pi f_0 t) \leftrightarrows \delta(f - f_0)$$

Rect(t) (unit length, unit amplitude pulse, centered on $t = 0$):

$$\text{rect}(t) \leftrightarrows \sin \phi f/\pi f$$

Gaussian distribution:

$$\exp(-\pi t^2) \leftrightarrows \exp(-\phi f^2)$$

Repeated and impulse (delta function) sampled waveforms:

$$\sum_{-\infty}^{\infty} u(t - nT) \leftrightarrows (1/T)U(f) \sum_{-\infty}^{\infty} \delta(f - n/T)$$

$$u(t) \sum_{-\infty}^{\infty} \delta(t - nT) \leftrightarrows (1/T) \sum_{-\infty}^{\infty} U(f - n/T)$$

Parseval's lemma:

$$\int_{-\infty}^{\infty} u(t)v^*(t)dt = \int_{-\infty}^{\infty} U(f)V^*(f)df$$

$$\int_{-\infty}^{\infty} |u(t)|^2 dt = \int_{-\infty}^{\infty} |U(f)|^2 df$$

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

2.44 LAPLACE TRANSFORMS

$$\bar{x}_s = \int_0^\infty x(t)\exp(-st)$$

Function	Transform	Remarks
$e^{-\alpha t}$	$\dfrac{1}{s+\alpha}$	
$\sin \omega t$	$\dfrac{\omega}{s^2+\omega^2}$	
$\cos \omega t$	$\dfrac{s}{s^2+\omega^2}$	
$\sinh \omega t$	$\dfrac{\omega}{s^2-\omega^2}$	
$\cosh \omega t$	$\dfrac{s}{s^2-\omega^2}$	
t^n	$n!s^{n+1}$	
$H(t)$	$1/s$	
$H(t-\tau)$	$\dfrac{1}{s}\exp(-s\tau)$	Heaviside step function
$x(t-\tau)H(t-\tau)$	$\exp(-s\tau)\tilde{x}(s)$	Shift in t
$\delta(t-\tau)$	$\exp(-s\tau)$	Dirac delta function
$\exp(-\alpha t)x(t)$	$\tilde{x}(s+\alpha)$	Shift in s
$\exp(-\alpha t)\sin \omega t$	$\dfrac{\omega}{(s+\alpha)^2+\omega^2}$	
$\exp(-\alpha t)\cos \omega t$	$\dfrac{(s+\alpha)}{(s+\alpha)^2+\omega^2}$	
$tx(t)$	$-\dfrac{d\bar{x}(s)}{ds}$	
$\dfrac{dx(t)}{dt} = x'(t)$	$s\bar{x}(s) - x(0)$	
$\dfrac{d^2x(t)}{dt^2} = x''(t)$	$s^2\bar{x}(s) - sx(0) - x'(0)$	
$\dfrac{d^nx(t)}{dx^n} = x^{(n)}(t)$	$s^n\bar{x}(s) - s^{n-1}x(0) - s^{n-2}x'(0) \ldots - sx^{(n-2)}(0) - x^{(n-1)}(0)$	

Convolution integral

$$\int_0^t x_1(\sigma)x_2(t-\sigma)d\sigma \rightarrow \bar{x}_1(s)\bar{x}_2(s)$$

PART 4

MATRICES AND DETERMINANTS

John Barron

2.45 LINEAR SIMULTANEOUS EQUATIONS

The set of equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_1 \\ &\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

may be written symbolically

$$\mathbf{Ax} = \mathbf{b}$$

in which **A** is the *matrix* of the coefficients a_{ij}, and **x** and **b** are the *column matrices* (or vectors) $(x_1 \cdots x_n)$ and $(b_1 \cdots b_n)$. In this case the matrix **A** is square $(n \times n)$. The equations can be solved unless two or more of them are not independent, in which case

$$\det \mathbf{A} = |\mathbf{A}| = 0$$

and there then exist non-zero solutions x_i only if $\mathbf{b} = 0$. If $\det \mathbf{A} \neq 0$, there exist non-zero solutions only if $\mathbf{b} \neq 0$. When $\det \mathbf{A} = 0$, **A** is *singular.*

2.46 MATRIX ARITHMETIC

If **A** and **B** are both matrices of m rows and n columns they are *conformable,* and

$$\mathbf{A} \pm \mathbf{B} = \mathbf{C} \text{ where } C_{ij} = A_{ij} \pm B_{ij}$$

Product

If **A** is an $m \times n$ matrix and **B** an $n \times l$, the product **AB** is defined by

$$(\mathbf{AB})_{ij} = \sum_{k=1}^{n} (\mathbf{A})_{ik}(\mathbf{B})_{kj}$$

In this case, if $l \neq m$, the product **BA** will not exist.

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

Transpose

The transpose of $\mathbf{A}$ is written $\mathbf{A}'$ or $\mathbf{A}^t$ and is the matrix whose rows are the columns of $\mathbf{A}$, i.e.,

$$(\mathbf{A}^t)_{ij} = (\mathbf{A})_{ji}$$

A square matrix may be equal to its transpose, and it is then said to be *symmetrical.* If the product $\mathbf{AB}$ exists, then

$$(\mathbf{AB})^t = \mathbf{B}^t\mathbf{A}^t$$

Adjoint

The *adjoint* of a square matrix $\mathbf{A}$ is defined as $\mathbf{B}$, where

$$(\mathbf{B})_{ij} = (A)_{ij}$$

and A_{ij} is the *cofactor* of a_{ji} in det $\mathbf{A}$.

Inverse

If $\mathbf{A}$ is non-singular, the *inverse* $\mathbf{A}^{-1}$ is given by

$$\mathbf{A}^{-1} = \text{adj } \mathbf{A}/\det \mathbf{A} \text{ and } \mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = 1$$

the *unit* matrix.

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$

if both inverses exist. The original equations $\mathbf{Ax} = \mathbf{b}$ have the solutions $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ if the inverse exists.

Orthogonality

A matrix $\mathbf{A}$ is orthogonal if $\mathbf{A}\mathbf{A}^t = 1$. If $\mathbf{A}$ is the matrix of a coordinate transformation $\mathbf{X} = \mathbf{AY}$ from variables y_i to variables x_i, then if $\mathbf{A}$ is orthogonal $\mathbf{X}^t\mathbf{X} = \mathbf{Y}^t\mathbf{Y}$, or

$$\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i^2$$

2.47 EIGENVALUES AND EIGENVECTORS

The equation

$$\mathbf{Ax} = \lambda\mathbf{x}$$

where $\mathbf{A}$ is a square matrix, $\mathbf{x}$ a column vector and λ a number (in general complex)

has at most n solutions ($\mathbf{x}$, λ). The values of λ are *eigenvalues* and those of $\mathbf{x}$ *eigenvectors* of the matrix $\mathbf{A}$. The relation may be written

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0$$

so that if $x \neq 0$, the equation $\mathbf{A} - \lambda\mathbf{I} = 0$ gives the eigenvalues. If $\mathbf{A}$ is symmetric and real, the eigenvalues are real. If $\mathbf{A}$ is symmetric, the eigenvectors are orthogonal. If $\mathbf{A}$ is not symmetric, the eigenvalues are complex and the eigenvectors are not orthogonal.

2.48 COORDINATE TRANSFORMATION

Suppose $\mathbf{x}$ and $\mathbf{y}$ are two vectors related by the equation

$$y = \mathbf{A}\mathbf{x}$$

when their components are expressed in one orthogonal system, and that a second orthogonal system has unit vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ expressed in the first system. The components of $\mathbf{x}$ and $\mathbf{y}$ expressed in the new system will be $\mathbf{x}'$ and $\mathbf{y}'$, where

$$\mathbf{x}' = \mathbf{U}'\mathbf{x}, \quad \mathbf{y}' = \mathbf{U}'\mathbf{y}$$

and $\mathbf{U}^t$ is the orthogonal matrix whose rows are the unit vectors $\mathbf{u}_1^t$, $\mathbf{u}_2^t$, etc. Then

$$\mathbf{y}' = \mathbf{U}'\mathbf{y} = \mathbf{U}^t\mathbf{A}\mathbf{x} = \mathbf{U}'\mathbf{A}\mathbf{x} = \mathbf{U}'\mathbf{A}\mathbf{u}\mathbf{x}'$$

or

$$y' = \mathbf{A}'\mathbf{x}'$$

where

$$\mathbf{A}' = \mathbf{U}'\mathbf{A}\mathbf{U}$$

Matrices $\mathbf{A}$ and $\mathbf{A}'$ are *congruent*.

2.49 DETERMINANTS

The determinant

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ q_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

is defined as follows. The first suffix in a_{rs} refers to the row, the second to the column which contains a_{rs}. Denote by M_{rs} the determinant left by deleting the rth row and sth column from D, then

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

$$D = \sum_{k=1}^{n} (-1)^{k+1} a_{1k} M_{1k}$$

gives the value of D in terms of determinants of order $n - 1$, hence by repeated application, of the determinant in terms of the elements a_{rs}.

2.50 PROPERTIES OF DETERMINANTS

If the rows of $|a_{rs}|$ are identical with the columns of $|b_{sr}|$, $a_{rs} = b_{sr}$ and $|a_{rs}| = |b_{sr}|$ that is, the *transposed* determinant is equal to the original.

If two rows or two columns are interchanged, the numerical value of the determinant is unaltered, but the sign will be changed if the permutation of rows or columns is odd.

If two rows or two columns are identical, the determinant is zero.

If each element of one row or one column is multiplied by k, so is the value of the determinant.

If any row or column is zero, so is the determinant.

If each element of the pth row or column of the determinant c_{rs} is equal to the sum of the elements of the same row or column in determinants a_{rs} and b_{rs}, then

$$|c_{rs}| = |a_{rs}| + |b_{rs}|$$

The addition of any multiple of one row (or column) to another row (or column) does not alter the value of the determinant.

Minor

If row p and column q are deleted from $|a_{rs}|$, the remaining determinant M_{pq} is called the *minor* of a_{pq}.

Cofactor

The *cofactor* of a_{pq} is the minor of a_{pq} prefixed by the sign which the product $M_{pq}a_{pq}$ would have in the expansion of the determinant, and is denoted by A_{pq}:

$$A_{pq} = (-1)^{p+qM} M_{pq}$$

A determinant a_{ij} in which $a_{ij} = a_{ji}$ for all i and j is called *symmetric,* whilst if $a_{ij} = -a_{ji}$ for all i and j, the determinant is *skew-symmetric.* It follows that $a_{ii} = 0$ for all i in a skewsymmetric determinant.

2.51 NUMERICAL SOLUTION OF LINEAR EQUATIONS

Evaluation of a determinant by direct expansion in terms of elements and cofactors is disastrously slow, and other methods are available, usually programmed on any existing computer system.

Reduction of Determinant or Matrix to Upper Triangular or to Diagonal Form

The system of equations may be written

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} x_1 \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

The variable x_1 is eliminated from the last $n - 1$ equations by adding a multiple $-a_{i1}/a_{11}$ of the first row to the ith, obtaining

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a'_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & a''_{nn} \end{bmatrix} x_1 \begin{bmatrix} x_1 \\ x_2 \\ \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b'_1 \\ \\ b''_n \end{bmatrix}$$

where primes indicate altered coefficients. This process may be continued by eliminating x_2 from rows 3 to n, and so on. Eventually the form will become

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a'_{22} & \cdots & a'_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & a''_{nn} \end{bmatrix} x_1 \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b'_2 \\ \vdots \\ b''_n \end{bmatrix}$$

x_n can now be found from the nth equation, substituted in the $(n - 1)$th to obtain x_{n-1} and so on.

Alternatively, the process may be applied to the system of equations in the form

$$\mathbf{Ax} = \mathbf{Ib}$$

where $\mathbf{I}$ is the unit matrix, and the same operations carried out upon $\mathbf{I}$ as upon $\mathbf{A}$. If the process is continued after reaching the upper triangular form, the matrix $\mathbf{A}$ can eventually be reduced to diagonal form. Finally, each equation is divided by the corresponding diagonal element of $\mathbf{A}$, thus reducing $\mathbf{A}$ to the unit matrix. The system is now in the form

$$\mathbf{Ix} = \mathbf{Bb}$$

and evidently $\mathbf{B} = \mathbf{A}^{-1}$. The total number of operations required is $O(n^3)$.

PART 5
DIFFERENTIAL EQUATIONS

R. Ken Livesley

A differential equation is an equation involving a *dependent* variable and its derivatives with respect to one or more *independent* variables. An *ordinary* differential equation is one in which there is only one independent variable—conventionally x or t. A *partial* differential equation is one in which there are several independent variables.

2.52 NOTATION AND DEFINITIONS

An ordinary differential equation with y as *dependent* variable and x as *independent* variable has the general form

$$f\left\{x;\, y,\, \frac{dy}{dx},\, \frac{d^2y}{dx^2},\, \ldots\right\} = 0$$

where $f\{\ \}$ represents some specified function of the arguments. *Solving* a differential equation involves obtaining an explicit expression for y as a known function of x.

The *order* of a differential equation is the order of the highest derivative appearing in it. Thus

$$\frac{d^2y}{dx^2} + 3\frac{dy}{dx} + 6y = 6$$

is a second-order equation. A differential equation of order n has a *general* solution containing n *arbitrary constants.* Specified values of the *dependent* variable and/or its derivatives which allow these arbitrary constants to be determined are called *boundary conditions* or (when the independent variable is t and the values are given at $t = 0$) *initial conditions.* Boundary conditions in which the dependent variable or its derivatives are assigned zero values are called *homogeneous* boundary conditions. A solution in which the arbitrary constants take definite values is called a *particular* solution.

A *linear* differential equation is one which is linear in the *dependent* variable and its derivatives, having the general form

$$p_n(x)\frac{d^ny}{dx^n} + \cdots + p_1(x)\frac{dy}{dx} + p_0(x)y = f(x) \tag{2.1}$$

where $p_0(x) \ldots p_n(x)$ and $f(x)$ are specified functions of x. If $f(x) \neq 0$ the differential equation is said to be *inhomogeneous.* If $f(x) = 0$, so that

$$p_n(x)\frac{d^ny}{dx^n} + \cdots + p_1(x)\frac{dy}{dx} + p_0(x)y = 0 \tag{2.2}$$

the differential equation is said to be *homogeneous.*

In a partial differential equation the independent variables are normally variables defining spatial position plus (possibly) time. A particular solution of a partial differential equation requires the definition of a *solution region* with a bounding curve or bounding surface, together with the specification of suitable boundary conditions on that curve or surface. A partial differential equation, like an ordinary differential equation, may be linear or nonlinear, and a linear partial differential equation may be homogeneous or inhomogeneous. Boundary conditions, specifying values of the dependent variable and/or its derivatives, may also be homogeneous or inhomogeneous.

2.53 ORDINARY DIFFERENTIAL EQUATIONS: ANALYTICAL SOLUTIONS

Simple analytical solutions exist for first-order linear differential equations and for linear equations of higher order with constant coefficients.

First-Order Linear Equations

A first-order linear differential equation has the general form $p_1(x)(dy/dx) + p_0(x)y = f(x)$, which can be written as

$$\frac{dy}{dx} + P(x)y = Q(x) \tag{2.3}$$

This equation has the general solution

$$ye^{\int P(x)dx} = \int Q(x)e^{\int P(x)dx}dx + C \tag{2.4}$$

where C is an arbitrary constant. The function $e^{\int P(x)dx}$ is known as the *integrating factor.*

Linear Equations with Constant Coefficients

Homogeneous Equations. A second-order homogeneous linear differential equation with constant coefficients has the general form

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0 \tag{2.5}$$

The general solution is

$$y = C_1e^{\lambda_1 x} + C_2e^{\lambda_2 x} \tag{2.6}$$

where λ_1, λ_2 are the roots of the *auxiliary equation* $a\lambda^2 + b\lambda + c = 0$, and C_1, C_2 are arbitrary constants.

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

If the roots of the auxiliary equation are complex, with values $\lambda_1 = \alpha + \int\beta$, $\lambda_2 = \alpha - i\beta$, it is more convenient to write the general solution of the differential equation in the form

$$y = e^{\alpha x}(C_1 \cos \beta x + C_2 \sin \beta x) \tag{2.7}$$

If the roots are equal, i.e., $\lambda_1 = \lambda_2 = \lambda$, say, then the general solution is

$$y = e^{\lambda x}(C_1 + C_2 x) \tag{2.8}$$

where again C_1, C_2 are arbitrary constants.

The solution of third- and higher-order homogeneous equations follows a similar pattern, the auxiliary equation being a polynomial equation in λ of appropriate degree.

Inhomogeneous Equations. A second-order inhomogeneous linear differential equation with constant coefficients has the general form

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x) \tag{2.9}$$

where $f(x)$ is a specified function. The general solution of equation (2.9) is the general solution of the homogeneous equation (2.5) containing two arbitrary constants (this solution is called the *complementary function)* plus a function (called the *particular integral*) which, when substituted into equation (2.9), gives the correct function f(x) on the right- hand side.

For many simple right-hand sides the particular integral can be found by replacing y in the differential equation by a trial solution containing one or more unknown parameters, here written as α, β, etc.

Right-hand side: $f(x)$	Trial solution: $y(x)$
constant	α
x^n(n integral)	$\alpha x^n + \beta x^{n-1} + \ldots$
e^{kx}	αe^{kx}
xe^{kx}	$(\alpha x + \beta)e^{kx}$
$x^n e^{kx}$	$(\alpha x^n + \beta x^{n-1} + \ldots)e^{kx}$
$\sin kx$, $\cos kx$	$\alpha \sin kx + \beta \cos kx$. (If only even differential coefficients occur in the differential equation then $\alpha \sin kx$ or $\beta \cos kx$ is sufficient.)
$e^{k_1 x} \sin k_2 z$, $e^{k_1 x} \cos k_2 x$	$e^{k_1 x}(\alpha \sin k_2 x + \beta \cos k_2 x)$

Equating the coefficients of the functions on the two sides of the equation gives

the values of the parameters. This technique can also be used to solve equations of third and higher orders.

If $f(x)$ has the same form as one of the terms in the complementary function then the substitution $y = uf(x)$ should be made, where u is an unknown function of x. This substitution generates a simple differential equation for $u(x)$.

Simultaneous Linear Differential Equations. The analysis of a linear mechanical or electrical system with several degrees of freedom may require the solution of a set of simultaneous linear differential equations, in which there is one independent variable (normally time) and several dependent variables. In cases where the equations have constant coefficients, as in the example

$$\frac{du}{dt} + 3\frac{dv}{dt} + u - v = t^2$$

$$2\frac{du}{dt} - \frac{dv}{dt} - 2u + 3v = 0$$

the equations can be solved by a procedure very similar to the elimination method for solving sets of linear algebraic equations. This procedure generates a linear differential equation (with order equal to the sum of the orders of the original equations) for one of the dependent variables: after solution of this equation the other dependent variables can be obtained by back-substitution.

Inserting the Initial or Boundary Conditions. A linear differential equation of order n has a general solution

$$y = \xi_0(x) + C_1\xi_1(x) + C_2\xi_2(x) + \cdots + C_n\xi_n(\mathrm{x}) \tag{2.10}$$

where $\xi_0(x)$ is the *particular integral* and $C_1\xi_1(x) + C_2\xi_2(x) \cdots + C_n\xi_n(x)$ is the *complementary function.* Once this general solution has been found, the values of the n constants can be obtained by imposing n boundary or initial conditions, i.e. n values of y and/or its derivatives at particular values of x. If all the boundary conditions are specified at a single value of x the problem is referred to as a *one-point boundary-value* problem or, if the independent variable is t and the conditions are specified at $t = 0$, as an *initial-value* problem. Initial value problems can also be solved by the use of Laplace transforms. The Laplace transform method determines a particular solution of a differential equation, with the initial conditions inserted, rather than the general solution (2.10).

Impulse and Frequency Responses: The Convolution Integral. The solution of the differential equation

$$a_n\frac{d^n y}{dt^n} + \cdots + a_1\frac{dy}{dt} + a_0 y = f(t) \tag{2.11}$$

for a general function of time $f(t)$ with homogeneous initial conditions

$$\frac{d^{n-1}y}{dt^{n-1}} = \frac{d^{n-2}y}{dt^{n-2}} = \cdots = \frac{dy}{dt} = y = 0 \text{ at } t = 0$$

can be obtained from the *impulse response* $g(t)$, which is the solution of the dif-

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

ferential equation with the same initial conditions when $f(t) = \delta(t)$. $\delta(t)$ is the Dirac δ-function, defined by the equations

$$\int_{-\infty}^{\infty} \delta(t)dt = 1;\ \delta(t) = 0 \text{ if } t \neq 0$$

The impulse response can be obtained by solving the homogeneous equation

$$a_n \frac{d^n y}{dt^n} + \cdots + a_1 \frac{dy}{dt} + a_0 y = 0 \tag{2.12}$$

with initial conditions $(d^{n-1}y)/(dt^{n-1}) = 1/a_n$, $(d^{n-2})/(dt^{n-2}) = \cdots = dy/dt = y = 0$ at $t = 0$. Alternatively, it can be found by the use of Laplace transforms.

The solution of equation (2.11) for an arbitrary right-hand side $f(t)$ is given in terms of the impulse response $g(t)$ by the *convolution integral*

$$y(t) = \int_0^1 g(\tau)f(t - \tau)d\tau \tag{2.13}$$

This integral is symmetric in the functions g and f, and can therefore be written in the alternative form

$$y(t) = \int_0^1 f(\tau)g(t - \tau)d\tau \tag{2.14}$$

If $f(t) = e^{i\omega t}$ and equation (2.11) represents a stable system (i.e. the complementary function has no exponential terms with positive real part) then as $t \to \infty$ the solution tends to the steady state form $y(t) = G(\omega)e^{i\omega t}$. The complex function $G(\omega)$ is called the *frequency response* of the system. It may be obtained from the differential equation by substituting the trial solution $y = \alpha e^{i\omega t}$ or from the impulse response by the use of equation (2.13). The latter derivation gives the result

$$G(\omega) = \int_{-\infty}^{\infty} g(\tau)e^{-i\omega\tau}d\tau \tag{2.15}$$

This equation states that the frequency response $G(\omega)$ is the *Fourier transform* of the impulse response $g(t)$

Linear Equations with Variable Coefficients

Second- and higher-order linear equations with variable coefficients do not, in general, have solutions which are expressible in terms of elementary functions. However, there are a number of second-order equations which occur frequently in applied mathematics and for which tables of solutions exist. Subroutines for generating these solutions are available on most scientific computers. Two of the most important of these equations are

$$\text{Bessel's equation:} \quad x^2 \frac{d^2y}{dx^2} + x \frac{dy}{dx} + (\lambda^2 x^2 - n^2)y = 0 \tag{2.16}$$

$$\text{Legendre's equation:} \quad (1 - x^2) \frac{d^2y}{dx^2} - 2x \frac{dy}{dx} + n(n + 1)y = 0 \tag{2.17}$$

In certain other cases an equation with variable coefficients can be converted into one with constant coefficients by means of a change of variable. In general, however, solutions of linear differential equations with variable coefficients can only be obtained by approximate methods.

2.54 ORDINARY DIFFERENTIAL EQUATIONS: APPROXIMATE SOLUTIONS

Approximate solutions of differential equations can be obtained by graphical, numerical or analytical methods.

A Graphical Method for First-Order Equations

A graphical solution of the general first-order equation $dy/dx = f(x,y)$ can be obtained as follows. A series of curves $f(x,y) = c_1, c_2, \ldots, c_i, \ldots$ (termed *isoclines*) are drawn in the x, y plane, where the c's are suitable constants. On each isocline line-segments are drawn with slope equal to the associated value of c_i: these segments give the direction of the solutions as they cross the isocline. The general form of these solutions can be obtained by joining up the segments to form continuous curves.

A simple example is shown in Figure (2.17), which illustrates the solution of the differential equation $dy/dx = -x/y$. The isoclines $-x/y = c_1, c_2, \ldots, c_i, \ldots$ are straight lines through the origin, and the segments which form part of the solutions are always perpendicular to the isoclines. It is clear from the figure that the solutions are circles centered on the origin: this is easily verified analytically.

Approximate Numerical Methods

Derivatives and Differences. If a continuous function $y(x)$ is sampled at a series of equally spaced points $x_0, \ldots, x_n, \ldots, x_N$ to give a set of values $y_0, \ldots, y_n, \ldots, y_N$ then it follows from the definition of a differential coefficient that

$$\left(\frac{dy}{dx}\right)_{n+1/2} \approx (y_{n+1} - y_n)/h, \qquad \left(\frac{dy}{dx}\right)_{n-1/2} \approx (y_n - y_{n-1})/h \tag{2.18}$$

or alternatively

$$\left(\frac{dy}{dx}\right)_n \approx (y_{n+1} - y_{n-1})/2h \tag{2.19}$$

where h is the sampling inverval, as shown in Figure 2.18. Taking the difference of the two equations (2.18) and dividing by h gives

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

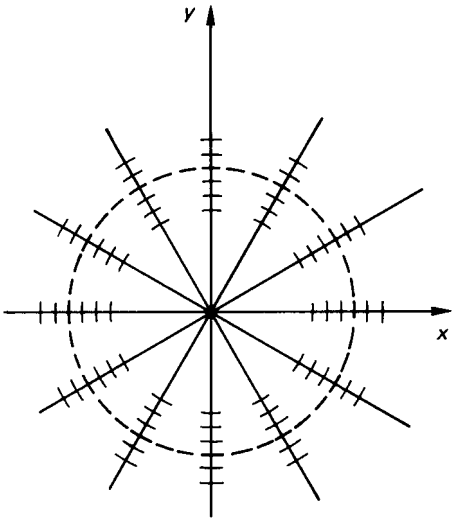


FIGURE 2.17 Isoclines for the differential equation $dy/dx = -x/y$.

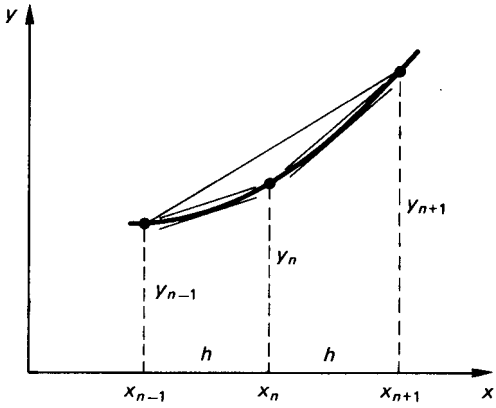


FIGURE 2.18 Approximate representations of dy/dx.

$$\left(\frac{d^2y}{dx^2}\right)_n \approx (y_{n+1} - 2y_n + y_{n-1})/h^2 \tag{2.20}$$

and the process can be continued in a similar way to give approximations to $(d^3y/dx^3)_{n+1/2}$, etc. The quantities $(y_1 - y_0), \ldots (y_{n+1} - y_n), \ldots (y_N - y_{N-1})$ are termed the *first differences* of the set of values y_n, the quantities $\ldots (y_{n+1} - 2y_n + y_{n-1}), \ldots$ the *second differences,* and so on. The role of differences in numerical analysis is similar to that of differential coefficients in calculus.

Two-Point Boundary-Value Problems. An approximate solution of the second-order linear differential equation

$$p_2(x)\frac{d^2y}{dx^2} + p_1(x)\frac{dy}{dx} + p_0(x)y = f(x) \tag{2.21}$$

with boundary conditions $y = y_0$ at $x = 0$, $y = y_N$ at $x = a$ can be found by dividing the solution range $0 \leqslant x \leqslant a$ into N equal intervals and replacing the continuous function $y(x)$ by a set of $N + 1$ quantities $y_n = y(x_n)$, $(n = 0, \ldots, N)$, where $x_n = nh$ and $h = a/n$. Replacing the differential coefficients in equation (2.21) by the approximations (2.19) and (2.20) gives

$$\begin{aligned} &p_2(x_n)(y_{n+1} - 2y_n + y_{n-1}) + hp_1(x_n)(y_{n+1} - y_{n-1})/2 \\ &\quad + h^2p_0(x_n)y_n = f(x_n) \quad (n = 1, \ldots, n - 1) \end{aligned} \tag{2.22}$$

Setting up an equation of this form at each of the points $x_1, \ldots, x_{N-1}$ produces a set of $n - 1$ simultaneous linear algebraic equations which can be solved for the unknown function values $y_1, \ldots, y_{N-1}$ (the values of y_0 and y_N which appear in these equations are known from the boundary conditions). Intermediate values of $y(x)$ can be found subsequently by interpolation.

Initial-Value Problems. The general first-order differential equation

$$\frac{dy}{dt} = f(t,y) \tag{2.23}$$

with initial condition $y = y_0$ at $t = t_0$ can be solved by a step-by-step procedure in which approximate function values $y_1, y_2, \ldots$ are computed successively at $t = t_1, t_2, \ldots$ The simplest step-by-step procedure is due to Euler and involves the replacement of the differential equation (2.23) by the approximation

$$y_{n+1} = y_n + hf(t_n,y_n) \quad (n = 0, 1, 2, \ldots) \tag{2.24}$$

where h is equal to the interval $t_{n+1} - t_n$. As shown in Figure 2.19, this procedure

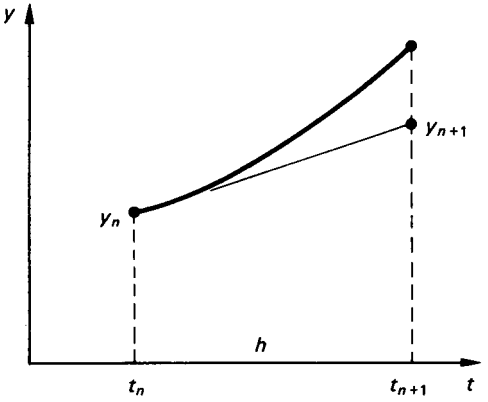


FIGURE 2.19 Euler's approximate integration procedure.

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

takes the tangent at each solution point as the solution over the next interval. The truncation error in a single step is $O(h^2)$. If the step-length h is kept constant over a given range $0 \leqslant t \leqslant T$ the number of steps is T/h, so that the truncation error over the range is $O(h)$. (The round-off error increases with the number of steps, so that there is an optimum value of h which minimizes the total error.)

The accuracy of the Euler procedure can be improved by using equation (2.24) as a predictor to obtain an approximate value y^*_{n+1}, which is then inserted in a suitable corrector formula to generate a more accurate value of y_{n+1}. A simple predictor/corrector pair is

$$\begin{aligned} &\text{Predictor} \quad y^*_{n+1} = y_n + hf(t_n, y_n) \\ &\text{Corrector} \quad y_{n+1} = y_n + h\{f(t_n, y_n) + f(t_{n+1}, y^*_{n+1})\}/2 \end{aligned} \tag{2.25}$$

One of the most popular predictor–corrector procedures is the Runge–Kutta. A single step of the procedure involves four evaluations of $f(t,y)$ in accordance with the formulae

$$a_1 = hf(t_n, y_n), \qquad a_2 = hf(t_n + h/2, y_n + a_1/2)$$

$$a_3 = hf(t_n + h/2, y_n + a_2/2), \qquad a_4 = hf(t_n + h, y_n + a_3)$$

the final value of y_{n+1} being

$$y_{n+1} = y_n + \{a_1 + 2a_2 + 2a_3 + a_4\}/6 \tag{2.26}$$

The error per step is $O(h^5)$, so that the error over a given range of t is $O(h^4)$. A computer subroutine for the Runge–Kutta procedure normally requires a user-supplied subroutine to evaluate $f(t,y)$ for specified values of t and y.

An initial-value problem involving a differential equation of second or higher order can be solved by reducing the differential equation to a set of first-order equations. For example, the third-order nonlinear equation

$$\frac{d^3y}{dt^3} + t\frac{d^2y}{dt^2} - 2t^2\left(\frac{dy}{dt}\right)^2 + y^2 = 10t$$

can be solved by introducing the additional variables u and v and writing the equation as

$$\frac{dy}{dt} = u; \qquad \frac{du}{dt} = v; \qquad \frac{dv}{dt} = 10t - tv + 2t^2u^2 - y^2$$

This set of first-order equations for the three variables u, v, and y can be solved by any of the methods described above, the step-by-step procedure being carried forward simultaneously for each of the variables.

Approximate Analytical Methods

An approximate solution of a linear differential equation can also be obtained by choosing a set of M *basis functions* $B_m(x)$ and expressing the unknown solution $y(x)$ as

$$y(x) \approx c_1 B_1(x) + \cdots + c_M b_M(x) = \sum_{m=1}^{M} c_m B_m(x) \tag{2.27}$$

There are a number of methods based on this approach. They may be classified according to the choice of basis functions $B_m(x)$ and the procedure used to find the constants c_m. The most important sets of basis functions are the integral powers of x (which generate power-series approximations) and the harmonic functions sin mx and cos mx (which generate Fourier approximations). In the following account the equation to be solved is written as

$$\mathfrak{L}y = w(x) \tag{2.28}$$

where $\mathfrak{L}$ represents a specified linear differential operator and $w(x)$ is a specified function of x. It is assumed that a solution is required in an interval $p \leqslant x \leqslant q$ and that sufficient homogeneous boundary conditions are specified at $x = p$ and $x = q$ to make the solution unique. It is further assumed that each of the approximating functions $B_1(x), \ldots, B_M(x)$ satisfies these boundary conditions.

In general the approximation (2.27) will not be capable of satisfying the differential equation (2.28) exactly, whatever values are assigned to the constants c_i: there will be an error function

$$\epsilon(x) = \mathfrak{L}\left\{\sum_{m=1}^{M} c_m b_m(x)\right\} - w(x)$$

which can be written in the form

$$\epsilon(x) = \sum_{m=1}^{M} c_m b_m(x) - w(x) \tag{2.29}$$

where $b_m(x) = \mathfrak{L}\{B_m(x)\}$.

Two procedures for finding sets of constants which make the error $\varepsilon(x)$ small are *collocation* and *Galerkin's method.* In the collocation method the constants c_m are obtained by making $\varepsilon(x)$ zero at a selected set of points x_k $(k = 1, \ldots, M)$ in the interval $p \leqslant x \leqslant q$. This generates a set of M simultaneous equations

$$\sum_{m=1}^{M} b_m(x_k) c_m = w(x_k) \quad (k = 1, \ldots, M) \tag{2.30}$$

which can be solved for the M constants. In Galerkin's method the constants c_m are obtained by making $\epsilon(x)$ orthogonal to the M basis functions $B(x)$,

$$\int_p^q B_k(x)\epsilon(x)dx = \int_p^q B_k(x)\left\{\sum_{m=1}^{M} c_m b_m(x) - w(x)\right\} dx = 0 \quad (k = 1, \ldots, M)$$

These equations can be written in the form

$$\sum_{m=1}^{M} \left\{\int_p^q B_k(x) b_m(x) dx\right\} c_m = \int_p^q B_k(x) w(x) dx \quad (k = 1, \ldots, M) \tag{2.31}$$

Equation (2.31), like equation (2.30), represents a set of M linear algebraic equa-

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

tions for the unknown constants c_m. If the differential operator $\mathfrak{L}$ is self-adjoint (a condition satisfied in most practical applications of the method) the coefficients

$$\int_p^q B_k(x) b_m(x) dx$$

form a symmetric matrix. If, in addition, the functions $B_m(x)$ are chosen to be the normalized eigenfunctions of the differential operator $\mathfrak{L}$, so that $\mathfrak{L}\{B_m(x)\} = b_m(x) = \lambda_m B_m(x)$, then equation (2.31) takes the simpler form

$$c_k = \int_p^q B_k(x) w(x) dx / \lambda_k \quad (k = 1, \ldots, M) \tag{2.32}$$

with each constant c_k depending only on the corresponding function $B_k(x)$.

2.55 PARTIAL DIFFERENTIAL EQUATIONS

Linear partial differential equations can be classified as elliptic, hyperbolic or parabolic. An elliptic differential equation is one in which the boundary conditions imposed on each segment of the boundary affect the solution at *all* points in the solution region or, conversely, one in which the solution at any point depends on the boundary conditions over the *whole* boundary, as shown in Figure 2.20(a). The commonest elliptic equation is Laplace's equation

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} = 0 \tag{2.33}$$

which is the equation governing gravitational fields in free space, steady heat and

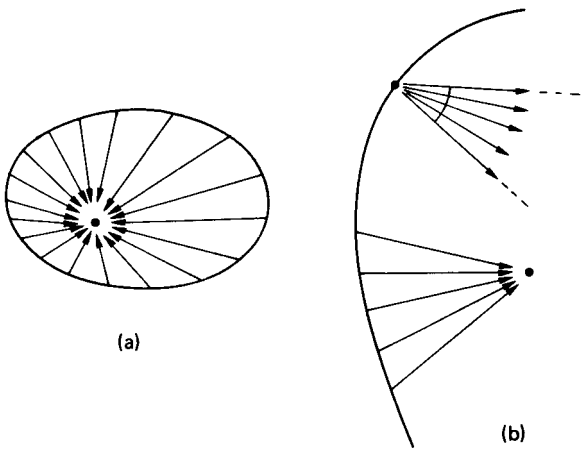


FIGURE 2.20 Partial differential equation types: (a) elliptic, (b) hyperbolic.

electrical conduction, seepage flow in soils, etc. The inhomogeneous form of Laplace's equation is Poisson's equation

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} = -\sigma \tag{2.34}$$

where σ is a known function of position. This equation governs gravitational fields in regions containing distributed matter, heat conduction in the presence of distributed heat sources, etc.

Another elliptic differential equation of interest is the bi-harmonic equation governing the bending of an initially flat plate:

$$\frac{\partial^4 \phi}{\partial x^4} + 2\,\frac{\partial^4 \phi}{\partial x^2 \partial y^2} + \frac{\partial^4 \phi}{\partial y^4} = -q/D \tag{2.35}$$

where ϕ is the transverse displacement of the plate, q is the known distribution of transverse load, and D is a constant representing the stiffness of the plate.

Equations (2.33)–(2.35) can also be written in the more general form $\nabla^2 \phi = 0$, $\nabla^2 \phi = -\sigma$, $\nabla^4 \phi = -a/D$, where ∇^2 is the Laplacian operator of vector calculus. This operator takes various forms, depending on the coordinate system (Cartesian, cylindrical polar, spherical polar, etc.) used to define the solution region.

A hyperbolic differential equation is one in which the boundary conditions on a segment of the boundary only affect a *part* of the solution region or, conversely, one in which the solution at any point only depends on the boundary conditions over *part* of the boundary, as shown in Figure 2.20(b). The commonest hyperbolic differential equation is the wave equation

$$\frac{\partial^2 \phi}{\partial x^2} = \frac{1}{a}\,\frac{\partial^2 \phi}{\partial t^2} \text{ or, more generally, } \nabla^2 \phi = \frac{1}{a^2}\,\frac{\partial^2 \phi}{\partial 7 t^2} \tag{2.36}$$

which governs the propagation of sound and other waves in both fluids and solids.

Another common partial differential equation is the diffusion equation

$$\frac{\partial^2 \phi}{\partial x^2} = \frac{1}{a^\delta}\,\frac{\partial \phi}{\partial t} \text{ or, more generally, } \nabla^2 \phi = \frac{1}{a}\,\frac{\partial \phi}{\partial t} \tag{2.37}$$

which governs, for example, the unsteady flow of heat in solids. The diffusion equation is an example of a parabolic differential equation. Such equations can be thought of as lying on the borderline between elliptic and hyperbolic forms.

Analytical Solutions: Separation of Variables

Simple analytical solutions exist for linear partial differential equations with constant coefficients. For example, Laplace's equation in two dimensions is satisfied by both the real and imaginary parts of any analytic function $f(z)$, where z is the complex variable $x + jy$. This fact allows many two- dimensional field problems to be solved by a technique known as *conformal mapping.* Similarly, the one-dimensional wave equation

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

$$\frac{\partial^2 \phi}{\partial x^2} = \frac{1}{a^2} \frac{\partial^2 \phi}{\partial t^2}$$

has solutions of the form $f(x \pm at)$, where f is an arbitrary differentiable function. These solutions represent waves of arbitrary shape traveling along the x axis.

Analytical solutions of linear partial differential equations can be obtained by using the method of *separation of variables.* For a differential equation whose dependent variable is ϕ and whose independent variables are x and y this method involves assuming a solution of the form $\phi = X(x)Y(y)$, where X is an unknown function of x only and Y is an unknown function of y only. Substitution of this solution into the differential equation yields ordinary differential equations for the functions X and Y.

Typical examples of separable solutions are the function

$$\phi = \begin{Bmatrix} \cos \\ \sin \end{Bmatrix} \lambda x \begin{Bmatrix} \cosh \\ \sinh \end{Bmatrix} \lambda y$$

which satisfies both the two-dimensional Laplace equation and the homogeneous plate bending equation and the function

$$\phi = e^{-a\lambda^2 t} \begin{Bmatrix} \cos \\ \sin \end{Bmatrix} \lambda x$$

which satisfies the one-dimensional diffusion equation.

Separable solutions always contain an arbitrary parameter λ called the *separation constant.* The imposition of boundary conditions on a solution may result in only certain values of λ being permissible. In such cases more general solutions can often be built up by combining a number of basic solutions involving these values of λ. For example, the solution of the one-dimensional diffusion equation given above implies the existence of a more general solution

$$\phi = \sum_{n=1}^{\infty} e^{-a\lambda_n^2 t}(A_n \cos \lambda_n x + B_n \sin \lambda_n x)$$

which can be made to fit a variety of boundary conditions by suitable choice of the constants A_n and B_n.

Numerical Solutions: The Finite-Difference Method

The finite-difference method for solving partial differential equations is similar to the numerical technique for solving ordinary differential equations with two-point boundary conditions. The following example shows how the method can be used to find the steady-state distribution of temperature within the L-shaped region shown in Figure 2.21 when the temperature variation on the boundary of the region is given. In this problem the temperature ϕ satisfies the two-dimensional Laplace equation

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0$$

with appropriate values of ϕ specified on the boundary.

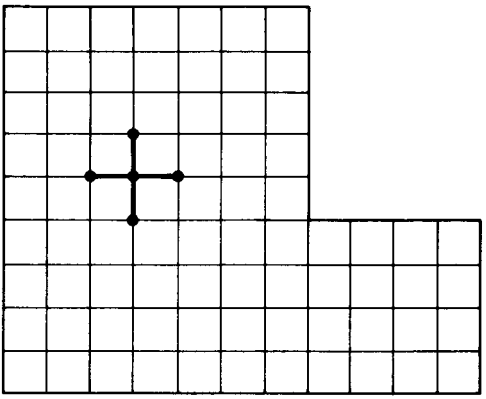


FIGURE 2.21 A finite-difference mesh.

The region is first covered with a uniform grid of squares, as shown in the figure. The intersections of the grid lines within the solution region are called *nodal points* and the values of ϕ at these points are called *nodal values:* it is these values which are determined by the method. At each nodal point the partial derivatives which make up the differential equation are replaced by differences, using an appropriately amended version of equation (2.20). This operation converts the partial differential equation into a linear algebraic equation involving the nodal values at the chosen nodal point and its four nearest neighbors. If these points are labeled as shown in Figure 2.22 then the linear equation associated with the point p is

$$(\phi_q + \phi_r + \phi_s + \phi_t - 4\phi_p)/h^2 = 0 \tag{2.38}$$

A similar equation can be constructed for each nodal point within the solution region (it is not necessary to construct equations for nodal points on the boundary). For nodal points adjacent to the boundary at least one of the values $\phi_q, \ldots, \phi_t$ will be known.

This procedure converts the partial differential equation into a set of n simultaneous linear equations, where n is the number of nodal points within the solution region. In precomputer days these equations were solved by an iterative process

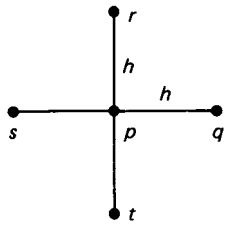


FIGURE 2.22 Nodal points associated with the difference equation for node p.

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

known as *relaxation.* Nowadays they are solved by a computer routine designed to take advantage of the sparse and banded nature of the coefficient matrix. Once the nodal values have been obtained, values of the solution at other points within the region can be found by interpolation.

Numerical Solutions: The Finite-Element Method

In recent years the finite-element method has largely replaced the finite-difference method as the standard numerical technique for solving problems of heat conduction and stress analysis in solid bodies. To assist in a comparison of the two approaches the following account considers the heat-conduction problem solved by finite differences.

The finite-element method also begins with the construction of a *mesh* covering the solution region. This mesh is commonly formed from triangles (the *elements* of the method) although quadrilaterals can also be used. The mesh need not be uniform—indeed, it is standard practice to grade the mesh so that it is finer in regions where the solution is likely to vary rapidly, as shown in Figure 2.23.

The finite-element method, like the finite-difference method, changes the problem of solving a partial differential equation into that of solving a system of linear algebraic equations for a set of nodal values. However, in contrast to the finite-difference method, in which the value of the solution is only defined at the nodal points, the finite-element method replaces the actual solution by an approximation which is linear (or, more generally, a low-order polynomial) within each element.

The first stage of the solution procedure involves the determination of the properties of each individual element. A typical triangular element with nodal values ϕ_1, ϕ_2, ϕ_3 is shown in Figure 2.24. If the temperature ϕ within the element is assumed to vary linearly with position then it can be expressed in terms of the nodal values as

$$\phi = \phi_1 n_1(x,y) + \phi_2 n_2(x,y) + \phi_{33}(x,y) \tag{2.39}$$

where n_1, n_2, n_3 are simple linear functions of x and y called *shape functions.* From

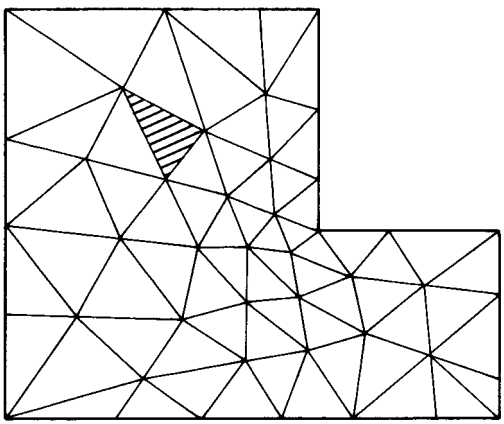


FIGURE 2.23 A finite-element mesh.

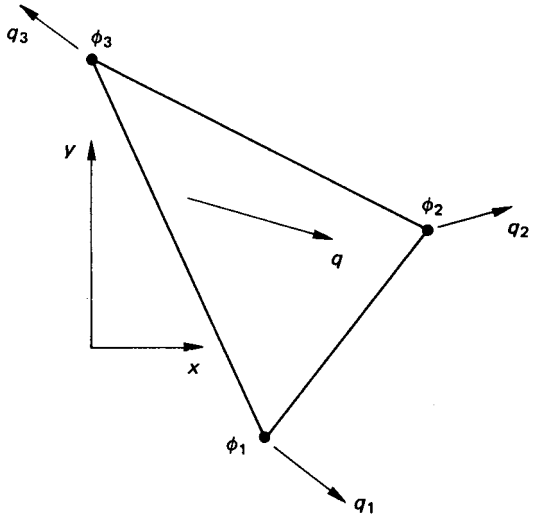


FIGURE 2.24 A typical finite element.

this expression for ϕ it is straightforward to obtain the density of heat flow **q** (constant within the element) and the amount of heat flowing across each side of the triangle as linear functions of the nodal temperatures ϕ_1, ϕ_2, ϕ_3. In preparation for the next part of the procedure these distributed boundary flows are replaced by equivalent concentrated flows q_1, q_2, q_3 at the vertices of the element, as shown in Figure 2.24, these concentrated flows being expressed as linear functions of the nodal temperatures ϕ_1, ϕ_2, ϕ_3.

The second stage of the solution procedure involves joining the elements together to form the solution region. This has two consequences. First, it imposes conditions of continuity on the temperature ϕ. If two elements have nodes p and q in common, as shown in Figure 2.25, then they share the same nodal values ϕ_p and ϕ_q. Furthermore, since ϕ is linear within each element, ϕ is also continuous on their common boundary pq. Second, it imposes conditions on the nodal heat flows. In the exact solution of a steady heat-flow problem the net outflow from any infinitesimal area within the solution region must be zero. In the finite-element method this condition is replaced by the condition that at each node within the solution region the equivalent concentrated nodal flows associated with the node must add up to zero. Since these nodal flows are known linear functions of the associated nodal values, this condition generates a linear equation which relates the nodal values at a group of neighboring nodes. Thus for the elements shown in Figure 2.25 the condition of zero net outflow at node p generates a linear equation involving ϕ_p and ϕ_q, . . . , ϕ_u. There is one such equation for each node within the solution region.

The final stage of the method is the solution of the nodal- flow equations for the nodal values. As with the finite-difference method, the coefficient matrix for these equations is both sparse and banded. After the equations have been solved, values of ϕ at points within elements can be found, if required, from equation (2.39).

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

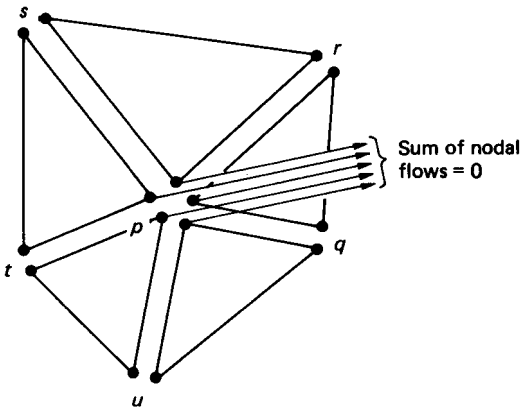


FIGURE 2.25 Adding up the nodal flows in a finite-element mesh.

This example has introduced the simplest form of finite element—the three-node triangle, within which the dependent variable varies *linearly.* Adding three additional nodes, one on each side of the triangle, allows a *quadratic* variation of the dependent variable within the triangle, giving improved accuracy. Four- and eight-noded quadrilaterals are also popular elements in the analysis of two-dimensional problems. In three-dimensional analyses the corresponding elements are tetrahedra, bricks, and wedges. Nowadays it is common for applications of the method to involve meshes with tens of thousands of nodes.

The application of the finite-element method to stress analysis follows similar lines, with (vector) displacements replacing (scalar) temperatures and (tensor) stresses replacing (vector) heat flow densities. Many commercial computer programs are now available for solving a wide range of stress and thermal analysis problems. The method can also be applied to fluid flow and electromagnetic field problems.

PART 6
STATISTICS

Fraidoon Mazda

2.56 INTRODUCTION

Data are available in vast quantities in all branches of engineering. This chapter presents the more commonly used techniques for presenting and manipulating data to obtain meaningful results.

2.57 AVERAGES

Arithmetic Mean

The arithmetic mean of n numbers $x_1, x_2, x_3, \ldots, x_n$ is given by

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n}$$

or

$$\bar{x} = \frac{\sum_{r=1}^{n} x_r}{n} \tag{2.40}$$

The arithmetic mean is easy to calculate and it takes into account all the figures. Its disadvantages are that it is influenced unduly by extreme values and the final result may not be a whole number, which can be absurd at times, e.g., a mean of 2½ people.

Median and Mode Median or "middle on" is found by placing all the figures in order and choosing the one in the middle, or, if there are an even number of items, the mean of the two central numbers. It is a useful technique for finding the average of items which cannot be expressed in figures, e.g., shades of a color. It is also not influenced by extreme values. However, the median is not representative of all the figures.

The mode is the most "fashionable" item, that is, the one which appears the most frequently.

Geometric Mean

The geometric mean of n numbers $x_1, x_2, x_3, \ldots, x_n$ is given by

$$x_g = \sqrt[n]{(x_1 \times x_2 \times x_3 \times \cdots \times x_n)} \tag{2.41}$$

This technique is used to find the average of quantities which follow a geometric

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

progression or exponential law, such as rates of changes. Its advantage is that it takes into account all the numbers, but is not unduly influenced by extreme values.

Harmonic Mean

The harmonic mean of n numbers $x_1, x_2, x_3, \ldots, x_n$ is given by

$$x_h = \frac{n}{\sum_{r=1}^{n} (1/x_r)} \tag{2.42}$$

This averaging method is used when dealing with rates or speeds or prices. As a rule when dealing with items such as A per B, if the figures are for equal As then use the harmonic mean but if they are for equal Bs use the arithmetic mean. So if a plane flies over three equal distances at speeds of 5 m/s, 10 m/s and 15 m/s the mean speed is given by the harmonic mean as

$$\frac{3}{1/5 + 1/10 + 1/15} = 8.18 \text{ m/s}$$

If, however, the plane were to fly for three equal times, of say, 20 seconds at speeds of 5 m/s, 10 m/s, and 15 m/s, then the mean speed would be given by the arithmetic mean as $(5 + 10 + 15)/3 = 10$ m/s.

2.58 DISPERSION

Range and Quartiles

The average represents the central figure of a series of numbers or items. It does not give any indication of the spread of the figures, in the series, from the average. Therefore, in Figure 2.26, both curves, A and B, have the same average but B has a wider deviation from the average than curve A.

There are several ways of stating by how much the individual numbers, in the series, differ from the average. The range is the difference between the smallest

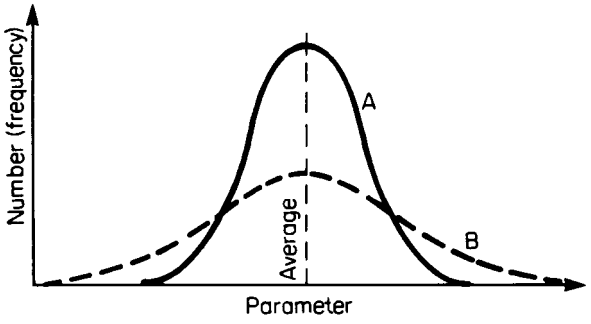


FIGURE 2.26 Illustration of deviation from the average.

and largest values. The series can also be divided into four quartiles and the dispersion stated as the interquartile range, which is the difference between the first and third quartile numbers, or the quartile deviation, which is half this value.

The quartile deviation is easy to use and is not influenced by extreme values. However, it gives no indication of distribution between quartiles and covers only half the values in a series.

Mean Deviation

This is found by taking the mean of the differences between each individual number in the series and the arithmetic mean, or median, of the series. Negative signs are ignored.

For a series of n numbers $x_1, x_2, x_3, \ldots, x_n$ having an arithmetic mean of $\bar{x}$ the mean deviation of the series is given by

$$\frac{\Sigma_{r=1}^{n} |x_r - \bar{x}|}{n} \tag{2.43}$$

The mean deviation takes into account all the items in the series. But it is not very suitable since it ignores signs.

Standard Deviation

This is the most common measure of dispersion. For this the arithmetic mean must be used and not the median. It is calculated by squaring deviations from the mean, so eliminating their sign, adding the numbers together and then taking their mean and then the square root of the mean. Therefore, for the series of equation 2.43 the standard deviation is given by

$$\sigma = \left(\frac{\Sigma_{r=1}^{n} (x_r - \bar{x})^2}{n} \right)^{1/2} \tag{2.44}$$

The unit of the standard deviation is that of the original series. So if the series consists of the heights of a group of children in meters, then the mean and standard deviation are in meters. To compare two series having different units, such as the height of children and their weights, the coefficient of variation is used, which is unitless:

$$\text{coefficient of variation} = \frac{\sigma}{\bar{x}} \times 100 \tag{2.45}$$

2.59 SKEWNESS

The distribution shown in Figure 2.26 is symmetrical since the mean, median, and mode all coincide. Figure 2.27 shows a skewed distribution. It has positive skewness; if it bulges the other way, the skewness is said to be negative.

There are several mathematical ways for expressing skewness. They all give a measure of the deviation between the mean, median and mode and they are usually

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

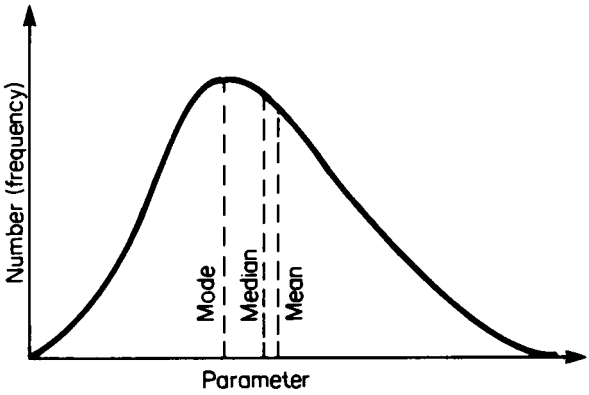


FIGURE 2.27 Illustration of skewness.

stated in relative terms, for ease of comparison between series of different units. The Pearson coefficient of skewness is given by

$$P_k = \frac{\text{mean} - \text{mode}}{\text{standard deviation}} \tag{2.46}$$

Since the mode is sometimes difficult to measure this can also be stated as

$$P_k = \frac{3(\text{mean} - \text{median})}{\text{standard deviation}} \tag{2.47}$$

2.60 COMBINATIONS AND PERMUTATIONS

Combinations

Combinations are the numbers of ways in which a proportion can be chosen from a group. Therefore the number of ways in which two letters can be chosen from a group of four letters A, B, C, D is equal to 6, i.e., AB, AC, AD, BC, BD, CD. This is written as

$$^4C_2 = 6$$

The factorial expansion is frequently used in combination calculations where

$$n! = n \times (n - 1) \times (n - 2) \times \cdots \times 3 \times 2 \times 1$$

Using this the number of combinations of n items from a group of n is given by

$$^nC_r = \frac{n!}{r!(n - r)!} \tag{2.48}$$

Permutations

Combinations do not indicate any sequencing. When sequencing within each combination is involved the result is known as a permutation. Therefore the number of permutations of two letters out of four letters A, B, C, D is 12, i.e., AB, BA, AC, CA, AD, DA, BC, CB, BD, DB, CD, DC. The number of permutations of r items from a group of n is given by

$$^{n}P_{r} = \frac{n!}{(n - r)!} \tag{2.49}$$

2.61 REGRESSION AND CORRELATION

Regression

Regression is a method of establishing a mathematical relationship between two variables. Several equations may be used to establish this relationship, the most common being that of a straight line. Figure 2.28 shows the plot of seven readings. This is called a scatter diagram. The points can be seen to lie approximately on the straight line AB.

The equation of a straight line is given by

$$y = mx + c \tag{2.50}$$

where x is the independent variable, y the dependent variable, m is the slope of the line and c its interception on the y-axis. c is negative if the line intercepts the y-axis on its negative part and m is negative if the line slopes the other way to that shown in Figure 2.28.

The best straight line to fit a set of points is found by the method of least squares as

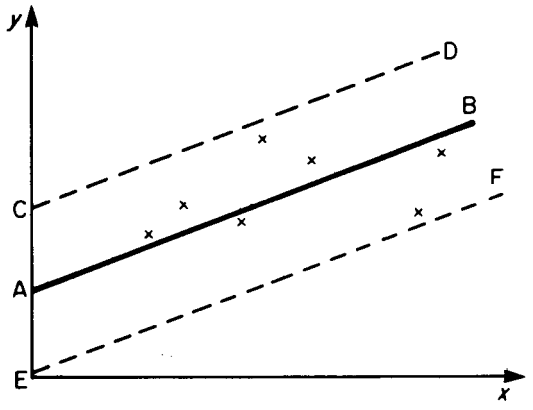


FIGURE 2.28 A scatter diagram.

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

$$m = \frac{\Sigma xy - (\Sigma x \Sigma y)/n}{\Sigma x^2 - (\Sigma x)^2/n} \tag{2.51}$$

and

$$c = \frac{\Sigma x \Sigma xy - \Sigma y \Sigma x^2}{(\Sigma x)^2 - n\Sigma x^2} \tag{2.52}$$

where n is the number of points. The line passes through the mean values of x and y, i.e., $\bar{x}$ and $\bar{y}$.

Correlation

Correlation is a technique for establishing the strength of the relationship between variables. In Figure 2.28 the individual figures are scattered on either side of a straight line and although one can approximate them by a straight line it may be required to establish if there is correlation between the x and y-readings.

Several correlation coefficients exist. The product moment correlation coefficient (r) is given by

$$r = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{n\sigma_x\sigma_y} \tag{2.53}$$

or

$$r = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{[\Sigma(x - \bar{x})^2\Sigma(y - \bar{y})^2]^{1/2}} \tag{2.54}$$

The value of r varies from $+1$, when all the points lie on a straight line and y increases with x, to -1, when all the points lie on a straight line but y decreases with x. When $r = 0$ the points are widely scattered and there is said to be no correlation between x and y.

The standard error of estimation in r is given by

$$S_y = \sigma_y(1 - r^2)^{1/2} \tag{2.55}$$

In about 95% of cases, the actual values will lie between plus or minus twice the standard error of estimated values given by the regression equation. This is shown by lines CD and EF in Figure 2.28. Almost all the values will be within plus or minus three times the standard error of estimated values.

It should be noted that σ_y is the variability of the y-values, whereas S_y is a measure of the variability of the y-values as they differ from the regression which exists between x and y. If there is no regression then $r = 0$ and $\sigma_y = S_y$.

It is often necessary to draw conclusions from the order in which items are ranked. For example, two judges may rank contestants in a contest and we need to know if there is any correlation between their rankings. This may be done by using the Rank correlation coefficient (R) given by

$$R = 1 - \frac{6\Sigma d^2}{n^3 - n} \tag{2.56}$$

where d is the difference between the two ranks for each item and n is the number

of items. The value of R will vary from $+1$ when the two ranks are identical to -1 when they are exactly reversed.

2.62 PROBABILITY

If an event A occurs n times out of a total of m cases then the probability of occurrence is stated to be

$$P(A) = n/m \tag{2.57}$$

Probability varies between 0 and 1. If $P(A)$ is the probability of occurrence then $1 - P(A)$ is the probability that event A will not occur and it can be written as $P(\overline{A})$.

If A and B are two events then the probability that either may occur is given by

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) \tag{2.58}$$

A special case of this probability law is when events are mutually exclusive, i.e., the occurrence of one event prevents the other from happening. Then

$$P(A \text{ or } B) = P(A) + P(B) \tag{2.59}$$

If A and B are two events then the probability that they may occur together is given by

$$P(A \text{ and } B) = P(A) \times P(B|A) \tag{2.60}$$

or

$$P(A \text{ and } B) = P(B) \times P(A|B) \tag{2.61}$$

$P(B|A)$ is the probability that event B will occur assuming that event A has already occurred and $P(A|B)$ is the probability that event A will occur assuming that event B has already occurred. A special case of this probability law is when A and B are independent events, i.e., the occurrence of one event has no influence on the probability of the other event occuring. Then

$$P(A \text{ and } B) = P(A) \times P(B) \tag{2.62}$$

Bayes' theorem on probability may be stated as

$$P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(\overline{A})P(B|\overline{A})} \tag{2.63}$$

As an example of the use of Bayes' theorem suppose that a company discovers that 80% of those who bought its product in a year had been on the company's training course. 30 percent of those who bought a competitor's product had also been on the same training course. During that year the company had 20% of the market. The company wishes to know what percentage of buyers actually went on its training course, in order to discover the effectiveness of this course.

If B denotes that a person bought the company's product and T that they went on the training course then the problem is to find $P(B|T)$. From the data $P(B) = 0.2$, $P(\overline{B}) = 0.8$, $P(T|B) = 0.8$, $P(T|\overline{B}) = 0.3$. Then from equation (2.63),

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

$$P(B|T) = \frac{0.2 \times 0.8}{0.2 \times 0.8 + 0.8 \times 0.3} = 0.4$$

2.63 PROBABILITY DISTRIBUTIONS

There are several mathematical formulae with well-defined characteristics and these are known as probability distributions. If a problem can be made to fit one of these distributions then its solution is simplified. Distributions can be discrete when the characteristic can only take certain specific values, such as 0, 1, 2, etc., or they can be continuous where the characteristic can take any value.

Binomial Distribution

The binomial probability distribution is given by

$$(p + q)^n = q^n + {}^nC_1 pq^{n-1} + {}^nC_2 p^2 q^{n-2} + \cdots + {}^nC_x p^x q^{n-x} + \cdots + p^n \quad (2.64)$$

where p is the probability of an event occurring, q $(= 1 - p)$ is the probability of an event not occurring and n is the number of selections.

The probability of an event occurring m successive times is given by the binomial distribution as

$$p(m) = {}^nC_m p^m q^{n-m} \quad (2.65)$$

The binomial distribution is used for discrete events and is applicable if the probability of occurrence p of an event is constant on each trial. The mean of the distribution $B(M)$ and the standard deviation $B(S)$ are given by

$$B(M) = np \quad (2.66)$$

$$B(S) = (npq)^{1/2} \quad (2.67)$$

Poisson Distribution

The Poisson distribution is used for discrete events and, like the binomial distribution, it applies to mutually independent events. It is used in cases where p and q cannot both be defined. For example, one can state the number of points which were scored in a football game, but not the goals which were not scored.

The Poisson distribution may be considered to be the limiting case of the binomial when n is large and p is small. The probability of an event occurring m successive times is given by the Poisson distribution as

$$p(m) = (np)^m \frac{e^{-np}}{m!} \quad (2.68)$$

The mean $P(M)$ and standard deviation $P(S)$ of the Poisson distribution are given by

$$P(M) = np \tag{2.69}$$

$$P(S) = (np)^{1/2} \tag{2.70}$$

Poisson probability calculations can be done by the use of probability charts as shown in Figure 2.29. This shows the probability that an event will occur at least m times when the mean (or expected) value np is known.

Normal Distribution

The normal distribution represents continuous events and is shown plotted in Figure 2.30. The x-axis gives the event and the y-axis the probability of the event occurring. The curve shows that most of the events occur close to the mean value and this is usually the case in nature. The equation of the normal curve is given by

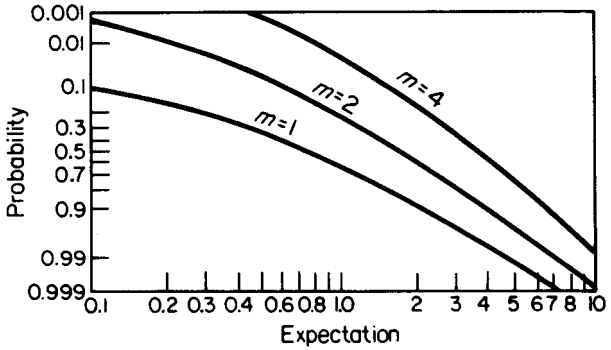


FIGURE 2.29 Poisson probability paper.

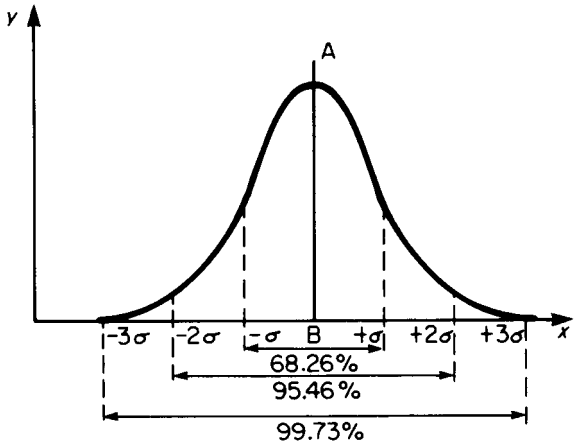


FIGURE 2.30 The normal curve.

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

$$y = \frac{1}{\sigma(2\pi)^{1/2}} e^{-(x-\bar{x})^2/(2\sigma^2)} \tag{2.71}$$

where $\bar{x}$ is the mean of the values making up the curve and σ is their standard deviation.

Different distributions will have varying mean and standard deviations but if they are distributed normally then their curves will all follow equation (2.71). These distributions can all be normalized to a standard form by moving the origin of their normal curve to their mean value, shown as B in Figure 2.30. The deviation from the mean is now represented on a new scale of units given by

$$\omega = \frac{x - \bar{x}}{\sigma} \tag{2.72}$$

The equation for the standardized normal curve now becomes

$$y = \frac{1}{(2\pi)^{1/2}} e^{-\omega^2/2} \tag{2.73}$$

The total area under the standardized normal curve is unity and the area between any two values of ω is the probability of an item from the distribution falling between these values. The normal curve extends infinitely in either direction but 68.26% of its values (area) fall between $\pm\sigma$, 95.46% between $\pm 2\sigma$, 99.73% between $\pm 3\sigma$ and 99.994% between $\pm 4\sigma$.

Table 2.1 gives the area under the normal curve for different values of ω. Since the normal curve is symmetrical the area from $+\omega$ to $+\infty$ is the same as from $-\omega$ to $-\infty$. As an example of the use of this table, suppose that 5000 street lamps have been installed in a city and that the lamps have a mean life of 1000 hours with a standard deviation of 100 hours. How many lamps will fail in the first 800 hours? From equation (2.72)

$$\omega = (800 - 1000)/100 = -2$$

Ignoring the negative sign, Table 2.1 gives the probability of lamps not failing as 0.977 so that the probability of failure is $1 - 0.977$ or 0.023. Therefore, 5000×0.023 or 115 lamps are expected to fail after 800 hours.

Exponential Distribution

The exponential probability distribution is a continuous distribution and is shown in Figure 2.31. It has the equation

$$y = \frac{1}{x} e^{-x/\bar{x}} \tag{2.74}$$

where $\bar{x}$ is the mean if the distribution. Whereas in the normal distribution the mean value divides the population in half, for the exponential distribution 36.8% of the population is above the average and 63.2% below the average. Table 2.2 shows the area under the exponential curve for different values of the ratio $K = x/\bar{x}$, this area being shown shaded in Figure 2.31.

TABLE 2.1 Area under the Normal Curve from $-\infty$ to ω

ω	0.00	0.02	0.04	0.06	0.08
0.0	0.500	0.508	0.516	0.524	0.532
0.1	0.540	0.548	0.556	0.564	0.571
0.2	0.579	0.587	0.595	0.603	0.610
0.3	0.618	0.626	0.633	0.640	0.648
0.4	0.655	0.663	0.670	0.677	0.684
0.5	0.692	0.700	0.705	0.712	0.719
0.6	0.726	0.732	0.739	0.745	0.752
0.7	0.758	0.764	0.770	0.776	0.782
0.8	0.788	0.794	0.800	0.805	0.811
0.9	0.816	0.821	0.826	0.832	0.837
1.0	0.841	0.846	0.851	0.855	0.860
1.1	0.864	0.869	0.873	0.877	0.881
1.2	0.885	0.889	0.893	0.896	0.900
1.3	0.903	0.907	0.910	0.913	0.916
1.4	0.919	0.922	0.925	0.928	0.931
1.5	0.933	0.936	0.938	0.941	0.943
1.6	0.945	0.947	0.950	0.952	0.954
1.7	0.955	0.957	0.959	0.961	0.963
1.8	0.964	0.966	0.967	0.969	0.970
1.9	0.971	0.973	0.974	0.975	0.976
2.0	0.977	0.978	0.979	0.980	0.981
2.1	0.982	0.983	0.984	0.985	0.985
2.2	0.986	0.987	0.988	0.988	0.989
2.3	0.989	0.990	0.990	0.991	0.991
2.4	0.992	0.992	0.993	0.993	0.993
2.5	0.994	0.994	0.995	0.995	0.995
2.6	0.995	0.996	0.996	0.996	0.996
2.7	0.997	0.997	0.997	0.997	0.997
2.8	0.997	0.998	0.998	0.998	0.998
2.9	0.998	0.998	0.998	0.998	0.999
3.0	0.999	0.999	0.999	0.999	0.999

Column 1 lists the ordinal values of ω or K and the corresponding values of area are presented in column 2. Interpolation between ordinal values can be achieved in steps of 0.02 by using the remaining 4 columns.

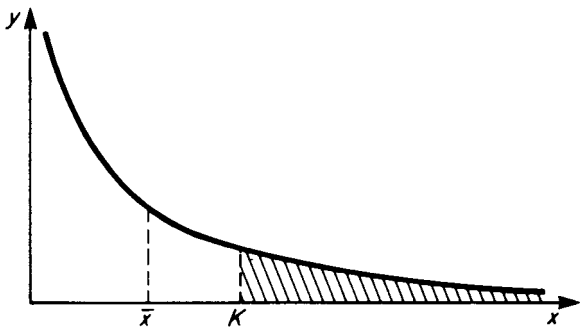


FIGURE 2.31 The exponential curve.

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

TABLE 2.2 Area under the Exponential Curve from K to $+\infty$

K	0.00	0.02	0.04	0.06	0.08
0.0	1.000	0.980	0.961	0.942	0.923
0.1	0.905	0.886	0.869	0.852	0.835
0.2	0.819	0.803	0.787	0.771	0.776
0.3	0.741	0.726	0.712	0.698	0.684
0.4	0.670	0.657	0.644	0.631	0.619
0.5	0.607	0.595	0.583	0.571	0.560
0.6	0.549	0.538	0.527	0.517	0.507
0.7	0.497	0.487	0.477	0.468	0.458
0.8	0.449	0.440	0.432	0.423	0.415
0.9	0.407	0.399	0.391	0.383	0.375

Column 1 lists the ordinal values of K and the corresponding values of area are presented in column 2. Interpolation between ordinal values can be achieved in steps of 0.02 by using the remaining 4 columns.

As an example suppose that the time between failures of a piece of equipment is found to vary exponentially. If results indicate that the mean time between failures is 1000 hours, then what is the probability that the equipment will work for 700 hours or more without a failure? Calculating K as $700/1000 = 0.7$ then from Table 2.2 the area beyond 0.7 is 0.497, which is the probability that the equipment will still be working after 700 hours.

Weinbull Distribution

This is a continuous probability distribution and its equation is given by

$$y = \alpha\beta(x - \gamma)^{\beta-1}e^{-\alpha(x-\gamma)^{\beta}} \tag{2.75}$$

where α is called the scale factor, β the shape factor and γ the location factor.

The shape of the Weibull curve varies depending on the value of its factors. β is the most important, as shown in Figure 2.32, and the Weibull curve varies from an exponential ($\beta = 1.0$) to a normal distribution ($\beta = 3.5$). In practice β varies from about ⅓ to 5. Because the Weibull distribution can be made to fit a variety of different sets of data, it is popularly used for probability distributions.

Analytical calculations using the Weibull distribution are cumbersome. Usually predictions are made using Weibull probability paper. The data are plotted on this paper and the probability predictions read from the graph.

2.64 SAMPLING

A sample consists of a relatively small number of items drawn from a much larger population. This sample is analyzed for certain attributes and it is then assumed that these attributes apply to the total population, within a certain tolerance of error.

Sampling is usually associated with the normal probability distribution and, based on this distribution, the errors which arise due to sampling can be estimated. Suppose a sample of n_s items is taken from a population of n_p items which are

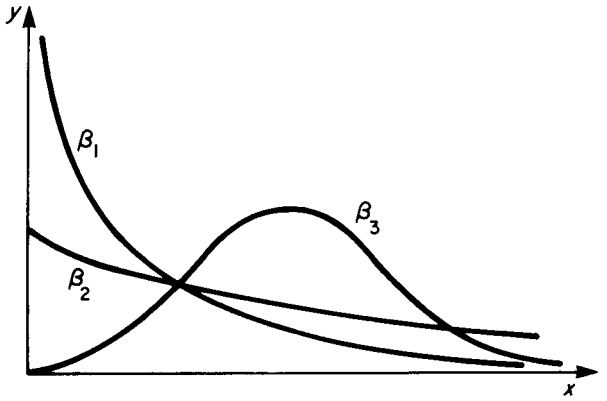


FIGURE 2.32 Weibull curves ($\alpha = 1$).

distributed normally. If the sample is found to have a mean of μ_s with a standard deviation of σ_s then the mean μ_p of the population can be estimated to be within a certain tolerance of μ_s. It is given by

$$\mu_p = \mu_s \pm \frac{\gamma\sigma_s}{n_s^{1/2}} \tag{2.76}$$

γ is found from the normal curve depending on the level of confidence we need in specifying μ_p. For $\gamma = 1$ this level is 68.26%; for $\gamma = 2$ it is 95.46% and for $\gamma = 3$ it is 99.73%.

The standard error of mean σ_e is often defined as

$$\sigma_e = \frac{\sigma_s}{n_s^{1/2}} \tag{2.77}$$

so equation (2.76) can be rewritten as

$$\mu_p = \mu_s \pm \gamma\sigma_e \tag{2.78}$$

As an example suppose that a sample of 100 items, selected at random from a much larger population, gives their mean weight as 20 kg with a standard deviation of 100 g. The standard error of the mean is therefore $100/(100)^{1/2} = 10$ g and one can say with 99.73% confidence that the mean value of the population lies between $20 \pm 3 \times 0.01$ or 20.03 kg and 19.97 kg.

If in a sample of n_s items the probability of occurrence of a particular attribute is p_s, then the standard error of probability p_e is defined as

$$p_e = \left(\frac{p_s q_s}{n_s}\right)^{1/2} \tag{2.79}$$

where $q_s = 1 - p_s$.

The probability of occurrence of the attribute in the population is then given by

$$p_p = p_s \pm \gamma \rho_e \tag{2.80}$$

where γ is again chosen to cover a certain confidence level.

As an example suppose a sample of 500 items shows that 50 are defective. Then the probability of occurrence of the defect in the sample is $50/500 = 0.1$. The standard error of probability is $(0.1 \times 0.9/500)^{1/2}$ or 0.0134. Therefore, we can state with 95.46% confidence that the population from which the sample was drawn has a defect probability of $0.1 \pm 2 \times 0.0134$, i.e. 0.0732 to 0.1268; or we can state with 99.73% confidence that this value will lie between $0.1 \pm 3 \times 0.0134$, i.e. 0.0598 to 0.1402.

If two samples have been taken from the same population and these give standard deviations of σ_{s1} and σ_{s2} for sample sizes of n_{s1} and n_{s2} then equation (2.77) can be modified to give the standard error of the difference between means as

$$\sigma_{de} = \left(\frac{\sigma_{s1}^2}{n_{s1}} + \frac{\sigma_{s2}^2}{n_{s2}} \right)^{1/2} \tag{2.81}$$

Similarly, equation (2.79) can be modified to give the standard error of the difference between probabilities of two samples from the same population as

$$p_{de} = \left(\frac{p_{s1} q_{s2}}{n_{s1}} + \frac{p_{s2} q_{s2}}{n_{s2}} \right)^{1/2} \tag{2.82}$$

2.65 TESTS OF SIGNIFICANCE

In taking samples we often obtain results which deviate from the expected. Tests of significance are then used to determine if this deviation is real or if it could have arisen due to sampling error.

Hypothesis Testing

In this system a hypothesis is formulated and is then tested at a given confidence level. For example, suppose a coin is tossed 100 times and it comes up heads 60 times. Is the coin biased or is it likely that this falls within a reasonable sampling error? The hypothesis is set up that the coin is not biased. Therefore one would expect that the probability of heads is 0.5, i.e. $p_s = 0.5$. The probability of tails, q_s, is also 0.5. Using equation (17.79) the standard error of probability is given by $p_e = (0.5 \times 0.5/100)^{1/2}$ or 0.05. Therefore, from equation (2.80) the population probability at the 95.45% confidence level of getting heads is $0.5 + 2 \times 0.05 = 0.6$. Therefore, it is highly likely that the coin is not biased and the results are due to sampling error.

The results of any significance test are not conclusive. For example, is 95.45% too high a confidence level to require? The higher the confidence level the greater the risk of rejecting a true hypothesis, and the lower the level the greater the risk of accepting a false hypothesis.

Suppose now that a sample of 100 items of production shows that five are defective. A second sample of 100 items its taken from the same production a few

months later and gives two defectives. Does this show that the production quality is improving? Using equation (2.82) the standard error of the difference between probabilities is given by $(0.5 \times 0.95/100 + 0.02 \times 0.98/100)^{1/2} = 0.0259$. This is less than twice the difference between the two probabilities, i.e., $0.05 - 0.02 = 0.03$, therefore the difference is very likely to have arisen due to sampling error and it does not necessarily indicate an improvement in quality.

Chi-Square Test

This is written as χ^2. If O is an observed result and E is the expected result then

$$\chi^2 = \sum \frac{(O - E)^2}{E} \tag{2.83}$$

The χ^2 distribution is given by tables such as Table 2.3, from which the probability can be determined. The number of degrees of freedom is the number of classes whose frequency can be assigned independently. If the data are presented in the form of a table having V vertical columns and H horizontal rows then the degrees of freedom are usually found as $(V - 1)(H - 1)$.

Returning to the earlier example, suppose a coin is tossed 100 times and it comes up heads 60 times and tails 40 times. Is the coin biased? The expected values for heads and tails are 50 each so that

$$\chi^2 = \frac{(60 - 50)^2}{50} + \frac{(40 - 50)^2}{50} = 4$$

The number of degrees of freedom is one since once we have fixed the frequency

TABLE 2.3 The Chi-Square Distribution

Degrees of freedom	Probability level				
	0.100	0.050	0.025	0.010	0.005
1	2.71	3.84	5.02	6.63	7.88
2	4.61	5.99	7.38	9.21	10.60
3	6.25	7.81	9.35	11.34	12.84
4	7.78	9.49	11.14	13.28	14.86
5	9.24	11.07	12.83	15.09	16.75
6	10.64	12.59	14.45	16.81	18.55
7	12.02	14.07	16.01	18.48	20.28
8	13.36	15.51	17.53	20.09	21.96
9	14.68	16.92	19.02	21.67	23.59
10	15.99	18.31	20.48	23.21	25.19
12	18.55	21.03	23.34	26.22	28.30
14	21.06	23.68	26.12	29.14	31.32
16	23.54	26.30	28.85	32.00	34.27
18	25.99	28.87	31.53	34.81	37.16
20	28.41	31.41	34.17	37.57	40.00
30	40.26	43.77	46.98	50.89	53.67
40	51.81	55.76	59.34	63.69	66.77

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

TABLE 2.4 Frequency Distribution of Accidents in a Factory During 24 Hours

Time (24-hour clock)	Number of accidents
0–6	9
6–12	3
12–18	2
18–24	6

for heads that for tails is defined. Therefore entering Table 2.3 with one degree of freedom the probability level for $\chi^2 = 4$ is seen to be above 0.025 (= 2.5%), i.e., there is a strong probability that the difference in the two results arose by chance and the coin is not biased.

As a further example suppose that over a 24-hour period the average number of accidents which occur in a factory is seen to be as in Table 2.4. Does this indicate that most of the accidents occur during the late night and early morning periods? Applying the χ^2 tests, the expected value, (if there was no difference between the time periods) would be the mean of the number of accidents, i.e., 5. Therefore, from equation (2.83),

$$\chi^2 = \frac{(9-5)^2}{5} + \frac{(3-5)^2}{5} + \frac{(2-5)^2}{5} + \frac{(6-5)^2}{5} = 6$$

There are three degrees of freedom, therefore from Table 2.3 the probability of occurrence of the result shown in Table 2.4 is seen to be greater than 10%. The conclusion would be that although there is a trend, as yet there are not enough data to show if this trend is significant or not. For example, if the number of accidents were each three times as large, i.e. 27, 9, 6, 18, respectively, then χ^2 would be calculated as 20.67 and from Table 2.3 it is seen that the results are highly significant since there is a very low probability, less than 0.5%, that it can arise by chance.

Significance of Correlation

The significance of the product moment correlation coefficient of equations (2.53) or (2.54) can be tested at any confidence level by means of the standard error of estimation given by equation (2.55). An alternative method is to use the Student t test of significance. This is given by

$$t = \frac{r(n-2)^{1/2}}{(1-r^2)^{1/2}} \tag{2.84}$$

where r is the correlation coefficient and n the number of items. Tables are then used, similar to Table 2.3, which give the probability level for $(n-2)$ degrees of freedom.

The Student t for the rank correlation coefficient is given by

$$t = R[(n-2)/(1-R^2)]^{1/2} \tag{2.85}$$

and the same Student t tables are used to check the significance of R.

PART 7

SI UNITS

Duncan T. Enright

The Systéme International d'Unités (SI) has been adopted and is defined by ISO 1000. Here the system is described and conversions to other commonly used systems are given.

SI comprises seven basic units from which a wide range of quantities can be derived in the form of products and quotients of these units which are shown in Table 2.5. The definitions of these units are as follows.

Meter (*m*). The meter is the length equal to 1,650,763.73 wavelengths in vacuum of the radiation corresponding to the transition between the levels $2p_{10}$ and $5d_5$ of the krypton-86 atom.

Kilogram (*kg*). The kilogram is the unit of mass; it is equal to the mass of the international prototype of the kilogram.

Second (*s*). The second is the duration of 9,192,631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the caesium-133 atom.

Ampere (*A*). The ampere is that constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross-section, and placed 1 m apart in vacuum, would produce between these conductors a force equal to 2×10^{-7} newtons per meter of length.

Kelvin (*K*). The kelvin, unit of thermodynamic temperature, is the fraction 1/273.16 of the thermodynamic temperature of the triple point of water.

Candela (*cd*). The candela is the luminous intensity, in the perpendicular direction, of a surface of 1/600,000 m^2 of a black body at the temperature of freezing platinum under a pressure of 101,325 newtons per square meter.

The supplementary base units are defined as follows:

Plane angle (*radian*). The angle subtended at the center of a circle of radius 1 m by an arc of length 1 m along the circumference.

TABLE 2.5 Basic SI Units

Quantity	Name of unit	Unit symbol
Length	meter	m
Mass	kilogram	kg
Time	second	5
Electric current	ampere	A
Thermodynamic temperature[a]	kelvin	K
Luminous intensity	candela	cd
Amount of substance	mole	mol

[a]Temperature difference is commonly expressed in degrees Celsius instead of degrees Kelvin. The unit of the temperature interval for these scales is the same: 0 K = −273.15°C; 273.15°K = 0°C.

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

Solid angle (*steradian*). The solid angle subtended at the center of a sphere of radius 1 m by an area of 1 m^2 on the surface.

Mole (*mol*) is the amount of substance of a system which contains as many elementary entities as there are atoms in 0.012 kg of carbon-12. The elementary entities must be specified and can be atom molecules, ion electrons, other particles, or specified groups of such particles.

2.66 DERIVED UNITS

SI is a rationalized and coherent system because, for any one physical quantity, it admits of only one measurement unit with its entire structure derived from no more than seven arbitrarily defined basic units. It is coherent because the derived units are always the products or quotients of two or more of these basic units. Thus the SI unit for velocity is m s^{-1} (meter per second) and for acceleration is m s^{-2} (meter per second every second). Special names (Table 2.6) have been given to some derived units as an aid to communication.

Although SI is complete in itself, certain non-SI units are recognized for use in conjunction with it where, for traditional, commercial or practical purposes, it is difficult to discard them. For example, it is impracticable to disregard the minute (in SI 60 seconds) and the hour (in SI 3600 seconds) which are noncoherent units.

2.67 GRAVITATIONAL AND ABSOLUTE SYSTEMS

There may be some difficulty in understanding the difference between SI and the Metric Technical System of units, which has been used principally in Europe. The

TABLE 2.6 Derived Units

Physical quantity	SI unit	Unit symbol
Force	newton	$N = kg\ ms^{-2}$
Work, energy quantity of heat	joule	$J = N\ m = kg\ m^2\ s^{-2}$
Power	watt	$W = J\ s^{-1} = kg\ m^2\ s^{-3}$
Electric charge	coulomb	$C = A\ s$
Electric potential	volt	$V = W\ A^{-1} = kg\ m^2\ A^{-1}\ s^{-3}$
Electric capacitance	farad	$F = A\ s\ V^{-1} = A^2s^4\ kg^{-1}\ m^{-2}$
Electric resistance	ohm	$\Omega = V\ A^{-1} = kg\ m^2\ s^{-3}$
Frequency	hertz	$Hz = s^{-1}$
Magnetic flux	weber	$Wb = kg\ m^2\ A^{-1}s^{-2}$
Magnetic flux density	tesla	$T = Wb\ m^{-2} = kg\ A^{-1}\ s^{-2}$
Inductance	henry	$H = kg\ m^2\ A^{-2}\ s^{-2}$
Luminous flux	lumen	$lm = cd\ sr$[a]
Illumination	lux	$lx = lm\ m^{-2}$

[a] One steradian (Sr) is the solid angle which, having its vertex at the center of a sphere, cuts off an area of the surface of the sphere equal to that of a square with sides of length equal to the radius of the sphere. The SI unit of electric dipole moment (A s m) is usually expressed as a coulomb meter (C m).

main difference is that while mass is expressed in kg in both systems, weight (representing a force) is expressed as kgf (a gravitational unit) in the MKSA system and as N in SI. An absolute unit of force differs from a gravitational unit of force because it induces unit acceleration in a unit mass whereas a gravitational unit imparts gravitational acceleration to a unit mass.

A comparison of the more commonly known systems and SI is shown in Table 2.7. It should be noted in particular how all energy and power, whether from a mechanical, electrical or heat source, share a common derived unit in the SI.

2.68 EXPRESSING MAGNITUDES OF SI UNITS

To express magnitudes of a unit, decimal multiples and submultiples are formed using the prefixes shown in Table 2.8. This method of expressing magnitudes ensures complete adherence to a decimal system.

2.69 RULES FOR USE OF SI UNITS AND THE DECIMAL MULTIPLES AND SUBMULTIPLES

1. The SI units are preferred, but it is impracticable to limit usage to these, therefore their decimal multiples and submultiples are also required. (For example, it is cumbersome to measure road distances or the breadth of a human hair in meters.)
2. In order to avoid errors in calculations it is preferable to use coherent units. Therefore, it is strongly recommended that in calculations only SI units themselves are used and not their decimal multiples and submultiples. (Example: use $N\ m^{-2} \times 10^6$ not $MN\ m^{-2}$ or $N\ mm^{-2}$ in a calculation.)
3. The use of prefixes representing 10 raised to a power which is a multiple of 3 is especially recommended. (Example: for length, km . . . m . . . mm. . . μm. Thus hm; dam; dm; cm are nonpreferred.)

TABLE 2.7 Commonly Used Units of Measurement

	SI (absolute)	FPS (gravitational)	FPS (absolute)	cgs (absolute)	Metric technical units (gravitational)
Length	meter (m)	ft	ft	cm	meter
Force	newton (N)	lbf	poundal (pdl)	dyne	kgf
Mass	kg	lb or slug	lb	gram	kg
Time	S	sec	sec	sec	sec
Temperature	°C K	°F	°F °R	°C K	°C K
Energy mech.	joule[a]	ft lbf	ft pdl	dyne cm = erg	kgf m
Energy heat		Btu	Btu	calorie	k cal.
Power mech.	watt	hp	hp	ergs	metric hp
Power elec.		watt	watt		watt
Electric current	amp	amp	amp	amp	amp
Pressure	$N\ m^{-2}$	$lbf\ ft^{-2}$	$pdl\ ft^{-2}$	$dyne\ cm^{-2}$	$kgf\ cm^{-2}$

[a] 1 joule = 1 newton meter or 1 watt second.

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

TABLE 2.8 The Internationally Agreed Multiples and Submutiples

Factor by which the unit is multiplied		Prefix	Symbol	Common examples
One million million	10^{12}	tera	T	
One thousand million	10^{9}	giga	G	gigahertz (GHz)
One million	10^{6}	mega	M	megawatt (MW)
One thousand	10^{3}	kilo	k	kilometer (km)
One hundred	10^{2}	hecto[a]	h	
Ten	10^{1}	deca[a]	da	decagram (dag)
UNITY	1			
One tenth	10^{-1}	deci[a]	d	decimeter (dm)
One hundredth	10^{-2}	centi[a]	c	centimeter (cm)
One thousandth	10^{-3}	milli	m	milligram (mg)
One millionth	10^{-6}	micro	μ	microsecond (μs)
One thousand millionth	10^{-9}	nano	n	nanosecond (ns)
One million millionth	10^{-12}	pico	p	picofarad (pF)
One thousand million millionth	10^{-15}	femto	f	
One million million millionth	10^{-18}	atto	a	

[a] To be avoided wherever possible.

4. When expressing a quantity by a numerical value of a unit it is helpful to use quantities resulting in numerical values between 0 and 1000. Examples:

   $12\ \text{kN} = 12 \times 10^{3}\ \text{N}$ instead of 12 000 N
   $3.94\ \text{mm} = 3.94 \times 10^{-3}\ \text{m}$ instead of 0.00394 m
   $14.01\ \text{kN m}^{-2} = 14.01 \times 10^{3}\ \text{N m}^{-2}$ instead of
   $14\ 010\ \text{N m}^{-2}$

5. Compound prefixes are not used. (Example: write nm not mμm.) Where, however, a name has been given to a product or a quotient of a basic SI unit (for example, the bar ($10^{5}\ \text{N m}^{-2}$)) it is correct practice to apply the prefix to the name (for example, millibar (10^{-3} bar)).

6. In forming decimal multiples and submultiples of a derived SI unit preferably only one prefix is used. The prefix should be attached to the unit in the numerator. (Example: MW m^{-2} not W mm^{-2}.) The exception is stress, where BSI recommend the use of N mm^{-2}.

7. Multiplying prefixes are printed immediately adjacent to the SI unit symbol with which they are associated. The multiplication of symbols is usually indicated by leaving a small gap between them. (Example: mN = millinewton. If written as m N this would indicate a meter newton.)

2.70 SI QUANTITIES, UNITS, AND SYMBOLS

Tables 2.9 to 2.25 contain lists of SI units used in certain fields together with other units or names of units which may be used and their values in SI units. The tables also show symbols typically used to describe quantities in the same fields.

The units shown in these tables generally obey the above guide rules. However, it is expected that a practical attitude should prevail in the interpretation of these recommendations, particularly as certain countries still use Metric Technical Units, which include such units as kgf cm^{-2}.

TABLE 2.9 Space and Time: Units

Quantity	SI units	Conversion factors and remarks on non-SI units
Plane angle[a]	radian (rad)	$1° = 0.017\ 453\ 3$ rad $1^+ = 2.908\ 88 \times 10^{-4}$ rad $1'' = 4.848\ 14 \times 10^{-6}$ rad 1 right angle = $\pi/2$ rad = 90°
Solid angle	sateradian (sr)	
Length[b]	meter (m)	1 inch = 25.4 mm 1 foot = 0.3048 m 1 yard = 0.9144 m 1 mile = 1.609.344 km 1 UK nautical mile = 1.853 18 km 1 angstrom (Å) = 10^{-10} meter or 10^{-1} nm
Area	square meter (m^2)	1 sq. in. = 645.16 mm^2 6.4516 cm^2 1 sq. ft = 0.092 903 m^2 1 sq. yd = 0.836 127 m^2 1 acre = 0.404 686 ha 1 sq. mile = 258.999 ha ha = hectare a = are
Volume[c]	cubic meter (m^5)	1 $in.^3$ = 16.3871 cm^3 1 ft^3 = 0.028 3168 m^3 1 yd^3 = 0.764 555 m^3 1 UK fl. oz = 28.4131 mlitre 1 gal = 4.5461 litre
Time	second (s)	1 day = 7.344 Ms 1 hour = 3600 s 1 minute = 60 s
Angular velocity	radian per second (rad s^{-1})	rev min^{-1} = 0.104 720 rad s^{-1} rev s^{-1} = 6.283 19 rad s^{-1} 1 degree s^{-1} = 0.017 4533 rad s^{-1}
Velocity	meter per second (m s^{-1})	1 km h^{-1} = 0.277 778 m s^{-1} 1 ft s^{-1} = 0.3048 m s^{-1} 1 mile h^{-1} = 0.447 04 m s^{-1} or 1.609 34 km h^{-1} knot (kn) 0.514 444 m s^{-1} 1 UK knot 0.514 773 m s^{-1}
Acceleration	meter per second squared (m s^{-2})	1 ft s^{-2} = 0.3048 m s^{-2}
Angular acceleration	radian per second squared (rad s^{-2})	1 in s^{-2} = 0.0254 m s^{-2}

[a] *Plane angle*. The Sumerian division of the circle in 360° (hence degrees) is retained for geometry, although dynamicists use the radian.

[b] *Length—the centimeter*. In many engineering disciplines the use of the centimeter is non-preferred. Sometimes the centimeter raised to a power (e.g., cm^2; cm^3; cm^4) is used to maintain a sensible range of numerical values in front of the unit. An instance of this concerns steel sections where the moduli of sections and moment of section may be given in steel tables of cm^3 and cm^4 respectively.

Where accuracy to the nearest millimeter is unwarranted, the centimeters can be used to imply a less precise dimension.

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

TABLE 2.9 Space and Time: Units (*Continued*)

[c] *Volume and capacity—the liter*. Before 1964 the 1901 liter was equal to 1.000 028 dm^3. At the XII International CGPM meeting on units (1964) the liter was redefined to equate exactly to 1 cubic decimeter. The same conference agreed that the liter should not be used to express the results of precise measurements, so as to make sure that where high precision was involved (say, greater than 1 part in 20 000) the possibility of confusion between the former (1901) liter and the new (1964) liter would be eliminated.

It is recommended that the results of precise measurements of volume be given only in terms of m^3, dm^3, cm^3, mm^3, etc. even though the milliter (ex. cm^1) and liter (ex dm^3) will still be used for operational and commercial purposes.

Because of the possible confusion of the symbol for the liter "I" with the figure "I," it is strongly recommended that the unit name be spelled in full.

Centiliter is sometimes used for arbitrary quantities implying a greater degree of tolerance.

TABLE 2.10 Space and Time: Symbols

Symbol	Quantity
θ (α, β, etc.)	plane angle
l	length
A	area
V	volume
t, T	time
ω	angular velocity
v, $\dot{s}$, ds/dt	velocity
a, $\ddot{s}$, d^2s/dt^2	acceleration
α, ω, $d\omega/dt$,	angular acceleration
s, x, y, z	distance or displacement

TABLE 2.11 Periodic and Related Phenomena: Units

Quantity	SI units	Units other than SI	Conversion factors and remarks on non-SI units
Frequency	hertz (Hz)		1 c/s (or c.p.s.) = 1 Hz
Rotational frequency[a]	reciprocal second (s^{-1})	rev min^{-1}	1 rev min^{-1} = 0.016 667 s^{-1}
Wavelength	meter (m)	Å[b]	1 Å = 10^{-10} m

[a] *Rotational frequency* (*rev/min*). The quantity rev/min is favored for rotating machinery.
[b] Electromagnetic radiative wavelengths are often quoted in angstroms (Å).

Mass and Weight

Confusion sometimes arises over the measuring of the terms *mass* and *weight*. Commonly, and in many branches of engineering, it has been the custom to refer to quantities of mass as weights, e.g., weight of coal in kilograms.

Weight, however, is dependent upon the gravitational force acting upon the mass. Thus for a mass (M), weight (W) = Mg, where g is the local acceleration due to gravity which varies slightly from point to point on the earth's surface. For practical purposes, an approximated figure of 9.81 or 9.807 meters per second squared (m s^{-2}) is used for g.

TABLE 2.12 Periodic and Related Phenomena: Symbols

Symbol	Quantity
T	periodic time
τ, (T)	time constant of an exponentially varying quantity
f, v	frequency
η	rotational frequency
ω	angular frequency
λ	wavelength
$\sigma(\dot{v})$	wavenumber
k	circular wavenumber
$\log_e (A_1/A_2)'$	natural logarithm of the ratio of two amplitudes
$10 \log_{10}(P_1/P_2)$	ten times the common logarithm of the ratio of two powers
δ	damping coefficient
Λ	logarithmic decrement
α	attenuation coefficient
β	phase coefficient
γ	propagation coefficient

The force unit in SI is the newton (N) and by using consistent units becomes the force applied to unit mass (kg) to impart unit acceleration (m s^{-2}) to the mass (as distinct from gravitational acceleration, which equals 9.806 65 m s^{-2}). Thus, it can be more readily understood by comparing the SI system with other systems for mass, weights, and measures as shown in Table 2.14.

Megagram

The tonne and kilogram are generally accepted as replacement units for ton and pounds. In particular, most lifting equipment already marked in tons can be considered as adequate for lifting the same number of tonnes because of the small excess (1.6%) of the ton over the tonne.

However, in soil mechanics the megagram (Mg) rather than the tonne is recommended. This is because with large masses involved in work on soil mechanics, confusion between the ton and the tonne could prove very expensive.

Pressure and Stress

The SI derived unit for force per unit area is the newton per square meter (N m^{-2}), referred to as the pascal (Pa), and this unit with suitable multiples is favored as the unit for stress. There are differences of opinion regarding the unit for pressure, but although some flexibility will have to be allowed in the expression of pressure values, the following practice should be adopted.

1. For the statement of stress property, use, without deviation, N m^{-2} and appropriate multiples of it (e.g., MN m^{-2}) or this, if preferred, expressed as N mm^{-2} or, if essential for nonmetallic materials, kN m^{-2}.
2. For pressure statements use either N m^{-2} (and suitable multiples and submultiples of it) or bar or mbar. In such cases the conversion 1 bar = 10^5 N m^{-2} will always be quoted for reference.

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

TABLE 2.13 Mechanics: Units

Quantity	SI units	Conversion factors and remarks on non-SI units
Mass	kilogram (kg)	1 ton = 1016.05 kg or 1.016 05 tonne 1 cwt = 50.8023 kg 1 lb = 0.453 592 37 kg 1 oz = 28.3495 g (avoir)
Mass density	kilogram per cubic meter (kg m^{-3})	1 lb/ft^3 = 16.0185 kg m^{-3} 1 lb/in^3 = 27.6799 g cm^{-3}
Specific volume	cubic meter per kilogram (m^3 kg^{-1})	1 ft^3/lb = 0.062 428 m^3 kg^{-1} 1 cm^3/g = 10^{-3} m^3 kg^{-1}
Momentum	kilogram meter per second (kg m s^{-1})	1 lb ft/s = 0.138 255 kg ms^{-1} 1 g cm/s = 10^{-5} kg m s^{-1}
Angular momentum	kilogram square meter per second (kg m^2 s^{-1})	1 lb ft^2/s = 0.042 140 1 kg m^2 s^{-1} 1 g cm^2/s = 10^{-7} kg m^2 s^{-1}
Moment of inertia	kilogram square meter (kg m^2)	1 lb ft^2 = 0.042 140 1 kg m^2 1 g cm^2 = 10^{-7} kg m^2
Force	newton (N)	1 tonf = 9.964 02 kN 1 lbf = 4.448 22 N 1 ozf = 0.278 014 N 1 pdl = 0.138 255 N 1 dyne = 10^{-5} N 1 kgf or kilopound = 9.806 65 N
Moment of force (torque)	newton meter (N m)	1 tonf ft = 3.037 03 kN m 1 lbf ft = 1.355 82 N m 1 pdl ft = 0.042 140 1 N m 1 lbf in. = 0.112 985 N m 1 dyne cm = 10^{-7} N m 1 kgf m = 9.806 65 N m
Mass per unit length	kilogram per meter (kg m^{-1})	1 ton/1000 yds = 1.111 16 kg m^{-1} 1 ton/mile = 0.631 342 kg m^{-1} 1 lb/in. = 17.8580 kg m^{-1} 1 lb/ft = 1.488 16 kg m^{-1} 1 lb/yd = 0.496 055 kg m^{-1}
Mass per unit area	kilogram per square meter (kg m^{-2})	1 lb/acra = 1.120 85 × 10^{-4} kg m^{-2} 1 ton/sq. mile = 3.922 98 × 10^{-4} kg m^{-2} 1 lb/1000 ft^2 = 4.882 43 kg m^{-2}
Mass rate of flow	kilogram per second (kg s^{-1})	1 lb/s = 0.453 592 kg s^{-1} 1 lb/h = 1.259 98 × 10^{-4} kg s^{-1} 1 UK ton/h = 0.282 235 kg s^{-1} or 1.016 05 tonne h^{-1}
Volume rate of flow	cubic meter per second (m^3 s^{-1})	1 ft^3/s (cusec) = 28.3168 × 10s-^{3} m^3 s^{-1} 1 gal/s = 4.546 09 × 10^{-3} m^3 s^{-1}
Mass flow rate per unit area	kilogram per square meter second (kg m^{-2} s^{-1})	1 lb/ft^2h = 1.356 23 × 10^{-3} kg m^{-2} s^{-1}

TABLE 2.13 Mechanics: Units (*Continued*)

Quantity	SI units	Conversion factors and remarks on non-SI units
Pressure and stress	newton per square meter ($N\ m^{-2}$) or pascal (Pa)	1 bar = 10^5 N m^{-2} 1 N/m^2 = 1 Pa 1 $lbf/in.^2$ = 6.894 76 kN m^{-2} 1 torr = 1.333 22 mbar = 133.322 N m^{-2} 1 in. Hg = 3386.39 N m^{-2} 1 in. W.G. = 2.490 89 mbar 1 kgf/cm^2 = 0.980 665 bar = 98.0665 kN m^{-2} 1 $tonf/in.^2$ = 15.4443 N mm^{-2} 1 pieze = 10^3 N m^{-2} 1 std atmosphere = 1013.25 × 10^2 N m^{-2} = 1.033 23 kgf cm^{-2} = 14.695 lbf in^{-2} = 760 torr+ = 29.921 3 in Hg
Second moment of area	meter to the power of four (m^4)	1 $in.^4$ = 41.6231 cm^4 1 ft^4 = 863 097 cm^4
Section modulus	cubic meter (m^3)	1 $in.^3$ = 16.3871 cm^3
Dynamic viscosity	newton second per square meter ($N\ sm^{-2}$)	1 P(poise) = 10^{-1} Ns m^{-2} 1 lbf s/ft^2 = 47.8803 N s m^{-2} or 47 880.3 cP 1 pdl s/ft^2 } = 1.488 16 N s m^{-2} or lb/ft s } or 1488.16 cP
Kinematic viscosity	square meter per second (m^2s^{-1})	1 St (stokes) = 10^{-4} m^2 s^{-1} 1 ft^2/h = 2.580 64 × 10^{-5} m^2 s^{-1} or 25.806 cSt 1 ft^2s = 0.092 903 m^2 s^{-1} or 9.2903 × 10^4 cSt
Surface tension	newton per meter ($N\ m^{-1}$)	1 lbf/ft = 14.5939 N m^{-1} 1 dyne/cm = 10^{-3} N m^{-1} 1 ft lbf = 1.355 82 J 1 erg = 10^{-7} J
Energy, work	joule (J = N m)	1 kgf m = 9.806 65 J 1 ft pdl = 0.042 140 1 J 1 hp h = 2.684 52 × 10^6 J 1 kcal = 4186.8 J 1 Btu = 1055.06 J 1 eV (electron volt) = (1.602 10 ± 0.00007) × 10^{-19} J 1 kWh = 3.6 × 10^6 J
Power	watt ($W = Js^{-1}$)	1 hp = 745.7 W 1 ft lbf/s = 1.355 82 W 1 metric hp = 735.499 W 1 kg m/s = 9.806 65 W 1 erg/s = 10^{-7} W
Impact strength	joule per square meter (J m^{-2})	
Fuel consumption		1 gal/mile = 2.825 liter km^{-1} 1 mile/gal = 0.354 km $liter^{-1}$
Specific fuel consumption	kilogram per joule (kgJ^{-1}) cubic meter per joule (m^3J^{-1})	1 lb/hp h = 0.168 97 kg MJ^{-1} 1 pint/hp h = 0.211 68 liter MJ^{-1}

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

TABLE 2.14 Mechanics: Symbols

Symbol	Quantity
m	mass
e, ρ	density (mass density)
d	relative density
v	specific volume
p	momentum
b, p_0, p_θ	moment of momentum (angular momentum)
I, J	moment of inertia (dynamic moment of inertia)
F	force
$G(P, W)$	weight
γ	specific weight (weight density)
M	moment of force
M	bending moment
T	torque, moment of a couple
p	pressure
σ	normal stress
τ	shear stress
e, ε	linear strain (relative elongation)
γ	shear strain (shear angle)
$\Theta\theta$	volume strain (bulk strain)
μ, v	Poisson's ratio Poisson's number
E	Young's modulus (modulus of elasticity)
G	shear modulus (modulus of rigidity)
K	bulk modulus (modulus of compression)
x, κ	compressibility (bulk compressibility)
I, I_a	second moment of area (second axial moment of area)
I_p, J	second polar moment of area
Z, $W\left(\frac{I}{v}\right)$	section modulus
$\mu(f)$	coefficient of friction (factor of friction)
$\eta(\mu)$	viscosity (dynamic viscosity)
γ	kinematic viscosity
$\sigma(\gamma)$	surface tension
A, W	work
E, W	energy
E_p, U, V, Φ	potential energy
E_k, K, T	kinetic energy
p	power

3. Pressures or pressure differences measured by manometer tube may often conveniently be expressed as a height of a column of fluid, the nature of the fluid being stated. Such readings must be converted to terms of N m^{-2} if they are to be used in calculations of flow, etc. On the other hand, manometers are sometimes used merely as indicators that a prescribed operating condition has been met. Judgement is therefore required as to when it can be of advantage to use mm H_20, mm Hg, etc. or when it is of advantage to calibrate and read manometers in a suitable multiple of N m^{-2} or in mbar. It is understood that manometers

TABLE 2.15 Systems of Weights and Measures

Quantity	Foot pound second	Metric, technical	SI
Mass	1 lb	1 kg	1 kg
Length	1 ft	1 m	1 m
Force	1 lbf	1 kgf	1 N
Definition of force	$\text{lbf} = \frac{\text{lb} \times \text{ft s}^{-2}}{\text{g}}$	$\text{kgf} = \frac{\text{kg} \times \text{m s}^{-2}}{\text{g}}$	$\text{N} = \text{kg} \times \text{m s}^{-2}$
Definition of weight (gravitational force)	1 lbf per 1 lb	1 kgf per 1 kg	9.806 65 N per 1 kg

TABLE 2.16 Gauge Vacuum and Pressure

	Vacuum			Gauge pressure			
Ins. Hg	30 in Hg	20 in Hg	10 in Hg	10 lbf/in^2	100 lbf/in^1	1000 lbf/in^2	2000 lbf/in^2
kN/m^2	−101.3	−67.73	−33.86	68.94	689.4	6894	13789.5
	−1.01325	−0.6773	−0.3386				
bar	(−1013.25 mbar)	(−677.3 mbar)	(−338.6 mbar)	0.6894	6.894	68.94	137.895
MN/m^2	−0.1013	−0.0677	−0.0338	0.0689	0.6894	6.894	13.7895

calibrated in mbar are becoming increasingly available and it is recommended that pressures expressed as a height of a column of fluid should progressively give place to a suitable multiple of the SI unit or to the millibar.

4. Pressure units themselves are often not modified to indicate whether the pressure value is "absolute" (i.e., above zero) or "gauge" (i.e., above atmospheric pressure). If, therefore, the context leaves any doubt as to which is meant, the word 'pressure' must be qualified appropriately.
   e.g., ". . . at a gauge pressure of 12.5 bar"
   or ". . . at a gauge pressure of 1.25 MN m^{-2}"
   or ". . . at an absolute pressure of 2.34 bar"
   or ". . . at an absolute pressure of 234 kN m^{-2}"

Table 2.16 illustrates some of these practices. Note that this table works on a gauge pressure basis (atmosphere = 0), thus vacuum is shown measured in negative millibars. This continues the custom of associating the higher numerical readings with greater vacuum.

Notwithstanding previous practices of referring to pump performances in terms of pressure, the pump total head should be specified in linear measure (meters).

Viscosity

The recognized derived SI units for dynamic and kinematic viscosity are Ns m^{-2} and m^2 s^{-1} respectively. However, the existing units, centipoise (cP), and centistoke (cSt), are so well established internationally, particularly for oils, that the operational use of these units will continue.

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

TABLE 2.17 Heat: Units

Quantity	SI units	Conversion factors and remarks on non-SI units
Absolute temperature	kelvin (K)	K = °C + 273.15 K = 1.8°R (Rankine)
Customary temperature		°C = 5/9 (°F − 32)
Temperature interval	kelvin (K)	1°C = 1 K = 1.8°F (alternative form 1 deg C = 1 deg K = 1.8 deg F)
Temperature coefficient (linear or volumetric)	($1\ K^{-1}$)	
Heat, quantity of heat, internal energy, enthalpy	joule (J)	1 Btu = 1055.06 J 1 cal (IT) = 4.1868 J 1 CHU = 1899.2 J 1 kWh = 3.6 MJ 1 therm = 105.506 MJ 1 erg = 10^{-7} J
Heat flow rate	watt (W)	1 Btu/h = 0.293 071 W 1 kcal/h = 1.163 W 1 cal/s = 4.1868 W 1 frigorie = 4.186 W
Density of heat flow rate	watt per square (meter (W m^{-2})	1 Btn/ft^2 h = 3.154 59 W m^{-2} 1 cal/cm^2 s = 41 868 W m^{-2} 1 CHU/ft^2 h = 5.678 W m^{-2}
Thermal conductivity	watt per meter kelvin (W m^{-1} K^{-1})	1 But/ft h °F = 1.730 73 W m^{-1} $°C^{-1}$ 1 kcal/m h °C = 1.163 W m^{-1} $°C^{-1}$ knows as *k* value
Coefficient of heat transfer	watt per square meter kelvin (W m^{-2} K)	1 Btu/ft^2 h °F = 5.678 26 W m^{-2} $°C^{-1}$ 1 cal/cm^2 s °C = 41 868 W m^{-2} $°C^{-1}$ 1 kcal/m^2 h °C = 1.163 W m^{-2} $°C^{-1}$ knows as *U* value
Heat capacity	joule per kelvin (J K^{-1})	1 Btu = deg R = 1899.11 J $°C^{-1}$ 1 cal/g °C = 4.1868 J K^{-1}
Specific heat capacity	joule per kilogram kelvin (J kg^{-1} K^{-1})	1 Btu/lb °F = 4.1868 kJ kg^{-1} $°C^{-1}$ 1 cal/g °C = 4.1868 J kg^{-1} $°C^{-1}$
Entropy	joule per kelvin (J K^{-1})	1 Btu/°R = 1899.11 J K^{-1}
Specific entropy	joule per kilogram per kelvin (J kg^{-1} K^{-1})	1 Btu/lb °F = 4.1868 kJ kg^{-1} K^{-1} 1 cal/g K = 4.1868 kJ kg^{-1} K^{-1}
Specific energy	joule per kilogram (J kg^{-1})	1 Btu/lb = 2.326 kJ kg^{-1} 1 cal/g = 4.1868 kJ kg^{-1}
Specific enthalpy, specific latent heat	joule per kilogram (J kg^{-1})	1 Btu/lb = 2.326 kJ kg^{-1})
Specific heat content (i) Mass basis	joule per kilogram (J kg^{-1})	1 kcal kg^{-1} = 4.1868 kJ kg^{-1} 1 Btu/lb = 2.326 kJ kg^{-1} 1 CHU/lb = 4.186 816 kJ kg^{-1} 1 therm/ton = 103.84 kJ kg^{-1}
(ii) Volume basis	joule per cubic meter (J m^{-3})	1 Btu/ft^3 = 37.2592 jK m^{-3} 1 Btu/gal = 0.232 08 kJ $liter^{-1}$ 1 therm/UK gal = 23.208 GJ m^{-3} 1 cal/cm^3 = 4.1868 MJ m^{-3} 1 kcal/m^3 = 4.1868 kJ m^{-3}
Heat release rate	watt pere cubic meter (W m^{-3})	1 Btu/ft^2 s = 37.2589 kW m^{-3} 1 cal/cm^3 h = 1.163 kW m^{-3}

TABLE 2.18 Heat: Symbols

Symbol	Quantity
T, Θ	thermodynamic temperature absolute temperature
t, θ	customary temperature
α, λ	linear expansion coefficient
α, β, γ	cubic expansion coefficient
β	pressure coefficient
Q	heat, quantity of heat
$\phi(q)$	heat flow rate
$q(\phi)$	density of heat flow rate
$\lambda(k)$	thermal conductivity
h, k, U, α	coefficient of heat transfer
$\sigma(\alpha, x, k)$	thermal diffusivity
C	heat capacity
c	specific heat capacity
c_p	specific heat capacity at constant pressure
c_V	specific heat capacity at constant volume
γ, x, k	ratio of the specific heat capacities
S	entropy
s	specific entropy
$U(E)$	internal energy
$H(I)$	enthalpy
F	free energy
G	Gibbs function
$u(e)$	specific internal energy
$h(i)$	specific enthalpy
f	specific free energy
g	specific Gibbs function
L	latent heat
l	specific latent heat

TABLE 2.19 Electricity and Magnetism: Units

Quantity	SI units	Conversion factors and remarks on non-SI units
Electric current	ampere (A)	1 emu = 10 A 1 esu = $\frac{1}{3} \times 10^{-9}$ A
Electric charge	coulomb (C)	1 Ah = 3600 C C = As
Charge density	coulomb per cubic meter (C m^{-3})	1 emu = 10^7 C m^{-3} 1 esu = $\frac{1}{3} \times 10^{-3}$ C m^{-3}
Surface density of charge	coulomb per square meter (C m^{-2})	
Electric field strength	coulomb per square meter (C m^{-2})	
	volt per meter (V m^{-1})	
Electric potential	volt (V)	
Displacement	coulomb per square meter (C m^{-2})	C m^{-2} = A s m^{-2}

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

TABLE 2.19 Electricity and Magnetism: Units (*Continued*)

Quantity	SI units	Conversion factors and remarks on non-SI units
Electric flux	coulomb (C)	
Capacitance	farad (F)	$F = A\ s\ V^{-1} = C\ V^{-1}$
Permittivity	farad per meter (F/m)	$\varepsilon_0 = 8.854 \times 10^{-12}$ F m^{-1}
Electric polarization	coulomb per square meter (C m^{-2})	
Electric dipole moment	coulomb meter (C m)	
Current density	ampere per square meter (A m^{-2})	
Linear current density	ampere per meter (A m^{-1})	
Magnetic field strength	ampere per meter (A m^{-1})	1 oersted = $10^3/4\pi$ A m^{-1}
Magnetic potential difference	(A)	1 gilbert = $10/4\pi$ A
Magnetic flux density	tesla (T)	Wb m^{-2} = T 1 gauss = 10^{-4} T
Magnetic flux	weber (Wb)	V s = Wb 1 maxwell = 10^{-8} Wb
Magnetic vector potential	weber per meter (Wb m^{-1})	1 maxwell cm^{-1} = 10^{-6} Wb m^{-1}
Mutual inductance, self-inductance	henry (H)	H = V s A^{-1}
Permeability	henry per meter (H m^{-1})	$\mu_0 = 4\pi \times 10^{-7}$ H m^{-1}
Magnetic moment	ampere square meter (A m^2)	
Magnetization	ampere per meter (A m^{-1})	1 oersted = $10^3/4\pi$ A m^{-1}
Magnetic polarization	tesla (T)	1 gauss = 10^{-4} T
Magnetic dipole moment	newton square meter per ampere (N m^2 A^{-1})	
Resistance	ohm (Ω)	
Conductance	reciprocal ohm ($1/\Omega$)	S = mho = Siemen
Resistivity	ohm meter (Ωm)	
Conductivity	reciprocal ohm meter (1 Ω^{-1} m^{-1})	
Reluctance	reciprocal henry (1 H^{-1})	
Permeance	henry (H)	
Impedance Reactance	ohm (Ω)	
Conductance	reciprocal ohm (1 Ω^{-1})	
Active power	watt (W)	
Apparent power	volt ampere (VA)	
Reactive power	var	
Electric stress	volt per meter (V m^{-1})	1 kV in.$^{-1}$ = 0.039 370 1 kV mm^{-1} or 1 V mil^{-1} = 39.370 1 kV m^{-1}

TABLE 2.20 Electricity and Magnetism: Symbols

Symbol	Quantity
I	electric current
Q	electric charge, quantity of electricity
e	volume density of charge, charge density
σ	surface density of charge
E, (K)	electric field strength
V, ϕ	electric potential
U, (V)	potential difference, tension
E	electromotive force
D	displacement (rationalized displacement)
D'	nonrationalized displacement
Ψ	electric flux, flux of displacement (flux of rationalized displacement)
Ψ'	flux of non-rationalized displacement
C	capacitance
ε	permittivity
ε_0	permittivity of vacuum
ε'	nonrationalized permittivity
ε'_0	nonrationalized permittivity of vacuum
ε_r	relative permittivity
χ_e	electric susceptibility
χ'_e	nonrationalized electric susceptibility
P	electric polarisation
p, (p_e)	electric dipole moment
J, (S)	current density
A, (α)	linear current density
H	magnetic field strength
H'	nonrationalized magnetic field strength
U_m	magnetic potential difference
F, F_m	magnetomotive force
B	magnetic flux density, magnetic induction
Φ	magnetic flux
A	magnetic vector potential
L	self-inductance
M, L_{12}	mutual inductance
k, (x, k)	coupling coefficient
σ	leakage coefficient
μ	permeability
μ_0	permeability of vacuum
μ'	nonrationalized permeability
μ'_0	nonrationalized permeability of vacuum
μ_r	relative permeability
x, k	magnetic suceptibility
x', k'	nonrationalized magnetic susceptibility
m	electromagnetic moment (magnetic moment)
H_i, (M)	magnetization
B_i, (J)	magnetic polarization
J'	nonrationalized magnetic polarization
ω	electromagnetic energy density
S	Poynting vector
c	velocity of propagation of electromagnetic waves in vacuo
R	resistance (to direct current)
G	conductance (to direct current)
ρ	resistivity

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

TABLE 2.20 Electricity and Magnetism: Symbols (*Continued*)

Symbol	Quantity
y, σ	conductivity
R, R_m	reluctance
A, (P)	permeance
N	number of turns in winding
m	number of phases
p	number of pairs of poles
ϕ	phase displacement
Z	impedance (complex impedance)
$[Z]$	modulus of impedance (impedance)
X	reactance
R	resistance
Q	quality factor
Y	admittance (complex admittance)
$[Y]$	modulus of admittance (admittance)
B	susceptance
G	conductance
P	active power
S, (P_s)	apparent power
Q (P_q)	reactive power

TABLE 2.21 Light: Units

Quantity	SI units	Conversion factors and remarks on non-SI units
Luminous intensity	candela (cd)	
Luminous flux	lumen (Im)	lm = cd sr (candela steradian)
Illuminance	lux (lx)	lx = lm m^{-2}
Luminance	candela per square meter (cd m^{-2})	stilb = 1 cd cm^{-2} apostilb = π^{-1} cd m^{-2} 1 cd/in^2 = 1550 cd m^{-2} 1 foot lambert = 3426 cd m^{-2} 1 lambert = 3183 cd m^{-2}

TABLE 2.22 Light: Symbols

Symbol	Quantity
E	illuminance
r	reflectance factor
V	Munsell value
l	luminous intensity
n	refractive index
ϕ	luminous flux
w	beamwidth

TABLE 2.23 Sound: Units

Quantity	SI units	Conversion factors and remarks on non-SI units
Sound intensity	watt per square meter ω m^{-2})	1 erg s^{-1} cm^{-2} = 10^{-3} W m^{-2}
Sound intensity (logarithmic)		1 decibel (dB) = 20 $1\log_{10}(P/P_0)$ where P = measured sound pressure and P_0 = reference sound pressure of 2×10^{-5}N m^{-2}
Loudness		1 phon
Attenuation		neper per meter (np m's-1)

TABLE 2.24 Sound: Symbols

Symbol	Quantity
T	period, periodic time
f, ω	frequency, frequency interval
ω	angular frequency, circular frequency
λ	wavelength
k	circular wave number
ρ	density (mass density)
P_s	static pressure
p	(instantaneous) sound pressure
ε, (x)	(instantaneous) sound particle displacement
u, v	(instantaneous) sound particle velocity
a	(instantaneous) sound particle acceleration
q, U	(instantaneous) volume velocity
c	velocity of sound
E	sound energy density
P, (N, W)	sound energy flux, sound power
I, J	sound intensity
Z_s, (W)	specific acoustic impedance
Z_a, (Z)	acoustic impedance
Z_m, (w)	mechanical impedance
L_p, (L_N, L_w)	sound power level
L_p, (L)	sound pressure level
δ	damping coefficient
Λ	logarithmic decrement
α	attenuation coefficient
β	phase coefficient
γ	propagation coefficient
δ	dissipation coefficient
r, τ	reflection coefficient
γ	transmission coefficient
α, (α_a)	acoustic absorption coefficient
R	sound reduction index sound transmission loss
A	equivalent absorption area of a surface or object
T	reverberation time
L_N, (Λ)	loudness level
N	loudness

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

Energy

The choice of a suitable commercial energy unit common to all energy-producing concerns has still to be resolved. The SI unit is the joule and its multiples. However, electrical interests favor the adoption of the kWh (3.6 MJ = 1 kWh). The following are the probable commercial field quantities:

Coal	tonne
Electricity	kWh
Gas	100 MJ (the therm)
Oil	litre; m^3 kg; tonne

Hardness Values

The hardness unit kgf is used to express the load applied by the indenter and this ensures that most hardness-testing machines and empirically based formulae are not made obsolete. Thus the Rockwell, Vickers, and Brinell hardness numbers are used. This number is arbitrary and dimensionless and is dependent upon the resistance offered by the material under test to a definite load.

Concentration

Concentration should preferably be expressed on a mass/mass basis (i.e., kg kg^{-1}; mg kg^{-1}) or a volume/volume basis (i.e. m^3 m^{-3}; liter m^{-3}; milliliter m^{-3}). It may also be expressed in parts per million (ppm) or as a percentage ''by mass'' or ''by volume,'' respectively.

pH Scale

This is a number based on the logarithm, to the base 10 of the reciprocal of the concentration of hydrogen ions in aqueous solution. It is used as a method of expressing small differences in the acidity or alkalinity of nearly neutral solutions in biological and electrolytic processes.

Quantities and Units of Light

The following definitions are based on the International Lighting Vocabulary.

Luminous flux (symbol ϕ): The light emitted by a source such as a lamp or received by a surface, irrespective of direction.

Lumen (abbreviation lm): The SI unit of luminous flux used in describing the total light emitted by a source or received by a surface. (A 100-watt incandescent lamp emits about 1200 lumens.)

Illumination: The process of lighting an object.

Illumination value (symbol E): The luminous flux incident on a surface, per unit area.

Lux (abbreviation lx): The SI unit of illumination value; it is equal to one lumen per square meter.

Lumen per square foot (abbreviation lm ft^{-2}): A nonmetric unit of illumination value, equal to 10.76 lux. (Previously called the foot-candle, a term still used in some countries.)

Service value of illumination: The mean value of illumination throughout the life of an installation and averaged over the working area.

Initial value of illumination: The mean value of illumination averaged over the working area before depreciation has started, i.e., when the lamps and fittings are new and clean and when the room is freshly decorated.

Mean spherical illumination (scalar illumination): The average illumination over the surface of a small sphere centered at a given point; more precisely, it is the flux incident on the surface of the sphere divided by the area of the sphere. The term *scalar* illumination is the lux: care is needed to avoid confusing the unit with the illumination on a plane which is measured in the same unit.

Illumination vector: A term used to describe the flow of light. It has both magnitude and direction. The magnitude is defined as the maximum difference in the value of illumination at diametrically opposed surface elements of a small sphere centered at the point under consideration. The direction of the vector is that of the diameter joining the brighter to the darker element.

Luminous intensity: The quantity which describes the illuminating power of a source in a particular direction. More precisely, it is the luminous flux emitted within a very narrow cone containing that direction divided by the solid angle of the cone.

Candela (abbreviation cd): The SI unit of luminous intensity. The term *candle power* designates a luminous intensity expressed in candelas.

PART 8

CONVERSION OF EXISTING IMPERIAL TERMS

Duncan T. Enright

If it is necessary to convert existing imperial terms to a metric equivalent, care should be taken to ensure that the converted value implies the same degree of accuracy. The conversion factor must convey the same order of precision as the original value. Thus to translate 1 in as 25.4 mm or 1000 ft as 304.8 m conveys a tolerance which, in most cases, would be too precise.

Particular care is needed when converting machined tolerances. With a simple dimension such as 0.836 in., it is reasonable to assume that this dimension can be met because an imperial micrometer can measure to 0.001 in. The conversion factor for 0.836 in. is 21.2344 mm. Thus 0.004 mm in our conversion represents an accuracy of 0.000 016 in., which is beyond the scope of most toolroom measuring devices. In such cases it should be borne in mind that a metric micrometer can measure to 0.01 mm and with a vernier attachment to 0.002 mm. Thus for the greatest possible accuracy our converted readings should be 21.234 mm. Table 2.25 gives some metric to other unit conversion factors.

TABLE 2.25 Metric Conversion Factors

SI units	Other units
SPACE AND TIME	
Length:	
1 m (micron)	$= 39.37 \times 10^{-6}$ in.
1 mm	= 0.039 370 1 in.
1 cm	= 0.393 701 in.
1 m	= 3.280 84 ft
1 m	= 1.093 61 yd
1 km	= 0.621 371 mile
Area:	
1 mm^2	$= 1.550 \times 10^{-3}$ $in.^2$
1 cm^2	= 0.1550 $in.^2$
1 m^2	= 10.7639 ft^2
1 m^2	= 1.195 99 yd^2
1 ha	= 2.471 05 acre
Volume:	
1 mm^3	$= 61.0237 \times 10^{-6}$ $in.^3$
1 cm^3	$= 61.0237 \times 10^{-3}$ $in.^3$
1 m^3	= 35.3147 ft^2
1 m^3	= 1.307 95 yd^3
Capacity:	
10^6 m^3	$= 219.969 \times 10^6$ gal
1 m^3	= 219.969 gal
1 litre (1)	= 0.219 969 gal = 1.759 80 pint

TABLE 2.25 Metric Conversion Factors (*Continued*)

SI units	Other units
Capacity flow:	
10^3 m^{-3} s^{-1}	= 791.9 × 10^6 gal h^{-1}
1 m^3 s^{-1}	= 13.20 × 10^3 gal min^{-1}
1 liter s^{-1}	= 13.20 gal min^{-1}
1 m^3 k^{-1} W h	= 219.969 gal k^{-1} W h
1 m^3 s^{-1}	= 35.3147 ft^3 s^{-1} (cusecs)
1 liter s^{-1}	= 0.588 58 × 10^{-3} ft^3 min^{-1} (cfm)
Velocity:	
1 m s^{-1}	= 3.280 84 ft s^{-1} = 2.236 94 mile h^{-1}
1 km h^{-1}	= 0.621 371 mile h^{-1}
Acceleration:	
1 m s^{-2}	= 3.280 84 ft s^{-2}
MECHANICS	
Mass:	
1 g	= 0.035 274 oz
1 kg	= 2.204 62 lb
1 t	= 0.984 207 ton = 19.6841 cwt
Mass flow:	
1 kg s^{-1}	= 2.204 62 lb s^{-1} = 7.936 64 klb h^{-1}
Mass density:	
1 kg m^{-3}	= 0.062 428 lb ft^{-3}
1 kg $liter^{-1}$	= 10.022 119 lb gal^{-1}
Mass per unit length:	
1 kg m^{-1}	= 0.671 969 lb ft^{-1} = 2.015 91 lb yd^{-1}
Mass per unit area:	
1 kg m^{-2}	= 0.204 816 lb ft^{-2}
Specific volume:	
1 m^3 kg^{-1}	= 16.0185 ft^3 lb^{-1}
1 liter $tonne^{-1}$	= 0.223 495 gal ton^{-1}
Momentum:	
1 kg m^2 s^{-1}	= 7.233 01 lbft s^{-1}
Angular momentum:	
1 kg m^2 s^{-1}	= 23.7304 $lbft^2$ s^{-1}
Moment of inertia:	
1 kg m^2	= 23.7304 $lbft^2$
Force:	
1 N	= 0.224 809 lbf
Weight (force) per unit length:	
1 N m^{-1}	= 0.068 521 8 lb ft^{-1} = 0.205 566 lbf yd^{-1}
Moment of force (or torque):	
1 Nm	= 0.737 562 lbf ft
Weight (force) per unit area:	
1 N m^{-2}	= 0.020 885 lbf ft^{-2}
Pressure:	
1 N m^{-2}	= 1.450 38 × 10^{-4} lbf $in.^{-2}$
1 bar	= 14.5038 lbf $in.^{-2}$
1 bar	= 0.986 923 atmosphere
1 mbar	= 0.401 463 in H_2O
	= 0.029 53 in Hg

TABLE 2.25 Metric Conversion Factors (*Continued*)

SI units	Other units
Stress:	
1 N mm^{-2}	= 6.474 90 × 10^{-2} tonf $in.^{-2}$
1 MN m^{-2}	= 6.474 90 × 10^{-2} tonf $in.^{-2}$
1 hbar	= 0.647 490 tonf $in.^{-2}$
Second moment of area:	
1 cm^4	= 0.024 025 $in.^4$
Section modulus:	
1 m^3	= 61 023.7 $in.^3$
1 cm^3	= 0.061 023 7 $in.^3$
Kinematic viscosity:	
1 m^2 s^{-1}	= 10.762 75 ft^2 s^{-1} = 10^6 cSt
1 cSt	= 0.038 75 ft^2 h^{-1}
Energy, work:	
1 J	= 0.737 562 ft lbf
1 MJ	= 0.3725 hph
1 MJ	= 0.277 78 kW h
Power:	
1 W	= 0.737 562 ft lbf s^{-1}
1 kW	= 1.3410 hp = 737.562 ft lbf s^{-1}
Fluid mass:	
(Ordinary) 1 kg s^{-1}	= 2.204 62 lb s^{-1} = 7936.64 lb h^{-1}
(Velocity) 1 kg m^{-2} s	= 0.204 815 lb ft^{-2} s
HEAT	
Temperature:	
(Interval) 1 degK	= 9/5 deg R (Rankine)
1 degC	= 9/5 deg F
(Coefficient) 1 $degR^{-1}$	= 1 deg F^{-1} = 5/9 deg C
1 $degC^{-1}$	= 5/9 deg F^{-1}
Quantity of heat:	
1 J	= 9.478 17 × 10^{-4} Btu
1 J	= 0.238 846 cal
1 kJ	= 947.817 Btu
1 GJ	= 947.817 × 10^3 Btu
1 kJ	= 526.565 CHU
1 GJ	= 526.565 × 10^3 CHU
1 GJ	= 9.478 17 therm
Heat flow rate:	
1 W(J s^{-1})	= 3.412 14 Btu h^{-1}
1 W m^{-2}	= 0.316 998 Btu ft^{-2} h^{-1}
Thermal conductivity:	
1 W m^{-1} $°C^{-1}$	= 6.933 47 Btu in ft^{-2} h^{-1} $°F^{-1}$
Heat transfer coefficient	
1 W m^{-3} $°C^{-1}$	= 0.176 110 Btu ft^{-2} h^{-1} $°F^{-1}$
Heat capacity:	
1 J ° g^{-1} C^{-1}	= 0.526 57 × 10^{-3} Btu $°R^{-1}$
Specific heat capacity:	
1 J $°C^{-1}$	= 0.238 846 Btu lb^{-1} $°F^{-1}$
1 kJ kg^{-1} $°C^{-1}$	= 0.238 846 Btu lb^{-1} $°F^{-1}$

TABLE 2.25 Metric Conversion Factors (*Continued*)

SI units	Other units
Entropy:	
1 J K^{-1}	= 0.526 57 × 10^{-3} Btu $°R^{-1}$
Specific entropy:	
1 J kg^{-1} $degC^{-1}$	= 0.238 846 × 10^{-3} Btu lb^{-1} $°F^{-1}$
1 J kg^{-1} $degK^{-1}$	= 0.238 846 × 10^{-3} Btu lb^{-1} $°R^{-1}$
Specific energy/specific latent heat:	= 0.429 923 Btu lb^{-1}
1 J g^{-1}	= 0.429 923 × 10^{-3} Btu lb^{-1}
1 J kg^{-1}	
Calorific value:	
1 kJ kg^{-1}	= 0.429 923 Btu lb^{-1}
1 kJ kg^{-1}	= 0.773 861 4 CHU lb^{-1}
1 J m^{-3}	= 0.026 839 2 × 10^{-3} Btu ft^{-3}
1 kJ m^{-3}	= 0.026 839 2 Btu ft^{-3}
1 kg $liter^{-1}$	= 4.308 86 Btu gal^{-1}
1 kJ kg^{-1}	= 0.009 630 2 therm ton^{-1}
ELECTRICITY	
Permeability:	
1 H m^{-t}	= $10^7/4\Pi$ μ_o
Magnetic flux density:	
1 tesla	= 10^4 gauss = 1 Wb m^{-2}
Conductivity:	
1 mho	= 1 reciprocal ohm
1 Siemen	= 1 reciprocal ohm
Electric stress:	
1 kV mm^{-1}	= 25.4 kV in^{-1}
1 kV m^{-1}	= 0.0254 kV in^{-1}

PART 9

ABBREVIATIONS

Duncan T. Enright

Table 2.26 gives an alphabetical list of commonly used abbreviations of units. Obsolete and rarely found units are also included. For definitions of these units see Tables 2.9–2.24 and other reference works listed at the end of this chapter.

TABLE 2.26 Common Abbreviations of Units

a	year (p.a. = per annum)
Å	ångstrom
A	ampere
asb	apostilb
AU	astronomical unit
AT	assay ton
b	barn
bar	bar
Bi	Biot (unit of current in electromagnetic CGS system)
Btu, BthU	British thermal unit
c	curie
C	coulomb
°C	degree Celsius
cal	calorie
cc	cubic centimeter
cd	candela
CHU	Centrigrade heat unit
Ci	curie
cl	centiliter
cm	centimeter
CM	carat
cP	centipoise
c/s	cycle per second
cSt	centistoke
ct	carat
cu. cm	cubic centimeter
cu ft	cubic foot
cu in	cubic inch
cusec	cubic foot per second
cwt	hundredweight
d	day
dB	decibel
dm	decimeter
dwt	pennyweight
dyn	dyne
e unit, E unit	X-ray dosage

TABLE 2.26 Common Abbreviations of Units (*Continued*)

erg	erg
eV	electronvolt
f	force
F	farad
°F	degree Fahrenheit
fc	foot candle
ft	foot
ft L	foot Lambert
ft lb	foot pound
g	gram
G	gauss
gal	gallon
Gb	gilbert
g cal	gram calorie
gl	gill
gm	gram
g.p.m.	gallons per minute
g.p.s.	gallons per second
gr	grain
Gs	gauss
h	hour
H	henry
ha	hectare
hp	horsepower
hp hr	horsepower hour
Hz	hertz
in.	inch
in. Hg	inch of mercury
J	joule
K	kelvin
kc	kilocycle
kcal	kilocalorie
kc/s	kilocycle per second
kg	kilogram
kgf	kilogram force
km	kilometer
kn, kt	knot
kV	kilovolt
kVA	kilovolt ampere
kW	kilowatt
kW h	kilowatt hour
L	lambert
l	liter
lb	pound
lbf	pound force
lea	league
lm	lumen
ly	light year
lx	lux
m	meter

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

TABLE 2.26 Common Abbreviations of Units (*Continued*)

m	molality molal concentration
M	molar concentration
mA	milliampere
mbar	millibar
mcps	mega cycles per second
MeV	mega electron volt
mF	millifarad
micron	length − 10^6 meter pressure − 10^{-3} mm Hg
mil	angular − 1/1000 rt. angle length 1/1000 inch volume − milliliter
min	minute (time)
mks	meter kilogram second
ml	milliliter
mL	millilambert
mm	millimeter
mm fd	micromicrofarad
mm Hg	millimeter of mercury
mmm	millimicrons
mol	mole (amount of substance)
mpg	miles per gallon
mpm	meters per minute
m/s mps	meters per second
mt	metric ton
mV	millivolt
mW h	megawatt hour
Mx	maxwell
N	newton
n. mile nm	nautical mile
Np	neper
nt	nit
ntm	net ton mile
n unit	neutron dose
Oe	oersted
oz	ounce (avoirdupois)
oz. t	ounce (troy)
p	perch
P	poise
P	phon
Pa	pascal
pc	parsec
pdl	poundal
ph	phot
psi	pounds per square inch
pwt	pennyweight

TABLE 2.26 Common Abbreviations of Units (*Continued*)

q, ql	quintal
qts	quart
r, R	Röntgen
R	Réaumier
°R	degree Rankine
rad	radian
rpm	revolutions per minute
rps	revolutions per second
s	second (time)
S	Siemen
S, St	stokes
sb	stilb
sn	sthéne
sr	steradian
T	tesla
t	tonne
th	thermie
V	volt
VA	volt ampere
W	watt
Wb	weber
yd	yard

PART 10

PHYSICAL AND CHEMICAL CONSTANTS

Duncan T. Enright

The following tables provide physical and chemical constants most of use in engineering disciplines. Table 2.27 lists universal constants. Table 2.28 is a table of elements. Table 2.29 presents the principal elements in order of valency. Table 2.30 gives the surface tension of some common metals, and Table 2.31 lists specific heats, melting points and densities of selected elements. Table 2.32 reports the sectional properties of metals.

2.71 ATOMIC NUMBER

Radon has an atomic number of 86 (International Atomic Weights Commission). Isotopes of radon will have the same atomic number but different atomic weights.

2.72 ATOMIC WEIGHT

Weights are related to an arbitrary reference value of 16 for oxygen. Atomic weights vary because of natural variation in composition (different isotopes in mixture affect the weight average) and because of relativistic effects.

2.73 DENSITY

Values are taken at 293 K.

2.74 MELTING POINT

Values are for 1 atm pressure (1.01325×10^5 Pa).

2.75 LINEAR COEFFICIENT OF EXPANSION

This is commonly known as the α value. It is worked out as a mm expansion over 100 m length using the equation $L = l(1 + \alpha t)$ for expansion and $l = L/(1 + \alpha t)$ for a contraction; where l = increase or decrease in length, L = original length, t = temperature change and α = linear coefficient of expansion.

TABLE 2.27 Universal Constants

Constant	Symbol	Numerical value	SI unit
Speed of light in vacuum	c	2.997 925(1)	10^{8} m s^{-1}
Gravitational constant	G	6.670(5)	10^{-11} N m^2 kg^2
Elementary charge	e	1.602 10(2)	10^{-19} C
Avogadro constant	N_A	6.022 52(9)	10^{26} $kmol^{-1}$
Mass unit	u	1.660 43(2)	10^{-27} kg
Electron rest mass	m_e	9.109 08(13)	10^{-31} kg
		5.485 97(3)	10^{-44} u
Proton rest mass	m_p	1.672 52(3)	10^{-27} kg
		1.007 276 63(8)	u
Neutron rest mass	m_n	1.674 82(3)	10^{-27} kg
Faraday constant	F	9.684 70(5)	10^{4} C mol^{-1}
Planck constant	h	6.625 59(16)	10^{-34} J s
	$h/2\pi$	1.054 494(25)	10^{-34} J s
Fine-structure constant	α	7.297 20(3)	10^{-3}
	$1/\alpha$	137.038 8(6)	
Charge-to-mass ratio for electron	e/m_e	1.758 796(6)	10^{11} C kg^{-1}
Quantum of magnetic flux	hc/e	4.135 56(4)	10^{-11} Wb
Rydberg constant	R_∞	1.097 373 1(1)	10^{7} m^{-1}
Bohr radius	a_0	5.291 67(2)	10^{-11} m
Compton wavelength of electron	$h/m_e c$	2.426 21(2)	10^{-12} m
	$\lambda C/2\pi$	3.861 44(3)	10^{-13} m
Electron radius	$e^2/m_e c^2 = r_e$	2.817 77(4)	10^{-15} m
Thomson cross-section	$8\pi r_c^2/3$	6.651 6(2)	10^{-29} m^2
Compton wavelength of proton	$\lambda_{C,p}$	1.321 398(13)	10^{-15} m
	$\lambda_{C,p}/2\pi$	2.103 07(2)	10^{-16} m
Gyromagnetic ratio of proton	γ	2.675 192(7)	10^{8} rad (s T)$^{-1}$
	$\gamma/2\pi$	4.257 70(1)	10^{7} Hz T^{-1}
(uncorrected for diamagnetism of H_2O)	γ'	2.675 123(7)	10^{8} rad (s T)$^{-1}$
	$\gamma'/2\pi$	4.257 59(1)	10^{7} Hz T^{-1}
Bohr magneton	μ_B	9.273 2(2)	10^{-24} J T^{-1}
Nuclear magneton	μ_N	5.050 50(13)	10^{-27} J T^{-1}
Proton magnetic moment	μ_p	1.410 49(4)	10^{-26} J T^{-1}
	μ_p/μ_N	2.792 76(2)	
(uncorrected for diamagnetism in H_2O sample)	μ_p/μ_N	2.792 68(2)	
Gas constant	R_0	8.314 34(35)	J K^{-1} mol
Boltzmann constant	k	1.380 54(6)	10^{-23} J K^{-1}
First radiation constant ($2\pi hc^2$)	c_1	3.741 50(9)	10^{-16} W m^{-2}
Second radiation constant (hc/k)	c_2	1.438 79(6)	10^{-2} m K
Stefan–Boltzmann constant	σ	5.669 7(10)	10^{-8} W m^{-2} K^4

TABLE 2.28 Table of Elements

Element	Symbol	Atomic no.	Atomic weight	Density (kg m^{-3})	Melting point (°C)	Linear coefficient of expansion/°C at normal temp. ($\times 10^{-6}$)	Heat conductivity (W m^{-1} C^{-1})	Electric resistivity (microhm cm)
Actinium	Ac	89	227.00	10 100	1 230	—	—	—
Aluminum	Al	13	26.98	2 700	657	24	217.7	2.655
Antimony	Sb	51	121.75	6 619	630.5	11.29	18.59	39
Argon	Ar	18	39.944	1.663 (liquid)	−187.9	—	0.017	—
Arsenic	As	33	74.92	5 733	813.8	3.86	—	35
Barium	Ba	56	137.34	3 500	710	—	—	60
Beryllium	Be	4	9.012	1 822	1 285	12.3	161	2.85
Bismuth	Bi	83	209.00	9 802	271	13.45	8.37	115
Boron	B	5	10.81	2 300	2 030	2	—	1.8×10^{12}
Bromine	Br	35	79.9	3 119	−28.3	—	—	—
Cadmium	Cd	48	112.40	8 652	321	29.8	90.86	7.59
Cadsium	Cs	55	132.91	1 899	26	97	—	20
Calcium	Ca	20	40.08	1 550	851.3	25	—	4.6
Californium	Cf	98	251	—	—	—	—	—
Carbon (Graphite)	C	6	12.01	2 220	3 500	1.2	23.87	1 000
Cerium	Ce	58	140.12	6 901	775.2	—	—	78
Chlorine	Cl	17	35.457	1 560 (liquid)	−101	11.44	0.007 2	10×10^{15}
Chromium	Cr	24	52.01	7 139	1 900	8.1	69.08	13.1
Cobalt	Co	27	58.94	8 904	1 490	12.08	69.08	9.7
Copper	Cu	29	63.54	8 941	1 082	16.42	386.5	1.682
Curium	Cm	96	247	—	—	—	—	—
Dysprosium	Dy	66	162.46	8 500	1 500	—	—	89
Erbium	Er	68	167.20	9 000	1 525	—	—	81
Europium	Eu	63	152.00	5 200	830	—	—	—
Fermium	Fm	100	257	—	—	—	—	
Fluorine	Fl	9	19.00	1 100 (liquid)	−223	—	—	—

TABLE 2.28 Table of Elements (*Continued*)

Element	Symbol	Atomic no.	Atomic weight	Density (kg m^{-3})	Melting point (°C)	Linear coefficient of expansion/°C at normal temp. ($\times 10^{-6}$)	Heat conductivity (W m^{-1} C^{-1})	Electric resistivity (microhm cm)
Gadolinium	Gd	64	156.90	7 900	1 320	—	—	
Gallium	Ga	31	69.72	5 910	29.79	18.3	—	126
Germanium	Ge	32	72.60	5 363	958.3	—	—	57.1
Gold	Au	79	197.20	19 310	1 063	14.4	296.1	89×10^3
Hafnium	Hf	72	178.60	11 400	1 700	—	—	2.42 29.6
Helium	He	2	4.003	0.166 3	2 000	—	0.139	—
Holmium	Ho	67	163.5	8 800	1 500	—	—	77
Hydrogen	H	1	1.008 1	0.083 8	−253	—	0.170	—
Indium	In	49	114.76	7 308	161.2	33	23.87	9
Iodine	I	53	126.92	4 927	113.5	93	0.043 5	1.3×10^{15}
Iridium	Ir	77	193.10	22 400	2 409	6.41	59	6.08
Iron	Fe	26	55.84	7 861	1 536	11.9	79.56	9.8
Krypton	Kr	36	83.70	2 160 (liquid)	−170.5	—	0.009	—
Lanthanum	La	57	138.92	6 146	826.4	—	—	59
Lead	Pb	82	207.21	11 320	327.3	29.5	34.75	20.65
Lithium	Li	3	6.94	534.3	186	56.0	71.14	8.5
Lutetium	Lu	71	175	9 870	1 700	—	—	54
Magnesium	Mg	12	24.32	1 739	651.3	25.7	154.9	4.46
Manganese	Mn	25	54.93	7 418	1 243	23	—	—
Mercury	Hg	80	200.61	13 540	−38.9	—	8.37	95.8
Molybdenum	Mo	42	95.95	10 190	2 620	5.49	146.6	4.77
Neodymium	Nd	60	144.27	7 058	1 024	—	—	79
Neon	Ne	10	20.183	0.839	−248.5	—	0.046	—
Neptunium	Np	93	239	1 900	—	—	—	—
Nickel	Ni	28	58.70	8 915	1 452	13.7	58.61	6.9

Niobium	Nb	41	92.91	8 571	2 420	7.2	52	15.2
Nitrogen	N	7	14.008	1 165	−209.5	—	0.025	—
Osmium	Os	76	190.2	22 480	3 000	5.7	—	9
Oxygen	O	8	16.000	1.332	−218	—	0.025	—
Palladium	Pd	46	106.7	11 990	1 555	11.60	67.41	10
Phosphorus	P	15	31.02	1 819 (yellow)	44.12	11.25	—	10^{17}
Platinium	Pt	78	195.23	21 420	1 774	8.8	69.5	9.83
Plutonium	Pu	94	239	19 800	640	—	—	150
Polonium	Po	84	209	9 320	254	—	—	—
Potassium	K	19	39.096	858.4	62.29	—	99.2	7
Praseodymium	Pr	59	140.92	6 616	940.1	8.3	—	88
Promethium	Pm	61	147	—	—	—	—	—
Protactinium	Pa	91	231	15 400	1 000	—	—	—
Radium	Ra	88	226.05	5 005	960	—	—	—
Radon	Rn	86	222	4 400	−71	—	—	—
Rhenium	Re	75	186.31	20 000	3 000	—	—	—
Rhodium	Rh	45	102.91	12 430	1 966	8.9	89.15	4.93
Rubidium	Rb	37	85.48	1 531	38.3	90.0	—	12.5
Ruthenium	Ru	44	101.7	12 210	2 300	8.5	—	10
Smarium	Sm	62	150.4	7 752	1 050	—	—	91.4
Scandium	Sc	21	45.10	2 434	1 204	—	—	50.5
Selenium	Se	34	78.96	4 816	220.1	87.0	—	8×10^{6}
Silicon	Si	14	28.08	2 408	1 427	—	83.74	85×10^{3}
Silver	Ag	47	107.88	10 520	960.5	18.9	407.9	1.62
Sodium	Na	11	22.997	969	97.52	71.0	135.1	4.6
Strontium	Sr	38	87.63	2 602	77.1	—	—	22.76
Sulphur	S	16	32.06	2 076	113	67.48	0.263 7	1.9×10^{17}

TABLE 2.28 Table of Elements (*Continued*)

Element	Symbol	Atomic no.	Atomic weight	Density (kg m^{-3})	Melting point (°C)	Linear coefficient of expansion/°C at normal temp. ($\times 10^{-6}$)	Heat conductivity (W m^{-1} C^{-1})	Electric resistivity (microhm cm)
Tantalum	Ta	73	180.88	16 620	3 017	6.5	54.43	15.5
Technetium	Tc	43	99	11 400	2 100	—	—	—
Tellurium	Te	52	127.61	6 200	452.3	16.8	6.016	1.6×10^5
Terbium	Tb	65	159.2	11 850	310	—	—	—
Thallium	Tl	81	204.39	11 850	303.8	28.0	38.9	18.1
Thorium	Th	90	232.12	11 520	1 700	12.3	—	18
Thulium	Tm	69	169.4	9 330	1 600	—	—	—
Tin	Sn	50	118.70	7 308	232	21	65.73	11.5
Titanium	Ti	22	47.90	4 512	1 680	7.14	—	—
Tungsten	W	74	183.92	19 320	3 370	4.0	199.3	5.48
Uranium	U	92	238.07	18 710	1 133	—	—	60
Vanadium	V	23	50.95	5 675	1 920	8	—	26
Xenon	Xe	54	131.3	152.7	−112	—	519.2	—
Ytterbium	Yb	70	173.04	6 900	824	—	—	27.7
Yttrium	Y	39	88.905	5 509	1 482	—	—	80
Zinc	Zn	30	65.38	7 142	419.5	30	112	5.5
Zirconium	Zr	40	91.22	6 366	1 850	6.3	—	41

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

TABLE 2.29 Table of Principal Elements Arranged in Order of Valency

Valency	Name	Symbol	Atomic weight
1 (Monovalent)	Bromine	Br	79.9
	Chlorine	Cl	39.457
	Fluorine	Fl	19.00
	Hydrogen	H	1.008 1
	Iodine	I	126.92
	Potassium	K	39.096
	Silver	Ag	107.88
	Sodium	Na	22.997
2 (Divalent)	Barium	Ba	137.34
	Cadmium	Cd	112.40
	Calcium	Ca	40.08
	Copper	Cu	63.54
	Magnesium	Mg	54.93
	Mercury	Hg	200.61
	Oxygen	O	16.000
	Zinc	Zn	65.38
3 (Trivalent)	Aluminum	Al	26.98
	Bismuth	Bi	209.00
	Boron	B	10.81
	Cobalt	Co	58.94
	Gold	Au	197.20
	Iron	Fe	55.84
	Nickel	Ni	58.70
4 (Tetravalent)	Lead	Pb	207.21
	Platinum	Pt	195.23
	Silicon	Si	28.08
	Tin	Sn	118.70
5 (Pentavalent)	Antimony	Sb	121.75
	Arsenic	As	74.92
	Nitrogen	N	14.008
	Phosphorus	P	31.02
6 (Hexavalent)	Chromium	Cr	52.01
	Manganese	Mn	54.93
	Sulphur	S	32.06

2.76 HEAT CONDUCTIVITY

This is sometimes called thermal conductivity and is known as the k factor. Values are taken at 293 K.

2.77 ELECTRICAL RESISTIVITY

Values are given at 293 K.

TABLE 2.30 Surface Tension of Some Common Metals

Metal	Temperature (°C)	Surface tension (mN m^{-1})
Antimony	640	350
Bismuth	269	378
Cadmium	320	630
Copper	1131	1103
Gold	1120	1128
Iron (acc. to C content)	1300–1420	1150–1500
Lead	327	452
Mercury	20	465
Silver	998	923
Tin	232	526
Zinc	419	758

Surface tension of water is 78.5 mN m^{-1} at 0°C and decreases by 0.152 mN m^{-1} for each degree rise of temperature.

TABLE 2.31 Specific Heats, Melting Points, and Densities

Material	Specific heat (kJ kg^{-1} $°C^{-1}$)	Melting point (°C)	Density (kg m^{-3}) (approx.)
Air (20°C)	0.996 8	—	1.2
Aluminum	0.914 9	657	2 700
Brass	0.395 3	850–950	8 400
Carbon	0.854 3	—	2 220
Cobalt	0.448	1 480	8 904
Copper	0.393 6	1 083	8 941
German silver	0.398	—	8 400
Graphite	0.837 4	—	2 300
Iron, cast	0.46–0.67	1 200	7 000
Lead (solid)	0.131 3	327.4	11 320
Lead (liquid)	0.197 2	—	—
Mercury	0.134	−38.9	13 540
Molybdenum	0.275 9	2 620	10 190
Nickel	0.454	1 452	8 915
Paraffin (solid)	2.6–2.9	38–56	900
Paraffin (liquid)	29.7	—	800
Pitch	—	—	1 100
Platinum	0.138 1	1 774	21 420
Silver	0.232 7	960.8	10 520
Solder	—	205–185	8 300
Tin (solid)	0.234 5	232	7 308
Tin (liquid)	0.268	—	—
Tungsten	0.146 6	3 370	19 320
Type metal	0.163 3	—	—
Water (20°C)	4.186 8	—	—
Zinc (cast)	0.39–0.50	419.5	7 142

ENGINEERING MATHEMATICS, UNITS, SYMBOLS, AND CONSTANTS

TABLE 2.32 Sectional Properties of Metals

Substance	E (MN m^{-2})	G (MN m^{-2})	σ	K (MN m^{-2})	Tensile strength (MN m^{-2})
Aluminum	70 300	26 100	0.345	75 500	90–150
Brass	101 000	37 300	0.350	111 800	280–730
Copper	129 800	48 300	0.343	137 800	120–400
Iron (cast)	152 000	60 000	0.270	109 000	100–230
Iron (wrought)	211 400	81 000	0.293	170 000	260–450
Lead	16 100	5 600	0.440	45 700	12–17
Magnesium	44 700	17 000	0.291	25 600	60–190
Silver	82 700	30 200	0.366	103 600	300
Platinum	168 000	61 000	0.377	228 000	330–370
Tantalum	185 700	69 200	0.342	196 300	800–1 100
Tin	49 900	18 400	0.357	58 200	20–35
Tungsten	411 000	160 000	0.280	311 500	1 500–3 500
Steel (mild)	211 900	82 200	0.291	169 200	430–690
Steel (hardened)	210 400	77 800	0.295	165 200	1 800–2 300

Notes: E is known as Young's modulus or longitudinal elasticity; G is known as shear or rigidity modulus; σ is known as Poisson's ratio; K is known as bulk modulus.

FURTHER READING

Avallone, E. A. and Baumeister III, T., *Marks' Standard Handbook for Mechanical Engineers,* 9th ed., McGraw-Hill, New York (1986).

Besterfield, D. H., *Quality Control,* Prentice-Hall, Englewood Cliffs, NJ (1979).

Blake, L. S. (ed.), *Civil Engineer's Reference Book,* 4th ed., Butterworths, London (1989).

Boyce, W. E. and DiPrima, R. C., *Elementary Differential Equations and Boundary Values,* 4th ed. Wiley, Chichester (1986).

Caplen, R. H., *A Practical Approach to Quality Control,* Business Books, London (1982).

Chalk, G. O. and Stick, A. W., *Statistics for the Engineer,* Butterworths, London (1975).

Cohen, S. S., *Practical Statistics,* Edward Arnold, London (1988).

David, H. A., *Order Statistics,* Wiley, Chichester (1981).

Doran, D. K. (ed.), *Construction Materials Reference Book,* Butterworth-Heinemann, Oxford (1992).

Dudewicz, E. J., and Mishra, S. N., *Modern Mathematical Statistics,* Wiley, New York (1988).

Dunn, R. A. and Ramsing, K. D., *Management Science: A Practical Approach to Decision Making,* Macmillan, London (1981).

Fitzsimmons, J. A., *Service Operations Management,* McGraw-Hill, New York (1982).

Grant, E. L. and Leavenworth, R. S., *Statistical Quality Control,* McGraw-Hill, New York (1980).

Hahn, W. C., *Modern Statistical Methods,* Butterworths, London (1979).

Jones, M. E. M., *Statistics,* Schofield & Sims, Huddersfield (1988).

Koshal, D. (ed.), *Manufacturing Engineer's Reference Book,* Butterworth-Heinemann, Oxford (1993).

Kreyszig, E., *Advanced Engineering Mathematics,* 5th ed., Wiley, Chichester (1983)

Laughton, M. A. and Say, M. G., *Electrical Engineer's Reference Book,* 14th ed., Butterworths, London (1985).

Livesley, R. K., *Finite Elements: An Introduction for Engineers,* Cambridge University Press, Cambridge (1983).

Lyons, S., *Handbook of Industrial Mathematics,* Cambridge University Press, Cambridge (1978).

Mazda, F. F., *Electronic Engineer's Reference Book,* 6th ed., Butterworths, London (1989)

Mazda, F. F., *Quantitative Techniques in Business,* Gee & Co., London (1979).

Nayler, G. H. F., *Dictionary of Mechanical Engineering,* Butterworths, London (1985).

Noltingk, B. E. (ed.), *Instrumentation Reterence Book*, Butterworths, London (1988).

Oberg, E., Jones, F. D., and Horton, H. L., *Machinery's Handbook,* 23rd ed., Industrial Press, New York (1988).

Siegel, A. F., *Statistics and Data Analysis,* Wiley, New York (1988)

Snow, D. A., *Plant Engineer's Reference Book,* Butterworth-Heinemann, Oxford (1991).

Tool and Manufacturing Engineer's Handbook, Society of Manufacturing Engineers, Dearborn, MI (1983–1988).

SECTION 3
MECHANICAL ENGINEERING PRINCIPLES

Section Editor: Mark Davies

PART 1

STATICS OF RIGID BODIES

Robert Paine

In general, the study of mechanics may be divided into two distinct areas. These are *statics,* which involves the study of bodies at rest, and *dynamics,* which is the study of bodies in motion. In each case it is important to select an appropriate mathematical model from which a "free body diagram" may be drawn, representing the system in space, with all the relevant forces acting on that system.

When a set of forces acts on a body, it gives rise to a resultant force or moment or a combination of both. The situation may be determined by considering three mutually perpendicular directions on the free body diagram and resolving the forces and moment in these directions. If the three directions are denoted by x, y, and z, then the sum of forces may be represented by ΣF_x, ΣF_y, and ΣF_z and the sum of the moments about respective axes by ΣM_x, ΣM_y, and ΣM_z. Then for equilibrium the following conditions must hold:

$$\Sigma F_x = \Sigma F_y = \Sigma F_z = 0 \tag{3.1}$$

$$\Sigma M_x = \Sigma M_y = \Sigma M_z = 0 \tag{3.2}$$

If the conditions in equations (3.1) and (3.2) are not satisfied, then there is a resultant force or moment, which is given by

$$F = [(\Sigma F_x)^2 + (\Sigma F_y)^2 + (\Sigma F_z)^2]^{1/2}$$

$$M = [(\Sigma M_x)^2 + (\Sigma M_y)^2 + (\Sigma M_z)^2]^{1/2}$$

The six conditions given in equations (3.1) and (3.2) satisfy problems in three dimensions. If one of these dimensions is not present (say, the z direction) the system reduces to a set of coplanar forces, and then

$$\Sigma F_x = \Sigma M_x = \Sigma M_y = 0$$

are automatically satisfied, and the necessary conditions of equilibrium in a two-dimensional system are

$$\Sigma F_x = \Sigma F_y = \Sigma M_z = 0 \tag{3.3}$$

If the conditions in equation (1.3) are not satisfied, then the resultant force or moment is given by

$$F = [(\Sigma F_x)^2 + (\Sigma F_y)^2]^{1/2}$$

$$M = \Sigma M_z$$

The above equations give solutions to what are said to be "statically determinate" systems. These are systems where there are the minimum number of constraints to maintain equilibrium (Mathesun 1987).

PART 2

STRENGTH OF MATERIALS

Robert Paine

Weight: The weight (W) of a body is that force exerted due to gravitational attraction on the mass (m) of the body: $W = mg$, where g is the acceleration due to gravity.

Center of gravity: This is a point, which may or may not be within the body, at which the total weight of the body may be considered to act as a single force. The position of the center of gravity may be found experimentally or by analysis. When using analysis the moment of each element of weight, within the body, about a fixed axis is equated to the moment of the complete weight about that axis:

$$\tilde{x} = \Sigma \delta mg \cdot x / \Sigma \delta mg,\ \tilde{y} = \Sigma \delta mg \cdot y / \Sigma \delta mg$$

$$\tilde{z} = \Sigma \delta mg \cdot z / \Sigma \delta mg$$

where δm is an element of mass at a distance of x, y or z from the respective axis, and $\tilde{x}$, $\tilde{y}$ and $\tilde{z}$ are the positions of the centers of gravity from these axes. Table 3.1 shows the position of the center of gravity for some standard shapes. (See Meriam and Kraige 1987 for a more comprehensive list.)

Shear force and bending moment: If a beam subject to loading, as shown in Figure 3.1, is cut, then in order to maintain equilibrium a shear force (Q) and a bending moment (M) must be applied to each portion of the beam. The magnitudes of Q and M vary with the type of loading and the position along the beam and are directly related to the stresses and deflections in the beam.

Relationship between shearforce and bending moment: If an element of a beam is subjected to a load w then the following relationship holds:

$$\frac{d^2M}{dx^2} = \frac{dF}{dx} = -w$$

Table 3.2 shows examples of bending moments, shear force, and maximum deflection for standard beams.

Bending equation: If a beam has two axes of symmetry in the xy plane then the following equation holds:

$$M_Z / I_Z = E / R_Z = \sigma / y$$

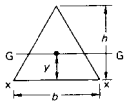
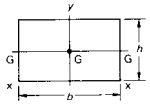
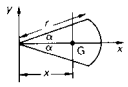
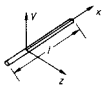
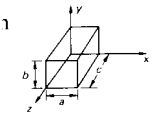
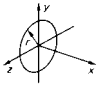
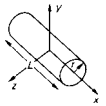
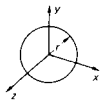
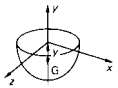
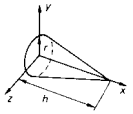
where M_Z is the bending moment, R_z the radius of curvature, I_z the moment of inertia, E the modulus of elasticity, y the distance from the principal axis, and σ the stress.

Torsion equation: If a circular shaft is subject to a torque (T) then the following equation holds:

$$T / J = \tau / r = G\theta / L$$

where J is the polar second moment of area, G the shear modulus, L the length, θ the angle of twist, τ the shear stress, and r the radius of the shaft.

TABLE 3.1 Centers of Gravity and Moments of Inertia or Second Moments of Area for Two-Dimensional Figures

Shape		G	I
Triangular area		$\bar{y} = h/3$	$I_{GG} = bh^3/36$ $I_{XX} = bh^3/12$
Rectangular-area			$I_{GG} = bh^3/12$ $I_{XX} = bh^3/3$
Circular sector		$\bar{x} = \frac{2r}{3}\frac{\sin\alpha}{\alpha}$	$I_{XX} = \frac{r^4}{4}\left(\alpha - \frac{1}{2}\sin^2\alpha\right)$ $I_{YY} = \frac{r^4}{4}\left(\alpha + \frac{1}{2}\sin^2\alpha\right)$
Slender rod			$I_{XX} = I_{ZZ} = ml^2/12$
Rectangular prism			$I_{XX} = m(b^2 + c^2)/12$ $I_{YY} = m(c^2 + a^2)/12$ $I_{ZZ} = m(a^2 + b^2)/12$
Thin disk			$I_{XX} = mr^2/2$ $I_{YY} = I_{ZZ} = mr^2/4$
Circular cylinder			$I_{XX} = mr^2/2$ $I_{YY} = I_{ZZ} = m(3r^2 + L^2)/12$
Sphere			$I_{XX} = I_{YY} = I_{ZZ} = 2mr^2/5$
Hemisphere		$\bar{y} = 3r/8$	$I_{YY} = 2mr^2/5$
Circular cone		$\bar{x} = h/4$	$I_{XX} = 3mr^2/10$ $I_{YY} = \frac{3mr^2}{20} + \frac{mh^2}{10}$

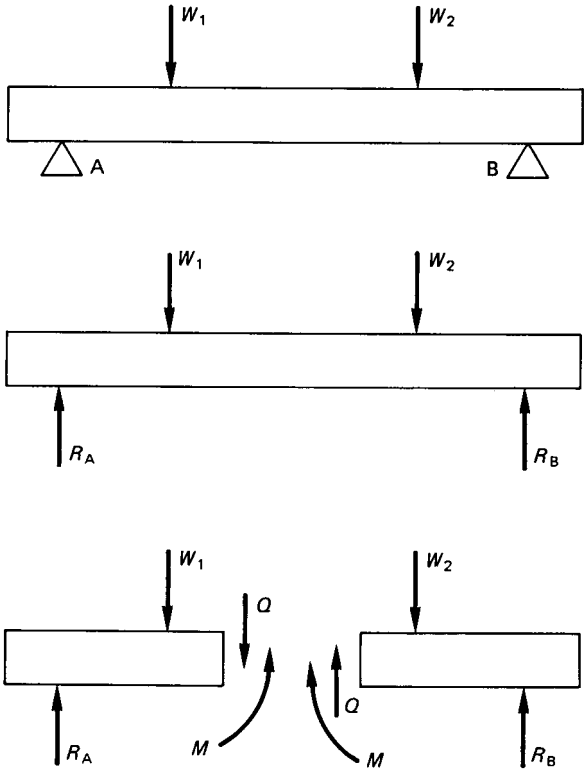
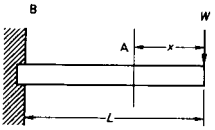
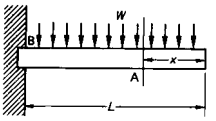


FIGURE 3.1

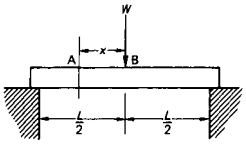
TABLE 3.2



One concentrated load W
M at $A = Wx$, Q at $A = W$
M greatest at B, and $= WL$
Q uniform throughout
Maximum deflection $= WL^3/3EI$ at the free end.



Uniform load of W
M at $A = Wx^2/2L$
Q at $A = Wx/L$
M greatest at $B = WL/2$
Q greatest at $B = W$
Maximum deflection $= WL^3/8EI$ at the free end.

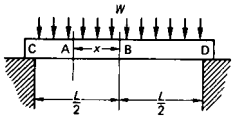


One concentrated load at the center of a beam

$$M \text{ at } A = \frac{W}{2}\left(\frac{L}{2} - x\right),$$

Q at $A = W/2$

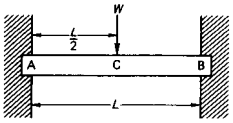
M greatest at $B = WL/4$
Q uniform throughout
Maximum deflection $= WL^3/48EI$ at the center



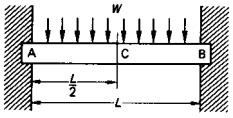
Uniform load W

$$M \text{ at } A = \frac{W}{2L}\left(\frac{L^2}{4} - x^2\right)$$

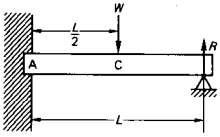
Q at $A = Wx/L$
M greatest at $B = WL/8$
Q greatest at C and $D = W/2$ maximum deflection at $B = 5WL^3/384EI$



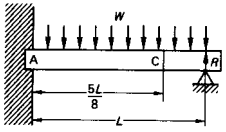
Beam fixed at ends and loaded at center.
M is maximum at A, B, and C and $= WL/8$.
Maximum deflection at $C = WL^3/192EI$



Beam fixed at ends with uniform load.
M maximum at A and B and $= WL/12$
Maximum deflection at $C = WL^3/384EI$



One concentrated load W
Reaction $R = 5W/16$
M maximum at A, and $= 3WL/16$
M at $C = 5WL/32$
Maximum deflection is $L/\sqrt{5}$ from the free end, and $= WL^3/107EI$



Uniform load W
Reaction $R = 3W/8$
M maximum at A, and $= WL/8$
M at $C = 9WL/128$
Maximum deflection is $3L/8$ from the free end, and $= WL^2/187EI$

PART 3

DYNAMICS OF RIGID BODIES

Robert Paine

3.1 BASIC DEFINITIONS

Newton's Laws of Motion

First Law. A particle remains at rest or continues to move in a straight line with a constant velocity unless acted on by an external force.

Second Law. The sum of all the external forces acting on a particle is proportional to the rate of change of momentum.

Third Law. The forces of action and reaction between interacting bodies are equal in magnitude and opposite in direction.

Newton's law of gravitation, which governs the mutual interaction between bodies, states:

$$F = Gm_1m_2/x^2$$

where F is the mutual force of attraction, G is a universal constant called the constant of gravitation, which has a value 6.673×10^{-11} m^3 kg^{-1} s^2, m_1 and m_2 are the masses of the two bodies, and x is the distance between the centers of the bodies.

Mass (m) is a measure of the amount of matter present in a body.

Velocity is the rate of change of distance (x) with time (t):

$$v = dx/dt \text{ or } \dot{x}$$

Acceleration is the rate of change of velocity (v) with time (t):

$$a = dv/dt \text{ or } d^2x/dt^2 \text{ or } \ddot{x}$$

Momentum is the product of the mass and the velocity. If no external forces are present then the momentum of any system remains constant. This is known as the conservation of momentum.

Force is equal to the rate of change of momentum (mv) with time (t):

$$F = d(mv)/dt$$

$$F = m \cdot dv/dt + v \cdot dm/dt$$

If the mass remains constant then this simplifies to $F = m \cdot dv/dt$, i.e., Force = mass $\times$ acceleration, and it is measured in newtons.

Impulse (I) is the product of the force and the time that force acts. Since $I = Ft = mat = m(v_2 - v_1)$, impulse is also said to be the change in momentum.

Energy: There are several different forms of energy which may exist in a system. These may be converted from one type to another but they can never be destroyed. Energy is measured in joules.

Potential energy (PE) is the energy which a body possesses by virtue of its position in relation to other bodies: $PE = mgh$, where h is the distance above some fixed datum and g is the acceleration due to gravity.

Kinetic energy (KE) is the energy a body possesses by virtue of its motion: $KE = \frac{1}{2}mv^2$.

Work (W) is a measure of the amount of energy produced when a force moves a body a given distance: $W = F \cdot x$.

Power (P) is the rate of doing work with respect to time and is measured in watts.

Moment of inertia (I): The moment of inertia is that property in a rotational system which may be considered equivalent to the mass in a translational system. It is defined about an axis xx as $I_{XX} = \Sigma \delta m x^2 = mk^2_{XX}$, where x is the perpendicular distance of an element of mass δm from the axis xx and k_{XX} is the radius of gyration about the axis xx. Table 3.1 gives some data on moments of inertia for standard shapes.

Angular velocity (ω) is the rate of change of angular distance (θ) with time:

$$= d\theta/dt = \dot{\theta}$$

Angular acceleration (α) is the rate of change of angular velocity (ω) with time:

$$= d\omega/dt \quad \text{or} \quad d^2\theta/dt^2 \quad \text{or} \quad \ddot{\theta}$$

Both angular velocity and accleration are related to linear motion by the equations $v = \omega x$ and $a = \alpha x$ (see Figure 3.2).

Torque (T) is the moment of force about the axis of rotation:

$$T = I_0 \alpha$$

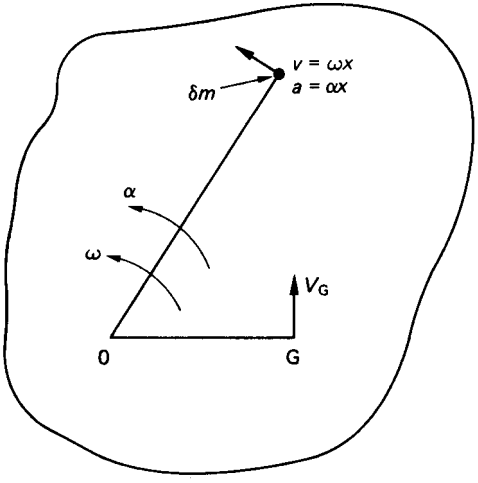


FIGURE 3.2

A torque may also be equal to a *couple,* which is two forces equal in magnitude acting some distance apart in opposite directions.

Parallel axis theorem: if I_{GG} is the moment of inertia of a body of mass m about its center of gravity, then the moment of inertia (I) about some other axis parallel to the original axis is given by $I = I_{GG} + mr^2$, where r is the perpendicular distance between the parallel axes.

Perpendicular axis theorem: If I_{XX}, I_{YY}, and I_{ZZ} represent the moments of inertia about three mutually perpendicular axes x, y, and z for a plane figure in the xy plane (see Figure 3.3) then $I_{ZZ} = I_{XX} + I_{YY}$.

Angular momentum (H_O) of a body about a point O is the moment of the linear momentum about that point and is ωI_{OO}. The angular momentum of a system remains constant unless acted on by an external torque.

Angular impulse is the product of torque by time, i.e. angular impulse $= Tt = I\alpha \cdot t = I(\omega_2 - \omega_1)$, the change in angular momentum.

Angular kinetic energy about an axis O is given by $\frac{1}{2}I_O\omega^2$.

Work done due to a torque is the product of torque by angular distance and is given by $T\theta$.

Power due to torque is the rate of angular work with respect to time and is given by $Td\theta/dt = T\omega$.

Friction: Whenever two surfaces, which remain in contact, move one relative to the other there is a force which acts tangentially to the surfaces so as to oppose motion. This is known as the force of friction. The magnitude of this force is μR, where R is the normal reaction and μ is a constant known as the coefficient of friction. The coefficient of friction depends on the nature of the surfaces in contact.

3.2 LINEAR AND ANGULAR MOTION IN TWO DIMENSIONS

Constant acceleration: If the accleration is integrated twice and the relevant initial conditions are used, then the following equations hold:

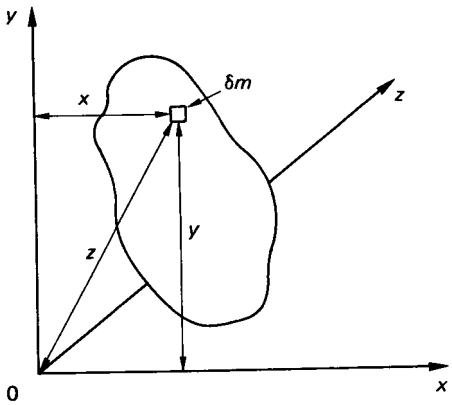


FIGURE 3.3

Linear motion	*Angular motion*
$x = v_1 t + \frac{1}{2}at^2$	$\theta = \omega_1 t + \frac{1}{2}\alpha t^2$
$v_2 = v_1 + at$	$\omega_2 = \omega_1 + \alpha t$
$v_2^2 = v_1^2 + 2ax$	$\omega_2^2 = \omega_1^2 + 2\alpha\theta$

Variable acceleration: If the acceleration is a function of time then the area under the acceleration time curve represents the change in velocity. If the acceleration is a function of displacement then the area under the acceleration distance curve represents half the difference of the square of the velocities (see Figure 3.4).

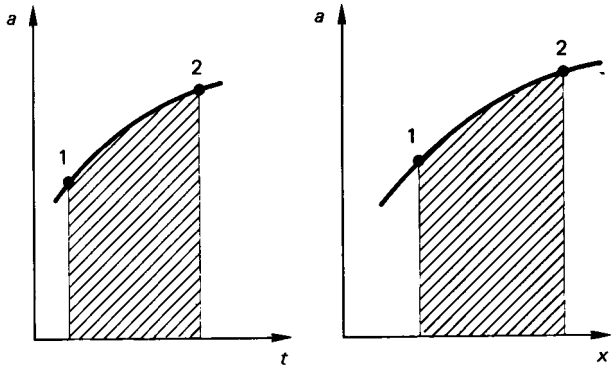
Curvilinear motion is when both linear and angular motions are present.

If a particle has a velocity v and an acceleration a then its motion may be described in the following ways:

1. *Cartesian components,* which represent the velocity and acceleration along two mutually perpendicular axes x and y (see Figure 3.5(a)):

$$v_X = v\cos\theta, \quad v_Y = v\sin\theta, \quad a_X = a\cos\phi$$

$$a_Y = a\sin\phi$$



$$a = \frac{dv}{dt} \text{ or } a\,dt = dv$$

$$\text{Area} = \int_1^2 a.dt$$

$$= v_2 - v_1$$

$$a = \frac{dx}{dt} \cdot \frac{dv}{dx}$$

$$a = v\frac{dv}{dx}$$

$$\text{or } a\,dx = v\,dv$$

$$\text{Area} = \int_1^2 a\,dx = \frac{v_2^2}{2} - \frac{v_1^2}{2}$$

FIGURE 3.4

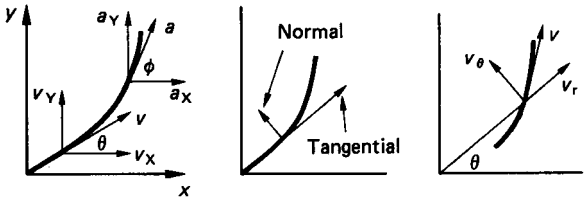


FIGURE 3.5

2. *Normal and tangential components:* see Figure 3.5(b):

$$v_t = v = r\dot{\theta} = r\omega, \quad v_n = 0$$

$$a_t = r\dot{\theta} + \dot{r}\dot{\theta} = r\alpha + \dot{r}\omega$$

$$a_n = v\dot{\theta} = r\omega^2$$

3. *Polar coordinates:* see Figure 3.5(c):

$$v_t = \dot{r}, \quad v_\theta = r\dot{\theta}$$

$$a_t = \ddot{r} - r\dot{\theta}^2, \quad a_\theta = r\ddot{\theta} + 2\dot{r}\dot{\theta}$$

3.3 CIRCULAR MOTION

Circular motion is a special case of curvilinear motion in which the radius of rotation remains constant. In this case there is an acceleration towards the center of $\omega^2 r$. This gives rise to a force towards the center known as the *centripetal force*. This force is reacted to by what is called the *centrifugal reaction*.

Velocity and acceleration in mechanisms: A simple approach to determine the velocity and acceleration of a mechanism at a point in time is to draw velocity and acceleration vector diagrams.

Velocities: If in a rigid link AB of length l the end A is moving with a different velocity to the end B, then the velocity of A relative to B is in a direction perpendicular to AB (see Figure 3.6).

When a block slides on a rotating link the velocity is made up of two components, one being the velocity of the block relative to the link and the other the velocity of the link.

Accelerations: If the link has an angular acceleration α then there will be two components of acceleration in the diagram, a tangential component αl and a centripetal component of magnitude $\omega^2 l$ acting towards A.

When a block slides on a rotating link the total acceleration is composed of four parts: first, the centripetal acceleration towards O of magnitude $\omega^2 l$; second, the tangential acceleration αl; third, the acceleration of the block relative to the link; fourth, a tangential acceleration of magnitude $2v\omega$ known as Coriolis acceleration.

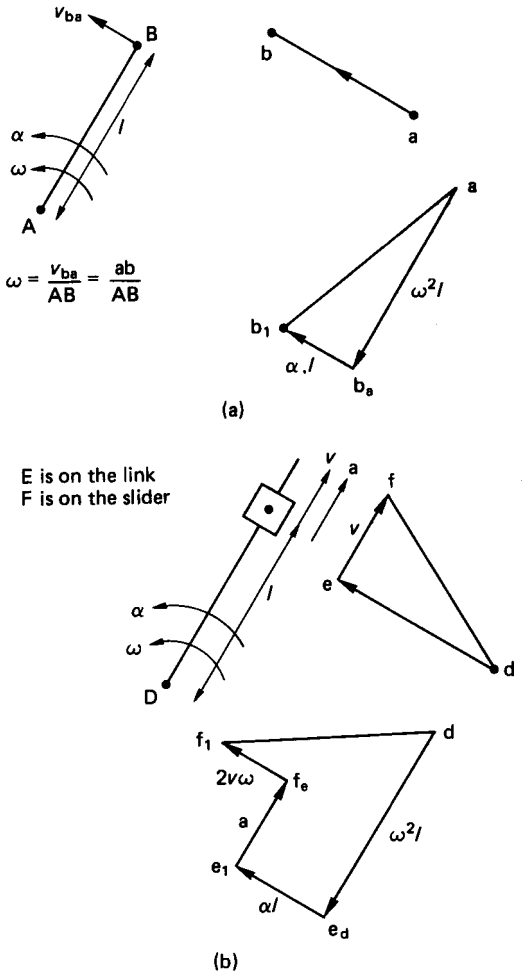


FIGURE 3.6

The direction of Coriolis acceleration is determined by rotating the sliding velocity vector through 90° in the direction of the link angular velocity ω.

3.4 LINEAR AND ANGULAR MOTION IN THREE DIMENSIONS

Motion of a Particle in a Moving Coordinate System

xyz is a moving coordinate system, with its origin at *O*, which has a position vector $\mathbf{R}$, a translational velocity vector $\dot{\mathbf{R}}$ and an angular velocity vector $\boldsymbol{\omega}$ relative to a

fixed coordinate system XYZ, origin at O'. Then the motion of a point P whose position vector relative to O is $\mathbf{P}$ and relative to O' is $\mathbf{r}$ is given by the following equations (see Figure 3.7):

$$\dot{\mathbf{r}} = \dot{\mathbf{R}} + \dot{\boldsymbol{\rho}}_r + \boldsymbol{\omega} \times \boldsymbol{\rho}$$

where $\boldsymbol{\rho}_r$ is the velocity of the point P relative to the moving system xyz and $\boldsymbol{\omega} \times \boldsymbol{\rho}$ is the vector product of $\boldsymbol{\omega}$ and $\boldsymbol{\rho}$:

$$\ddot{\mathbf{r}} = \ddot{\mathbf{R}} + \dot{\boldsymbol{\omega}} \times \boldsymbol{\rho} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{\rho}) + 2\boldsymbol{\omega} \times \dot{\boldsymbol{\rho}}_r + \ddot{\boldsymbol{\rho}}_t$$

where $\boldsymbol{\rho}_r$ is the acceleration of the point P relative to the moving system. Thus $\dot{\boldsymbol{r}}$ is the sum of:

1. The relative velocity $\ddot{\boldsymbol{\rho}}_r$
2. The absolute velocity $\mathbf{R}$ of the moving origin O
3. The velocity $\boldsymbol{\omega} \times \boldsymbol{\rho}$ due to the angular velocity of the moving axes xyz

and $\ddot{\mathbf{r}}$ is the sum of:

1. The relative acceleration $\ddot{\boldsymbol{\rho}}_r$
2. The absolute acceleration $\ddot{\mathbf{R}}$ of the moving origin O
3. The tangential acceleration $\dot{\boldsymbol{\omega}} \times \boldsymbol{\rho}$ due to the angular acceleration of the moving axes xyz
4. The centripetal acceleration $\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{\rho})$ due to the angular velocity of the moving axes xyz
5. Coriolis component acceleration $2\boldsymbol{\omega} \times \boldsymbol{\rho}_r$ due to the interaction of coordinate angular velocity and relative velocity

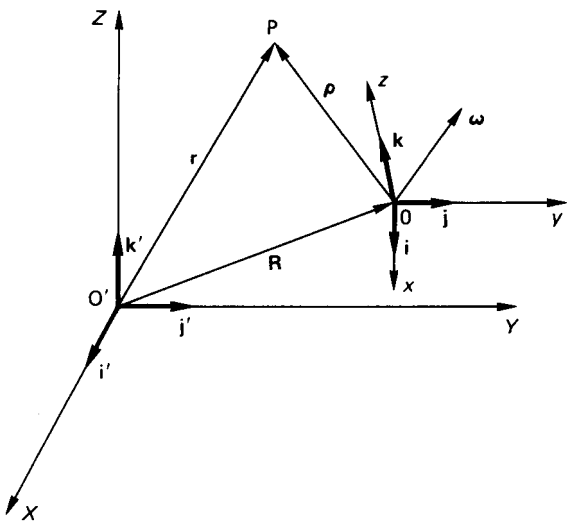


FIGURE 3.7

In all the vector notation a right-handed set of coordinate axes and the right-hand screw rule is used.

Gyroscopic Effects

Consider a rotor which spins about its geometric axis (see Figure 3.8) with an angular velocity ω. Then two forces F acting on the axle to form a torque T, whose vector is along the x-axis, will produce a rotation about the y axis. This is known as precession, and it has an angular velocity Ω. It is also the case that if the rotor is precessed then a torque T will be produced, where T is given by $T = I_{XX}\omega\Omega$. When this is observed it is the effect of gyroscopic reaction torque that is seen, which is in the opposite direction to the gyroscopic torque (Scarborough 1958).

3.5 BALANCING

In any rotational or reciprocating machine where accelerations are present, unbalanced forces can lead to high stresses and vibrations. The principle of balancing is such that by the addition of extra masses to the system the out-of-balance forces may be reduced or eliminated.

3.6 BALANCING OF ROTATING MASSES

Single Out-of-Balance Mass

One mass (m) at a distance r from the center of rotation and rotating at a constant angular velocity ω produces a force $m\omega^2 r$. This can be balanced by a mass M placed diametrically opposite at a distance R, such that $MR = mr$.

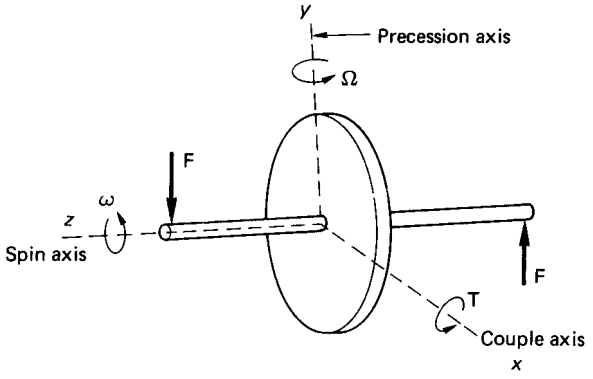


FIGURE 3.8

Several Out-of-Balance Masses in One Transverse Plane

If a number of masses (m_1, m_2, . . .) are at radii (r_1, r_2, . . .) and angles (θ_1, θ_2, . . .) (see Figure 3.9), then the balancing mass M must be placed at a radius R such that MR is the vector sum of all the mr terms.

Masses in Different Transverse Planes

If the balancing mass in the case of a single out-of-balance mass were placed in a different plane then the centrifugal force would be balanced. This is known as *static balancing*. However, the moment of the balancing mass about the original plane would lead to what is known as *dynamic unbalance.*

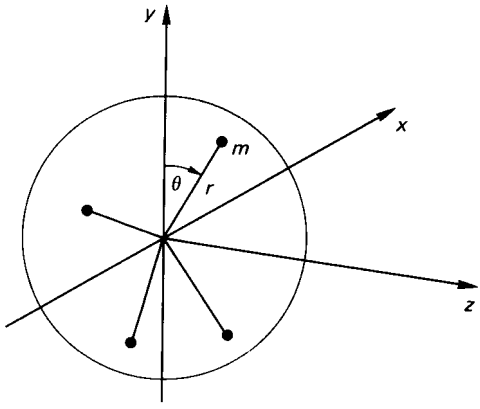
To overcome this, the vector sum of all the moments about the reference plane must also be zero. In general, this requires two masses placed in convenient planes (see Figure 3.10).

Balancing of Reciprocating Masses in Single-Cylinder Machines

The acceleration of a piston as shown in Figure 3.11 may be represented by the equation (Wilson 1959)

$$\ddot{x} = -\omega^2 r[\cos\theta + (1/n)\cos 2\theta + (1/4n)(\cos 2\theta - \cos 4\theta) + \ldots .*$$

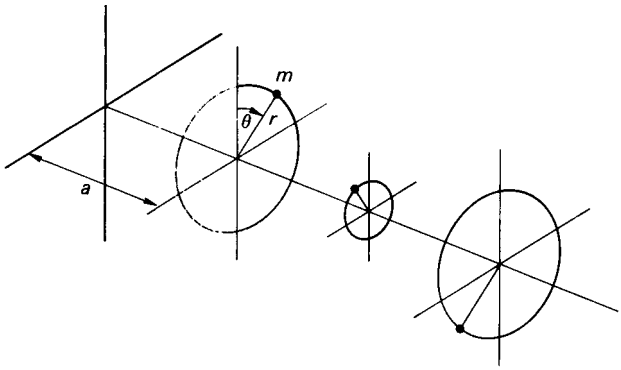
where $n = 1/r$. If n is large, then the equation may be simplified and the force given by



$\Sigma F_X = \Sigma m\omega^2 r \sin\theta = 0$
$\Sigma F_Y = \Sigma m\omega^2 r \cos\theta = 0$

FIGURE 3.9

*This equation forms an infinite series in which higher terms are small, and they may be ignored for practical situations.



$\Sigma F_X = \Sigma m\omega^2 r \sin\theta = 0$ and $\Sigma F_Y = \Sigma m\omega^2 r \cos\theta = 0$
as in the previous case, also
$\Sigma M_X = \Sigma m\omega^2 r \sin\theta \,.\, a = 0$
$\Sigma M_Y = \Sigma m\omega^2 r \cos\theta \,.\, a = 0$

FIGURE 3.10

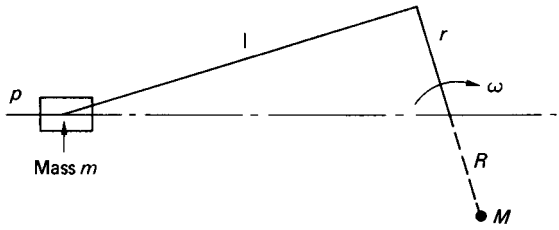


FIGURE 3.11

$$F = m\ddot{x} \simeq -m\omega^2 r[\cos\theta + (1/n)\cos 2\theta]$$

The term $m\omega^2 r \cos\theta$ is known as the *primary force* and $(1/n)m\omega^2 r \cos 2\theta$ as the *secondary force.* Partial primary balance is achieved in a single-cylinder machine by an extra mass M at a radius R rotating at the crankshaft speed. Partial secondary balance could be achieved by a mass rotating at 2ω. As this is not practical this is not attempted. When partial primary balance is attempted a transverse component $M\omega^2 R \sin\theta$ is introduced. The values of M and R are chosen to produce a compromise between the reciprocating and the transverse components.

Balancing of Reciprocating Masses in Multicylinder Machines

When considering multicylinder machines, account must be taken of the force produced by each cylinder and the moment of that force about some datum. The conditions for primary balance are

$$F = \Sigma m\omega^2 r \cos\theta = 0, \quad M = \Sigma m\omega^2 r \cos\theta \cdot a = 0$$

where a is the distance of the reciprocating mass m from the datum plane.

MECHANICAL ENGINEERING PRINCIPLES

In general, the cranks in multicylinder engines are arranged to assist primary balance. If primary balance is not complete, then extra masses may be added to the crankshaft but these will introduce an unbalanced transverse component. The conditions for secondary balance are

$$F = \Sigma m\omega^2(r/n)\cos 2\theta = \Sigma m(2\omega)^2(r/4n)\cos 2\theta = 0$$

and

$$M = \Sigma m(2\omega)^2(r/4n)\cos 2\theta \cdot a = 0$$

The addition of extra masses to give secondary balance is not attempted in practical situations.

PART 4

VIBRATIONS

Robert Paine

Christopher Beards

3.7 SINGLE-DEGREE-OF-FREEDOM SYSTEMS

The term *degrees of freedom* in an elastic vibrating system is the number of parameters required to define the configuration of the system. To analyze a vibrating system, a mathematical model is constructed, which consists of springs and masses for linear vibrations. The type of analysis then used depends on the complexity of the model.

Rayleigh's method: Rayleigh showed that if a reasonable deflection curve is assumed for a vibrating system, then by considering the kinetic and potential energies* an estimate to the first natural frequency could be found. If an inaccurate curve is used then the system is subject to constraints to vibrate it in this unreal form, and this implies extra stiffness such that the natural frequency found will always be high. If the exact deflection curve is used then the natural frequency will be exact.

Transverse Vibration of Beams

Consider a beam of length (l), weight per unit length (w), modulus (E), and moment of inertia (I). Then its equation of motion is given by

*Consider the equation of motion for an undamped system (Figure 3.13):

$$m \cdot \frac{d^2x}{dt^2} + kx = 0 \tag{3.4}$$

but

$$\frac{d^2x}{dt^2} = \frac{d}{dt}\left(\frac{dx}{dt}\right) = \frac{dx}{dt} \cdot \frac{d}{dx}\left(\frac{dx}{dt}\right) = \frac{1}{2} \cdot \frac{d}{dx}\left(\frac{dx}{dt}\right)^2$$

Therefore, equation (3.4) becomes

$$\frac{d}{dx}\left[\frac{1}{2}m\left(\frac{dx}{dt}\right)^2\right] + kx = 0$$

Integrating gives

$$\frac{1}{2}m\left(\frac{dx}{dt}\right)^2 + \frac{1}{2}kx^2 = \text{Constant}$$

the term $\frac{1}{2}m(dx/dt)^2$ represents the kinetic energy and $\frac{1}{2}kx^2$ the potential energy.

MECHANICAL ENGINEERING PRINCIPLES

$$EI\frac{d^4y}{dx^4} - w\omega^2 y/g = 0$$

where ω is the natural frequency. The general solution of this equation is given by

$$y = A\cos\beta x + B\sin\beta x + C\cosh\beta x + D\sinh\beta x$$

where $\beta^4 = w\omega^2/gEI$.

The four constants of integration A, B, C, and D are determined by four independent end conditions. In the solution trigonometrical identities are formed in β which may be solved graphically, and each solution corresponds to a natural frequency of vibration. Table 3.3 shows the solutions and frequencies for standard beams (Young and Feglar 1949).

Dunkerley's empirical method is used for beams with multiple loads. In this method the natural frequency (f_1) is found due to just one of the loads, the rest being ignored. This is repeated for each load in turn and then the natural frequency of vibration of the beam due to its weight alone is found (f_0). Then the natural frequency of vibration of the complete system (f) is given by

$$\frac{1}{f^2} = \frac{1}{f_0^2} + \frac{1}{f_1^2} + \frac{1}{f_2^2} + \frac{1}{f_3^2} + \cdots \frac{1}{f_n^2}$$

(see Cole 1950 for a more detailed explanation).

Whirling of shafts: If the speed of a shaft or rotor is slowly increased from rest there will be a speed where the deflection increases suddenly. This phenomenon is known as whirling. Consider a shaft with a rotor of mass m such that the center of gravity is eccentric by an amount e. If the shaft now rotates at an angular velocity ω then the shaft will deflect by an amount y due to the centrifugal reaction (see Figure 3.12). Then

$$m\omega^2(y + e) = ky$$

where k is the stiffness of the shaft. Therefore,

$$y = \frac{e}{(k/m\omega^2 - 1)}$$

When $(k/m\omega^2) = 1$, y is then infinite and the shaft is said to be at its critical whirling speed ω_c. At any other angular velocity ω the deflection y is given by

$$y = \left(\frac{\omega^2}{\omega_c^2 - \omega^2}\right) \cdot e$$

When $\omega < w_c$, y is the same sign as e and as ω increases towards ω_c the deflection theoretically approaches infinity. When $\omega > \omega_c$, y is opposite in sign to e and will eventually tend to $-e$. This is a desirable running condition with the center of gravity of the rotor mass on the static deflection curve. Care must be taken not to increase ω too high, as ω might start to approach one of the higher modes of vibration (Thompson 1983).

Torsional vibrations: The following section deals with transverse vibrating systems with displacements x and masses m. The same equations may be used for

TABLE 3.3

End conditions	Trigometric equation	Solutions		
		$\beta_1 l$	$\beta_2 l$	$\beta_3 l$
$x = 0,\ y = 0,\ y' = 0$ $x = 1,\ y = 0,\ y' = 0$	$\cos \beta l \cdot \cosh \beta l = 1$	4.730	7.853	10.966
$x = 0,\ y = 0,\ y' = 0$ $x = 1,\ y'' = 0,\ y''' = 0$	$\cos \beta l \cdot \cosh \beta l = -1$	1.875	4.694	7.855
$x = 0,\ y = 0,\ y'' = 0$ $x = 1,\ y = 0,\ y'' = 0$	$\sin \beta l = 0$	3.142	6.283	9.425
$x = 0,\ y = 0,\ y' = 0$ $x = 1,\ y = 0,\ y'' = 0$	$\tan \beta l = \tanh \beta l$	3.927	7.069	10.210

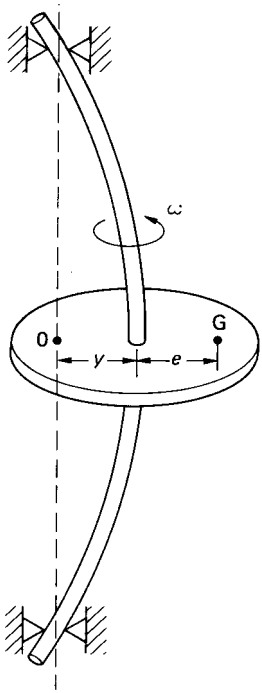


FIGURE 3.12

torsional vibrating systems by replacing x by θ the angular displacement and m by I, the moment of inertia.

Undamped Free Vibrations

The equation of motion is given by $m\ddot{x} + kx = 0$ or $\ddot{x} + \omega_n^2 x = 0$, where m is the mass, k the stiffness and $\omega_n^2 = k/m$, which is the natural frequency of vibration of the system (see Figure 3.13). The solution to this equation is given by

$$x = A \sin(\omega_n t + \alpha)$$

where A and α are constants which depend on the initial conditions. This motion is said to be *simple harmonic* with a time period $T = 2\pi/\omega_n$.

Damped Free Vibrations

The equation of motion is given by $m\ddot{x} + c\dot{x} + kx = 0$ (see Figure 3.14), where c is the viscous damping coefficient, or $\ddot{x} + (c/m)\dot{x} + \omega_n^2 x = 0$. The solution to this equation and the resulting motion depends on the amount of damping. If $c > 2m\omega_n$ the system is said to be overdamped. It will respond to a disturbance by slowly

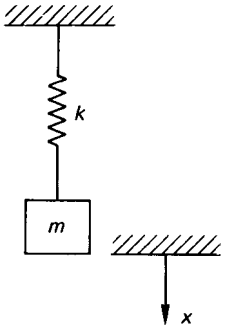


FIGURE 3.13

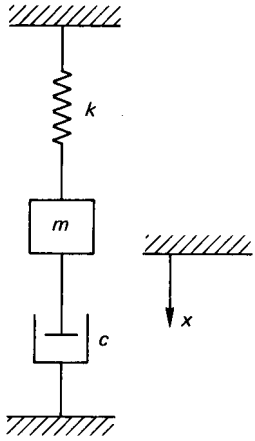


FIGURE 3.14

returning to its equilibrium position. The time taken to return to this position depends on the degree of damping (see Figure 3.15(c)). If $c = 2m\omega_n$ the system is said to be critically damped. In this case it will respond to a disturbance by returning to its equilibrium position in the shortest possible time. In this case (see Figure 3.15(b))

$$x = e^{-(c/2m)t(A+Bt)}$$

where A and B are constants. If $c < 2m\omega_n$, the system has a transient oscillatory motion given by

$$x = e^{-(c/2m)t}[C \sin(\omega_n^2 - c^2/4m^2)^{1/2}t + D(\cos \omega_n^2 - c^2/4m^2)^{1/2}t]$$

where C and D are constants. The period

$$T = \frac{2\pi}{(\omega_n^2 - c^2/4m^2)^{1/2}}$$

(see Figure 3.15(a)).

Logarithmic Decrement. A way to determine the amount of damping in a system is to measure the rate of decay of successive oscillations. This is expressed by a term called the *logarithmic decrement* (δ), which is defined as the natural logarithm of the ratio of any two successive amplitudes (see Figure 3.16):

$$\delta = \log_e(x_1/x_2)$$

where x is given by

$$x = e^{cT/m} \sin\left[\left(\omega^2 - \frac{c^2}{4m^2}\right)^{1/2} + \phi\right]$$

Therefore,

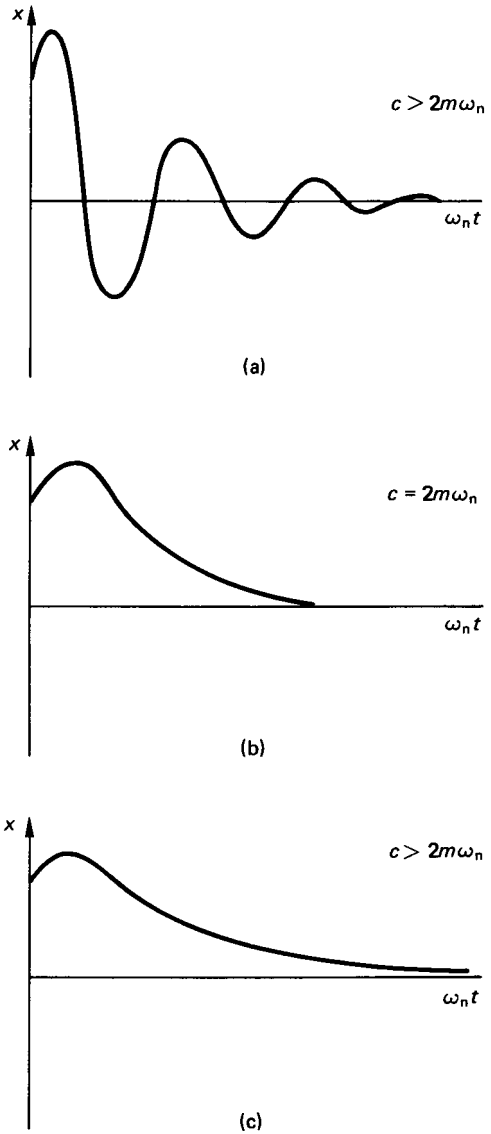


FIGURE 3.15

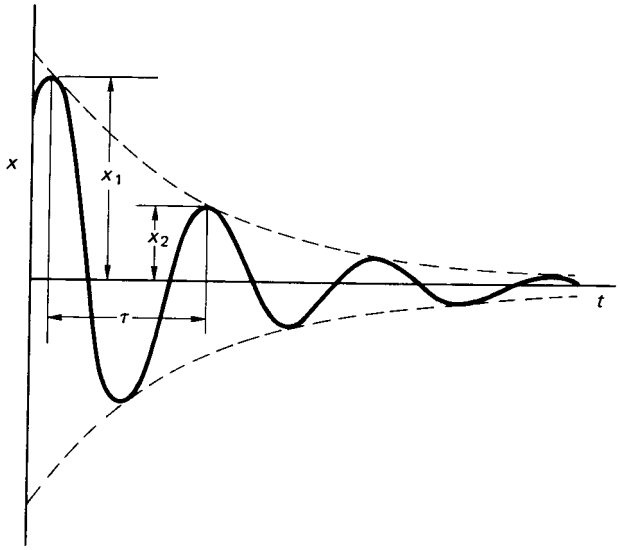


FIGURE 3.16

$$\delta = \log_e(e^{-cT/2m}/e^{-c(T+\tau)/2m})$$
$$= c\tau/2m$$

where τ is the period of damped oscillation.

If the amount of damping present is small compared to the critical damping, τ approximates to $2\pi/\omega$, and then

$$\delta = c\pi/m\omega_n$$

Forced Undamped Vibrations

The equation of motion is given by (see Figure 3.17)

$$m\ddot{x} + kx = F_0 \sin \omega t$$

or

$$\ddot{x} + \omega_n^2 = (F_0/m) \sin \omega t$$

The solution to this equation is

$$x = C \sin \omega_n t + D \cos \omega_n t + F_0 \cos \omega t[m(\omega_n^2 - \omega^2)]$$

where ω is the frequency of the forced vibration. The first two terms of the solution are the transient terms, which die out, leaving an oscillation at the forcing frequency of amplitude

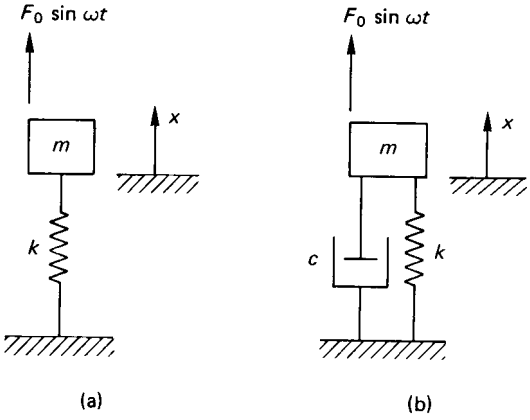


FIGURE 3.17

$$F_0/[m(\omega_n^2 - \omega^2)]$$

or

$$\frac{F_0}{k}\left(\frac{\omega_n^2}{\omega_n^2 - \omega^2}\right)$$

The term $\omega_n^2/(\omega_n^2 - \omega^2)$ is known as the dynamic magnifier, and it gives the ratio of the amplitude of the vibration to the static deflection under the load F_0. When $\omega = \omega_n$ the amplitude becomes infinite and resonance is said to occur.

Forced Damped Vibrations

The equation of motion is given by (see Figure 3.17(b))

$$m\ddot{x} + c\dot{x} + kx = F_0 \sin \omega t$$

or

$$\ddot{x} + (c/m)\dot{x} + \omega_n^2 = (F_0/m) \sin \omega t$$

The solution to this equation is in two parts: a transient part as in the undamped case which dies away, leaving a sustained vibration at the forcing frequency given by

$$x = \frac{F_0}{m} \frac{1}{[(\omega_n^2 - \omega^2)^2 + (c(\omega/m)^2]^{1/2}} \sin(\omega t - \alpha)$$

The term

$$\frac{\omega_n^2}{[(\omega_n^2 - \omega^2)^2 + (c\omega/m)]^2]^{1/2}}$$

is called the dynamic magnifier. Resonance occurs when $\omega \simeq \omega_n$. As the damping

is increased the value of ω for which resonance occurs is reduced. There is also a phase shift as ω increases tending to a maximum of π radians. It can be seen in Figure 3.18(a) that when the forcing frequency is high compared to the natural frequency the amplitude of vibration is minimized.

Forced Damped Vibrations Due to Reciprocating or Rotating Unbalance

Figure 3.19 shows two elastically mounted systems, (a) with the excitation supplied by the reciprocating motion of a piston, and (b) by the rotation of an unbalanced rotor. In each case the equation of motion is given by

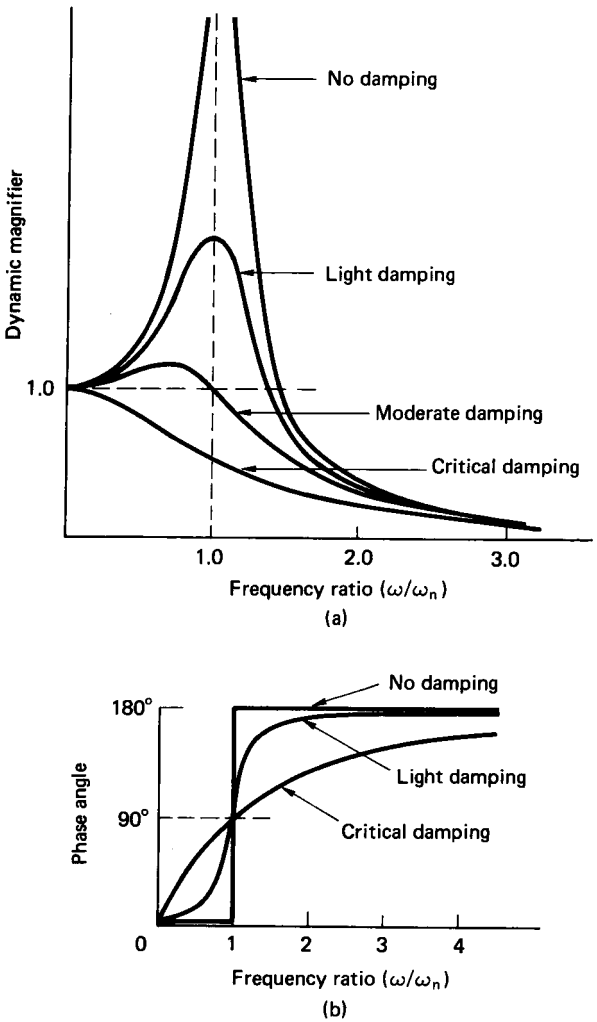


FIGURE 3.18

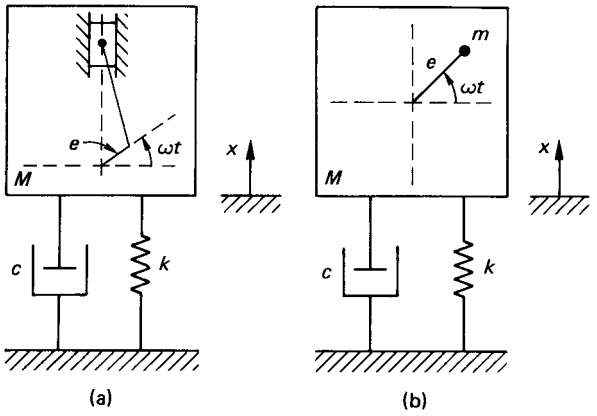


FIGURE 3.19

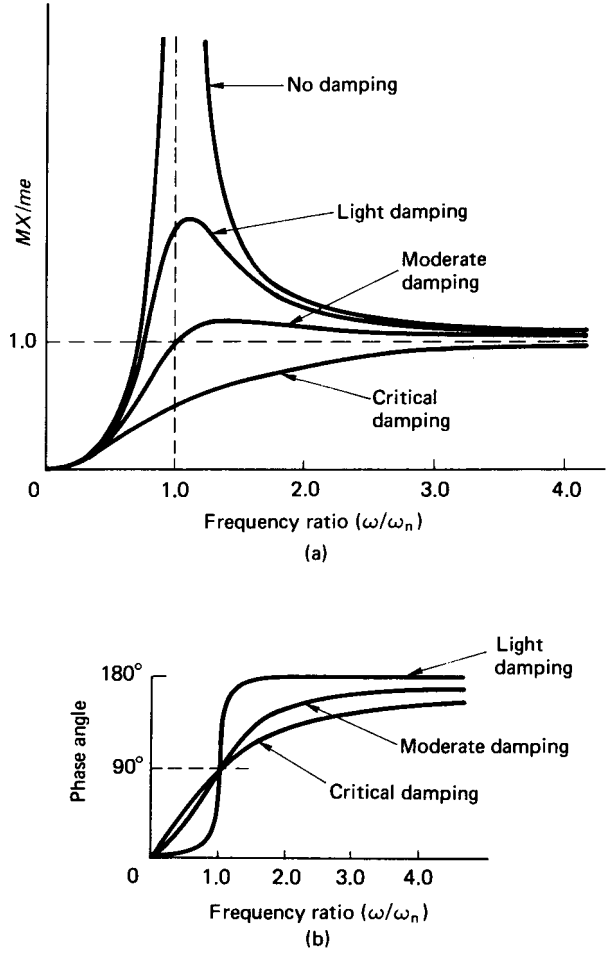


FIGURE 3.20

$$(M - m)\ddot{x} + c\dot{x} + kx = (me\omega^2) \sin \omega t$$

The solution of this equation is a sinusoid whose amplitude, X, is given by

$$X = \frac{me\omega^2}{\sqrt{[(K - M\omega^2)^2 + (c\omega)^2]}}$$

In representing this information graphically it is convenient to plot MX/me against ω/ω_n for various levels of damping (see Figure 3.20(a)). From this figure it can be seen that for small values of ω the displacement is small, and as ω is increased the displacement reaches a maximum when ω is slightly greater than ω_n. As ω is further increased the displacement tends to a constant value such that the center of gravity of the total mass M remains stationary. Figure 3.20(b) shows how the phase angle varies with frequency.

Forced Damped Vibration. If a system as shown in Figure 3.21 has a sinusoidal displacement applied to its base of amplitude, y, then the equation of motion becomes

$$m\ddot{x} + c\dot{x} + kx = ky + c\dot{y}$$

The solution of this equation yields

$$\frac{x}{y} = \sqrt{\left[\frac{k^2 + (c\omega)^2}{(k - m\omega^2)^2 + (c\omega)^2}\right]}$$

where x is the amplitude of motion of the system.

When this information is plotted as in Figure 3.22, it can be seen that for very small values of ω the output amplitude X is equal to the input amplitude Y. As ω is increased towards ω_n the output reaches a maximum. When $\omega = \sqrt{2}\ \omega_n$ the curves intersect and the effect of damping is reversed.

The curves in Figure 3.22 may also be used to determine the amount of sinusoidal force transmitted through the springs and dampers to the supports, i.e., the axis (X/Y) may be replaced by (F_t/F_0) where F_0 is the amplitude of applied force and F_t is the amplitude of force transmitted.

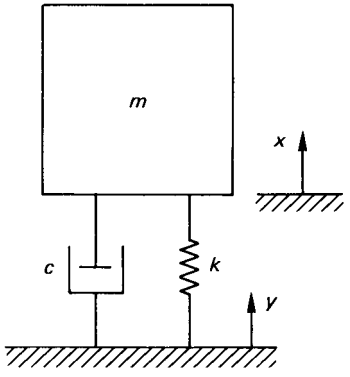


FIGURE 3.21

MECHANICAL ENGINEERING PRINCIPLES

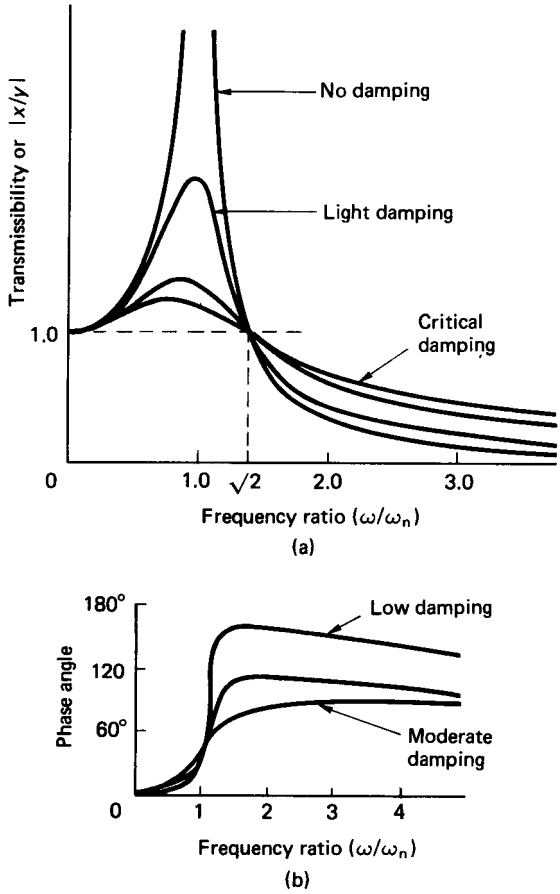


FIGURE 3.22

3.8 MULTI-DEGREE-OF-FREEDOM SYSTEMS

Normal Mode Vibration

The fundamental techniques used in modeling multi-degree-of-freedom systems may be demonstrated by considering a simple two-degree-of-freedom system as shown in Figure 3.23. The equations of motion for this system are given by

$$m_1\ddot{x}_1 + (k_1 + k_2)x_1 - k_2x_2 = 0$$

$$m_2\ddot{x}_2 + (k_3 + k_2)x_2 - k_2x_1 = 0$$

or in matrix form:

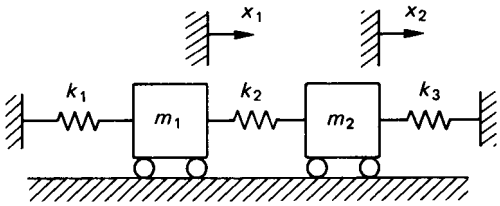


FIGURE 3.23

$$\begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \begin{Bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{Bmatrix} + \begin{bmatrix} (k_1 + k_2) & -k_2 \\ -k_2 & (k_3 + k_2) \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$$

Assuming the motion of every point in the system to be harmonic, then the solutions will take the form

$$x_1 = A_1 \sin \omega t$$

$$x_2 = A_2 \sin \omega t$$

where A_1 and A_2 are the amplitudes of the respective displacements. By substituting the values of x_1, x_2, $\ddot{x}_1$, and $\ddot{x}_2$ into the original equations the values of the natural frequencies of vibration may be found along with the appropriate mode shapes. This is a slow and tedious process, especially for systems with large numbers of degrees of freedom, and is best performed by a computer program.

The Holtzer Method

When only one degree of freedom is associated with each mass in a multimass system, then a solution can be found by proceeding numerically from one end of the system to the other. If the system is being forced to vibrate at a particular frequency, then there must be a specific external force to produce this situation. A frequency and a unit deflection is assumed at the first mass and from this the inertia and spring forces are calculated at the second mass. This process is repeated until the force at the final mass is found. If this force is zero, then the assumed frequency is a natural frequency. Computer analysis is most suitable for solving problems of this type.

Consider several springs and masses as shown in Figure 3.24. Then with a unit deflection at the mass m_1 and an assumed frequency ω there will be an inertia

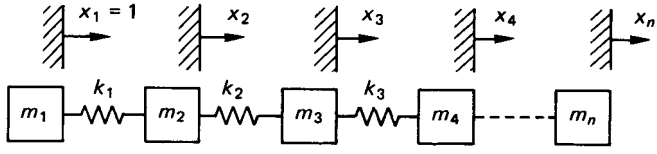


FIGURE 3.24

force of $m_1\omega^2$ acting on the spring with stiffness k_1. This causes a deflection of $m_1\omega^2/k_1$, but if m_2 has moved a distance x_2 then $m_1\omega^2/k_1 = 1 - x_2$ or $x_2 = 1 - m_1\omega^2/k_1$. The inertia force acting due to m_2 is $m_2\omega^2x_2$, thus giving the total force acting on the spring of stiffness k_2 as $[m_1\omega^2 + m_2\omega^2x_2]/k_2$. Hence the displacement at x_3 can be found and the procedure repeated. The external force acting on the final mass is then given by

$$\sum_{i=1}^{n} m_i\omega^2 x_i.$$

If this force is zero, then the assumed frequency is a natural one.

3.9 RANDOM VIBRATIONS

Introduction

If the vibration response parameters of a dynamic system are accurately known as functions of time, the vibration is said to be *deterministic*. However, in many systems and processes responses cannot be accurately predicted; these are called *random processes*. Examples of a random process are turbulence, fatigue, the meshing of imperfect gears, surface irregularities, the motion of a car running along a rough road, and building vibration excited by an earthquake (Figure 3.25).

A collection of sample functions $x_1(t)$, $x_2(t)$, $x_3(t)$, . . . , $x_n(t)$ which make up the random process $x(t)$ is called an *ensemble* (Figure 3.26). These functions may comprise, for example, records of pressure fluctuations or vibration levels, taken under the same conditions but at different times.

Any quantity which cannot be precisely predicted is nondeterministic and is known as a *random variable* or a *probabilistic quantity*. That is, if a series of tests are conducted to find the value of a particular parameter, x, and that value is found to vary in an unpredictable way that is not a function of any other parameter, then x is a random variable.

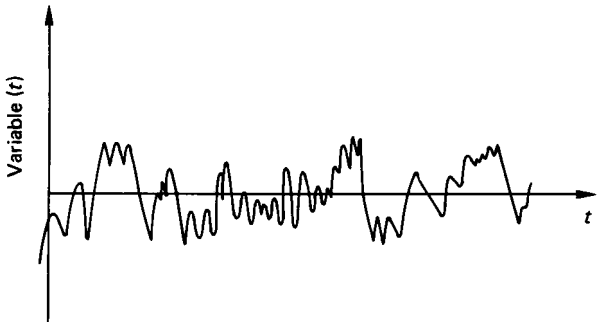


FIGURE 3.25 Example random process variable as $f(t)$.

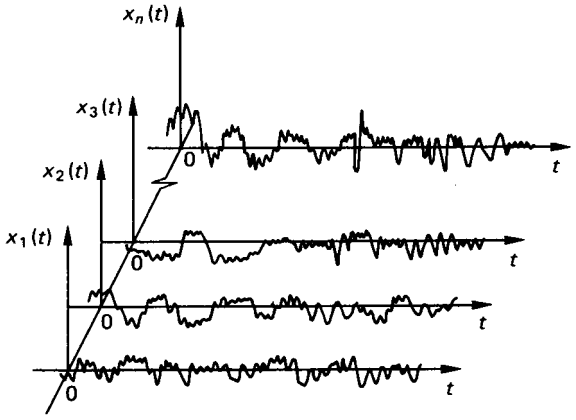


FIGURE 3.26 Ensemble of a random process.

Probability Distribution

If n experimental values of a variable x are $x_1, x_2, x_3, \ldots . x_n$, the probability that the value of x will be less than x' is n'/n, where n' is the number of x values which are less than or equal to x'. That is,

$$\text{Prob}(x \leqslant x') = n'/n$$

When n approaches ∞ this expression is the probability distribution function of x, denoted by $P(x)$, so that

$$P(x) = \underset{n\to\infty}{\text{Lt}} \; (n'/n)$$

The typical variation of $P(x)$ with x is shown in Figure 3.27. Since $x(t)$ denotes a physical quantity,

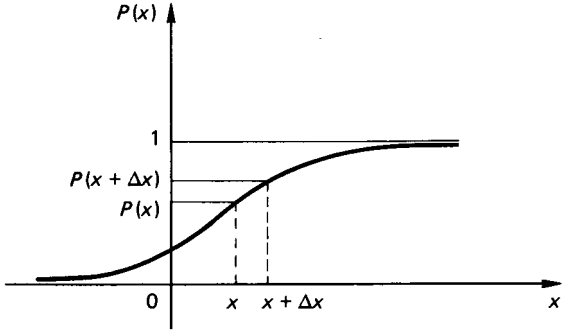


FIGURE 3.27 Probability distribution function as $f(x)$.

MECHANICAL ENGINEERING PRINCIPLES

$$\text{Prob}(x < -\infty) = 0, \text{ and } \text{Prob}(x < +\infty) = 1$$

The *probability density function* is the derivative of $P(x)$ with respect to x and this is denoted by $p(x)$. That is,

$$p(x) = \frac{dP(x)}{dx}$$

$$= \underset{\Delta x \to 0}{\text{Lt}} \left[\frac{P(x + \Delta x) - P(x)}{\Delta x} \right]$$

where $P(x + \Delta x) - P(x)$ is the probability that the value of $x(t)$ will lie between x and $x + \Delta x$ (Figure 3.27). Now

$$p(x) = \frac{dP(x)}{dx}$$

so that

$$P(x) = \int_{-\infty}^{x} p(x)dx$$

Hence,

$$P(\infty) = \int_{-\infty}^{\infty} p(x)dx = 1$$

so that the area under the probability density function curve is unity.

A random process is *stationary* if the joint probability density

$$p(x)(t_1), x(t_2), x(t_3), \ldots)$$

depends only on the time differences $t_2 - t_1$, $t_3 - t_2$, and so on, and not on the actual time instants. That is, the ensemble will look just the same if the time origin is changed. A random process is *ergodic* if every sample function is typical of the entire group.

The expected value of $f(x)$, which is written $E[f(x)]$ or $\overline{f(x)}$ is

$$E[f(x)] = \overline{f(x)} = \int_{-\infty}^{\infty} f(x)p(x)dx$$

so that the expected value of a stationary random process $x(t)$ is

$$E[x(t_1)] = E[x(t_1 + t)]$$

for any value of t.

If $f(x) = x$, the expected value or *mean value* of x, $E[x]$ or $\bar{x}$, is

$$E[x] = \bar{x} = \int_{-\infty}^{\infty} xp(x)dx$$

In addition, if $f(x) = x^2$, the *mean square value* of x, $\bar{x}^2$ is

$$E[x^2] = \bar{x}^2 = \int_{-\infty}^{\infty} x^2 \cdot p(x)dx$$

The *variance* of x, σ^2 is the mean square value of x about the mean, that is,

$$\sigma^2 = E[(x - \bar{x})^2] = \int_{-\infty}^{\infty} (x - \bar{x})^2 p(x)dx = \overline{(x^2)} - (\bar{x})^2$$

σ is the *standard deviation* of x, hence

$$\text{Variance} = (\text{Standard deviation})^2$$
$$= \{\text{Mean square} - (\text{Mean})^2\}$$

If two or more random variables x_1 and x_2 represent a random process at two different instants of time, then

$$E[f(x_1,x_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1,x_2)p(x_1,x_2)dx_1dx_2$$

and if t_1 and t_2 are the two instants of time,

$$E[x(t_1),x(t_2)] = R(t_1,t_2)$$

which is the *autocorrelation function* for the random process (Figure 3.28).

For random processes which are stationary,

$$E[x(t_1),x(t_2)] = R(t_1,t_2) = R(t_2 - t_1) = R(\tau), \text{ say,}$$

since the average depends only on time differences. If the process is also ergodic, then

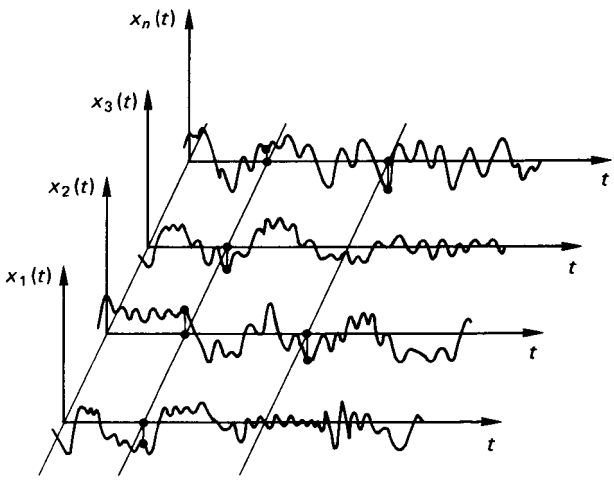


FIGURE 3.28 Random processes.

$$R(\tau) = \underset{T\to\infty}{\text{Lt}} \frac{1}{2\pi} \int_{-T}^{T} x(t)x(t+\tau)dt$$

It is worth noting that

$$R(0) = E[x(t)^2] = \underset{T\to\infty}{\text{Lt}} \frac{1}{2T} \int_{-T}^{T} x^2(t)dt$$

which is the average power in a sample function.

Random Processes

The most important random process is the *Gaussian* or *normal random process*. This is because a wide range of physically observed random waveforms can be represented as Gaussian processes, and the process has mathematical features which make analysis relatively straightforward.

The probability density function of a Gaussian process $x(t)$ is

$$p(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-1/2[(x-\bar{x})/\sigma]^2]}$$

where σ is the standard deviation of x and $\bar{x}$ is the mean value of x. The values of σ and $\bar{x}$ may vary with time for a non-stationary process but are independent of time if the process is stationary.

One of the most important features of the Gaussian process is that the response of a linear system to this form of excitation is usually another (but still Gaussian) random process. The only changes are that the magnitude and standard deviation of the response may differ from those of the excitation.

A Gaussian probability density function is shown in Figure 3.29. It can be seen to be symmetric about the mean value $\bar{x}$, and the standard deviation σ controls the spread.

The probability that $x(t)$ lies between $-\lambda\sigma$ and $+\lambda\sigma$, where λ is a positive number, can be found since, if $\bar{x} = 0$,

$$\text{Prob}\{-\lambda\sigma \le x(t) \le +\lambda\sigma\} = \int_{-\lambda\sigma}^{+\lambda\sigma} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-1/2(x^2/\sigma^2)}dx$$

Figure 3.30 shows the Gaussian probability density function with zero mean. This integral has been calculated for a range of values of λ and the results are given in Table 3.4. The probability that $x(t)$ lies outside the range $-\lambda\sigma$ to $+\lambda\sigma$ is 1 minus the value of the above integral. This probability is also given in Table 3.4.

Spectral Density

The spectral density $S(\omega)$ of a stationary random process is the Fourier transform of the autocorrelation function $R(\tau)$, and is given by

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R(\tau)e^{-i\omega\tau}d\tau$$

The inverse, which also holds true, is

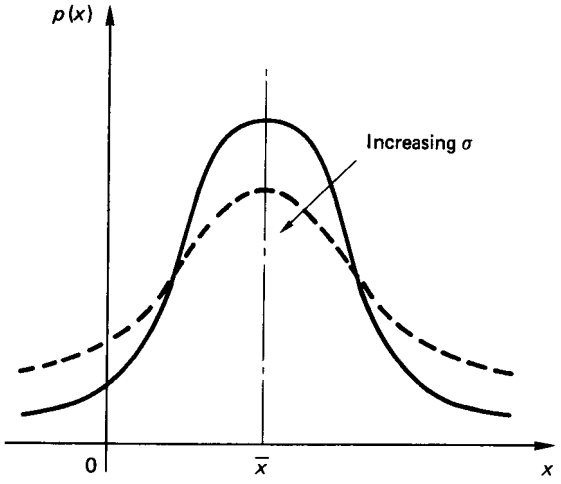


FIGURE 3.29 Gaussian probability density function.

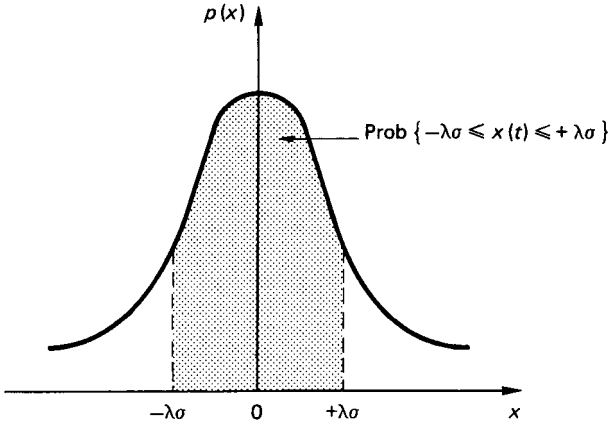


FIGURE 3.30 Gaussian probability density function with zero mean.

$$R(\tau) = \int_{-\infty}^{\infty} S(\omega) e^{-i\omega\tau} d\omega$$

If $\tau = 0$

$$R(0) = \int_{-\infty}^{\infty} S(\omega) d\omega$$

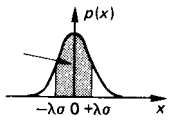
$$= E[x^2]$$

That is, the mean square value of a stationary random process x is the area under

TABLE 3.4

Value of λ	Prob$[-\lambda\sigma \leq x(t) \leq \lambda\sigma]$	Prob$[\lvert x(t)\rvert > \lambda\sigma]$
0	0	1.0000
0.2	0.1585	0.8415
0.4	0.3108	0.6892
0.6	0.4515	0.5485
0.8	0.5763	0.4237
1.0	0.6827	0.3173
1.2	0.7699	0.2301
1.4	0.8586	0.1414
1.6	0.8904	0.1096
1.8	0.9281	0.0719
2.0	0.9545	0.0455
2.2	0.9722	0.0278
2.4	0.9836	0.0164
2.6	0.9907	0.0093
2.8	0.9949	0.0051
3.0	0.9973	0.0027
3.2	0.9986	0.00137
3.4	0.9993	0.00067
3.6	0.9997	0.00032
3.8	0.9998	0.00014
4.0	0.9999	0.00006

Prob$[-\lambda\sigma \leq x(t) \leq +\lambda\sigma]$



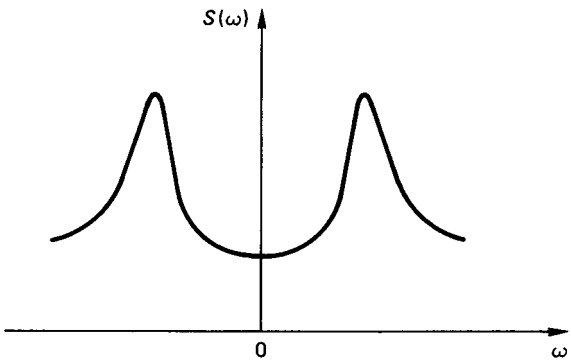


FIGURE 3.31 Typical spectral density function.

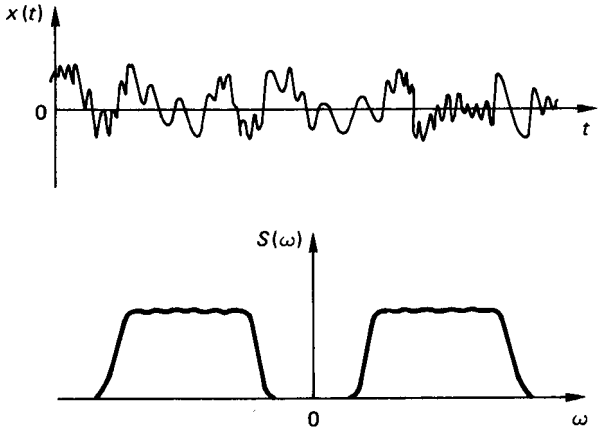


FIGURE 3.32 Wide-band process.

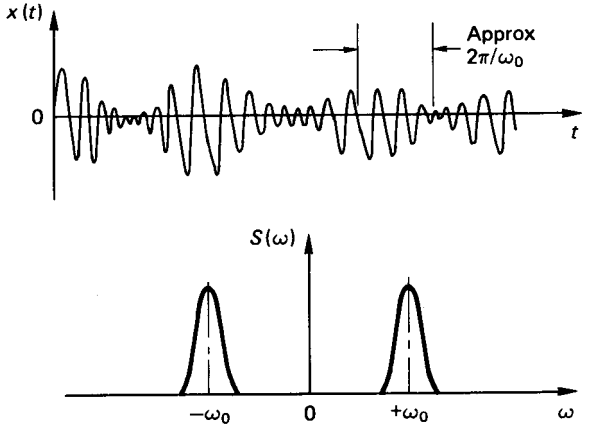


FIGURE 3.33 Narrow-band process.

the $S(\omega)$ against frequency curve. A typical spectral density function is shown in Figure 3.31.

A random process whose spectral density is constant over a very wide frequency range is called *white noise*. If the spectral density of a process has a significant value over a narrower range of frequencies, but one which is nevertheless still wide compared with the center frequency of the band, it is termed a *wide-band* process (Figure 3.32). If the frequency range is narrow compared with the centre frequency it is termed a *narrow-band process* (Figure 3.33). Narrow-band processes frequently occur in engineering practice because real systems often respond strongly to specific exciting frequencies and thereby effectively act as a filter.

PART 5
MECHANICS OF FLUIDS

Peter Tucker

3.10 INTRODUCTION

Fluid is one of the two states in which matter can exist, the other being solid. In the fluid state the matter can flow; it will, in general, take the shape of its container. At rest a fluid is not able to sustain shear forces.

Some ''solids'' may flow over a long period (glass window panes thicken at the base after a long time in a vertical position). The substances considered in fluid mechanics are those that are continuously fluid.

Fluid mechanics is a study of the relationships between the effects of forces, energy and momentum occurring in and around a fluid system. The important properties of a fluid in fluid mechanics terms are *density, pressure, viscosity, surface tension,* and, to some extent, *temperature,* all of which are intensive properties.

Density is the mass per unit volume of the substance. *Pressure* is the force per unit area exerted by the fluid on its boundaries. *Viscosity* is a measure of the fluid's resistance to flow and may be considered as internal friction. The higher the coefficient of viscosity, the greater the resistance. *Surface tension* is a property related to intermolecular attraction in the free surface of a liquid resulting in the apparent presence of a very thin film on the surface. The meniscus at the intersection of a liquid and its container wall and capillarity are further examples of intermolecular attraction.

Temperature is more relevant to thermodynamics than to fluid mechanics. It indicates the state of thermal equilibrium between two systems or, more loosely, the level of thermal energy in a system.

3.11 FLUID STATICS

Pressure at a Depth

The variation of pressure p and depth h in a fluid of density ρ is given by

$$\int_{p_1}^{p_2} dp = \int_{h_1}^{h_2} \rho g dh \tag{3.5}$$

Most liquids are assumed to be of constant density ρ. In such a liquid the pressure at a depth h below a free surface is given by

$$p = p_0 + \rho g h \tag{3.6}$$

where p_0 is the pressure above the free surface.

For gases equation (3.5) may be solved only if the relationship between ρ and h is known. A typical case is the atmosphere, where the relationship may be taken as polytropic or isothermal, depending on the altitude. Tables relating the properties of the atmosphere to altitude are readily available as the International Atmosphere.

Pressure Measurement

Pressure may be expressed as a pressure p in Pa, or as a pressure head h in m of the fluid concerned. For a fluid of density ρ, $p = \rho g h$. There are various instruments used to measure pressure.

Manometers. Manometers are differential pressure-measuring devices, based on pressure due to columns of fluid. A typical U-tube manometer is shown in Figure 3.34(a). The difference in pressure between vessel A containing a fluid of density ρ_A and vessel B containing fluid of density ρ_B is given by

$$p_A - p_B = \rho_B g Z_B + (\rho_m - \rho_B) g h - \rho_A g Z_A \tag{3.7}$$

where h is the difference in the levels of the manometer fluid of density ρ_m and $\rho_m > \rho_A$ and $\rho_m > \rho_B$. If $\rho_A = \rho_B = \rho$, then the difference in pressure head is

$$\frac{p_A - p_B}{\rho g} = Z_B - Z_A + \left(\frac{\rho_m}{\rho} - 1\right) h \tag{3.8}$$

If $\rho_m < \rho_A$ and $\rho_m < \rho_B$ then an inverted U-tube manometer is used as shown in Figure 3.34(b). In this case the pressure difference is

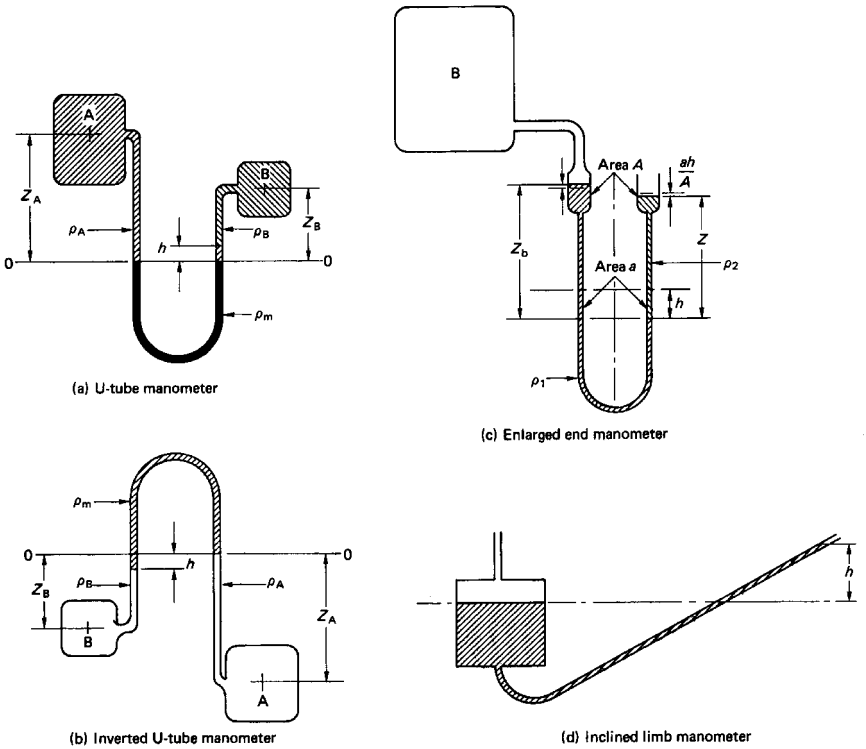


FIGURE 3.34

$$p_A - p_B = \rho_A g Z_A + (\rho_B - \rho_m) g h - \rho_B g Z_B \tag{3.9}$$

and if $\rho_A = \rho_B = \rho$ then the difference in pressure head is

$$\frac{p_A - p_B}{\rho g} = Z_A - Z_B + \left(1 - \frac{\rho_m}{\rho}\right) h \tag{3.10a}$$

of if $\rho_m << \rho$ (the manometer fluid a gas and A and B containing liquid),

$$\frac{p_A - p_B}{\rho g} = Z_A - Z_B + h \tag{3.10b}$$

The accuracy of a U-tube manometer may be increased by sloping one of the legs to increase the movement of the fluid interface along the leg for a given difference in vertical height. This may be further enhanced by replacing the vertical leg by a reservoir and the inclined leg by a small-bore tube (Figure 3.34(d)).

Another method is to increase the cross-sectional area of the ends of the legs (or one of the legs), as shown in Figure 3.34(c), so that a small movement of the free surfaces in the enlarged ends results in a large movement of the surface of separation.

Dial Gauges. Most pressure dial guages make use of a *Bourdon tube*. This is a curved tube with an oval cross section. Increase in pressure causes the tube to straighten, decrease makes it bend. The movement of the free end turns a pointer over a scale, usually via a rack-and-pinion mechanism. The scale may be calibrated in the required pressure units.

Diaphragm Gauges. In these gauges the pressure changes produce a movement in a diaphragm which may be detected by a displacement transducer, or by the output from strain gauges attached to the diaphragm surface.

Piezoelectric Transducers. A piezoelectric crystal produces a voltage when deformed by an external force. This induced charge is proportional to the impressed force and so the output can be used to supply a signal to a measuring device which may be calibrated in pressure units.

Fortin Barometer. Barometers are used to measure the ambient or atmospheric pressure. In the Fortin barometer a column of mercury is supported by the atmospheric pressure acting on the surface of the mercury reservoir. The height h of the column above the reservoir surface, usually quoted as millimeters of mercury (mm Hg), may be converted to pressure units p_0 by

$$\begin{aligned} p_0 &= \rho g h = 13.6 \times 9.81 h \\ &= 133.42 h \text{ Pa} \end{aligned} \tag{3.11}$$

Aneroid Barometer. In this device the atmospheric pressure tends to compress an evacuated bellows against the elasticity of the bellows. The movement of the free end of the bellows drives a pointer over a dial (or a pen over a drum graph) to indicate (or record) atmospheric pressure variations.

Buoyancy

When a body is immersed in a fluid the difference in pressure over the depth of the body produces a displacement force on the body. The first recognition of this is attributed to Archimedes.

Displacement Force. The buoyancy or displacement force F_B on a body fully or partially immersed in a fluid is equal to the weight of the volume of the fluid equivalent to the immersed volume of the body (the weight of the displaced volume V_D of the fluid):

$$F_B = \rho g V_D \tag{3.12}$$

This buoyancy force acts vertically upwards through the centroid of the displaced volume, which is known as the *center of buoyancy* (B). If the buoyancy force is equal to the weight of the body then the body will float in the fluid. If the weight of the body is greater than the buoyancy force then the body will sink. If the buoyancy force is greater than the weight of the body then the body will rise.

In a liquid, for example, a body will sink until the volume of liquid displaced has a weight which is equal to that of the body. If the body is more dense than the liquid then the body will not float at any depth in the liquid. A balloon will rise in air until the density of the air is such that the weight of the displaced volume of air is equal to the weight of the balloon.

Stability of a Floating Body. Figure 3.35 shows bodies in various stages of equilibrium. A body is in *stable* equilibrium if a small displacement produces a restoring force or moment as for the ball in the saucer in Figure 3.35(a) or the floating bodies in (d) and (g). A body is in *unstable* equilibrium if a small displacement produces a disturbing force or moment as for the ball in Figure 3.35(b) or the floating bodies (e) and (h). A body is in *neutral* equilibrium if a small displacement produces no force or moment as for the ball in Figure 3.35(c) or the floating bodies in (f) and (i).

For a partially immersed body, the point at which the line of action of the buoyancy force F_B cuts the vertical center line of the floating body in the displaced position is known as the *metacenter* (M). For a floating body to be stable M must lie above the body's center of gravity, G. If M lies below G the body is unstable; if M lies on G the body is in neutral equilibrium. The distance GM is known as the *metacentric height*. The distance of the metacenter above the center of buoyancy

$$BM = \frac{I}{V_D} \tag{3.13}$$

where I second moment of area of the body at the water line (liquid surface) about its central axis normal to the direction of displacement. (See Table 3.5.)

Period of Oscillation of a Stable Floating Body. A floating body oscillates with the periodic time T of a simple pendulum of length k^2/GM, where k is the radius of gyration of the body about its axis of rotation. The periodic time is given by

$$T = 2\pi \left(\frac{k^2}{GMg} \right)^{0.5} \tag{3.14}$$

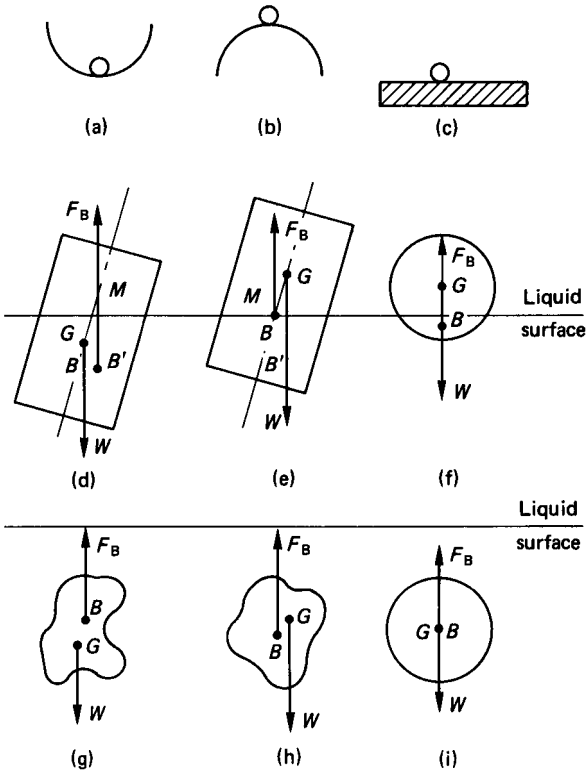


FIGURE 3.35 Stability.

3.12 FLUID FLOW

Definitions

Continuity. For almost all analysis, a fluid is considered to be a *continuum*, that is, with nondiscontinuities or cavities in the flow stream. Cavitation, two-phase flow, "bubbly" flow, etc. are special cases with nonstandard relationships.

Therefore for a continuum, by considering the flow through an elemental cuboid the *continuity equation* in three dimensions may be shown to be

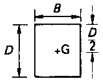
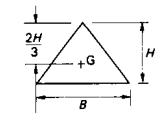
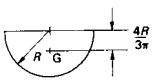
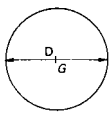
$$\frac{\partial}{\partial x}(\rho v_x) + \frac{\partial}{\partial y}(\rho v_y) + \frac{\partial}{\partial z}(\rho v_z) = 0 \tag{3.15}$$

where v_x is the fluid velocity in the x direction, etc. For a fluid of constant density

$$\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} = 0 \tag{3.16}$$

That is, the velocity of an incompressible fluid flow cannot increase in all three directions at the same time without producing discontinuity or cavitation.

TABLE 3.5 Second Moments of Area

	Parallel axis theorem $I_X = I_G + A\bar{x}^2$	
	Area A	I_G
	BD	$\frac{BD^3}{12}$
	$\frac{BH}{2}$	$\frac{BD^3}{36}$
	$\frac{\pi R^2}{2}$	$0.1102R^4$
	$\frac{\pi D^2}{4}$	$\frac{\pi D^4}{64}$

For two-dimensional flow:

$$\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} = 0 \tag{3.17}$$

For one-dimensional flow the continuity equation may be linked with the *conservation of mass,* which states that for steady flow conditions mass flow rate, $\dot{m}$, is constant throughout a flow system:

$$\dot{m} = \rho A v \tag{3.18}$$

where A is the cross-sectional area normal to the direction of flow.

***Circulation* Γ.** Circulation is defined as the line integral of the tangential velocity around a closed contour:

$$\Gamma = \oint v_s ds \tag{3.19}$$

Γ is positive if the closed contour is on the left.

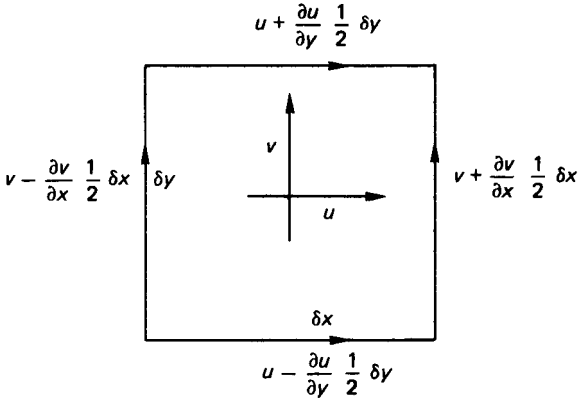
Vorticity ζ. Vorticity is defined as the circulation per unit area, and by considering the circulation around the element in Figure 3.36(a) it can be shown that

$$\zeta = \frac{\Gamma}{\partial x \partial y} = \frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \tag{3.20}$$

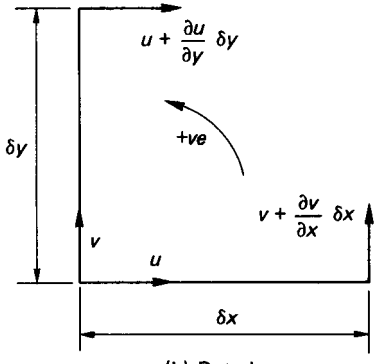
Rotation ω. Rotation is defined as the instantaneous mean angular velocity of two mutually perpendicular lines in a plane of the flow field. By considering the angular velocities of the two lines *OA* and *OB* in Figure 3.36(b) it can be shown that

$$\omega = \frac{1}{2}\left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y}\right) \tag{3.21}$$

or the rotation is equal to half the vorticity.



(a) Vorticity



(b) Rotation

FIGURE 3.36

Stream Lines. The stream line is a line drawn in a flow stream which is everywhere tangential to the direction of flow. A family of stream lines may be described mathematically by a stream function ψ, where ψ fn(x,y). Each stream line has the same function with a value of ψ peculiar to that line.

Stream Tubes. Since a line has no thickness, there can be no flow along a stream line. The stream tube is a concept introduced to enable flow along a stream line to be studied. It is a tube of infinitely small cross-section with a stream line as its axis.

Energy. Energy is the stored form of *heat* and *work*. The basic concepts applied in fluid mechanics are:

- The conservation of energy
- That energy is transferred only as heat or work
- That energy in a fluid flow system is stored only as *internal energy, kinetic energy,* or *potential energy*

Other forms of energy (electrical, magnetic, chemical, etc.) may have to be taken into account in some circumstances, but are not usually included in general fluid mechanics relationships.

Enthalpy and *entropy* need to be considered for gas flow analysis. The basic energy-flow equation is the *steady-flow energy equation:*

$$\dot{Q} + \dot{W} = \dot{m}\Delta\left(h + \frac{v^2}{2} + gZ\right) \tag{3.22}$$

where $\dot{Q}$ = the rate of heat transfer
$\dot{W}$ = the rate of work transfer (power)
h = the specific enthalpy (if e is the specific internal energy, p the pressure and ρ the fluid density, then $h = e + (p/\rho)$)
Z = the height above some datum
v = the mean velocity of flow

Specific means "per unit mass." For non-steady flow conditions, either *quasi-steady* techniques or the integration of infinitely small changes may be employed.

Momentum. Momentum is the product of mass and velocity (mv). Newton's laws of motion state that the force applied to a system may be equated to the rate of change of momentum of the system, in the direction of the force. The change in momentum may be related to time and/or displacement. In a steady flow situation the change related to time is zero, so the change of momentum is usually taken to be the product of the mass flow rate and the change in velocity with displacement. Hence, the force applied across a system is

$$F = \dot{m}\Delta v \tag{3.23}$$

where Δv is the change in velocity in the direction of the force F.

For flow in two or three dimensions the resultant force may be obtained by resolving the forces in the usual way. The flow round an expanding bend shown in Figure 3.37 is a typical example. The force in the x direction, F_x, and the force in the y direction, F_y, are given by

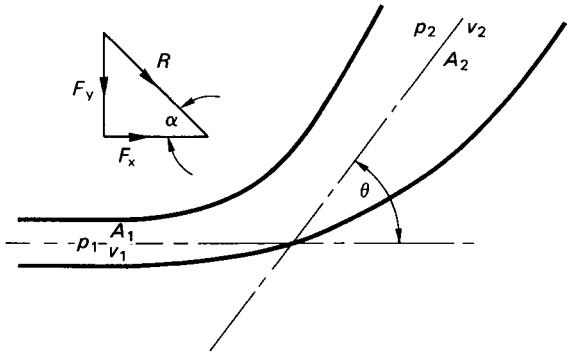


FIGURE 3.37 Expanding bend.

$$F_x = p_1A_1 + \dot{m}v_1 - (p_2A_2 + \dot{m}v_2)\cos\theta \tag{3.23a}$$

$$F_y = -(p_2A_2 + \dot{m}v_2)\sin\theta \tag{3.23b}$$

from which the resultant force F_R and its angle of inclination α can be found:

$$F_R = \{F_x^2 + F_y^2\}^{0.5} \tag{3.24}$$

and

$$\alpha = \tan^{-1}\left(\frac{F_y}{F_x}\right) \tag{3.25}$$

In pipe flow the pressure forces pA must not be forgotten.

Equations of Motion

Application of the momentum equation in three dimensions to an irrotational, inviscid fluid flow leads to the *Euler equation:*

$$-\frac{Dv}{Dt} = \frac{1}{\rho}\nabla p + g\nabla h \tag{3.26}$$

which for steady flow along a stream tube becomes:

$$\frac{1}{\rho}dp + vdv + gdZ = 0 \tag{3.27}$$

Integration for a constant-density fluid gives:

$$\frac{p}{\rho} + \frac{v^2}{2} + gZ = \text{constant} \tag{3.28}$$

These energy per unit mass terms may be converted to energy per unit weight terms, or heads, by dividing by g to give:

$$\frac{p}{\rho g} + \frac{v^2}{2g} + Z = \text{constant} \tag{3.29}$$

which is the *Bernoulli* (or constant head) equation.

These equations are the generally more useful simplifications of the *Navier–Stokes* equation:

$$\frac{Dv}{Dt} = \rho B - \nabla p + \nabla\{u(\nabla v + \nabla \epsilon)\} \tag{3.30}$$

where B is the body force and ϵ the rate of expansion.

Incompressible Pipe Flow

Flow Regimes. The two major flow regimes are *laminar* and *turbulent*. Laminar flow may be fairly accurately modelled mathematically. The fluid moves in smooth layers and the velocity is everywhere tangential to the direction of motion. Any perturbations are quickly dampened out by the fluid viscosity.

In turbulent flow the mathematical models usually need to be empirically modified. Viscous damping may not be sufficient to control the perturbations, so that the fluid does not move in smooth layers and the instantaneous velocity may have components at an angle to the direction of motion.

The ratio of inertia forces to viscous forces in a fluid flow is known as *Reynolds' number* (*Re*). In a pipe diameter D, with a fluid of density ρ and dynamic viscosity η flowing with velocity v, Reynolds' number $Re = \rho Dv/\eta$.

A high value of $Re > 2300$ indicates relatively low damping, predicting turbulent flow. A low value of $Re < 2000$ indicates relatively high damping, predicting laminar flow. These values were suggested in an historical experiment by Osborne Reynolds.

Pipe Losses (friction). Liquids (and gases under small pressure changes) flowing through pipes usually behave as incompressible fluids. Within the flow there is a relationship between the shear stress in the fluid and the gradient of the change of velocity across the flow. In most light liquids and gases, the relationship approximates to the *Newtonian* one:

$$\tau = \eta \frac{dv}{dy} \tag{3.31}$$

where τ is the shear stress in the fluid, dv/dy the gradient of the velocity distribution across the pipe and η the dynamic viscosity.

The viscosity of the fluid produces not only the velocity variation across the flow but also a loss of useful energy along the pipe usually regarded as a friction loss. The force associated with this loss of energy appears as a shear force in the fluid at the pipe wall. A relationship between the shear stress at the pipe wall τ_0 and the *friction coefficient, f* is:

$$\tau_0 = \frac{1}{2} \rho v^2 f \tag{3.32}$$

where v is the average flow velocity.

For use in pipe flow problems with viscous fluids the Bernoulli equation (3.29) may be adapted to include a head loss term, h_L. Applied between two positions (1) and (2) in a pipe in a gravitational field of acceleration g, the head equation gives:

$$\frac{p_1}{\rho g} + \frac{v_1^2}{2g} + Z_1 = \frac{p_2}{\rho g} + \frac{v_2^2}{2g} + Z_2 + h_L \tag{3.33}$$

where the head loss term h_L is the loss of energy per unit *weight* of fluid flowing.

Note that if a pump, say, is introduced between (1) and (2) an energy gain per unit weight term h_w, equivalent to the output of the pump written as a head, should be added to the left-hand side of the equation to give

$$\frac{p_1}{\rho g} + \frac{v_1^2}{2g} + Z_1 + h_w = \frac{p_2}{\rho g} + \frac{v_2^2}{2g} + Z_2 + h_L \tag{3.34}$$

The relationship used to determine the head loss in a pipe depends on the flow regime in operation as well as the type and surface finish of the pipe wall.

A mathematical analysis of laminar flow may be used to obtain an expression for the head loss along a pipe in terms of the fluid properties, pipe dimensions and flow velocity. Relating the pressure change along a length, L, of pipe of diameter, D, to the change in shear force across the flow produces *Poiseuille's equation:*

$$h_L = 32\,\frac{\eta v L}{\rho g D^2} \tag{3.35}$$

If the flow regime is turbulent, then the relationships in the flow cannot be easily described mathematically, but the head loss may be derived by equating the shear force at the pipe wall to the change in pressure force along the pipe. This gives the *D'Arcy equation:*

$$h_L = \frac{4fL}{D}\,\frac{v^2}{2g} \tag{3.36}$$

This relationship may also be established using dimensional analysis.

Unfortunately, the friction coefficient, f is *not* a constant but depends on the type of flow and the roughness of the pipe walls. There are general relationships between f and Re which may be expressed as equations of varying complexity or as charts. For smooth pipes:

$$\frac{1}{\sqrt{f}} = 4\,\log_{10}\,(2R_e\,\sqrt{f}) - 1.6 \tag{3.37}$$

For rough pipes with a roughness size k this becomes:

$$\frac{1}{\sqrt{f}} = 4\,\log_{10}\left(\frac{D}{2k}\right) + 3.48 \tag{3.38}$$

The *Colebrook* and *White equation* is a general or universal friction equation:

$$\frac{1}{\sqrt{f}} = 3.47 - 4\,\log_{10}\left(\frac{2r}{D} + \frac{9.35}{Re\sqrt{f}}\right) \tag{3.39}$$

It is, however, usually more useful to obtain values of f from a chart such as Figure

3.38. (*Note:* the value of f used in American equations for head losses is *four times* that used in the United Kingdom, so if values of f are obtained from American texts they should be moderated accordingly or the corresponding American equation used.)

An empirical relationship widely used in water pipe work is the *Hazen–Williams equation,* usually written as:

$$v = 1.38\ C\ m^{0.63} \left(\frac{h_L}{L}\right)^{0.54} \tag{3.40}$$

where m is the ratio of the cross-sectional area of flow to the wetted perimeter known as the *hydraulic mean diameter* and C is a coefficient which depends on the condition of the pipe wall.

Pipe Losses (Changes in Section). When a fluid flows through a sharp (sudden) change in the cross-section of a pipe, energy is dissipated in the resulting turbulent eddies at the edge of the flow stream, producing a loss of head (or energy per unit weight). If the flow is from a smaller area to a larger one (sudden enlargement), the head loss is

$$h_L = \frac{(v_1 - v_2)^2}{2g} \tag{3.41}$$

When the flow is from a larger area to a smaller area (sudden contraction), the narrowed flow stream entering the smaller pipe is known as a *vena contracta*. The loss of head is assumed to be that due to a sudden enlargement from the vena contracta to the full area of the smaller pipe:

$$h_L = \frac{v_2^2}{2g}\left(\frac{1}{C_c} - 1\right)^2 \tag{3.42}$$

The *contraction coefficient*, C_c, is the ratio of the vena contracta area to that of the smaller pipe area. A typical value of C_c is 0.6, which gives

$$h_L \simeq = 0.5\ \frac{v_2^2}{2g} \tag{3.43}$$

which is also the head loss at the sharp entry to a pipe from a reservoir. Energy dissipation at changes in section, and pipe entry and exit, may be reduced by making the changes smooth and gradual, though this may be relatively costly.

Other pipe fittings, such as valves, orifice plates and bends, produce varying values of head loss, usually quoted as a fraction of the velocity head ($v^2/2g$).

Pipe Networks. A system of pipes may be joined together either in series (one after the other) or parallel (all between the same point). The friction head loss across a system of pipes in series is the sum of the losses along each pipe individually. The flow rate through each pipe will be the same. Using D'Arcy's head loss equation:

$$h_L = 4f_1 \frac{L_1 v_1^2}{D_1 2g} + 4f_2 \frac{L_2 v_2^2}{D_2 2g} + \ldots 4f_n \frac{L_n v_n^2}{D_n 2g} \tag{3.44}$$

and

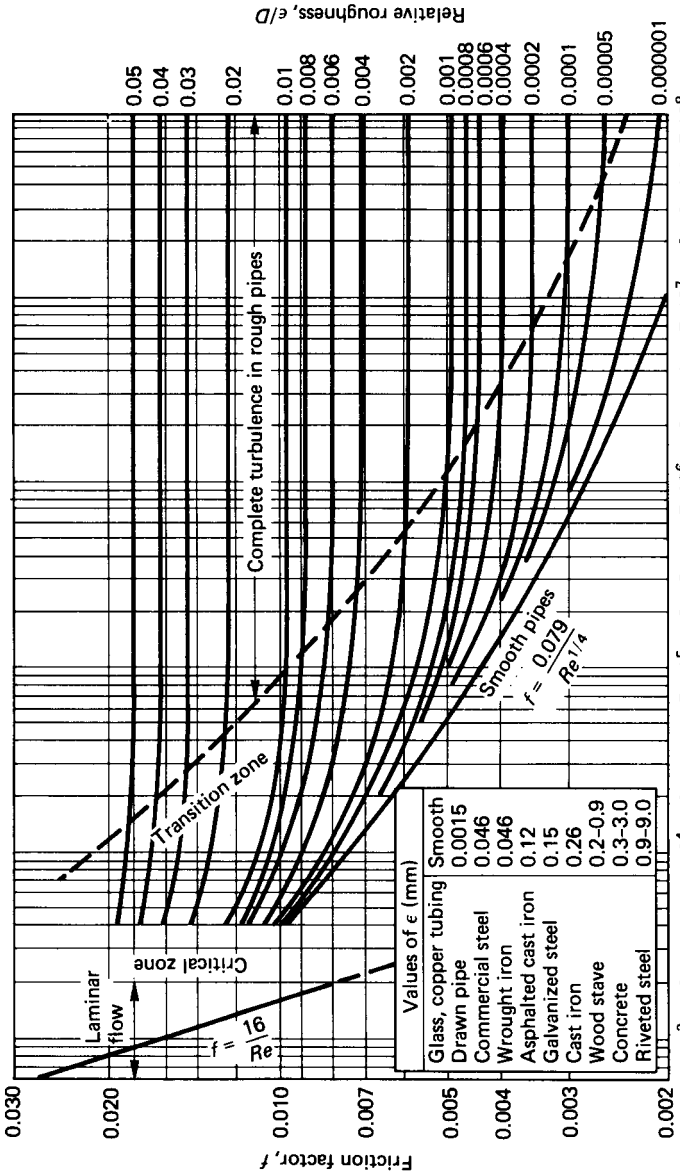


FIGURE 3.38

$$\dot{V} = v_1 A_1 = v_2 A_2 = \cdots = v_n A_n \tag{3.45}$$

If the system of pipes is connected in parallel the head loss across the system is equal to the head loss along any one of the pipes, when the flow has settled down to steady. The flow rate through the system is the sum of the flow rates along each pipe. Again using the D'Arcy equation:

$$h_L = 4f_1 \frac{L_1 v_1^2}{D_1 2g} = 4f_2 \frac{L_2 v_2^2}{D_2 2g} = \cdots = 4f_n \frac{L_n v_n^2}{D_n 2g} \tag{3.46}$$

$$\dot{V} = v_1 A_1 + v_2 A_2 + \ldots v_n A_n \tag{3.47}$$

In addition, the rate of flow into each junction of a network, either in series or parallel, is equal to the rate of flow out of it.

Pipe network problems are thus solved by setting up a number of such equations and solving them simultaneously. For a large number of pipes a computer program may be needed to handle the number of variables and equations. An example of a pipe network computer solution is given in Douglas et al. (1986).

3.13 FLOW MEASUREMENT

Pipe Flow

Orifices and Nozzles (see Figure 3.39(a)). Another basic flow measurement technique is to introduce some restriction into the flow passage and calibrate the resulting pressure changes against known flow rates.

Often the restriction in a pipe is in the form of an *orifice plate* (a plate with a hole) or a *nozzle*. A simple application of the Bernoulli equation may be used for the design calculations, but it is always advisable to calibrate any measurement device in conditions as close to the required operating conditions as possible.

Bernoulli and the continuity equations give the flow rate:

$$\dot{V} = C_d A_o \left\{ \frac{2(p_p - p_o)}{\rho[1 - (A_o/A_p)^2]} \right\}^{0.5} \tag{3.48}$$

where A_o = the orifice (or nozzle throat) area
A_p = the upstream pipe area
p_p = the upstream pressure
p_o = the pressure at the orifice or the nozzle throat
C_d = a discharge coefficient which takes account of losses and contraction of the flow stream through the device

Recommended orifice and nozzle dimensions, values of C_d and methods of operation are contained in BS 1042. It is most important to place the orifice or nozzle so that its operation is not affected by perturbations in the upstream flow caused by valves, bends or other pipe fittings.

Venturi Meters (see Figure 3.39(b)). The introduction of any restriction, particularly a sharp-edged orifice or nozzle, in a pipe will result in a loss of head. If it is required to keep this loss to a minimum, a *Venturi meter* may be used. The flow

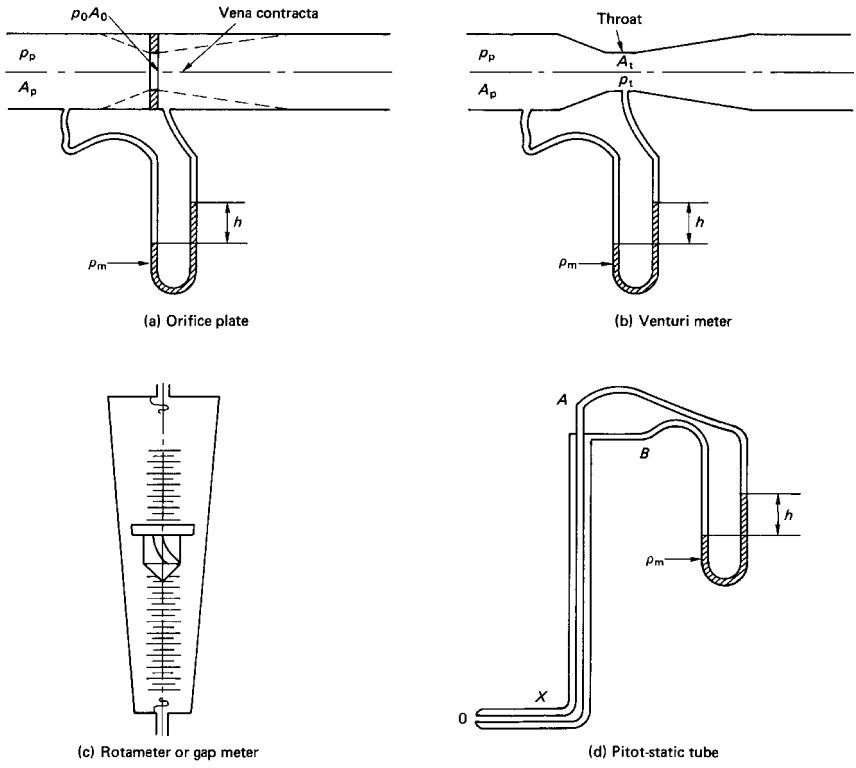


FIGURE 3.39 Flow meters.

passage in a Venturi is gradually and smoothly reduced to a throat followed by a controlled expansion to full pipe section. In this way the head loss across the meter is greatly reduced, but the cost of producing a Venturi meter is much higher than that of an orifice. Equation (3.48) may be used to calculate the flow rate $\dot{V}$, but the value of C_d will now be approximately 0.98 for a well-designed Venturi meter. Again, BS 1042 should be consulted for recommended dimensions, values of C_d and methods of operation.

Rotameter or Gap Meter (see Figure 3.39(c)). If, somewhere within the system, it is acceptable to tolerate flow up a vertical section of piping, then a *rotameter or gap meter* may be used. This instrument depends on the balancing of the weight of a rotating float in a tapered glass tube with the drag forces in the annular passage surrounding the float. The drag forces depend on the flow rate and the corresponding area of the annulus. As the flow rate increases, the annulus area which will produce a drag force equal to the weight of the float also increases. Therefore the float moves up the tapered tube until the annulus area is such that the forces again balance. As the flow rate decreases the float descends to a reduced annulus area to again achieve a balance of forces.

Velocity Meter. These are devices that measure velocity and not flow rate directly. *Pitot* and *Pitot-static* tubes are examples of such velocity-measuring instruments, making use of the pressure difference between the undisturbed flow stream and a point where the flow velocity is zero. They consist of two concentric tubes bent into an L shape as in Figure 3.39, with the outer tube joined to the inner at the toe of the *L*, at *O*. This end is usually spherical with a hole through to the inner tube. The undisturbed flow is assumed to be in the region of the holes round the periphery of the outer tube at *X*. The velocity is assumed zero at the spherical end presented to the flow, at *O*.

The flow velocity, v may be calculated by applying Bernoulli's equation between the two points *O* and *X* to give

$$v = C_v \left[2 \frac{(p_o - p_x)}{\rho} \right]^{0.5} \tag{3.49}$$

where p_o is connected to *O* via the inner tube to the tapping at *A*, p_x is connected to *X* via the outer tube to the tapping at *B* and C_v is a coefficient to cater for losses and disturbances not
accounted for in Bernoulli's equation. C_v is often taken to be unity. The pressure difference may be measured using a manometer and then written into equation (3.49) as a head, h, to give

$$v = \left[2gh \left(\frac{\rho_m}{\rho} - 1 \right) \right]^{0.5} \tag{3.50}$$

As usual, it is advisable to calibrate the tube and obtain a calibration curve or an accurate value for C_v.

Care should be taken when a pitot-static tube is used to measure pipe flow, since the velocity will vary across the pipe. As a rough guide to the flow rate, the maximum velocity, which is at the center of the pipe, may be taken to be twice the average velocity. Alternatively, the velocity at half the radius may be taken to be equal to the average velocity in the pipe. For an accurate evaluation the velocity distribution curve may be plotted and the flow rate through the pipe found by integration. This may be approximated to by dividing the cross-section into a series of concentric annuli of equal thickness, measuring the velocity at the middle of each annulus, multiplying by the corresponding annulus area, and adding to give the total flow rate.

Current meters, torpedo-shaped devices with a propeller at the rear, may be inserted into pipes. The number of rotations of the propeller are counted electrically. This number together with coefficients peculiar to the propeller are used in empirical equations to determine the velocity.

Velometers, vaned anemometers, and *hot wire anemometers* are not usually used to measure the velocities of incompressible fluids in pipes, and will be discussed below under Gas Flow.

3.14 BOUNDARY LAYER FLOW

When a fluid flows over a solid boundary there is a region close to the boundary in which the fluid viscosity may be assumed to have an effect. Outside this region the fluid may be assumed inviscid. The viscous effect within the region is evidenced

by a reduction in velocity as the boundary is approached. Outside the region the velocity is constant. The region is known as a *boundary layer*.

It is usual to assume that at the solid surface the fluid velocity is zero and at the boundary layer outer edge it is equal to the undisturbed flow velocity v_s. This defines the boundary layer thickness δ. (In practice, δ may be taken to be the distance from the boundary surface at which the velocity is 99% of the undisturbed velocity, or 0.99 v_s.)

When a flow stream at a velocity v_s passes over a flat plate, the boundary layer thickness δ is found to increase with the distance x along the plate from the leading edge. Near the leading edge the flow inside the boundary layer may be assumed to be laminar, but as x increases the flow becomes turbulent and the rate of increase of δ with x also increases, as shown in Figure 3.40.

Within even a turbulent boundary layer there is a narrow region close to the plate surface where the flow is laminar. This is known as the *viscous sublayer* and has thickness δ_b. The reduction in velocity across the boundary layer is associated with a shear force at the plate surface, usually known as the *drag force*.

Application of the momentum equation produces *Von Karman's momentum integral*, in which the drag force per unit width, F_D, becomes

$$F_D = \rho v_s^2 \int_0^\delta \frac{v}{v_s}\left(1 - \frac{v}{v_s}\right) dy \tag{3.51}$$

where v is the velocity within the boundary layer at a distance y above the plate surface. (The integral

$$\int_0^\delta \frac{v}{v_s}\left(1 - \frac{v}{v_s}\right) dy$$

may be defined as the *momentum thickness* (θ) and the integral

$$\int_0^\delta \left(1 - \frac{v}{v_s}\right) dy$$

as the *displacement thickness* (δ^*) so that

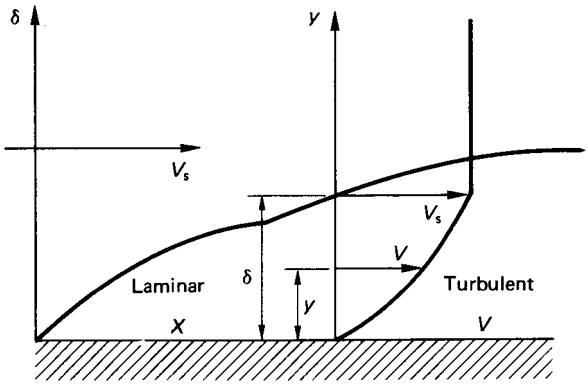


FIGURE 3.40 Boundary layer.

$$F_D = \rho v_s^2 \theta \tag{3.52}$$

In order to solve the Von Karman integral equation (3.51) or equation (3.52) it is necessary to know the value of δ and the relationship between v and y, the velocity distribution. Both of these are dependent on each other and the flow regime, laminar or turbulent, within the boundary layer.

Laminar Boundary Layers. A flat plate laminar boundary is normally assumed if $Re_x < 500\ 000$. (Re_x is Reynolds' number based on x or $\rho v_s x/\eta$.) For laminar boundary layers various simplified velocity distribution relationships may be used, such as linear, sinusoidal, or cosinsusoidal. The generally accepted most accurate relationship is, however, that obtained by the reduction of a four-term polynomial, which gives

$$\frac{v}{v_s} = \frac{3y}{2\delta} - \frac{1}{2}\left(\frac{y}{\delta}\right)^3 \tag{3.53}$$

From this the shear stress at the plate surface, τ_o, may be found for Newtonian fluids:

$$\tau_o = \eta \left(\frac{dv}{dy}\right)_{y=0} \tag{3.54}$$

The shear force

$$F_D = \int_0^x \tau_0 dx \tag{3.55}$$

Substitution from equation (3.53) in equations (3.51) and (3.54) and equating F_D from equations (3.51) and (3.55) leads via a separation of variables technique to

$$\frac{\delta}{x} = 4.64\ Re_x^{-0.5} \tag{3.56}$$

The drag force is usually quoted in terms of a *drag coefficient,* C_D:

$$C_D = \frac{2F_D}{\rho v_s^2 x} \tag{3.57}$$

or

$$F_D = \frac{1}{2} C_D \rho v_s^2 x \tag{3.58}$$

By manipulation of the above equations an equation for C_D for a laminar boundary over the whole length L of the plate:

$$C_D = 1.29 Re_L^{-0.5} \tag{3.59}$$

is obtained. The drag force on the whole plate surface of area A is found from:

$$F = \frac{1}{2} C_D v_s^2 A \tag{3.60}$$

Turbulent Boundary Layers

For Re_x 500 000, a flat plate boundary layer is assumed to be turbulent. In a turbulent boundary layer the velocity distribution is often written in a power form:

$$\frac{v}{v_s} = \left(\frac{y}{\delta}\right)^{1/n} \tag{3.61}$$

The index n varies between 6 and 9, depending on Re_x.

Because of the presence of the laminar sublayer, the turbulent regime is not continuous down to the plate surface, and $(dv/dy)_{y=0}$ does not give a useful result.

The equation used for τ_0 is

$$\tau_o = 0.0225\rho v_s^2 R_\delta^{-0.25} \tag{3.62}$$

based on work on smooth pipes by Blasius.

Taking $n = 7$ and using the same techniques as for laminar boundary layers gives:

$$\frac{\delta}{x} = 0.37\ Re_x^{-0.2} \tag{3.63}$$

and

$$C_D = 0.072\ Re_L^{-0.2} \tag{3.64}$$

This result assumes that the turbulent boundary layer obtains over the whole length of the plate to L.

Prandtl suggested a more realistic expression which takes into account the presence of a laminar boundary layer near the leading edge:

$$C_D = 0.074\ Re_L^{-0.2} - 1700\ Re_L^{-1} \tag{3.65}$$

This may be used for $5 \times 10^5 < Re_L < 10^7$. For $10^7 < Re_L < 10^9$ Schlichting (1960) suggests a logarithmic velocity distribution and

$$C_D = 0.44(\log_{10} Re_L)^{-2.58} = 3.91(\ln Re_L)^{-2.58} \tag{3.66}$$

Again, equation (3.60) may be applied to find the drag force on the whole plate.

Viscous Sublayers

The analyses above assume that the plate surface is smooth or at least *hydraulically smooth*. A surface is regarded as hydraulically smooth if the average roughness height k is less than the laminar sublayer thickness δ_b. For a turbulent layer with a velocity distribution power index of $n = 1/7$, the laminar sublayer thickness at a point at a distance x along the plate from the leading edge is given by

$$\frac{\delta_b}{\delta} = 199\ Re_x^{-0.7} \tag{3.67}$$

Thus δ_b may be compared with the roughness height, k, if the boundary layer thickness, δ, is known.

3.15 PRESSURE TRANSIENTS

Pressure transients may cause damage to piping systems such as aircraft fuel supply systems. It can be much more serious on a larger scale, where high-pressure rises over short periods may cause severe damage. Similar effects due to valve closure can be analysed on different levels of sophistication.

The simplest is the *rigid column* theory, which assumes that the fluid is incompressible, and that the valve is closed relatively slowly.

Slow Valve Closure

When a fluid flowing through a pipe with a velocity v_0 undergoes a change in velocity there is an associated change in pressure. Equating the force due to the pressure change to the rate of change of momentum during closure gives the resulting pressure rise Δp over a length of pipe L:

$$\Delta p = -\rho L \frac{dv}{dt} \tag{3.68}$$

The solution to this equation depends on a knowledge of the relationship between v and t (the valve closure rate in terms of the flow velocity).

Equation (3.68) is only applicable to relatively slow valve closure rates in which the closure time should not be less than $2L/C$ (where C is the speed of sound in the fluid).

Time to Establish Flow

The rigid-column theory is also often used to calculate the time required to establish flow in a pipe on opening a valve. The theory implies that the time required to fully establish the flow is infinite and so the time t to achieve 99% of the final velocity v_0 is usually accepted:

$$t = 2.646 \frac{Lv_0}{gH} \tag{3.69}$$

where H is the supply head to the pipe entrance. The time t_x required to reach $x\%$ of the final velocity is given by

$$t_x = \frac{Lv_0}{2gH} \ln\left(\frac{1 + 0.01x}{1 - 0.01x}\right) \tag{3.70}$$

Rapid Valve Closure

When a fluid is brought to rest *instantaneously* from a velocity of v_0 by the closure of a valve at the exit of a pipe of diameter D, there will be a relatively high pressure rise at the valve. If the valve closure time is *less than* $2(L/c)$ then the resultant pressure rise is as if it were instantaneous; c is the speed at which the pressure wave travels through the fluid, which is the sonic velocity.

On such a rapid valve closure the kinetic energy of the flow is converted into strain energy in both the pipe material and the fluid (even liquids are acknowledged as compressible in this context). The resulting pressure wave is transmitted through the fluid away from the valve as shown in Figure 3.41. The pressure rise produced is

$$\Delta p = \rho c v_0 \tag{3.71}$$

For a fluid of bulk modulus G, in a pipe of wall thickness x, of a material with a Young's modulus E and Poisson's ratio σ, the velocity of the pressure wave is

$$c = \left[\rho \left\{ \frac{1}{G} + \frac{D}{Ex} (1.25 - \sigma) \right\} \right]^{-0.5} \tag{3.72}$$

or

$$c = \left[\rho \left(\frac{1}{G} + \frac{D}{Ex} \right) \right]^{-0.5} \tag{3.73}$$

if longitudinal stress is small compared to hoop stress.

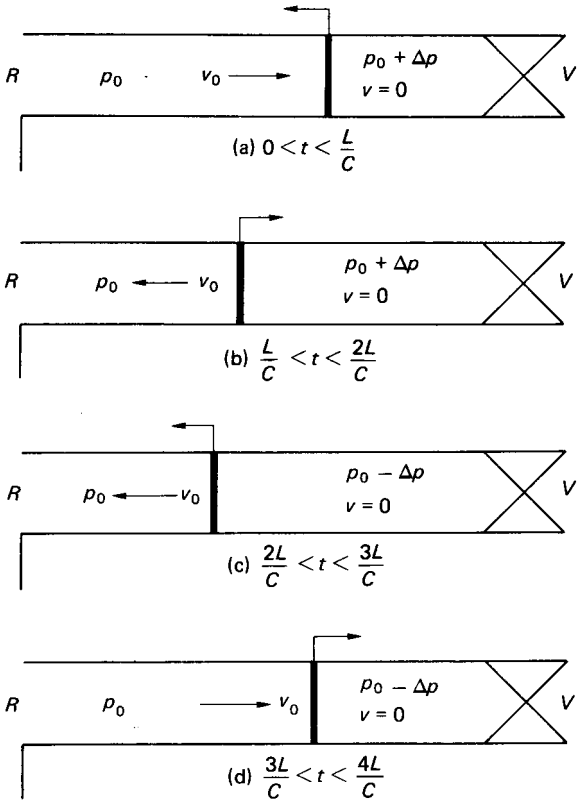


FIGURE 3.41 Progress of a pressure wave.

Equation (3.72) is often written as $c = [G_e/\rho]^{0.5}$, where G_e is the effective bulk modulus of the fluid and pipe combinations:

$$\frac{1}{G_e} = \frac{1}{G} + \frac{D}{Ex}(1.25 - \sigma) \tag{3.74}$$

If the valve is at the entrance to the pipe, then rapid valve closure results in a rarefaction (pressure drop) at the valve. In other words, the pressure change is

$$\Delta p = -\rho c v_0 \tag{3.75}$$

The Progress of a Pressure Wave

Assuming no friction and no cavitation in the fluid, the progress of a pressure wave along a pipe between a valve and a reservoir following valve closure is as shown in Figure 3.41. The fluid in the pipe is successively brought to rest by the passage of the pressure wave.

At a time $t = L/c$ after valve closure, the pressure wave reaches the reservoir. The whole of the fluid in the pipe is at rest at a pressure $p = p_0 + \Delta p$, which at the reservoir end instantaneously drops to reservoir pressure p_0. The resulting pressure wave travels along the pipe towards the valve and the fluid at the higher pressure in the pipe flows towards the reservoir at its initial velocity v_0.

At $t = 2L/c$, the situation is the same as for a rapid closure of a valve downstream of the flow, producing an instantaneous pressure drop to $p_0 - \Delta p$ and a rarefaction which travels towards the reservoir. The passage of the rarefaction successively brings the fluid to rest along the pipe.

At $t = 3L/c$, the rarefaction reaches the reservoir and the pressure instantaneously rises to reservoir pressure p_0. The resulting pressure wave travels towards the valve and fluid flows away from the reservoir at velocity v_0.

At $t = 4L/c$, the situation is the same as when the valve first closed at $t = 0$, and the cycle is repeated.

In practice, friction quickly dampens out the pressure waves and cavitation reduces the pressure decrease during the rarefactions.

A typical plot of pressure against time at a valve following rapid valve closure is superimposed on the theoretical plot in Figure 3.42.

3.16 GAS FLOW

General Relationships

The behavior of gases during processes involving thermal energy interactions and exchanges fits more properly into a study of thermodynamics. However, if only the flow mechanics are considered, the thermal and temperature effects may be restricted to those mainly relating to pressure and density. Applications of this approach are found in aircraft and rocket engines.

The most straightforward approach is to consider zero thermal energy transfer (heat transfer) to or from the fluid, or adiabatic flow. If, in addition, the changes in the fluid's properties are assumed to be reversible, then the flow becomes isentropic

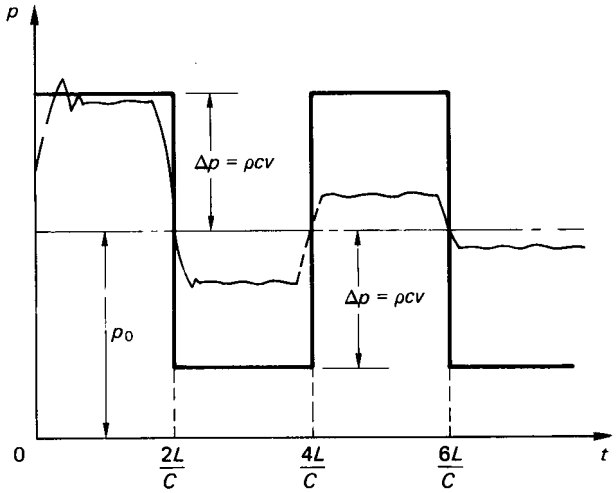


FIGURE 3.42 Pressure versus time at a valve.

and the relationship between pressure and density the simple and well-known one for an ideal gas:

$$\frac{p}{\rho^{\gamma}} = \text{constant} \tag{3.76}$$

Also, the *ideal gas law:*

$$\frac{p}{\rho} = RT \tag{3.77}$$

applies. Other useful relationships are:

1. The ratio of specific heats $\gamma = c_p/c_v$ (3.78)

2. The gas constant $R = c_p - c_v$ (3.79)

3. The *universal gas constant* $R_o = R \times$ molecular mass
$= 8.3143 \text{ kJ kg}_{\text{mol}}^{-1} \text{ K}^{-1}$

The terms *stagnation* or *total* temperature T_0 and pressure p_0 are often applied as the datum temperature and pressure of a fluid flow, even when stagnation conditions (zero velocity) do not exist in the particular situation under consideration. In gas flow the relationships between T_0 and the temperature T and p_0 and the pressure p at some point in the flow is often given in terms of the *Mach number* (M), the ratio of the flow velocity v to that of sound c, i.e.,

$$M = \frac{v}{c} \tag{3.80}$$

and

$$c = \left[\gamma \frac{p}{\rho}\right]^{0.5} = [\gamma RT]^{0.5} \tag{3.81}$$

In these terms T_0 and p_0 may be found from Euler's equation to be

$$T_0 = T\left[1 + \frac{(\gamma - 1)}{2} M^2\right] \tag{3.82}$$

$$p_0 = p\left[1 + \frac{(\gamma - 1)M^2}{2}\right]^{\gamma(\gamma-1)} \tag{3.83}$$

Flow in Ducts and Nozzles

Ducts. The analysis of gas flow in ducts is based on the Euler equation (3.26) and the one-dimensional continuity equation (3.18). Consideration of the differential forms of these equations will demonstrate that for subsonic flow ($M < 1$) the velocity will increase as the cross-sectional area of the duct decreases (in the converging entrance to a convergent/divergent nozzle, for example). For supersonic flow ($M > 1$) the velocity will increase as the cross-sectional area increases (in the diffuser of the convergent/divergent nozzle).

The properties of the fluid at a position in the flow stream where the local Mach number is unity are often denoted by a superscript $*$ (p^*, ρ^*, T^*) and used as a datum, so that

$$M^* = 1; \quad \text{and} \quad v^* = c^* = [\gamma RT^*]^{0.5} \tag{3.84}$$

The ratios of the properties at any position in the flow stream to those at the * position are:

$$1.\ \frac{p}{p^*} = \left[\frac{\gamma + 1}{2 + (\gamma - 1)M^2}\right]^{\gamma/(\gamma-1)} \tag{3.85}$$

$$2.\ \frac{\rho}{\rho^*} = \left[\frac{\gamma + 1}{2 + (\gamma - 1)M^2}\right]^{1/(\gamma-1)} \tag{3.86}$$

$$3.\ \frac{T}{T^*} = \frac{\gamma + 1}{2 + (\gamma - 1)M^2} \tag{3.87}$$

$$4.\ \frac{v}{v^*} = M\left[\frac{\gamma + 1}{2 + (\gamma - 1)M^2}\right]^{1/2} \tag{3.88}$$

$$5.\ \frac{A}{A^*} = \frac{1}{M}\left[\frac{2 + (\gamma - 1)M^2}{\gamma + 1}\right]^{(\gamma+1)/2(\gamma-1)} \tag{3.89}$$

For air the ratios may be calculated by substituting $\gamma = 1.4$, or obtained from published tables and charts (Houghton and Brock 1961).

Nozzles. A nozzle is an example of a duct with a smoothly decreasing cross-sectional area, followed in some cases by an increasing area (convergent/divergent nozzle) (see Figure 3.43). Since the velocity in the throat (minimum cross-section)

MECHANICAL ENGINEERING PRINCIPLES

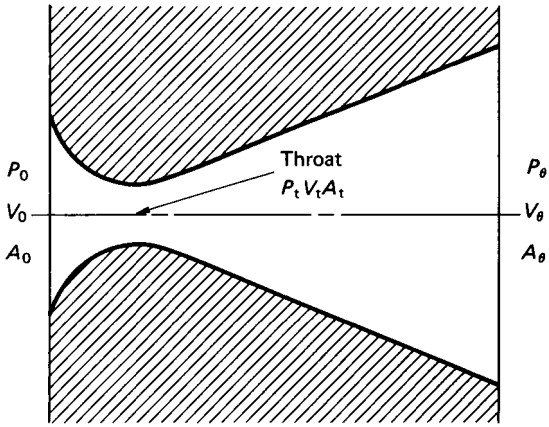


FIGURE 3.43 Convergent/divergent nozzle.

is often sonic, the approach velocity may be negligible. The throat velocity v_t and the exit velocity v_e are found by applying the Euler equation (3.26) between the upstream (entry conditions suffix$_0$) and throat and exit, respectively:

$$v_t = \left\{ \frac{2}{\gamma - 1} \frac{p_0}{\rho_0} \left[1 - \left(\frac{p_t}{p_0} \right)^{(\gamma-1)/\gamma} \right] \right\}^{1/2} \tag{3.90}$$

For v_e the pressure term p_e replaces p_t in equation (3.90). The mass flow rate through the nozzle is usually found at the throat by substituting v_t in the mass flow equation (3.19) to give

$$\dot{m} = C_d A_t \left\{ \frac{2}{\gamma - 1} \frac{p_0}{\rho_0} \left[\left(\frac{p_t}{p_0} \right)^{2/\gamma} - \left(\frac{p_t}{p_0} \right)^{(\gamma+1)/\gamma} \right] \right\}^{1/2} \tag{3.91}$$

where C_d is a discharge coefficient which depends on the nozzle design. For a well-designed nozzle C_d will be close to unity.

The mass flow rate will be the same at the exit as at the throat. It may be calculated from the exit conditions by substituting A_e and P_e for A_t and p_t, respectively, in equation (3.91).

Nozzles are usually designed for maximum mass flow rate. This will occur when the throat velocity is sonic ($v_t = c$). The pressure ratio which produces this situation is known as the *critical pressure ratio,* given by

$$\left(\frac{p_t}{p_0} \right)_{\text{crit}} = \left(\frac{2}{\gamma + 1} \right)^{\gamma/(\gamma-1)} \tag{3.92}$$

For many light diatomic gases such as air, where γ is approximately 1.4, $(pt_t)/(p_0)_{\text{crit}} = 0.528$.

The throat area will be that which gives the required mass flow rate through the throat at sonic velocity for critical pressure ratio. The exit area will be that which gives the calculated exit velocity for the given mass flow rate at the exit conditions. For convergent nozzles the throat also becomes the exit.

If nozzles, orifices, or Venturi meters are used to measure gas flow rates through a pipe then the approach velocity may be signficant and the mass flow rate given by

$$\dot{m} = C_d A_t \rho_0 \left\{ \frac{2(\rho_0/\rho_0)[\rho_t/\rho_0)^{2/r} - (\rho_t/\rho_0)^{(\gamma+1)/r}}{(\gamma - 1)[1 - (\rho_t/\rho_0)^{2/r}(\rho_t/\rho_0)^2} \right\}^{1/2} \tag{3.93}$$

Shock Waves

Under normal design conditions the flow in the nozzle downstream of the throat will be supersonic. The velocity of the gas at exit will depend on the external pressure p_b (back pressure). If the back pressure is greater than the theoretical exit pressure $p_e (p_b > p_e)$ then *shock waves* will be set up in the nozzle. These are discontinuities similar to standing waves in open-channel flow. The shock waves set up in such a way are *normal shock waves,* normal to the direction of flow.

If $p_b < p_e$ then the expansion will continue outside the nozzle (over-expansion).

If conditions upstream of a normal shock wave are denoted by suffix 1 and downstream by suffix 2, then it can be shown that the product of the up- and downstream velocities is equal to the square of the sonic velocity at $M = 1$:

$$v_1 v_2 = C^{*2} \tag{3.94}$$

and since $M_1 > 1$, then $M_2 < 1$. Also,

$$1.\ M_2 = \left[\frac{2 + (\gamma - 1)M_1^2}{2\gamma M_1^2 - (\gamma - 1)} \right]^{0.5} \tag{3.95}$$

$$2.\ \frac{p_2}{p_1} = \frac{2\gamma M_1^2}{(\gamma + 1)} - \frac{\gamma - 1}{(\gamma + 1)} \tag{3.96}$$

$$3.\ \frac{\rho_2}{\rho_1} = \frac{v_1}{v_2} = \frac{(\gamma + 1)M_1^2}{2 + (\gamma - 1)M_1^2} \tag{3.97}$$

These are known as the *Rankine–Hugoniot* relationships. Values for air may be obtained by putting $\gamma = 1.4$ or by the use of published tables (Houghton and Brock 1961).

The strength of a shock wave may be defined as the ratio of the pressure change across the wave to the upstream pressure, or in terms of the upstream Mach number:

$$\text{Shock wave strength} = \frac{p_2 - p_1}{p_1} = \left(\frac{2}{\gamma + 1} \right)(M_1^2 - 1) = 1.167(M_1^2 - 1), \text{ for air.} \tag{3.98}$$

Oblique shock waves, at an angle β to the upstream flow direction, are produced when a supersonic gas flow is turned through an angle θ by an obstruction such as an aircraft's nose, wing or tail, the inside walls of a duct, etc. The relationships between the up- and downstream Mach numbers and the angles β and θ are published in tables and charts (Houghton and Brock 1961).

In some cases both the up- and downstreams will be supersonic and subsequent shock waves produced, for example, at the leading and trailing edges of a wing. The effects of such shock waves produced by aircraft in flight, say, may be noted at ground level as *sonic booms*.

Gas Flow Measurement

Gas flow rates through ducts will normally be measured using devices and techniques similar to those used for incompressible fluids, namely orifice plates, venturi meters and nozzles. However, for gases the flow rate is usually quoted as a mass flow rate. Equations (3.90) and (3.92) may be used with orifices and Venturis as well as nozzles. Relevant values of C_d for each device will be found in BS 1042, in addition to operational advice.

When pitot-static tubes are used to measure gas flow velocities equation (3.49) may be acceptable for low flows with low pressure differences. At high velocities the compressibility must be taken into account and equation (3.90) used with stream conditions at X replacing those at the throat t.

For accurate velocity measurement with little disturbance to the flow *hot wire anemometers* may be used. The resistance of an electrically heated wire is related to the temperature of the wire, which in turn is related to the velocity of the fluid flow past the wire. The wire resistance measured on a bridge may be calibrated against a known velocity, to give either direct readout or (more usually) a calibration curve. The fine wire of the anemometer is suceptible to fluid contamination.

Other velometers and anemometers (depending on the relationship between the speed of rotation of a set of blades and the velocity (or speed) of the gas flow) may be used in very large cross-sectional ducts or to measure wind speed in the open air. They may depend on the rotation to generate a small electrical current, which can be calibrated as a speed, or the number of revolutions may be inserted into an empirical formula. A typical example is the three-vaned meteorological anemometer. Systems are now widely available for the measurement of velocity over a two dimensional field. Examples are laser doppler anemometry and particle image velocimetry, both of which require the seeding of the measured flow.

3.17 IDEAL FLUID FLOW

The concept of using idealized conditions to establish the shape of the mathematical models of real situations is common in engineering science studies. These models may then be modified to accommodate observed relationships, for application to real situations.

Ideal fluid (or *potential*) flow is such a concept. It may be used to set up flow patterns in the region of a flow stream outside the boundary layers. The combination of ideal flow and the boundary layer effects may be used to predict the performance of a real situation, so long as the limitations of both are recognized. The fluid is assumed to be inviscid and the flow steady, continuous, and irrotational. This means that there are no cavities or discontinuities in the flow stream, and that the fluid particles do not rotate about their own axes, even though the flow may be circular.

The continuity equation (3.16) applies, and may be modified to

$$\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} = 0 \tag{3.99}$$

if required for two-dimensional flow. For irrotation in two dimensions equation (3.21) becomes

$$\frac{\partial v_x}{\partial y} - \frac{\partial v_y}{\partial x} = 0 \tag{3.100}$$

The Stream Function. From the definitions of the stream line and stream function the equation to a stream line may be shown to be

$$v_x dy - v_y dx = 0 \tag{3.101}$$

and since the stream function ψ is an equation which describes a family of stream lines, then, for example,

$$\psi = 2x - y$$

represents a family of parallel straight lines with a variable intercept ψ.

For unit thickness in the z direction the volumetric flow rate $\dot{V}$ between two stream lines 1 and 2 is

$$\dot{V} = \psi_1 - \psi_2 \tag{3.102}$$

or

$$d\dot{V} = d\psi \tag{3.103}$$

Stream functions may be *superposed* so that if

$$\psi_1 = \text{fn}(x,y) \text{ describes flow pattern } A$$

and

$$\psi_2 = \text{fn}(x,y) \text{ describes flow pattern } B$$

then $\psi = \psi_1 + \psi_2$ describes the flow pattern produced by the combination of A and B.

The x and y components of the flow velocity v are given by

$$v_x = -\frac{\partial \psi}{\partial y}; \quad \text{and} \quad v_y = \frac{\partial \psi}{\partial x} \tag{3.104}$$

or

$$v_x = \frac{\partial \psi}{\partial y}; \quad \text{and} \quad v_y = -\frac{\partial \psi}{\partial x}$$

depending on sign convention. In *polar coordinates* the components are

$$\text{Radial velocity } v_r = \frac{-\partial \psi}{r \partial \theta} \tag{3.105}$$

$$\text{Tangential velocity } v_\theta = \frac{\partial \psi}{\partial r} \tag{3.106}$$

The Velocity Potential

In a gravitational field there is a property the change in which is independent of the path of the change: *potential energy*. In a continuous, irrotational flow field there is also a property the change in which is independent of the path of the change. This property is the *velocity potential* (ϕ). It can be shown that

$$v_x = -\frac{\partial \phi}{\partial x}; \quad \text{and} \quad v_y = -\frac{\partial \phi}{\partial y} \tag{3.107}$$

or

$$v_\theta = -\frac{\partial \phi}{r \partial \theta}; \quad \text{and} \quad v_t = -\frac{\partial \phi}{\partial r} \tag{3.108}$$

Lines of constant ϕ are known as *velocity potential* lines with an equation

$$v_x dx + v_y dy = 0 \tag{3.109}$$

which intercept the stream lines at right angles to form an orthogonal network of characteristic pattern for each flow field.

From equations (3.104) and (3.107) it can be seen that

$$\frac{\partial \psi}{\partial y} = \frac{\partial \phi}{\partial x}; \quad \text{and} \quad \frac{\partial \psi}{\partial x} = -\frac{\partial \phi}{\partial y} \tag{3.110}$$

which are the *Cauchy–Riemann* equations. In such flow fields the *Laplace* equations for ϕ and ψ must both be satisfied:

$$\nabla^2 \phi = 0; \quad \text{and} \quad \nabla^2 \psi = 0 \tag{3.111}$$

Because the fluid is ideal, the Bernoulli equation (3.27) may be readily applied between points in the field, both along and across the stream lines.

Flow Patterns

Examples of simple flow patterns are shown in Figure 3.44 with the equations to their stream functions and velocity potentials. A *source* is mathematically a point at which fluid appears and flows radially outwards. A *sink* is a negative source at which fluid flows radially inwards to disappear at a point (similar to the plug hole in a domestic sink, where, however, there is a vortex superposed to produce a spiral vortex or whirlpool). A *vortex* is flow in concentric circles with no radial flow. A *doublet* is the superposition of a source and a sink of equal strength m, initially a distance $2a$ apart brought infinitely close together so that the product of their strength and the distance between them remains a constant k. k is the strength of the doublet and is equal to $2am$.

This is a mathematical concept which is apparently impractical but yields a useful flow pattern. It is often used in combination with other simple patterns.

Modeling

Since there can be no flow across a stream line and the fluid is assumed inviscid, any stream line in a flow pattern may be replaced by a solid surface with no effect on the rest of the pattern. A stream line forming a closed contour may be replaced by a solid body to model the flow pattern around a body of the same shape. This provides a method of writing mathematical models to describe the flow streams around various shapes.

The limitations of the model must be understood, and the effects of rotation and viscosity particularly considered, when applying the analysis to *real* situations. For

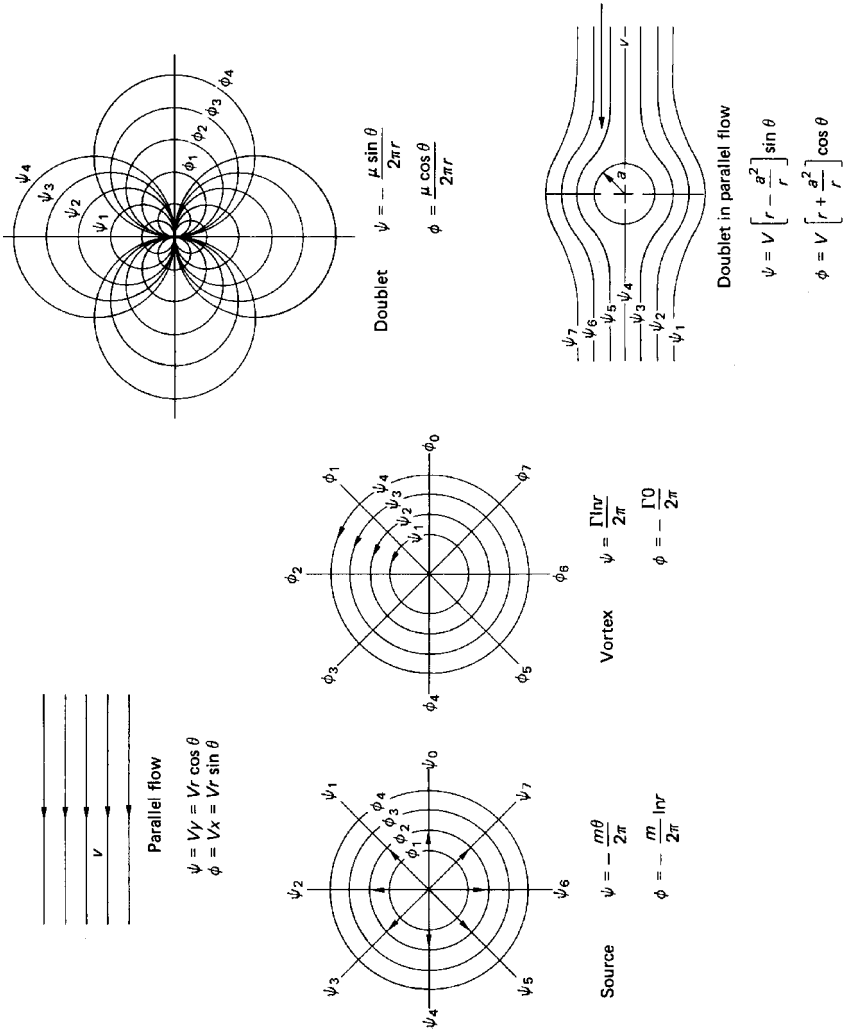


FIGURE 3.44 Simple flow patterns.

example, the flow round a cylinder may be modeled by a combination of doublet and parallel flow. The drag and lift forces calculated by integrating the resulting pressure forces at the cylinder surface appear to be zero. This is obviously not correct.

In the real situation the viscosity of the fluid produces a boundary layer at the cylinder surface, which, because of curvature, separates from the surface to form a wake. The presence of the wake disturbs the downstream flow pattern and the form drag force is a consequence. Viscosity also introduces a surface friction drag.

The shedding of vortices from the cylinder surface into the wake also produces alternate positive and negative lift forces, which are not predicted by the ideal flow analysis, although telegraph wires may often be observed vibrating in the wind.

3.18 CONCLUSION

Fluid mechanics is often regarded as an empirical subject which makes use of formulae based only on observed experimental results. This misconception is further compounded by the extensive use of coefficients (discharge especially) to account for effects which are difficult to model mathematically. However, almost all fluid mechanics equations in common use are based on the conservation of energy, the conservation of momentum or the fact that the rate of change of momentum may be equated to an applied force, usually a pressure force. The experimental checking and empirical amendment to derived formulae is just *good engineering practice.*

This section does not provide rigorous derivations of the various equations quoted. Some appreciation of such derivations may be required in order to establish the limitations and modifications necessary for the application of the equations, especially to nonstandard situations. For this, standard textbooks on fluid mechanics as listed below should be consulted.

PART 6

PRINCIPLES OF THERMODYNAMICS

Dennis H. Bacon

3.19 INTRODUCTION

Thermodynamics is concerned with energy transfers in processes. Two modes of transfer are recognized: *work* (transfer) and *heat* (transfer). Before proceeding, the terminology used in thermodynamics must be defined.

The particular part of the *working substance* under consideration is called the *system,* and this is separated from the *surroundings* by a *boundary.* In the *closed* or *nonflow* system the mass of working substance is constant, but in the *open* or *flow* system there is a mass flow rate across the boundary. Some processes in reciprocating plant may be considered by nonflow analysis but in steam plant, for example, most are considered by flow analysis. The *state* of a system is defined by the *properties* (pressure, temperature, etc.). Properties are normally expressed *specifically* (i.e., per unit mass) to enable charts or tables to be used. The state of simple substances can be described by two independent properties, but complex ones such as mixtures need more definition. A *change of state* is achieved by a *process* which is idealized as *reversible* with no losses. Reversible processes can be described by mathematical equations and enable analysis to be made to give answers for ideal situations. Real processes have losses and are described as *irreversible,* and the ideal results are multiplied by a coefficient or efficiency (based on measurement or experience) to predict real performance.

3.20 THE LAWS OF THERMODYNAMICS

The First Law of Thermodynamics

This is a law of energy conservation. When applied to a process we write

$$Q - W = \Delta E \quad \text{or} \quad q - w = \Delta e$$

where Q is the heat transfer (kJ) or q is the specific heat transfer (kJ/kg), W is the work transfer (kJ) or w is the specific work transfer (kJ/kg), and ΔE is the energy change (kJ) or Δe is the specific energy change (kJ/kg).

The change symbol Δ means final value minus initial value. ΔE embraces all forms of energy, but in the nonflow process it is usual to find that the only significant change is in the internal energy (U,u) and we write the nonflow energy equation

$$Q - W = \Delta U \quad \text{or} \quad q - w = \Delta u$$

For the *steady flow* system we write

MECHANICAL ENGINEERING PRINCIPLES

$$\dot{Q} - \dot{W}_x = \dot{m}\Delta\left(h + \frac{V^2}{2} + gz\right) \quad \text{or} \quad q - w_x = \Delta\left(h + \frac{V^2}{2} + gz\right)$$

where $\dot{Q}$ and $\dot{W}_x$ are the energy transfer rates and $\dot{m}$ is the steady mass flow rate across the boundary (in and out), Δh is the change in specific enthalpy ($h = u + pv$), $\Delta V^2/2$ is the change in specific kinetic energy, and Δgz is the change in specific potential energy. The suffix x is used on the work transfer to denote that this is the useful work from the system as the flow work is included in the enthalpy term. In flow problems it will also be necessary to use the continuity equation

$$\dot{m} = \rho A V$$

where ρ is the density and A is the area normal to the velocity V. Analysis of nonsteady flow may also be made, in which case energy terms to allow for the storage of energy in the system will be added.

Warning: A sign convention for work and heat is built into the equations above. Positive work means work obtained from the system and positive heat means heat put into the system. Care should be taken to be clear about the symbol V, which may appear as velocity or volume in many equations.

In order to allow continuous energy transfers a *cycle* is defined in which a series of processes brings the working substance back to the initial state so that the cycle can be repeated continuously. If we apply the first law to a cycle it follows that ΔE is zero and

$$\sum_{\text{cycle}} Q = \sum_{\text{cycle}} W$$

The Second Law of Thermodynamics

It might be thought that the first law of thermodynamics permits all the heat transfer to a cycle to be returned as work transfer, but unfortunately the second law places restraints on the achievement of this desirable situation. The restraint takes the practical form of demanding that some of the heat transfer to the cycle *must* be rejected as a heat transfer to a lower temperature. Thus when we build a *heat engine* it has to exchange heat with (at least) two reservoirs in order to produce work (Figure 3.45). Since work is the objective, the amount produced per unit heat input is vital information and we define the *thermal efficiency* of a heat engine as

$$\eta_{\text{thermal}} = \frac{\text{Net work transfer from the cycle}}{\text{Heat transfer to the cycle}} = \frac{W}{Q_1}$$

Since the first law states $Q_1 - Q_2 = W$ we see that efficiency is less than unity.

The second law makes further investigations and determines the maximum possible efficiency of a heat engine using reversible isothermal processes to transfer heat from two reservoirs as

$$\eta_{\text{thermal maximum}} = 1 - (T_{\min}/T_{\max})$$

This efficiency is known as the Carnot efficiency and is not attainable due to losses. It is also found that constant temperature processes, except during phase change, are not practical and real processes of heat transfer take place at approximately

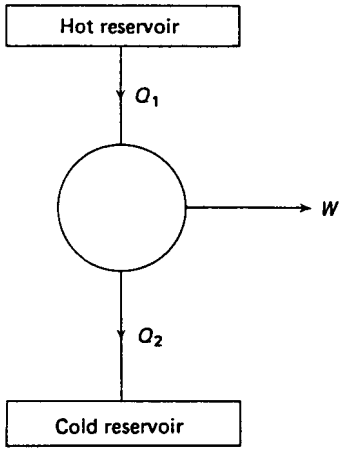


FIGURE 3.45 A heat engine.

constant volume or constant pressure. One positive product of the second law is that it tells the engineer that thermal efficiency will be increased by making the maximum cycle temperature as high as possible (a materials constraint) and by making the minimum cycle temperature as low as possible (ambient conditions).

The losses in a real cycle are due to internal fluid friction and the necessity of having a temperature difference to cause a heat transfer. The larger the temperature difference, the greater the losses. The fluid friction losses in a work-producing process are defined by the process efficiency

$$\eta_{\text{process}} = \frac{\text{Actual work produced}}{\text{Ideal work produced}}$$

which is inverted for work-absorbing processes (compression). The most common application of this efficiency is in steady flow *adiabatic* processes. Ideally, these are processes with no heat transfer which are often used as models for real processes in which the heat transfers are negligible compared with the work transfers (turbine expansion). In such processes the efficiency is known as the isentropic efficiency because the ideal adiabatic process has constant entropy.

Losses due to fluid friction and losses due to heat transfer across finite temperature differences are found to result in an increase in the value of the entropy that would be expected in a reversible process. Thus, an expected increase would be larger and an expected decrease would be smaller. It is not easy to define entropy except mathematically. In practical use as the abscissa of charts it enables work transfers in ideal adiabatic processes to be represented as vertical lines if enthalpy is used as an ordinate, and in this guise is a valuable visual method of presentation.

3.21 THERMOECONOMICS

When a more detailed study, see Kotas (1985) and Bejan (1982), of a flow process is made by the second law of thermodynamics it is found that specific entropy (J/

MECHANICAL ENGINEERING PRINCIPLES

kgK) appears as part of a property known as availability. In a flow process we write $b = h - T_0 s$, in which b is the specific *availability* function, h is the specific enthalpy, T_0 is the temperature (absolute) of the surroundings, and s is the specific entropy. The second law shows that the maximum work potential or *exergy* of any state in surroundings at state 0 is given by $b - b_0$. Thus for a change of state in a flow process from 1 to 2 the maximum specific work obtainable is given by the exergy change, $w_{x_{max}} = (b_1 - b_0) - (b_2 - b_0) = (b_1 - b_2) = -\Delta b$. If we measure or predict by analysis the actual work achieved it is possible to determine numerically the lost work or irreversibility in the process. If engineering plant is to be designed to the best advantage it is clear that processes should be chosen to minimize this loss. The lost work may be associated with costs and we move into the developing field of thermoeconomics. Clearly, this is a complex subject but it is important in that it unites thermodynamics with costs and can help in the design of long-life expensive plant, such as aircraft engines.

3.22 WORK, HEAT, PROPERTY VALUES, PROCESS LAWS AND COMBUSTION

To deploy the laws of thermodynamics outlined above we need more information. To perform simple cycle analysis the data below are vital.

Work

In a nonflow process work transfer can be determined from $w = \int p dv$. The mathematical relation for the process is known as the process law (qv). In most flow processes used in engineering cycles the adiabatic approximation is used so that the steady flow energy equation, neglecting changes in kinetic and potential energy, gives

$$w_x = \Delta h$$

Heat

This is usually an unknown quantity and is found by the application of the energy equation. As stated earlier, many processes are approximately adiabatic so that heat transfer is zero and in others heat transfer is obtained from combustion data or, if a heat exchange process, by heat exchanger efficiency.

Property Values

These are found in tables or from charts for common substances. Computer formulations are now widely available.

Process Laws

This is a particularly important step in thermodynamic analysis because an idealized reversible process has to be chosen to represent as closely as possible the real

process in order to calculate energy changes. When the working substance is a gas, it is convenient in an elementary analysis to use perfect gas laws with the process calculation. These are

$$pv = RT \quad \text{or} \quad pV = mRT$$

$$u = c_v \Delta T \quad \text{and} \quad h = c_p \Delta T$$

where c_v and c_p are the specific heat capacities at constant volume and constant pressure, respectively, which are related as follows:

$$c_p - c_v = R \quad \text{and} \quad c_p / c_v = \gamma$$

where R is the specific gas constant and γ is the isentropic index.

Ideal processes commonly used are *constant pressure, constant volume, constant temperature* (which for a perfect gas becomes pv = constant) together with two other more general relations: the *adiabatic process,* pv^k = constant (which for a perfect gas becomes pv^γ = constant), and the *polytropic process,* pv^n = constant. The last process is a general relation between pressure and volume which is used if none of the other clearly special cases are considered valid. Usually $1 < n < 1.4$.

It is possible (by using the gas laws) in adiabatic and polytropic gas processes to rearrange the relations to involve pressure and temperature or temperature and volume to yield very useful relations:

$$\frac{T_1}{T_2} = \left(\frac{p_1}{p_2}\right)^{(\gamma-1)/\gamma}; \frac{T_1}{T_2} = \left(\frac{V_2}{V_1}\right)^{\gamma-1}: \frac{T_1}{T_2} = \left(\frac{p_1}{p_2}\right)^{(n-1)/n}; \frac{T_1}{T_2} = \left(\frac{V_2}{V_1}\right)^{n-1}$$

Processes may be represented on property diagrams to enable cycle visualization (Figures 3.46–3.51).

Combustion

To avoid involving complex chemical equations, engineers often use the calorific value of a fuel coupled with a combustion efficiency to estimate the energy transfers in combustion processes. Thus, the rate of energy input by combustion is

$$\dot{E} = \dot{m}_f \cdot CV \cdot \eta_{\text{comb}}$$

where $\dot{m}_f$ is the fuel mass flow rate, CV the calorific value of the fuel, and η_{comb} the combustion efficiency.

3.23 CYCLE ANALYSIS

One example will be given of the simple analysis of the ideal Joule cycle for a gas turbine plant (Figure 3.51). The cycle consists of four flow processes described in Table 3.6 and analyzed by the steady flow energy equation.

From the data in the table it can be seen that the specific work $w = c_p(T_3 - T_4) - c_p(T_2 - T_1)$ and the thermal efficiency

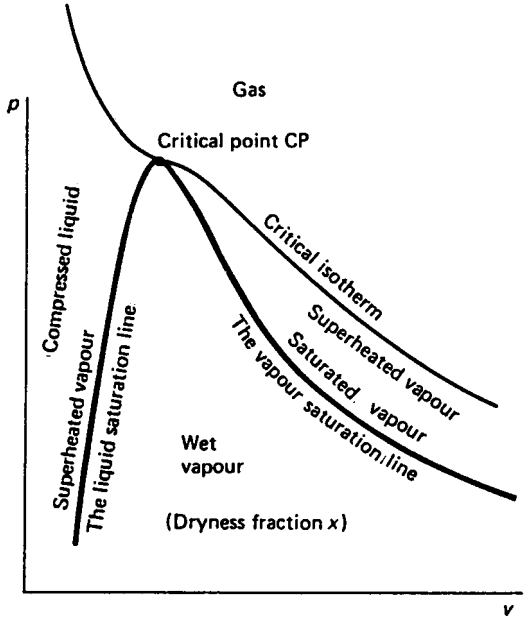


FIGURE 3.46 Substance phases and definitions.

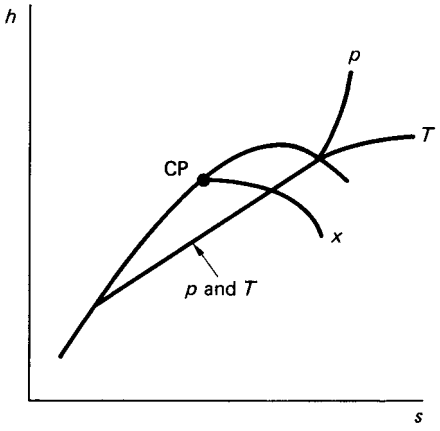


FIGURE 3.47 Steam processes on an *h–s* diagram.

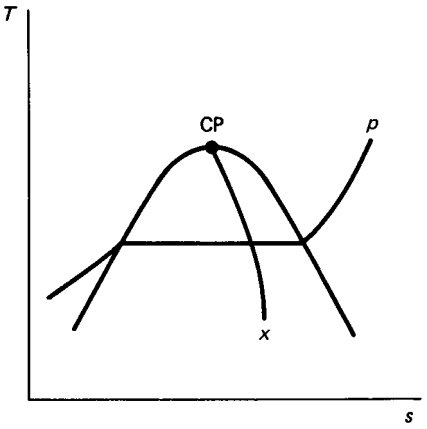


FIGURE 3.48 Steam processes on a T–s diagram.

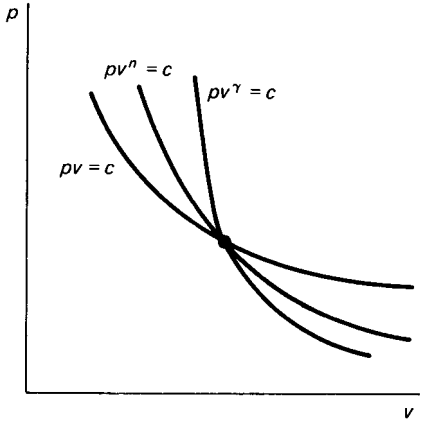


FIGURE 3.49 Gas processes on a p–v diagram.

$$\eta_{\text{thermal}} = \frac{c_p(T_3 - T_4) - c_p(T_2 - T_1)}{c_p(T_3 - T_2)}$$

If allowance is made for the isentropic efficiency of the compression and expansion processes the cycle diagram is changed to show the associated entropy increases but the expressions for work and efficiency above are still valid with the changed values of T_2 and T_4 (Figure 3.52). These values are determined from the use of the reversible adiabatic process relation and the isentropic efficiency as

$$T_2 - T_1 = T_1(r_p^{(\gamma-1)/\gamma} - 1)/\eta_c$$

and

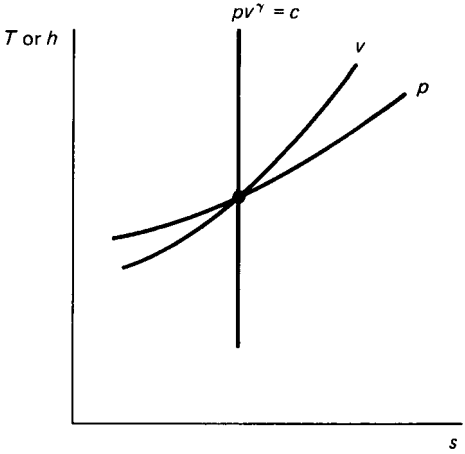


FIGURE 3.50 Gas processes on a T–s diagram.

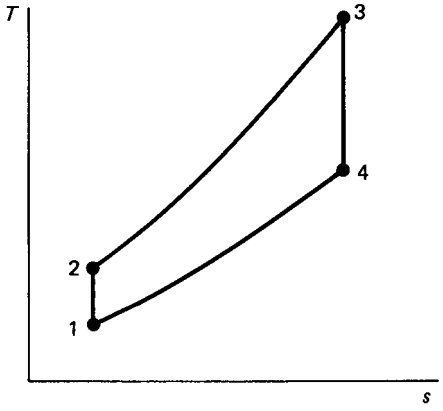


FIGURE 3.51 The Joule cycle.

$$T_3 - T_4 = \eta_t T_3 (1 - 1/r_p^{(\gamma-1/\gamma)})$$

where r_p is the cycle pressure ratio,

$$\eta_c = \frac{T_2' - T_1}{T_2 - T_1}$$

and

$$\eta_t = \frac{T_3 - T_4}{T_3 - T_4'}$$

η_c and η_t being the isentropic efficiencies of compression and expansion. If these values are substituted into the work and thermal efficiency expressions they become

TABLE 3.6

Process	Description	$q - w_x = h$
1 to 2	Reversible adiabatic compression $pv^\gamma = \text{constant}$	$-w_x = h_z - h_1 = c_p(T_2 - T_1)$
2 to 3	Reversible constant pressure heat transfer *to* the cycle	$q = h_3 - h_2 = c[p(T_3 - T_2)$
3 to 4	Reversible adiabatic expansion $pv^\gamma = \text{constant}$	$-w_x = h_4 - h_3\ w_x = c_p(T_3 - T_4)$
4 to 1	Reversible constant pressure heat transfer *from* the cycle	$q = h_1 - h_4 = c_p(T_1 - T_4)$

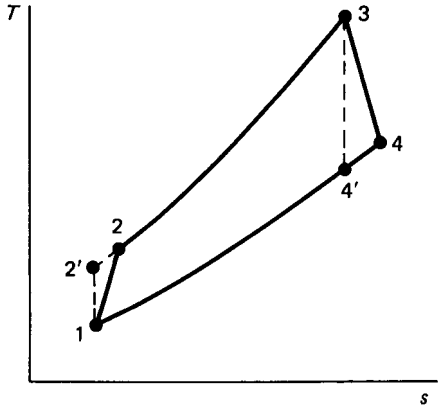


FIGURE 3.52 The effect of isentropic process efficiency on the Joule cycle.

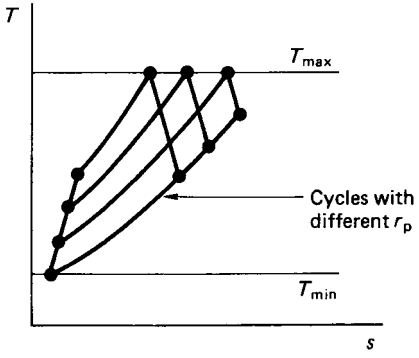


FIGURE 3.53 The effect of pressure ratio in a cycle with fixed T_{max} and T_{min}.

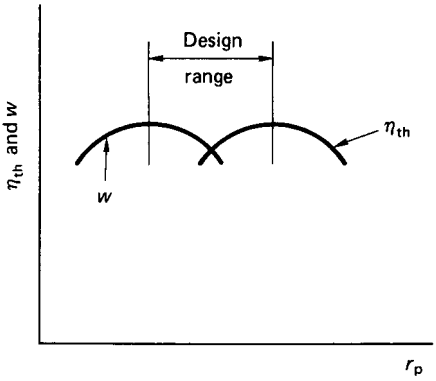


FIGURE 3.54 Thermal efficiency and specific work transfer variation in a Joule cycle with allowance for isentropic process efficiency.

more useful in that they involve the thermodynamically significant maximum and minimum cycle temperatures which are fixed by material and ambient conditions respectively, so that the only variable is the cycle pressure ratio (Figure 3.53). If the expressions are differentiated with respect to this pressure ratio it is possible to find the pressure ratio for maximum work and that for maximum efficiency. The cycle designer then has a choice, depending on the proposed application and Figure 3.54 shows that it would be expected that the chosen ratio would fall between these two maxima. Obviously, this simple approach is not the complete answer to gas turbine cycle analysis, but it illustrates the use of the laws of thermodynamics, and similar work may be done for other plant cycles.

PART 7
HEAT TRANSFER

Dennis H. Bacon

3.24 INTRODUCTION

Whenever a temperature difference occurs there is an energy flow from the higher temperature to the lower. A study of heat transfer is concerned with the determination of the instananeous rates of energy flow in all situations. We determine heat transfer rates in watts. These rates will be constant in situations where the temperature difference remains constant but variable (transient) when the temperature difference varies either due to the heat transfers or to other energy changes such as internal chemical reaction.

There are three modes of heat transfer:

1. *Conduction,* which is of greatest interest in solid bodies but also occurs in fluids, where it is often overshadowed by convection
2. *Convection,* which occurs in fluids when energy is transferred due to the motion of the fluid
3. *Radiation,* which occurs between two systems at different temperatures which need not be in contact provided any intervening medium is transparent to the radiation.

In practice, all three modes may occur simultaneously and it is necessary to draw up a balance at a boundary. For example, energy may be conducted to the surface of an electric storage heater and is then convected and radiated to the surroundings. Thus calculations can become complex, and in this particular case where energy is added at certain times this is a continuously varying situation.

Three approaches to heat transfer will be discussed below:

1. A simple method suitable for many estimations
2. A more detailed appraisal of the field
3. Comments on the use of computers

3.25 BASIC PRINCIPLES OF HEAT TRANSFER (*White 1984; Özişik 1985; Kreith and Bohn 1986*)

Conduction

Fourier's law for conduction states:

$$q = -k \frac{dT}{dx}$$

The thermal conductivity $k(\mathrm{Wm^{-1}K^{-1}})$ is a property of the material which varies

with temperature but for small temperature ranges is usually considered constant. Typical values are shown in Table 3.7. With constant k Fourier's equation can be integrated for four common situations.

Plane Surfaces. Integration gives

$$\dot{Q} = k_{12}A\frac{(T_1 - T_2)}{x_2 - x_1} = \Delta T/\left(\frac{\Delta x}{Ak}\right)_{12}$$

(see Figure 3.55). The quantity $(\Delta x/kA)$ is known as the thermal resistance in KW^{-1}. Thermal resistances can be added in a similar way to electrical resistances so that for a multilayer plane surface there are a number of resistances in series (Figure 3.56). Thus we can write

$$Q = (T_1 - T_4)/\sum\frac{\Delta x}{kA}$$

Cylindrical Surfaces. For tubes it is more convenient to evaluate heat transfer rates per unit length, and integration gives

$$\dot{Q} = (T_1 - T_2)/\frac{\ln r_2/r_1}{2\pi k_{12}L} = \Delta T/\frac{\ln r_2/r_1}{2\pi k_{12}L}$$

(see Figure 3.57), and in this case the thermal resistance is

TABLE 3.7

Substance	Thermal conductivity, $Wm^{-1}\ K^{-1}$, at 20°C
Aluminum	204
Iron	52
Water	0.597
Air	0.026 (100 kPa)
Glass wool	0.04

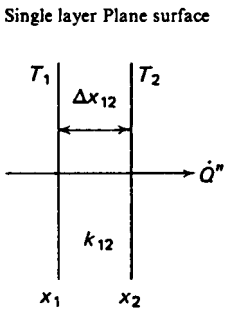


FIGURE 3.55 One-dimensional conduction through a single-layer plane wall.

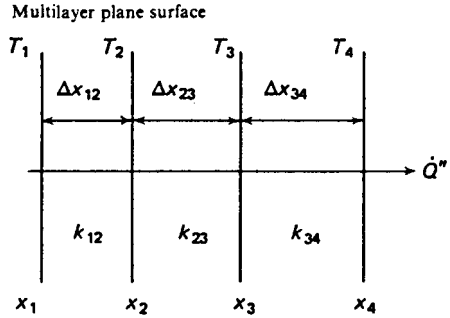


FIGURE 3.56 One-dimensional conduction through a multilayer plane.

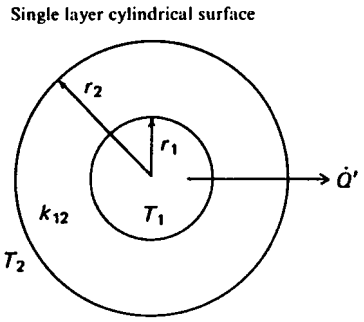
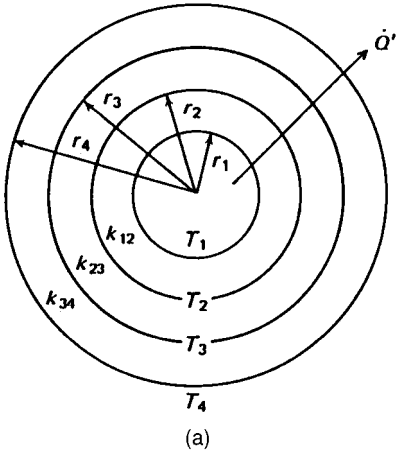


FIGURE 3.57 Conduction through a single-layer cylindrical surface.



(a)

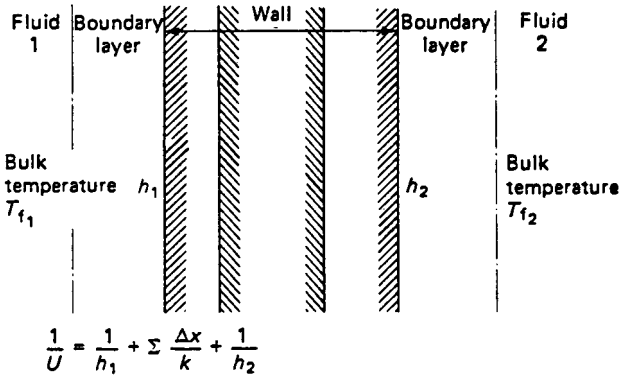
FIGURE 3.58a Conduction through a multilayer cylindrical surface; *U*-value for a plane surface.

$$\left(\frac{\ln r_2/r_1}{2\pi k_{12}L}\right) \text{ in KW}^{-1}$$

For a multilayer tube, thermal resistances are added to give

$$\dot{Q}' = (T_1 - T_4)/\sum\left(\frac{\ln (r_{\text{outer}}/r_{\text{inner}})}{2\pi kL}\right)$$

(see Figure 3.58).



(b)

FIGURE 3.58b *U*-value for a cylindrical surface.

Convection

The fundamental equation for convective heat transfer at a solid–fluid interface is

$$\dot{Q}'' = h\theta$$

where θ is the temperature difference between surface and fluid. The surface heat transfer coefficient $h(\mathrm{Wm^{-2}K^{-1}})$ is not a property of the fluid or the surface but depends on the flow rate, the fluid properties, and the surface shape. The coefficient has to be determined for each situation and can vary considerably (Table 3.8). Although the determination of h is crucial to convection calculations, it is an extremely difficult process, and accurate prediction of convective heat transfer is not always possible.

If we express the convection equation in a thermal resistance form suitable for plane surfaces,

$$\dot{Q} = \frac{\theta}{nA}$$

it can be seen that the thermal resistance is $(2/nA)\mathrm{KW^{-1}}$. For tubular surfaces it is again more convenient to work per unit length, so that

$$\dot{Q}' = \theta \Big/ \left(\frac{1}{2\pi rh}\right)$$

and the thermal resistance is $(1/2\pi rh)\mathrm{mKW^{-1}}$.

Overall Heat Transfer Coefficients

A common heat transfer situation is a solid wall separating two fluids, and for this problem the thermal resistances for conduction and convection can be added to enable the heat transfer rate to be determined in terms of the two fluid temperatures.

For a plane surface (Figure 3.58(b),

$$\dot{Q}'' = (T_{f_1} - T_{f_2}) \Big/ \left(\frac{1}{h_1} + \sum \frac{\Delta x}{k} + \frac{1}{h_2}\right)$$

For a tubular surface (Figure 3.59),

TABLE 3.8 Range of values of surface heat transfer coefficient ($\mathrm{W\ m^{-2}K^{-1}}$)

Free convection	Gases	0.5 to 500
	Liquids	50 to 2000
Forced convection	Gases	10 to 700
	Liquids	100 to 10 100

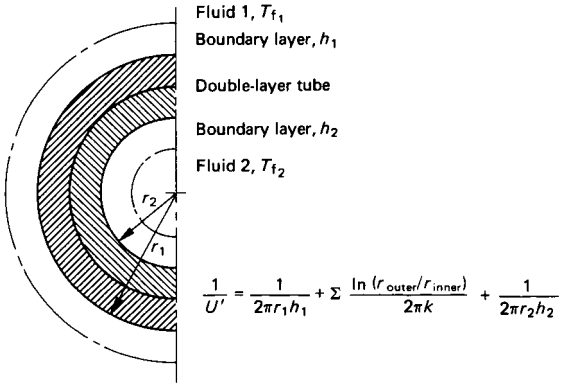


FIGURE 3.59 U'-value for a cylindrical surface.

$$\dot{Q}' = (T_{f_1} - T_{f_2}) \Big/ \left(\frac{1}{2\pi r_1 h_1} \right) + \sum \left(\frac{\ln(r_{\text{outer}}/r_{\text{inner}})}{2\pi k} \right) + \frac{1}{2\pi r_2 h_2}$$

It can be seen that the added resistances may be inverted to give an overall *conductance* which is known as a *U*-value or overall heat transfer coefficient.

For a plane surface,

$$\frac{1}{U} = \frac{1}{h_1} + \sum \frac{\Delta x}{k} + \frac{1}{h_2}$$

For a tubular surface,

$$\frac{1}{U'} = \frac{1}{2\pi r_1 h_1} + \sum \left(\frac{\ln(r_{\text{outer}}/r_{\text{inner}})}{2\pi k} \right) + \frac{1}{2\pi r_2 h_2}$$

The heat transfer rate is then simply written

For a plane surface,

$$\dot{Q} = UA(T_{f_1} - T_{f_2}) \text{ where } A \text{ is the area}$$

For a tubular surface

$$\dot{Q} = U'l(T_{f_1} - T_{f_2})$$

where l is the length.

The situation in which this technique is commonly used is in heat exchanger design. It should suffice for simple calculations provided suitable values of the surface heat transfer coefficients for convection can be obtained. (*CIBS Guide* 1980). If the temperature difference is not constant then a mean value should be used. A suitable equation for a mean can be found in the heat exchanger section which follows.

MECHANICAL ENGINEERING PRINCIPLES

Radiation

Radiation is of central importance in space application where all rejected heat is radiated to space. The rate of energy emitted by an ideal black body is given by the Stefan–Boltzmann law, in which the absolute temperature (Celsius + 273) is raised to the fourth power:

$$\dot{E}''_b = \sigma T^4$$

where σ is the Stefan–Boltzmann constant $5.67 \times 10^{-8}\text{Wm}^{-2}\text{K}^{-4}$ and the subscript b refers to the ideal black body. Real bodies emit less radiation, and the monochromatic emissivity is defined by

$$\varepsilon_\lambda = \left[\frac{\dot{E}''_\lambda}{\dot{E}''_{b\lambda}}\right]_T$$

The value of ε varies with λ and T because real bodies are selective emitters, but for simple calculations it is often assumed that emissivity is constant. The calculations associated with this assumption are based on grey body theory, for which the rate of energy emission is given by

$$\dot{E}''_g = \varepsilon\sigma T^4$$

It would be unwise to estimate unknown emissivities, and measurements would need to be made unless suitable data could be found.

Radiation incident on a body may be absorbed, reflected or transmitted. Thus we write $\alpha + \rho + \tau = 1$ where α, ρ and τ are the absorptivity, reflectivity, and transmissivity, respectively. Ideal black bodies absorb all incident radiation but real bodies do not. Gases are often assumed to transmit all radiation, but this is not always true, particularly with hydrocarbon combustion products and atmospheric transmission. Solids have a transmissivity of zero. With these simple ideas it is necessary to know the values of only α and ρ. It can be shown that a gray body has absorbtivity equal to emissivity, $\alpha = \varepsilon$. Thus, provided the transmissivity is zero, a knowledge of the grey body emissivity enables reflectivity to be determined, since

$$\rho = 1 = \varepsilon$$

The only simple radiation problem that can be solved with the simple approach above is that of a gray body in large surroundings (see Figure 3.60). The word 'large' implies that radiation not incident on the grey body which will be incident on the surroundings and will therefore be reflected will not be re-incident on the gray body. Thus, the surroundings are effectively black. (This might be true with a linear size factor greater than 10.) In this simple case it can be shown that the heat transfer rate is

$$\dot{Q} = \varepsilon\sigma A(T^4 - T_s^4)$$

where ε is the emissivity of the body, A is the area of the body, T is the temperature of the body in K and T_s is the temperature of the large (black) surroundings in K.

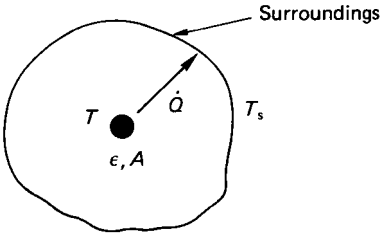


FIGURE 3.60 A small gray body in large (black) surroundings.

Simple Transient Problems

If a body is being cooled or heated by convection or radiation and the thermal conductivity is large so that the rapid heat transfer rates within the body enable it to be assumed that the body temperature distribution is uniform, then the situation is known as a lumped capacity system. For such a system the complex methods of transient heat transfer are not required and a simple energy balance equation may be drawn up and integrated. The most common case is quenching, a convective boundary problem for which in time dt a small heat transfer δQ occurs when the body temperature changes from T by an amount dT (Figure 1.61). Thus,

$$\delta Q = \rho c_p V dT = -hA(T - T_f)$$

where T_f is the fluid temperature, A is the body surface area, V is the body volume, ρ is the body density, c_p is the body specific heat, and h is the surface heat transfer coefficient. Integration gives

$$\frac{\theta}{\theta_0} = e^{-(hAt/\rho c_p V)}$$

where θ_0 is the initial temperature difference between fluid and body and θ is the temperature difference at any future time t. The quantity $(\rho c_p V/hA)$ may be regarded as the time constant of the system.

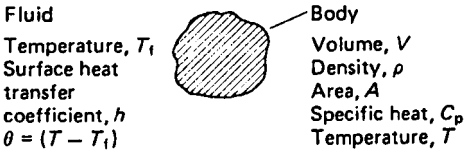


FIGURE 3.61 A lumped-capacity system with convection.

3.26 ANALYSIS OF HEAT TRANSFER

Conduction

By considering the thermal equilibrium of a small, three-dimensional element of solid, isotropic material it can be shown that for a rectangular coordinate system

$$\frac{\partial T}{\partial t} = \alpha \left[\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} \right] + \frac{\dot{Q}'''}{\rho c_p}$$

where $\partial T/\partial t$ is the rate of change of temperature with time, α is the thermal diffusivity of the material $\alpha = k/\rho c_p$ and $\dot{Q}'''$ is the internal heat generation rate per unit volume, which may be due, for example, to electric current flow for which $\dot{Q}''' = i^2 r$, where i is the current density and r the resistivity. The solution to this equation is not easy, and numerical approximation methods are often used. One such method is the finite difference technique, in which continuously varying temperatures are assumed to change in finite steps. Consider the three planes shown a distance Δx apart (Figure 3.62). At

$$A \frac{\partial T}{\partial x} = \frac{T_1 - T_2}{\Delta x}$$

and at

$$B \frac{\partial T}{\partial x} = \frac{T_2 - T_3}{\Delta x}$$

so that at

$$C \frac{\partial^2 T}{\partial x^2} = \frac{T_1 + T_3 - 2T_2}{\Delta x^2}$$

Similarly, $\partial T/\partial t$ may be written

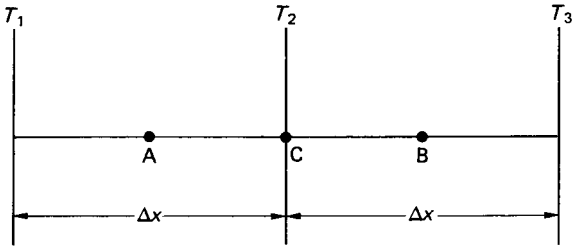


FIGURE 3.62 One-dimensional finite difference formulation.

$$\frac{T_{n,1} - T_{n,0}}{\Delta t}$$

where $T_{n,1}$ is the temperature at layer n at time 1 and $T_{n,0}$ is the temperature at layer n at time 0.

For steady state situations $\partial T/\partial t = 0$ and a two-dimensional plane surface will be used for illustration requiring a solution of

$$\alpha \left[\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \right] + \frac{\dot{Q}'''}{\rho c_p} = 0$$

Consider the surface to be divided by a grid (Figure 3.63). It is then found that the solution for any point in the plane for steady state conduction without heat generation is

$$T_1 + T_2 + T_3 + T_4 - 4T_0 = 0$$

or with heat generation is

$$T_1 + T_2 + T_3 + T_4 - 4T_0 = -\frac{a^2 \dot{Q}'''}{k}$$

At the boundary of the plane conditions are usually isothermal, in which case $T =$ constant, or convective when an energy balance yields for a straight boundary (Figure 3.64)

$$\frac{T_1}{2} + T_2 + \frac{T_3}{2} + \frac{ha}{k} \cdot T_f - T_0 \left(2 + \frac{ha}{k} \right) = 0$$

where h is the surface heat transfer coefficient, T_f is the fluid temperature, and a is the grid size. Similar expressions can be derived for corners, curves, etc. at the boundary.

For a large number of grid points, a large number of simultaneous equations are obtained which can be solved by iteration or Gaussian elimination. Computer programs may be used to advantage. The solution obtained will be the temperature

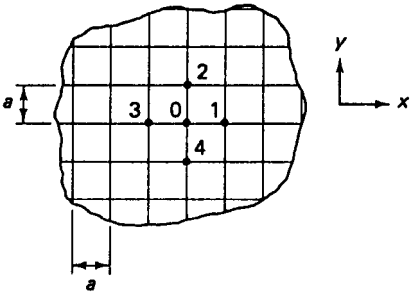


FIGURE 3.63 The two-dimensional grid concept.

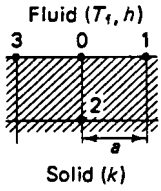


FIGURE 3.64 Convective surface nomenclature.

distribution in the plane, and the heat transfer rates may be found at the boundary (Figure 3.65):

With an isothermal boundary $\dot{Q} = \Sigma k(T_m - T_{wall})$

With a convective boundary $\dot{Q} = \Sigma ha(T_f - T_m)$

For transient heating or cooling for which $\partial T/\partial t \neq 0$, a one-dimensional illustration is used. The equation to be solved is

$$\frac{\partial T}{\partial t} = \alpha \frac{\partial^2 T}{\partial x^2}$$

when there is no heat generation (Figure 3.66).expressed in finite difference form this becomes

$$T_{n,1} = F\left(T_{n-1,0} + T_{n+1,0} + T_{n,0}\left(\frac{1}{F} - 2\right)\right)$$

where F is the nondimensional grid size Fourier number $F = \alpha \Delta t/a^2$. The only unknown in this equation is $T_{n,1}$, the temperature at layer n after one time interval Δt. Thus from a knowledge of the initial conditions successive temperatures in each layer can be found directly for each time interval. This is the *explicit* method and is used for tabular or graphical (Schmidt method) solutions. If $F > 0.5$ the solution

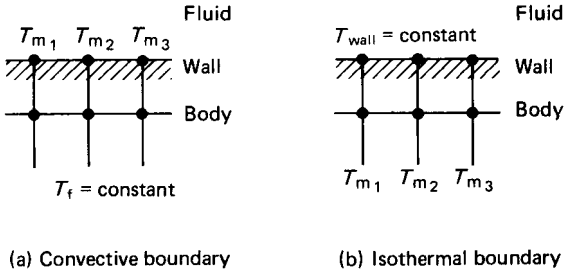


FIGURE 3.65 Calculation of heat transfer rate.

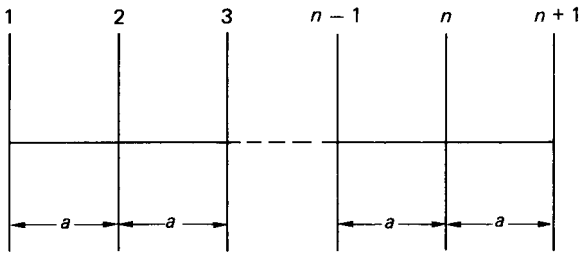


FIGURE 3.66 One-dimensional transient conduction formulation.

is unstable and in three dimensions the criterion becomes severe. The boundary conditions may be isothermal or convective and in the latter case the solution is

$$T_{n,1} = 2F \cdot T_{n-1,0} + T_{n,0}[1 - 2F - 2F \cdot B] + 2F \cdot BT_f$$

where B is the nondimensional grid Biot number, $B = ha/k$. For this case the solution is unstable if $(F + FB) > 0.5$. The solutions obtained give the temperature distribution in the one-dimensional plane and the heat transfer is found at the boundary

$$Q = \sum_t kA \left(\frac{\partial T}{\partial x}\right)_{1,t} \Delta t$$

or from the temperature profile

$$Q = \sum_{\text{layer}} mc_p(T_{\text{final}} - T_{\text{initial}})$$

The stability problems of the explicit method can be overcome by the use of *implicit* methods for which there is no direct solution, but a set of simultaneous equations are obtained which may be solved by Gaussian elimination. A computer program may be used to advantage. A satisfactory implicit method is that due to Crank and Nicolson. The importance of a stable solution is that if the choice of F is limited then the grid size and time interval cannot be freely selected, leading to excessive calculations for solution. The implicit method releases this constraint but care is still needed to ensure accuracy.

Although the finite difference method has been chosen for demonstration because the method is easy to understand, most modern computer programs are based on the finite element technique. However, the mathematical principles are involved, and would not lend themselves to simple programming. Before the availability of computer software analytical solutions were obtained and presented as graphs of transient solutions for slabs, cylinders, and spheres. These graphs enable solutions for other shapes to be obtained by superposition methods. Such methods should be used to avoid or validate computer solutions.

Warning: If fiber-reinforced materials are used in which the lay-up is arranged to give directional structural strength it will be found that the thermal conductivity has directional variation and the methods above will need considerable amendment.

Convection

A knowledge of the surface heat transfer coefficient h is essential in determining heat transfer rates. Fluid flow over a solid surface is a boundary layer problem, and the heat transfer depends on boundary layer analysis. This analysis may be by differential or integral approach, but solution is difficult and the modeling of turbulence is complex. Computer solutions based on numerical approximations may be used to advantage, but simple approaches have been used for many years and are still extremely useful. These methods are based on Reynolds' analogy (modified by later workers) and dimensional analysis backed by experimentation.

Convection may be free or forced. In *forced convection* it is found that the heat transfer coefficient can be included in a nondimensional relation of the form

MECHANICAL ENGINEERING PRINCIPLES

$$Nu = \phi(Re,Pr) = \text{constant} \cdot Re^a \cdot Pr^b$$

where Nu is the Nusselt number ($Nu = hl/k$), Re is the Reynolds number ($Re = \rho Vl/\mu$), and Pr is the Prandtl number ($Pr = \mu c_p/k$). In these relations l is a representative length dimension (diameter for a pipe and some chosen length for a plate), V is the bulk or free stream velocity outside the boundary layer. The values of the constants a and b depend on whether the flow is laminar or turbulent and on the geometry of the situation, and are usually found by experiment.

The determination of whether flow is laminar or turbulent is by the value of the Reynolds number;

For plates, $Re < 500\,000$, flow is laminar: $Re > 500\,000$, flow is turbulent

For tubes, $Re < 2000$, flow is laminar: $Re > 4000$, flow is turbulent

(between these two values there is a transition zone). There are many relations to be found in texts which allow for entry length problems, boundary conditions, etc. and it is not feasible to list them all here. Two relations are given below which give average values of Nusselt number over a finite length of plate or tube in forced, turbulent flow with Mach number less than 0.3 using total plate length and diameter for representative length dimension. Care must be taken in any empirical relation to use it as the author intended.

Plate: $$Nu = 0.036Re^{0.8}Pr^{0.33}$$

In this relation fluid properties should be evaluated at the film temperature, $T_{\text{film}} = (T_{\text{wall}} + T_{\text{bulk}})/2$.

Tube: $$Nu = 0.023Re^{0.8}Pr^{0.4}$$

In this relation fluid properties should be evaluated at the bulk temperature, $0.6 < Pr < 160$ and $(l/d) > 60$.

It should be noted that the index of Reynolds number of 0.8 is characteristic of turbulent flow; in laminar flow 0.5 is found.

It must be emphasized that reference to other texts in all but these simple cases is essential to estimate heat transfer coefficients. It should also be pointed out that the values obtained from such relations could give errors of 25%, and a search of the literature might reveal equations more suited to a particular situation. However, an estimate within 25% is better than no knowledge, and is a suitable starting point which may be modified in the light of experience.

For complex heat exchange surfaces such as turbine blade cooling in an aircraft engine, empirical information is usually presented graphically (on these graphs the nondimensional group St (Stanton number) may appear:

$$St = \frac{Nu}{RePr} = \frac{h}{\rho Vc_p}$$

In *free convection* the relationship used is $Nu = \phi(Pr \cdot Gr)$, where Gr is the Grashof number, $\rho^2\beta g\theta l^3/\mu^2$ in which β is the coefficient of cubical expansion of the fluid and θ is a temperature difference (usually surface to free stream temperature). The transition from laminar to turbulent flow is determined by the product $(Pr \cdot Gr)$ known as the Rayleigh number, Ra. As a simple example, for plane or cylindrical vertical surfaces, it is found that

For $Ra < 10^9$, flow is laminar and $Nu = 0.59(Pr \cdot Gr)^{0.25}$

For $Ra > 10^9$, flow is turbulent and $Nu = 0.13(Pr \cdot Gr)^{1/3}$

The representative length dimension is height and the resulting heat transfer coefficients are average values for the whole height. Film temperature is used for fluid properties. Warnings similar to those given for forced convection apply to the use of these equations.

Phase change convection heat transfer (condensing and evaporation) shows coefficients that are, in general, higher than those found in single-phase flow. They are not discussed here but information may be found in standard texts.

Radiation

In space technology applications heat transfer systems are designed for conduction and radiation transfer only. When an emitting body is not surrounded by the receiver, the spatial distribution of energy from the radiating point needs to be known. To determine this distribution the intensity of radiation i_ϕ is defined in any direction ϕ as

$$i_\phi = \left(\frac{d\dot{E}''}{d\omega}\right)_\phi$$

where $d\omega$ is a small solid angle subtended at the radiating point by the area intercepting the radiation, $d\omega = dA/r^2$ (Figure 3.67) (the solid angle represented by a sphere is 4π steradians). Lambert's law of diffuse radiation states that $i_\phi = i_n \cos \phi$ where i_n is the normal intensity of radiation which can be determined for black and gray bodies;

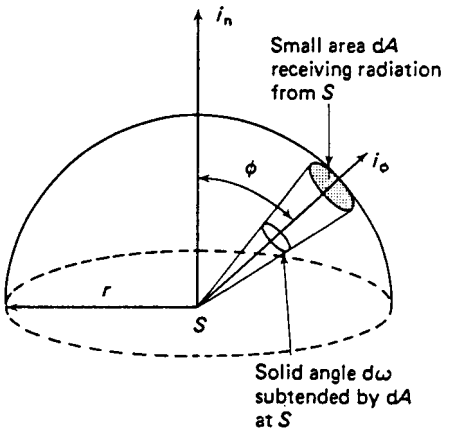


FIGURE 3.67 Spatial distribution of radiation.

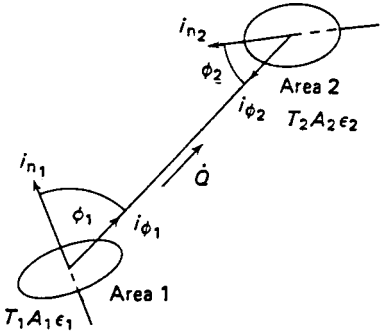


FIGURE 3.68 Heat transfer by radiation between two arbitrarily disposed gray surfaces.

MECHANICAL ENGINEERING PRINCIPLES

Black $i_n \sigma T^4/\pi$ Gray $i_n\ \varepsilon\sigma T^4/\pi$

With this knowledge of the radiation intensity in any direction it is only necessary to determine the amount that any body can see of any other body to calculate the heat transfer rate. For this purely geometric problem mathematical analysis (Figure 3.68) suggests a quantity variously known as the geometric, configuration or shape factor, which is defined as the fraction of the energy emitted per unit time by one surface that is intercepted by another surface. The geometric factor is given by

$$F_{12} = \frac{1}{A_1}\int_{A_1}\int_{A_2}\frac{\cos\phi_1\ \cos\phi_2\ dA_1 dA_2}{\pi x^2}$$

It can be seen that $A_1F_{12} = A_2F_{21}$, a useful reciprocal relation. It is also clear that the equation will require skill to solve in some situations, and to overcome this problem geometric factors are available for many situations in tables or on graphs (Hottel charts). The charts can give more information than anticipated by the use of shape factor algebra, which enables factors to be found by addition, subtraction, etc. (Figure 3.69).

Having established the intensity of radiation and the geometric factor, problems may be solved by an electrical analogy using the radiosity of a surface. Radiosity is defined as the total emitted energy from a gray surface:

$$\dot{J}'' = \dot{E}''_g + \rho\dot{G}''$$

where $\dot{J}''$ is the radiosity and $\rho\dot{G}''$ the reflected portion of the incident radiation $\dot{G}''$. Since $\dot{E}''_g = \varepsilon\dot{E}''_0$ the net rate of radiation leaving a gray surface of area A becomes

$$\frac{\dot{E}''_b - \dot{J}''}{\rho/A\varepsilon}$$

which may be envisaged as a potential difference, $\dot{E}''_b - \dot{J}''$, divided by a resistance, ρ/AE. A similar geometric resistance of $1/AF$ can be established to enable complete circuits to be drawn up. Thus for a three-body problem we may sketch the analogous electrical circuit (Figure 3.70) and apply Kirchhoff's electric current law to each J'' node to obtain three simultaneous equations of the form

$$\frac{\dot{E}''_{b_1} - \dot{J}''_1}{\rho_1/A_1\varepsilon_1} + \frac{\dot{J}''_2 - \dot{J}''_1}{1/A_1F_{12}} + \frac{\dot{J}''_3 - \dot{J}''_1}{1/A_1F_{13}} = 0$$

If there are more than three bodies sketching becomes complex and the equation above can be rearranged and generalized. For N surfaces ($j = 1$ to N) there will be N equations, the ith of which ($i = 1$ to N) will be

$$\dot{J}''_i - (1 - \varepsilon_i)\sum_{j=1}^{N} F_{ij}\dot{J}''_j = \varepsilon_i\dot{E}''_{b_i}$$

When $j = i$, F_{ii} will be zero unless the the surface is concave and can see itself. This set of N simultaneous equations may be solved by Gaussian elimination for which a computer program may be used. The output of the solution will be N

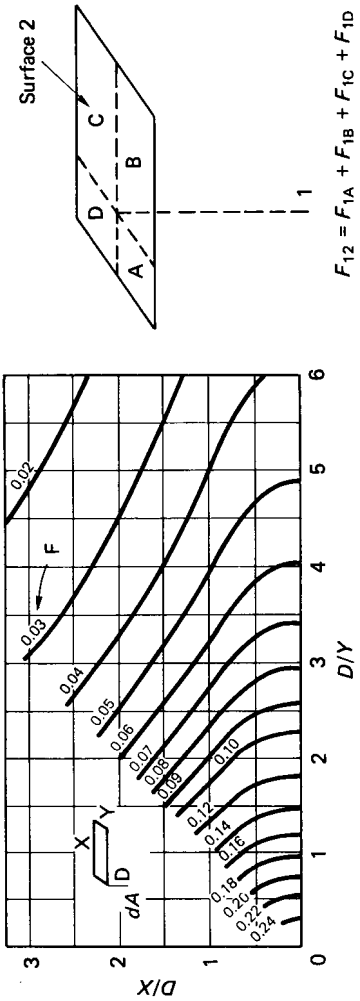


FIGURE 3.69 The geometric factor chart.

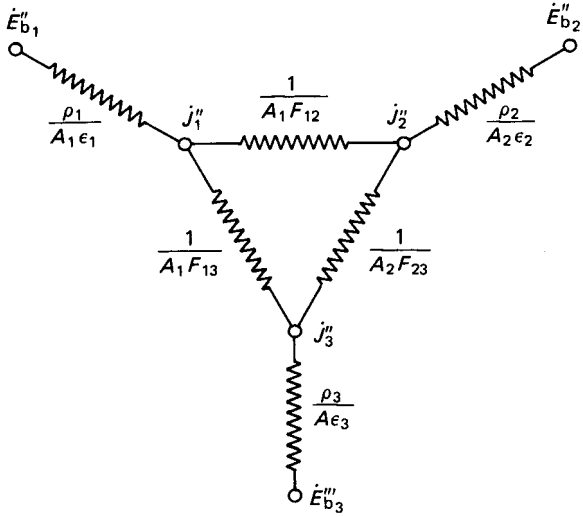


FIGURE 3.70 The electrical analogy for gray-body radiation problems.

values of $\dot{J}''$ and any required heat flux can be found. In the three-body problem, for example, the heat transfer from 1 to 2

$$\dot{Q}_{12} = \frac{\dot{J}''_1 - \dot{J}''_2}{1/A_1F_{12}}$$

or the total heat transfer from body 1 is

$$\frac{\dot{E}''_1 - \dot{J}''_1}{\rho_1/A_1\varepsilon_1}$$

Special cases are:

1. If $N - 1$ bodies are in large surroundings then for this Nth body $\dot{J}'' = \dot{E}''_b$.
2. An insulated or refractory surface has no black body potential but contributes to the heat transfer by taking up an equilibrium temperature T given by $\sigma T^4 = \dot{J}''$
3. Radiation shield problems will show six resistances in series rather than the series-parallel circuits used previously.

In all the discussion on radiation above no account has been taken of the selective emitter for which emissivity is not constant. Additional techniques are required to solve these real problems and care should be exercised if widely varying emissivity is encountered. In new situations it will be necessary to determine emissivity by experiment before proceeding. No account has been taken of intervening media for which transmissivity is not unity. Gas absorption and radiation needs further information. For gases such as oxygen, nitrogen and hydrogen with symmetric, diatomic molecules, the above work is adequate, but asymmetric molecular struc-

tures cause problems. In particular, hydrocarbon fuel combustion products (H_2O, CO_2, CO, SO_2) are important in engineering calculations and account must be taken of their radiation properties. Solar radiation problems also need special consideration.

Finned Surfaces

In many heat exchange problems involving the determination of a *U*-value it is found that the surface heat transfer coefficient on one side of the solid interface is much smaller than that on the other. The smaller coefficient will dictate the heat transfer rate achieved, and in order to overcome this problem fins may be added to this poor convection surface to increase the area for heat transfer. This problem will occur in liquid to gas exchangers on the gas side. The addition of fins will alter the flow pattern so that a new coefficient should if possible be determined. It is also possible that there may be variation of coefficient over the fin surface.

Simple fin theory in which conduction along the fin is balanced with convection from the surface can be used to determine the temperature distribution and heat transfer rate of the fin. For example, when a long fin of constant cross-sectional area is examined (Figure 3.71) it is found that the temperature distribution is

$$\frac{\theta}{\theta_0} = \frac{\cosh(m(l - x))}{\cosh(ml)}$$

and the heat transfer rate is

$$\dot{Q} = mkA\theta_0 \tanh(ml)$$

where θ is the temperature difference between fin and fluid and θ_0 is the difference at the fin root, $m = hp/kA$, p is the fin perimeter and A is the cross-sectional area.

A fin efficiency is then defined to compensate for the varying temperature difference along the fin as

$$\eta_{\text{fin}} = \frac{\text{Actual heat transfer rate}}{\text{Heat transfer rate if whole fin were at the wall temperature}}$$

which for the simple case above is $\eta_{\text{fin}} = \tanh(ml)/ml$. Fins are usually fitted in

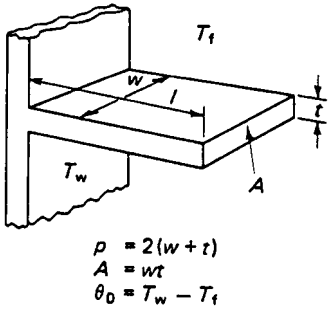


FIGURE 3.71 Simple rectangular fin nomenclature.

MECHANICAL ENGINEERING PRINCIPLES

arrays and the efficiency of a fin system can be established in the form of an area weighted fin efficiency, η'

$$\eta' = \eta_{\text{fin}}\beta + 1 - \beta$$

where

$$\beta = \frac{\text{total fin area}}{\text{total area including fins}} = \frac{A_{\text{fin}}}{A}$$

The U value based on the enhanced area A is then

$$\frac{1}{U_A} = \frac{1}{(A_1/A)h_1} = \sum \frac{\Delta x}{R} + \frac{1}{\eta' h_2}$$

For complete surfaces β is supplied by the manufacturer.

When fins of more complex shape are used (tapered fins or annular fins) the cross-sectional area is not constant and fin efficiency data are obtained from graphs. Care should be taken in the interpretation of such graphs since the equations above may not agree with the definitions used for the graphs.

Heat Exchangers

It is possible to design a heat exchanger with the information above and obtain a basic idea of size and configuration of simple tubular structures. It is first necessary to realize that temperature differences change along a heat exchanger, and that flow may be parallel or counter in direction. The latter is to be preferred, as it leads to smaller sizes. To allow for the changing temperature difference a log mean temperature difference is used (Figure 3.72):

$$\theta_{\text{LMTD}} = \frac{\theta_1 - \theta_2}{\ln(\theta_1/\theta_2)}$$

and to allow for varying flow patterns (which are neither counter nor parallel flow) graphs are available to give a factor F to modify the (usually) counterflow value of mean temperature difference. Thus, for any heat exchanger,

$$\theta_{\text{mean}} = F\theta_{\text{LMTD}}$$

If an estimated mean U-value for the surface is then determined the heat exchange equation is

$$\dot{Q} = UAF\theta_{\text{LMTD}}$$

Thus the area is determined. There are many solutions to this equation to satisfy all the constraints which will include:

$$\text{Energy for each stream } \dot{Q} = (\dot{m}\Delta h)_{\text{hot}} = (\dot{m}\Delta h)_{\text{cold}} \quad (\Delta h = \text{enthalpy change})$$

$$\text{Continuity of each stream } \dot{m} = \rho AV$$

$$\text{Heat transfer area } A = \pi dl$$

The above equations will involve options with varying tube numbers and di-

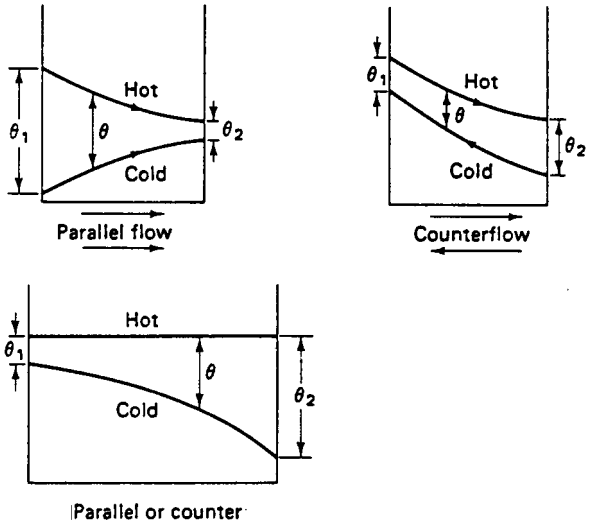


FIGURE 3.72 The mean temperature difference in heat exchangers.

ameters which will affect the determination of heat transfer coefficients (and hence the *U*-value), and a number of solutions will be obtained. The optimum choice will involve allowable pressure drops, velocities and exit temperatures. It is not a difficult calculation, but, because of the choice, a computer program may be used. There are short cuts to this approach based on the interrelation between pressure drop and heat transfer (modified Reynolds' analogy) to determine heat transfer coefficients and a method based on graphs of effectiveness (E), capacity ratio (c), and number of transfer units (NTU) is sometimes used.

By whatever method a design is achieved, it will, unless the application is very simple, be necessary to consult a professional heat exchanger designer with experience and full computer programs.

3.27 USE OF COMPUTERS

Computers may be used as an aid to heat transfer calculations at various levels of skill. At the simplest level, computer programs may be written to determine, for example, heat losses in buildings, heat inputs from pipes and radiators, etc. This will save repetitive calculations and build a small library of useful programs. To aid those whose heat transfer and computing skills are slight, Bacon (1989) gives a simple approach to problems. For more detail on the mathematics of finite difference techniques in heat transfer Myers (1971) is useful.

An alternative approach to numerical approximations in computer work is to use finite element methods. For heat transfer applications Myers (1971) is again useful. Very few people will find the need for finite element programming skills, as there

is an ever-increasing range of software aimed at the solution of many engineering problems, including heat transfer. Many of these are PC packages in which there is integration between design, drawing, manufacture and analysis of stress, vibration, heat transfer, etc.

Whenever numerical approximation techniques are used, large-scale analysis leads to considerable demand on data storage and computing time. It is therefore essential to do as much as possible with simple methods before becoming committed to large-scale finite element packages. If it is decided that the use of such a package is necessary it is vital to be sure of the requirements of the problem, for volume of input data and output results will be large and unpalatable. To assist with this problem, preprocessing packages are used for data input and mesh generation with graphic display, and post-processing packages are used for graphic display of the results. For example, the temperature distribution in a combustion chamber or piston displayed by color graphics enables easier identification of problem areas.

Finally, it must be emphasized that computing is not a substitute for understanding heat transfer problems. It is an aid to enable a more detailed investigation to be achieved and presented in a fashion to enable engineers to improve their designs.

3.28 HEAT TRANSFER: NOMENCLATURE

A	area
B	Biot number
c_p	Specific heat capacity at constant pressure
$\dot{E}''_b$	Black-body emissive power
$\dot{E}'_g$	Gray-body emissive power
F	Fourier number, geometric factor, mean temperature factor
g	gravitational constant
Gr	Grashof number
h	surface heat transfer coefficient
i_ϕ	intensity of radiation in direction ϕ
i_h	normal intensity of radiation
$\dot{J}''$	radiosity
k	thermal conductivity
l	length
m	mass, fin parameter
$\dot{m}$	mass flow rate
Nu	Nusselt number
p	perimeter
Pr	Prandtl number
Q	heat transfer
$\dot{Q}$	heat transfer rate
$\dot{Q}'$	heat transfer rate per unit length

$\dot{Q}''$	heat transfer rate per unit area
$\dot{Q}'''$	heat transfer rate per unit volume
r	radius
Ra	Rayleigh number
Re	Reynolds number
St	Stanton number
T	temperature
t	time
U	overall heat transfer coefficient per unit area
U'	overall heat transfer coefficient per unit length
V	velocity
x, y, z	rectangular coordinates

Greek Letters

α	thermal diffusivity, absorbtivity
β	coefficient of cubical expansion
Δ	change in
η	efficiency
ε	emissivity
λ	wavelength
μ	viscosity
ϕ	angle
ρ	density, reflectivity
σ	Stefan–Boltzmann constant
Σ	summation
θ	temperature difference
τ	transmissivity
φ	solid angle

REFERENCES

Adams, J. A. and Rogers, D. F. 1973. *Computer-Aided Heat Transfer Analysis*, McGraw-Hill, New York.

Anvoner, S. 1972. *Solution of Problems in Mechanics of Machines,* vol. 2, Pitman, London.

Bacon, D. H. 1989. *Basic Heat Transfer,* Butterworths, London.

Bacon, D. H. 1983. *BASIC Thermodynamics and Heat Transfer,* Butterworths, London.

Bejan, A. 1982. *Entropy Generation through Heat and Fluid Flow,* Wiley, Chichester.

CIBS Guide. 1980. Section A3, Thermal Properties of Building Structures; Section A5, Thermal Response of Buildings. CIBS, London.

Cole, E. B. 1950. *Theory of Vibrations,* Crosby Lockwood, London.

Douglas, J. F., Gasiorek, J. M. and Swaffield, J. A. 1986. *Mchanics of Fluids,* 3d ed., Pitman, London.

Haywood, R. W. 1976. *Thermodynamic Tables in SI* (*Metric*) *Units,* Cambridge University Press, Cambridge.

Houghton, F. L. and Brock, A. E. 1961. *Tables for the Compressible Flow of Dry Air,* Edward Arnold, London.

Kern, D. Q. 1984. *Process Heat Transfer,* McGraw-Hill, New York.

Kotas, T. J. 1965. *The Exergy Method of Thermal Plant Analysis*, Butterworths, London.

Kreith, F. and Bohn, M. S. 1986. *Principles of Heat Transfer,* Harper & Row, New York.

Mathesun, J. A. L. 1971. *Hyperstatic Structures: An Introduction to the Theory of Statically Indeterminate Structures,* Butterworths, London.

Meriam, J. L. and Kraige, L. G. 1987. *Engineering Mechanics,* vol. 2, *Dynamics,* John Wiley, Chichester.

Myers, G. E. 1971. *Analytical Methods in Conduction Heat Transfer,* McGraw-Hill, New York.

Özişik, M. N. 1984. *Heat Transfer,* McGraw-Hill, New York.

Rogers, G. F. C. and Mayhew, Y. R. 1986. *Engineering Thermodynamics, Work and Heat Transfer,* Longman, Harlow.

Scarborough, J. 1958. *The Gyroscope: Theory and Applications,* Interscience, New York.

Schlichting, H. 1960. *Boundary Layer Theory,* 4th ed., McGraw-Hill, New York.

Thompson, W. T. 1983. *Theory of Vibrations with Applications,* 2nd ed., George Allen & Unwin, London.

White, F. M. 1984. *Heat Transfer,* Addison-Wesley, Reading, MA.

Wilson, W. K. 1959. *Vibration Engineering: A Practical Treatise on the Balancing of Engines, Mechanical Vibration and Vibration Isolation,* Griffin, London.

Young, D. and Feglar, R. P., Jr. 1949. *Tables of Characteristic Functions Representing Normal Modes of Vibration of a Beam,* University of Texas Publication Number 4913 (July).

FURTHER READING

Crandall, S. H., *Random Vibration,* Technology Press and John Wiley, Chichester (1958).

Crandall, S. H. and Mark, W. D., *Random Vibration in Mechanical Systems,* Academic Press, London (1963).

Davenport, W. B., *Probability and Random Processes,* McGraw-Hill, New York (1970).

Gorman, D. J., *Free Vibration Analysis of Beams and Shafts,* Wiley, Chichester (1975).

Harker, R., *Generalised Methods of Vibration Analysis,* Wiley, Chichester (1983).

Hatter, D., *Matrix Computer Methods of Vibration Analysis,* Butterworths, London (1973).

Helstrom, C. W., *Probability and Stochastic Processes for Engineers,* Macmillan, London (1984).

Johnston, E. R. and Beer, F. P., *Mechanics for Engineers,* Volume 1, *Statics;* Volume 2, *Dynamics,* McGraw-Hill, New York (1987).

Meriam, J. L. and Kraige, L. G., *Engineering Mechanics,* vol. 1, *Statics,* 2nd Wiley, Chichester (1987).

Nestorides, E. J., *A Handbook of Torsional Vibration,* Cambridge University Press, Cambridge (1958).

Newland, D. E., *An Introduction to Random Vibrations and Spectral Analysis,* 2nd ed., Longman, Harlow (1984).

Nigam, N. C., *Introduction to Random Vibrations,* MIT Press (1983).

Nikravesh, P. E., *Computer Aided Analysis of Mechanical Systems,* Prentice-Hall, Englewood Cliffs, NJ (1988).

Piszek, K. and Niziol, J., *Random Vibration of Mechanical Systems,* Ellis Horwood, Chichester (1986).

Robson, J. D., *An Introduction to Random Vibration,* Edinburgh University Press (1963).

Tse, F. S., Morse, I. E., and Hinkle, R. T., *Mechanical Vibrations: Theory and Applications,* 2nd ed., Allyn and Bacon, New York (1979).

General Fluid Mechanics

Ireland, J. W., *Mechanics of Fluids,* Butterworths, London (1971).

Walshaw, A. C. and Jobson, D. A., *Mechanics of Fluids,* 3d ed., Longman, London (1979).

Reference Tables

Rogers, G. F. C. and Mayhew, Y. R., *Thermodynamic and Transport Properties of Fluids,* 3rd ed., Blackwell, Oxford (1980).

Additional Computer Solutions

Smith, P. D., *BASIC Hydraulics,* Butterworths, London (1982).

Ideal Fluid Flow

O'Neill, M. E. and Chorlton, F., *Ideal and Incompressible Fluid Dynamics,* Ellis Horwood, Chichester (1986).

For Students Preparing for Examinations

Brasch, D. J. and Whyman, D., *Problems in Fluid Flow,* Edward Arnold, London (1986).

Douglas, J. F., *Solutions of Problems in Fluid Mechanics,* Parts 1 and 2, 3d ed., Pitman, London (1987).

SECTION 4
ELECTRICAL AND ELECTRONIC PRINCIPLES

Section Editor: Mark Davies
Author: Charles Fraser

4.1 BASIC ELECTRICAL TECHNOLOGY

Flux and Potential Difference

The concept of flux and potential difference enables a unified approach to be adopted for virtually all the field type of problems. Generally, the flowing quantity is termed the flux and the quantity that drives the flow is called the potential difference. This consistency of method is equally applicable to problems in fluid flow, heat transfer, electrical conduction, electrostatics, and electromagnetism, to name but a few.

In general terms, the flux may be written as

$$\text{Flux} = \frac{(\text{Field characteristic}) \times (\text{Cross-sectional area}) \times (\text{Potential difference})}{(\text{Length})} \tag{4.1}$$

In specific terms, for the flow of an electric current through a conducting medium, equation (4.1) takes the form:

$$I = \frac{\sigma a V}{l} \tag{4.2}$$

where I = the current in amperes (A)
σ = is the conductivity of the medium (siemens/m), i.e., the field characteristic
a = the cross-sectional area of the medium (m^2)
l = the length of the medium (m)
V = the applied potential difference, or voltage (V)

The group $(\sigma a/l)$ is termed the conductance, denoted by G and measured in siemens, thus:

$$I = GV \tag{4.3}$$

The reciprocal of conductance is referred to as the resistance, R, and is measured in ohms (Ω). Hence,

$$I = V/R \tag{4.4}$$

Equation (4.4) is the familiar Ohm's law, which defines a linear relationship between voltage and current in a conducting medium. If the resistance, R, varies with the magnitude of the voltage, or the current, then the resistance is nonlinear. Rectifiers constitute one particular class of nonlinear resistors.

Comparing equations (4.4) and (4.2) gives:

$$R = l/(\sigma a) \tag{4.5}$$

It is more usual, however, to quote the resistivity as opposed to the conductivity, and resistance is generally written as:

$$R = \rho l/a \tag{4.6}$$

where ρ is the resistivity of the conductor in ohm-meters.

The resistance of all pure metals is temperature dependent, increasing linearly for moderate increases in temperature. Other materials, including carbon and many insulators, exhibit a decreasing resistance for an increase in temperature.

Simple Resistive Circuits

The effective total resistance of a series arrangement is the algebraic sum of all the resistances in series, i.e.,

$$R_t = R_1 + R_2 + R_3 \tag{4.7}$$

where R_t is the total resistance of the circuit.

For resistors in parallel the effective total resistance obeys an inverse summation law, i.e.,

$$\frac{1}{R_t} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} \tag{4.8}$$

Electromotive Force and Potential Difference

In a metallic conductor that has a potential difference applied across opposite ends, free electrons are attracted to the more positive end of the conductor. It is this drift of electrons which constitutes the electric current, and the effect is simply nature's attempt to redress an energy imbalance. Although the negatively charged electrons actually drift towards the positive end of the conductor, traditional convention gives the direction of the current flow from positive to negative. There is nothing really at issue here, since it is only a simple sign convention that was adopted long before the true nature of the atom and its associated electrons were postulated.

A current of 1 A is associated with the passage of 6.24×10^{18} electrons across any cross-section of the conductor per second. The quantity of charge is the coulomb, Q, and

$$Q = I \cdot t \tag{4.9}$$

where 1 coulomb of charge is passed when a current of 1 A flows for a period of 1 s.

The electromotive force (e.m.f.) is that which tends to produce an electric current in a circuit, and is associated with the energy source. Potential difference is simply the product of current and resistance across any resistive element in a circuit, irrespective of the energy source. For circuit elements other than purely resistive, the potential difference across the element becomes a time-dependent function.

Power and Energy

Power is the rate at which energy is expended, or supplied. The potential difference across any two points in a circuit is defined as the work done in moving unit charge from a lower to a higher potential. Thus the work done in moving Q coulombs of charge across a constant potential difference of V volts is:

$$W = Q \cdot V \tag{4.10}$$

Therefore,

$$\text{Power} = \frac{dW}{dt} = \frac{dQ}{dt} V$$

From equation (4.9) $(dQ/dt) = I$. Thus

$$\text{Power} = IV \tag{4.11}$$

Using Ohm's law, the power dissipated across a simple resistive circuit element is

$$\text{Power} = IV = I(I \cdot R) = I^2\text{R} \tag{4.12}$$

Network Theorems

A network consists of a number of electrical elements connected up in a circuit. If there is no source of electromotive force in the circuit it is said to be passive. When the network contains one or more sources of electromotive force it is said to be active.

A number of well-established theorems have been developed for the analysis of complex resistive networks (Hughes 1987; Bell and Whitehead 1987; Bell 1984) and are listed below:

Kirchhoff's first law: The algebraic sum of the currents entering $(+ve)$ and leaving $(-ve)$ a junction is zero.

Kirchhoff's second law: The algebraic sum of potential differences and e.m.f.'s around any closed circuit is zero.

Superposition theorem: In a linear resistive network containing more than one source of e.m.f. the resultant current in any branch is the algebraic sum of the currents that would be produced by each e.m.f. acting on its own while the other e.m.f.'s are replaced with their respective internal resistances. *Thevenin's theorem:* The current through a resistor R connected across any two points in an

active network is obtained by dividing the potential difference between the two points, with R disconnected, by $(R + r)$, where r is the resistance of the network between the two connection points with R disconnected and each e.m.f. replaced with its equivalent internal resistance. An alternative statement of Thevenin's theorem is: "Any active network can be replaced at any pair of terminals by an equivalent e.m.f. in series with an equivalent resistance." The more concise version of Thevenin's theorem is perhaps a little more indicative of its power in application.

Norton's theorem: Any active network can be replaced at any pair of terminals by an equivalent current source in parallel with an equivalent resistance. It may appear that Norton's theorem is complementary to Thevenin's theorem and both can be equally well used in the analysis of resistive networks.

Other useful network analysis techniques include mesh analysis, which incorporates Kirchhoff's first law, and nodal analysis, which is based on Kirchhoff's second law. Mesh and nodal analysis are also essentially complementary techniques.

Double-Subscript Notation

To avoid ambiguity in the direction of current, e.m.f., or potential difference, a double-subscript notation has been adopted. Figure 4.1 shows a source of e.m.f. that is acting from D to A. The e.m.f. is therefore E_{da}. The current flows from A to B, by traditional convention, and is designated I_{ab}. From this simple circuit it is apparent that $I_{ab} = I_{bc} = I_{cb} = I_{da}$.

The potential difference across the load R is denoted V_{bc} to indicate that the potential at B is more positive than that at C. If arrowheads are used to indicate the potential difference, then they should point towards the more positive potential.

Electrostatic Systems

Electrostatic systems are quantified by the physical behavior of the charge. Fortunately, the unified field approach lends itself well to the quantification of electrostatic systems.

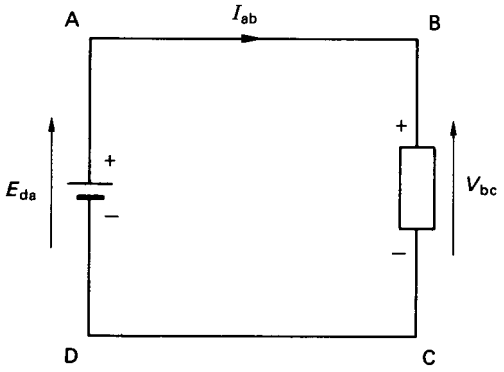


FIGURE 4.1 Double-subscript notation.

Figure 4.2 shows two parallel, conducting metal plates separated by an evacuated space. A potential difference is applied across the plates such that they become charged at equal magnitude but opposite sign. For the electrostatic system, equation (4.1) is written

$$Q = \frac{\varepsilon_0 a V}{l} \tag{4.13}$$

where Q = the total charge in coulombs
ε_0 = the permittivity of free space in Farads/m, i.e., the field characteristic
a = the cross-sectional area of the plates
l = the distance separating the plates
V = applied potential difference

The group $(\varepsilon_0 a/l)$ is termed the capacitance of the system. It is usually denoted by C, and is measured in farads (F). Thus

$$Q = C \cdot V \tag{4.14}$$

It is more common to use the microfarad (μF) or the picofarad (pF) as the unit of measurement.

$$\text{NB: } 1\ \mu\text{F} = 10^{-6}\ \text{F: } 1\ \text{pf} = 10^{-12}\ \text{F}$$

If the plates are separated by an insulating medium other than free space, then these so-called dielectric media have a different value of permittivity. The actual permittivity is related to the permittivity of free space by the relative permittivity of the dielectric, i.e.,

$$\varepsilon = \varepsilon_0 \cdot \varepsilon_\tau \tag{4.15}$$

where ε_τ is the relative permittivity of the dielectric. The permittivity of free space, ε_0, is numerically equal to $(1/36\pi) \times 10^{-9}$. The relative permittivity of some of the more common dielectric materials are listed in Table 4.1.

Simple Capacitive Circuits

For three capacitors connected in a simple parallel arrangement, the equivalent total capacitance is given as the algebraic sum of all the capacitances in the circuit, i.e.,

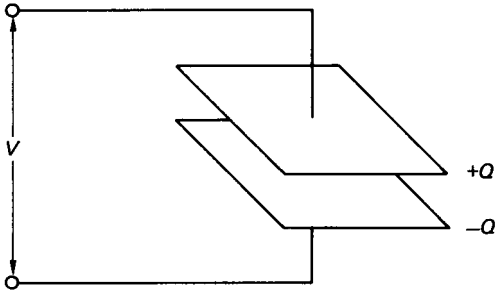


FIGURE 4.2 Electrostatic system.

TABLE 4.1 Relative Permittivities of Some Typical Dielectric Materials

Material	Relative permittivity
Air	1
Paper	2–2.5
Porcelain	6–7
Mica	3–7

$$C = C_1 + C_2 + C_3 \tag{4.16}$$

where C is the total capacitance. For a series capacitance arrangement of three capacitors, the total equivalent capacitance is related through the inverse summation given as

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} \tag{4.17}$$

Equations (4.16) and (4.17) can be used to reduce series and parallel capacitor circuits to a single equivalent capacitor.

Composite capacitors, involving different dielectric media, may also be treated in the same manner as a series capacitor arrangement.

Charging a Capacitor

Figure 4.3 shows a parallel plate capacitor that is connected in series with a resistor to a source of e.m.f. (say, a battery) through a switch. Initially, the capacitor is uncharged before the switch is closed. When the switch is closed a charging current will flow until such time that the potential difference across the capacitor is equal to the e.m.f. available from the source. The charging process consists of taking electrons from plate A and transferring them through the external wiring to plate

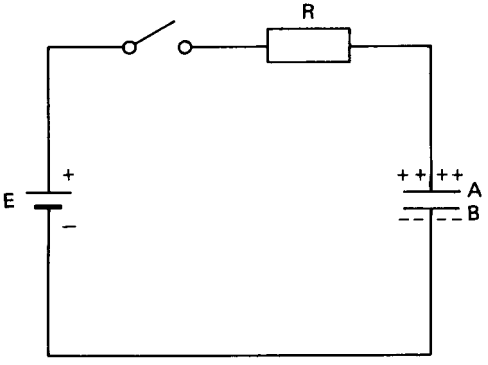


FIGURE 4.3 Charging a capacitor.

B. The energy required to do this is derived from the battery. The build-up of electrons from the negative terminal of the battery to plate *B* of the capacitor induces a dielectric flux between the plates and a balancing positive charge is developed on plate *A*. As long as the dielectric flux is changing, a current will flow externally. Eventually a state of equilibrium will be reached. Note that no electrons can pass through the dielectric since it is an insulator.

The instantaneous current during charging is

$$i = dQ/dt$$

From equation (4.14), this may be written for a capacitor as

$$i = dQ/dt = C(dv/dt) \tag{4.18}$$

where v is the instantaneous voltage. The instantaneous power is therefore

$$p = iv = Cv(dv/dt)$$

The energy supplied over the time period, dt, is

$$Cv(dv/dt)dt = Cvdv$$

Hence, the total energy supplied is

$$\int_0^V Cvdv = \frac{1}{2} CV^2 \tag{4.19}$$

Dielectric Strength

If the potential difference across opposite faces of a dielectric material is increased above a particular value, the material breaks down. The failure of the material takes the form of a small puncture, which renders the material useless as an insulator. The potential gradient necessary to cause break-down is normally expressed in kilovolts/millimeter and is termed the *dielectric strength*. The dielectric strength of a given material decreases with increases in the thickness. Table 4.2 gives approximate values for some of the more common dielectric materials.

Electromagnetic Systems

The magnetic field can be defined as the space in which a magnetic effect can be detected, or observed. An obvious magnetic field is observable around a straight length of conductor carrying a current. In particular, exactly the same magnetic field as that produced by a bar magnet is observed when the current-carrying conductor is formed into a helical type coil. The equipotential loops describe the path of the magnetic flux, ϕ, and although the flux lines have no physical meaning, they provide a convenient vehicle to quantify various magnetic effects.

The direction of the magnetic flux is governed by the so-called right-hand screw rule. This states that the direction of the magnetic field produced by a current corresponds with the direction given by turning a right-hand screw thread. The direction of the current corresponds with the translational movement of the screw.

TABLE 4.2 Dielectric Strength of Some Common Insulators

Material	Thickness (mm)	Dielectric strength (kV/mm)
Air	0.2	5.75
	0.6	4.92
	1.0	4.36
	10.0	2.98
Mica	0.01	200
	0.10	115
	1.00	61
Waxed paper	0.10	40–60

Magnetic Field of a Toroid

Figure 4.4 shows a toroidal coil, of N turns, which is wound round an annular former. A resultant magnetic flux, shown as broken lines in the figure, is generated when the coil carries a current. For the magnetic field, equation (4.1) takes the general form:

$$\phi = \frac{\mu a F}{l} \tag{4.20}$$

where ϕ = the magnetic flux (in webers)
μ = the permeability of the medium (in henrys/m)
a = the cross-sectional area of the flux path in the toroid
l = the length of the flux path
F = the magnetic potential difference, or magnetomotive force (in amperes)

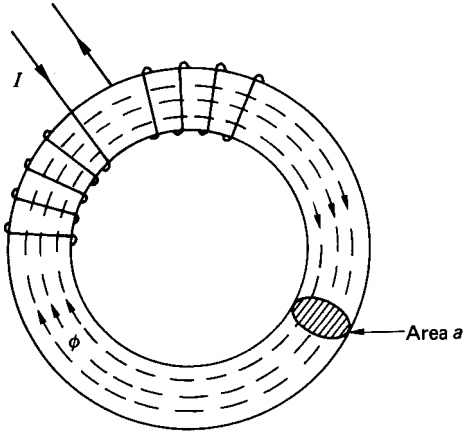


FIGURE 4.4 Toroid.

ELECTRICAL AND ELECTRONIC PRINCIPLES

The magnetomotive force (m.m.f.) is equal to the product of the number of turns on the coil and the current carried, i.e.,

$$F = IN \tag{4.21}$$

Note that the m.m.f. is descriptively expressed in ampere-turns. Since the number of turns is already a dimensionless quantity, the accepted unit of magnetomotive force is the ampere (A).

The group $(\mu a/l)$ is termed the permeance and the inverse of permeance is the reluctance, S. Thus, equation (4.20) may be rewritten as

$$\phi = F/S \tag{4.22}$$

Equation (4.22) represents an electromagnetic version of Ohm's law.

Alternatively, equation (4.20) can be expressed as

$$\frac{\phi}{a} = \mu \frac{F}{1}$$

or

$$B = \mu H \tag{4.23}$$

where $B = \phi/a$ is the magnetic flux density (in webers/m^2, or Tesla (T)) and $H = F/l$ is the magnetic intensity (in A/in).

Permeability

The permeability of free space, μ_0, is numerically equal to $4\pi \times 10^{-7}$. The absolute permeability of other materials is related to the permeability of free space by the relative permeability, i.e.,

$$\mu = \mu_0 \cdot \mu_r \tag{4.24}$$

For air and other nonmagnetic materials, the absolute permeability is the same constant. For magnetic materials, absolute permeability is not a fixed constant but varies nonlinearly with the flux density. The nonlinear variation of permeability is conveniently displayed as a functional plot of magnetic flux density, B, against magnetic intensity, H. Figure 4.5 illustrates a number of B–H curves for some common materials.

Also shown in Figure 4.5 is the B–H curve for air, the only straight-line relationship in the diagram. It is apparent that for an applied magnetic intensity, the magnetic flux developed in a coil with a ferrous core is many times greater than that through a similar coil with an air core. In most practical systems, therefore, a ferrous core is normally used, since it greatly facilitates the establishment of a magnetic flux.

Faraday's Law

Faraday's law states that the e.m.f. induced in a magnetic circuit is equal to the rate of change of flux linkages in the circuit, and is given as

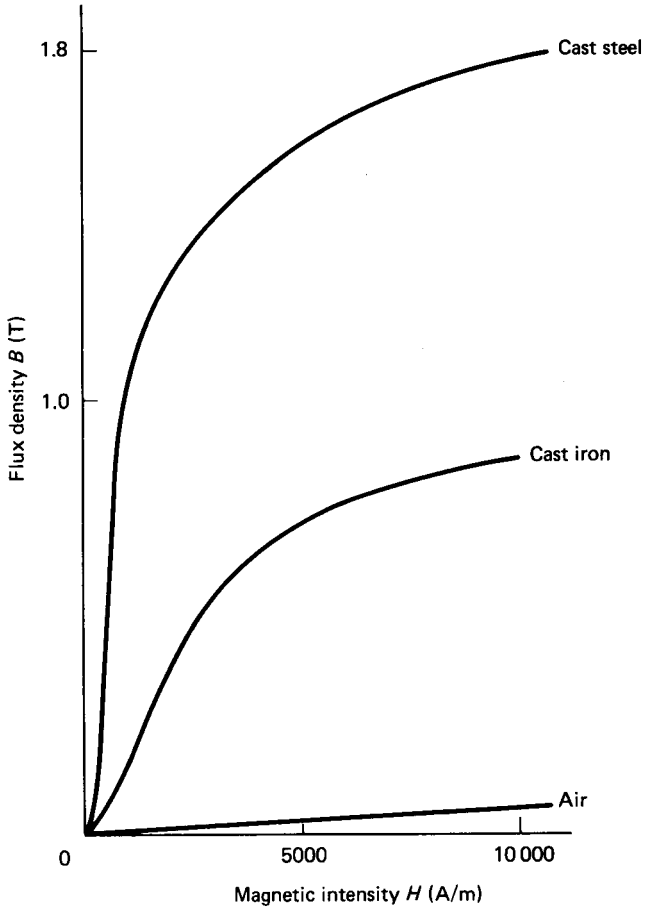


FIGURE 4.5 *B–H* curves for some common materials.

$$e = N(d\phi/dt) \tag{4.25}$$

where e is the instantaneous induced e.m.f.

Equation (4.25) forms the basis of all electrical power generation machines and is a statement of the fact that an electric current can be produced by the movement of magnetic flux relative to a coil. In all rotating electrical generators it is actually the coil that is moved relative to the magnetic field. The net result, however, is exactly the same.

The direction of the induced e.m.f. is always such that it tends to set up a current to oppose the motion (or the change of magnetic flux) which was responsible for inducing the e.m.f. This is essentially a statement of Lenz's law. In many texts, therefore, the right-hand side of equation (4.25) is often shown as a negative quantity.

The motion, or change of flux, is associated with the application of a mechanical force which ultimately provides the torque required to drive the electric generator.

Figure 4.6 shows a single conductor of length l meters, carrying an induced current I and lying perpendicular to a magnetic field of flux density, B T.

The force applied causes the conductor to move through a distance dx meters. The mechanical work done is therefore $F \cdot dx$. The electrical energy produced is given as the product of the power developed and the time duration, i.e., $e \cdot I \cdot dt$. For no external losses, the mechanical work done is converted into electrical energy. Thus,

$$e \cdot I \cdot dt = F \cdot dx \tag{4.26}$$

Using equation (4.25), the induced e.m.f. is equal to the rate of change of flux linkage. For a single conductor, $N = 1$, and in consequence

$$e = (B \cdot l \cdot dx)/dt$$

Therefore,

$$(B \cdot l \cdot dx/dt) \cdot I \cdot dt = F \cdot dx$$

i.e.,

$$F = B \cdot l \cdot I \tag{4.27}$$

Equation (4.27) relates the applied force to the corresponding current generated in a conductor moving through a magnetic field. The equation applies equally to an electric generator or, conversely, to a motor, in which case the electrical power supplied is converted into a mechanical torque via the electromagnetic effect.

Self-Induced e.m.f.

If a current flows through a coil, a magnetic flux links that coil. If, in addition, the current is a time-varying quantity, then there will be a rate of change of flux linkages associated with the circuit. The e.m.f. generated will oppose the change in flux linkages.

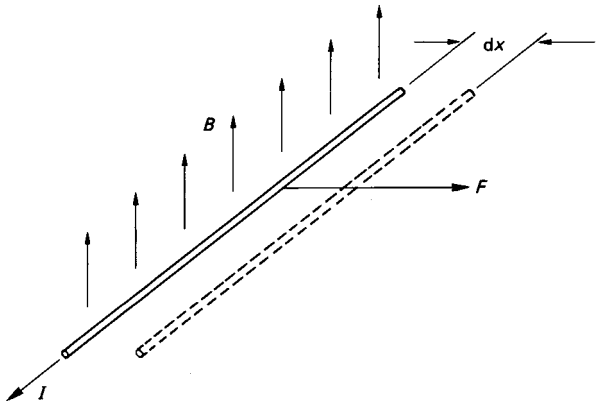


FIGURE 4.6 Generation of e.m.f.

When dealing with electric circuits it is convenient if the voltage across individual elements can be related to the current flowing through them. Figure 4.7 shows a simple circuit comprising a coil having N turns and resistance R, connected in series with a time-varying voltage. The voltage drop across the terminals A and B can be split into two components. First, there is the voltage drop due solely to the resistance of the coiled element. Second, there is a voltage drop, which is a consequence of the self-induced e.m.f. generated through the electromagnetic effect of the coil. Thus,

$$\begin{aligned} v &= vr + vl \\ &= iR + N\frac{d\phi}{dt} \end{aligned} \tag{4.28}$$

From equations (4.20) and (4.21),

$$\phi = \frac{\mu aF}{l} = \frac{\mu aiN}{l}$$

Therefore,

$$\begin{aligned} v &= iR + N\frac{d}{dt}\left[\frac{\mu aiN}{l}\right] \\ &= iR + N^2\left[\frac{\mu a}{l}\right]\frac{di}{dt} \end{aligned} \tag{4.29}$$

The group $N^2(\mu a/l)$ is called the self-inductance of the coil and is denoted by L. The unit of self-inductance is the henry (H). Therefore,

$$v = iR + L\frac{di}{dt} \tag{4.30}$$

By comparing equations (4.28) and (4.30), it is apparent that

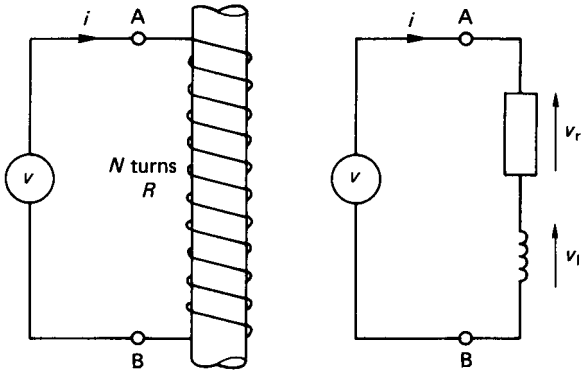


FIGURE 4.7 Self-induced e.m.f.

$$L \frac{di}{dt} = N \frac{d\phi}{dt}$$

Integration then gives

$$L = N\phi / i \tag{4.31}$$

The nature of the self-induced e.m.f. (i.e., Ldi/dt) is such that it will oppose the flow of current when the current is increasing. When the current is decreasing the self-induced e.m.f. will reverse direction and attempt to prevent the current from decreasing.

Energy Stored in an Inductor

$$\text{Instantaneous power} = vi$$

$$\text{Energy stored} = W = \int_0^t vidt$$

$$= \int_0^t L \frac{di}{dt} idt$$

$$= L \int_0^I idi = \frac{1}{2} LI^2 \tag{4.32}$$

Mutual Inductance

Two coils possess mutual inductance if a current in one of the coils produces a magnetic flux that links the other coil. Figure 4.8 shows two such coils sharing a common magnetic flux path in the form of a toroid. The mutual inductance between the two coils is

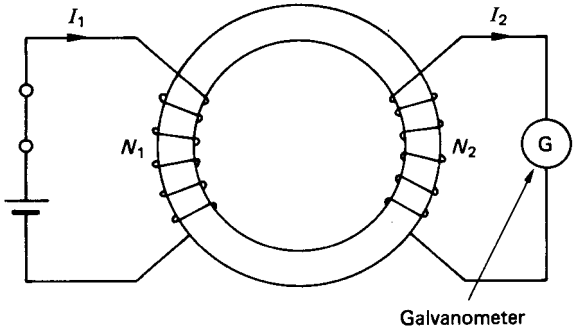


FIGURE 4.8 Mutual inductance.

$$M = \frac{N_2 \phi}{I_1} \text{ (H)} \tag{4.33}$$

where N_2 = the number of turns on coil 2
I_1 = current through coil 1
ϕ = the magnetic flux linking coils 1 and 2

The mutual inductance effect finds great application both to electrical transformers and to rotating electrical machines.

Hysteresis in Magnetic Circuits

Hysteresis can be described with reference to a toroidal coil wound on an iron core (see Figure 4.4). The current supplied to the coil can be imagined to be taken through a cyclic process where it is increased from 0 to $+I$ A, back through 0 to $-I$ A, and again back through 0 to $+I$ A. Measurement of the flux density in the core, as the current varies, results in a B–H curve as depicted in Figure 4.9.

The behavior of the B–H relationship is termed a *hysteresis loop*. This behavior is typical for ferrous cores and is an illustration of the fact that all the electrical energy supplied to magnetize an iron core is not returned when the coil current is reduced to zero. The loss of energy is called *hysteresis loss*, and it is manifested as heat in the iron core.

Hysteresis is characterized by two parameters, the *remanent flux density* (or *remanence*) and the *coercive force*. The remanent flux density is the flux density that remains in the core when the magnetic intensity (i.e., the coil current) has been reduced to zero. The remanent flux density is represented by line OA in Figure 4.9. The coercive force is the magnetic intensity required to reduce the remanent flux density to zero, and is represented by line OC in Figure 4.9.

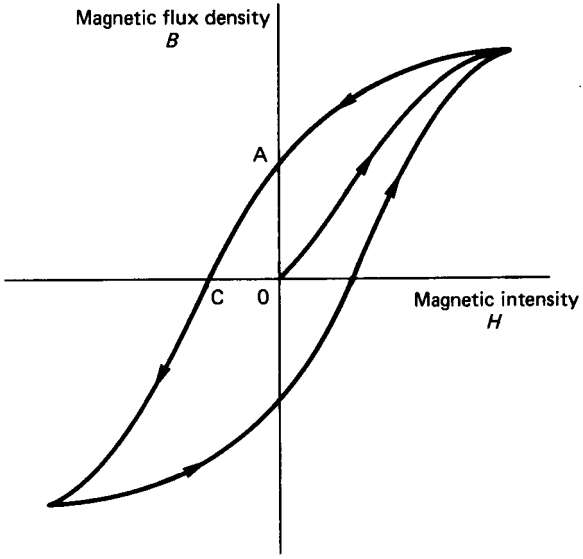


FIGURE 4.9 Hysteresis loop for an iron-cored toroid.

ELECTRICAL AND ELECTRONIC PRINCIPLES

Eddy Current Loss

Faraday's law (equation (4.25)) shows that a time-varying magnetic flux will induce an e.m.f. in a coil. If the ends of the coil are connected and form a closed circuit, then the induced voltage will circulate a current around the closed loop. Consider now an iron core, in which a time-varying magnetic flux exists. Since iron is a conductor, there will be a multitude of arbitrary closed paths within the iron matrix. These closed paths constitute effective conduction routes, and the varying magnetic flux will generate a flow of current round them. The currents are called *eddy currents* and, because of the ohmic resistance of the core, the end result is an energy loss as the eddy currents are dissipated as heat.

Eddy current losses can be greatly reduced by building the iron core in the form of laminations that are insulated from one another. The laminated assembly confines the path lengths for the eddy currents to each respective lamination. The cross-sectional area of the eddy current path is also reduced, and the eddy current loss is approximately proportional to the square of the thickness of the laminations. A practical minimum thickness for any lamination is about 0.4 mm. Increasing manufacturing costs could not justify the use of much thinner laminations.

Kirchhoff's Laws and the Magnetic Circuit

Figure 4.10 shows a magnetic circuit in which a magnetizing coil is wound on one of the limbs and another limb incorporates the usual feature of an air gap. Using the analogy between the magnetic and the conduction circuits, the magnetic circuit can be represented in terms of an energy source (or m.m.f.) and each limb of the magnetic circuit is written in terms of the appropriate reluctance, S. This is illustrated in Figure 4.11.

Given all the relevant dimensions and material properties, the problem is resolved to one of calculating the current required to establish a prescribed magnetic

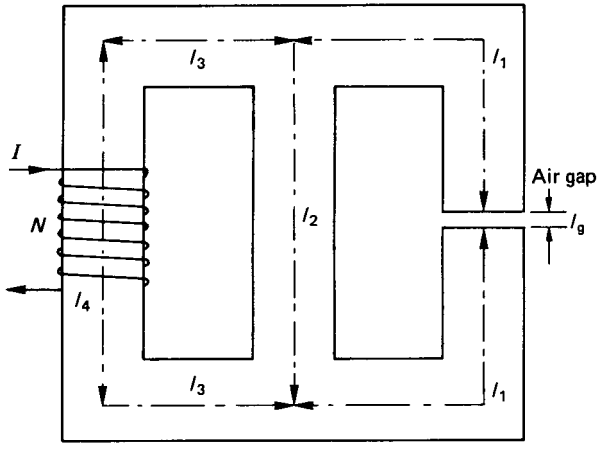


FIGURE 4.10 Magnetic circuit.

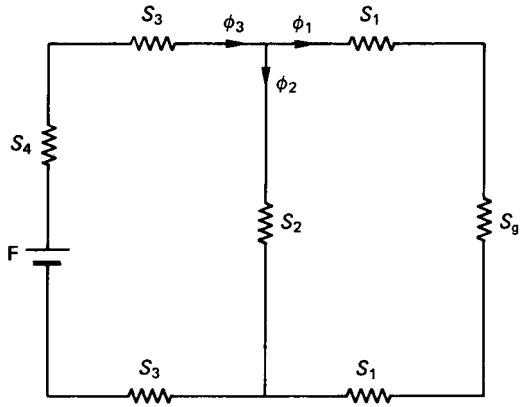


FIGURE 4.11 Representation of the magnetic circuit of Figure 4.10.

flux density in the air gap. The solution invokes the use of Kirchhoff's laws as they apply to magnetic circuits:

First law: At any instant in time, the sum of the fluxes flowing into a node is equal to the sum of the fluxes flowing out.

Second law: Around any closed magnetic circuit the total magnetomotive force is equal to the sum of all the m.m.f.s round the circuit.

Manipulation of equations (4.20) and (4.24) then yields the required solution. The self-inductance of the coil (if required) may be calculated from equation (4.31), or from the definition

$$L = N^2(\mu a/l) = N^2/S \tag{4.34}$$

It has already been shown that the lowest permeability is that of air, and that the m.m.f. required to produce a flux density in air is many times greater than that required to produce the same flux density in a ferrous material. It may reasonably be questioned therefore why air gaps are used at all in iron-cored magnetic circuits. The only function of the air gap is to provide a measure of linearity to the magnetic system such that the inductance remains reasonably constant over a range of operating currents.

Alternating Quantities

If an electrical quantity varies with time but does not change its polarity, it is said to be a direct current (d.c.) quantity. If the quantity alternates between a positive and a negative polarity, then it is classified as an alternating current (a.c.) quantity.

The period, T, is the time interval over which one complete cycle of the alternating quantity varies. The inverse of the period is the frequency, f, in hertz (Hz). Circular frequency, ω, in radians per second is also commonly used.

ELECTRICAL AND ELECTRONIC PRINCIPLES

Instantaneous values of the quantities encountered in electrical systems are usually denoted by lower-case letters. Since the instantaneous values are difficult to measure and quantify, a.c. quantities are usually expressed as root mean square (r.m.s.) values. For a periodically varying a.c. quantity, the r.m.s. value is given by

$$\text{r.m.s.} = \left[1/t \int_0^t (\text{quantity})^2 dt \right]^{1/2} \tag{4.35}$$

Many electrical quantities vary in a sinusoidal manner, and it can easily be shown that the r.m.s. value is simply related to the maximum value by

$$\text{r.m.s.} = \max/(\sqrt{2}) = 0.707 \text{ max} \tag{4.36}$$

Relationship between Voltage and Current in *R, L* and *C* Elements

For a simple resistive element, current is directly proportional to voltage. The current waveform will therefore be essentially the same shape as the voltage waveform.

For an inductive coil with negligible resistance, the relation between voltage and current is given by equation (4.30), i.e.,

$$v = L \frac{di}{dt}$$

Thus,

$$i = \frac{1}{L} \int v dt \tag{4.37}$$

The relation between voltage and current for a capacitive element is given by equation (4.18), i.e.,

$$i = C \frac{dv}{dt}$$

For the capacitive element it can be seen that a current will flow only when the voltage is changing. No current can flow if the voltage is constant since dv/dt will then be equal to zero. The capacitor then, will block any steady d.c. input and indeed is sometimes used for this express purpose.

RL and *RC* Circuits under Transient Switching Conditions

Circuits involving a single resitor, capacitor, or inductance are rare. It is more usual to find circuits involving some or other combination of these elements in both d.c. and ac. applications. Figure 4.12 illustrates two simple RL and RC circuits.

RL Circuit. With the switch open there is no flow of current in the circuit. At the instant of switching, the current will rise and eventually reach a steady-state value of V_s/R. The transient period is governed by equation (4.30), which represents a

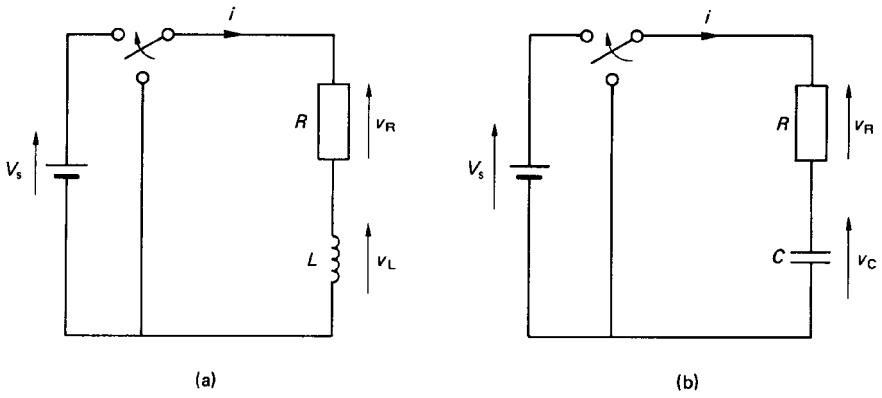


FIGURE 4.12 Simple *RL* and *RC* circuits under transient switching conditions.

first-order, ordinary differential equation in *i*. The solution of equation (4.30) involves separating the variables to allow integration. The general solution is

$$i = I[1 - \exp(-Rt/L)] \tag{4.38}$$

Equation (4.38) shows that the current growth in the circuit will rise exponentially to reach a steady state value as time, *t*, increases. It may also be shown that

$$v = L\frac{di}{dt} = V_s \cdot \exp(-Rt/L) \tag{4.39}$$

The time constant, *T*, for the *RL* circuit is *L/R*.

RC Circuit. In Figure 4.12(b), with the switch open there is zero potential difference across the capacitor. On closing the switch the voltage across the capacitor will rise in an asympototic manner, reaching a steady-state value of V_s. From Kirchhoff's second law:

$$V_s = iR + v_c \tag{4.40}$$

where v_c is the instantaneous voltage across the capacitor.

From equation (4.18) we can write:

$$V_s = RC\frac{dv_c}{dt} + v_c \tag{4.41}$$

Equation (4.41) shows that the instantaneous voltage across the capacitor also conforms to a first-order system. The solution gives

$$v_c = V_s[1 - \exp(-t/RC)] \tag{4.42}$$

The time constant for the simple *RC* circuit is

ELECTRICAL AND ELECTRONIC PRINCIPLES

$$T = RC \tag{4.43}$$

Both the simple *RL* and *RC* circuits are first-order systems with a generalized form of transient behavior. In circuits containing both inductive and capacitive elements the transient behavior is governed by a second-order ordinary differential equation. The transient behavior of these circuits is, however, less important than their response to sinusoidally varying inputs.

Steady-State Alternating Currents

In most practical applications in electrical engineering the voltages and currents are sinusoidal. A simple series *RLC* circuit is depicted in Figure 4.13. Since the current is common to each of the circuit elements, it is used for reference purposes. The instantaneous current is defined as

$$i = I_m \sin(\omega t) \tag{4.44}$$

where I_m is the maximum (or peak) value of the current and ω is the angular, or circular frequency in radians/is.

The voltage drop across the resistor is

$$v_R = iR = I_m R \sin(\omega t) \tag{4.45}$$

Equation (4.45) indicates that the voltage drop across the resistor is in phase with the current. In other words, v_R reaches a positive maximum at the same instant as the current, i. The voltage drop across the inductor is

$$\begin{aligned} v_L = L\frac{di}{dt} &= L\frac{d}{dt}[I_m \sin(\omega t)] \\ &= LI_m\omega \cos(\omega t) \\ &= \omega L I_m \sin(\omega t + 90) \end{aligned} \tag{4.46}$$

The relationship between current and voltage drop across the inductor is shown in Figure 4.14.

It can be seen that there is a phase difference between the voltage drop and the current through the inductor. In fact, v_L reaches a positive maximum "before" i

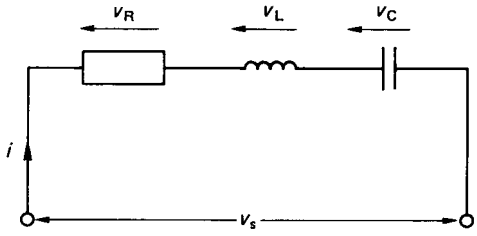


FIGURE 4.13 Series *RLC* circuit.

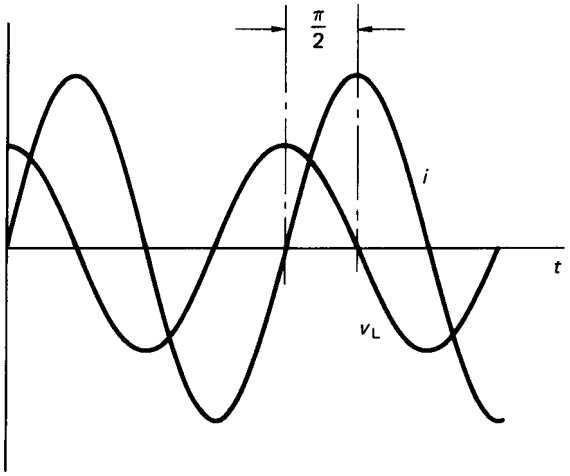


FIGURE 4.14 Current and voltage drop across an inductor.

and v_L is said to "lead" the current by 90°. For the capacitor, the voltage drop is given by

$$\begin{aligned} v_c &= 1/C \int idt = 1/C \int I_m \sin(\omega t)dt \\ &= -\frac{I_m}{\omega C}\cos(\omega t) \\ &= +\frac{I_m}{\omega C}\sin(\omega t - 90) \end{aligned} \tag{4.47}$$

The voltage drop across the capacitor therefore reaches its positive maximum after that of *i*. In general terminology, v_c "lags" i by 90°.

Equations (4.45)–(4.47) are all of similar form in that they can be expressed as

$$\text{Voltage drop} = \text{constant} \times \text{current}$$

In equation (4.46) the constant ωL is termed the *inductive reactance* and is denoted by X_L. In equation (4.47) the constant $(1/\omega C)$ is the *capacitive reactance*, which is denoted as X_c. Both of these reactances have units of ohms.

The total voltage drop across the three circuit elements is

$$\begin{aligned} v &= v_R + v_L + v_C \\ &= iR + L\frac{di}{dt} + \frac{1}{C}\int idt \end{aligned}$$

Therefore,

$$v = I_m R \sin(\omega t) + \omega L I_m \sin(\omega t + 90°) + \frac{I_m}{\omega C} \sin(\omega t - 90°) \qquad (4.48)$$

While equation (4.48) defines the total instantaneous voltage drop in mathematical terms, it is rather cumbersome to deal with. To simplify the analysis, the addition of a.c. voltages is conveniently performed using a graphical technique involving phasors.

Phasor Diagrams

Any sinusoidally varying quantity can be represented as a phasor, which is a vector quantity. The length of the phasor is proportional to the magnitude of the product of the reactance and the maximum current. The direction of the phasor is determined by the phase angle and its relation to some common reference.

For the *RLC* circuit of Figure 4.13 the voltage drop across the inductance may be arbitrarily assumed greater than that across the capacitor. The total voltage drop in the circuit is then given as the phasor addition of the three individual potential difference components. This is illustrated in Figure 4.15.

The vector addition of the three phasors shows that the source voltage leads the current by an angle of ϕ degrees, i.e.,

$$\overline{V} = V_m \sin(\omega t + \phi) \qquad (4.49)$$

The circuit is therefore essentially inductive and, using the standard notation, the total phasor voltage is designated by a capital letter with an overbar.

Complex Notation

Since inductive and capacitive elements in a.c. circuits involve a phase shift of $+90°$ and $-90°$, respectively, the complex number notation is used extensively to manipulate phasor quantities. The complex operator j (defined as $\sqrt{-1}$) is a unit

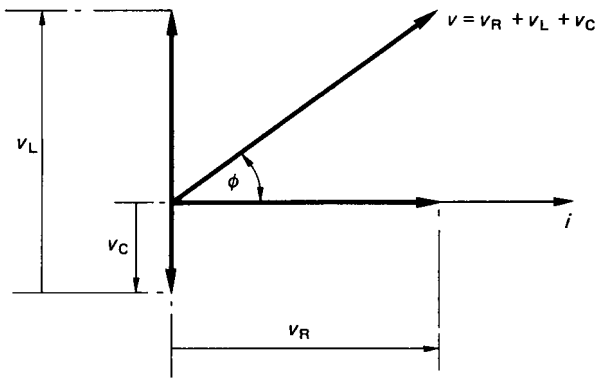


FIGURE 4.15 Phasor diagram for series *RLC* circuit.

operator that, when multiplying a phasor, shifts it by 90° in an anti-clockwise direction. Thus, for the series *RLC* circuit,

$$\bar{V}_R = \bar{I}R,\ \bar{V}_L = j\bar{I}X_L \text{ and } \bar{V}_c = -j\bar{I}X_c$$

where $\bar{I}$ can be taken as the r.m.s. value of the current. The voltage drop across the complete circuit can then be written as

$$\begin{aligned}\bar{V} &= \bar{I}R + j\bar{I}X_L - j\bar{I}X_C \\ &= \bar{I}[R + j(W_L - X_C)] \end{aligned} \tag{4.50}$$

The term in the square brackets is called the *impedance* of the circuit and is denoted by $\bar{Z}$. Thus,

$$\bar{V} = \bar{I} \cdot \bar{Z} \tag{4.51}$$

Equation (4.51) represents Ohm's law for a.c. circuits. The phase angle between the source voltage and the current is

$$\phi = \tan^{-1}[(X_L - X_C)/R] \tag{4.52}$$

The Parallel *RLC* Circuit

A parallel *RLC* circuit is shown in Figure 4.16. The applied voltage is common to all the circuit elements and it is therefore chosen as the reference.

Using Ohm's law, the currents through each of the circuit elements are

$$\bar{I}_R = \bar{V}/R;\ \bar{I}_L = \bar{V}/X_L;\ \bar{I}_C = \bar{V}/X_C$$

Applying Kirchhoff's first law, the total current is the vector sum of the three currents I_R, I_L, and I_C. The magnitude and phase of the total current may subse-

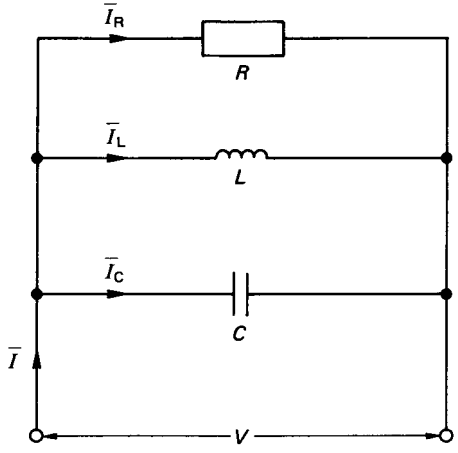


FIGURE 4.16 Parallel *RLC* circuit.

quently be determined from a phasor diagram, or calculated using the complex number notation. Using the latter and noting that the current through an inductor lags the voltage while the current through a capacitor leads the voltage, it may be shown that

$$\begin{aligned}\bar{I} &= \bar{I}_R + \bar{I}_L + \bar{I}_C \\ &= \bar{V}\left[\frac{1}{R} - j\left(\frac{1}{X_C} - \frac{1}{X_L}\right)\right] \end{aligned} \tag{4.53}$$

and the phase angle,

$$\phi = \tan^{-1}\left[\frac{R(X_L - X_C)}{X_L \cdot X_C}\right] \tag{4.54}$$

Power and Power Factor in a.c. Circuits

Denoting the phase angle between the voltage and the current as ϕ, it may be shown (Bell and Whitehead 1987) that the average power is

$$P_{\text{av}} = \frac{V_m}{\sqrt{2}}\frac{I_m}{\sqrt{2}}\cos(\phi).$$

In terms of r.m.s. values:

$$P_{\text{av}} = VI\cos(\phi) \tag{4.55}$$

where $\cos(\phi)$ is called the *power factor*.

Power factor is an important parameter when dealing with electrical transformers and generators. All such machines are rated in terms of kilovolt amperes (kVA), which is a measure of the current-carrying capacity for a given applied voltage. The power that can be drawn depends both on the kVA rating and the power factor of the load. Figure 4.17 shows the relationship between kVA, kilowatts (kW), and

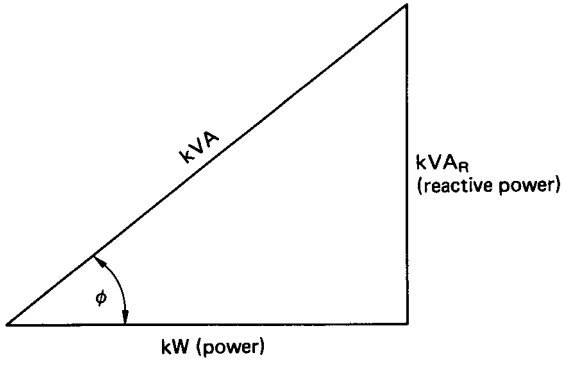


FIGURE 4.17 Power triangle.

power factor, sometimes referred to as the power triangle. It can readily be seen that

$$\text{kW} = \text{kVA}\cos(\phi) \tag{4.56}$$

and

$$\text{kVA}_R = \text{kVA}\sin(\phi) \tag{4.57}$$

where kVA_R is the reactive power. Thus, knowing the kVA rating and the power factor of a number of various loads, the power requirements from a common supply may be determined.

When quoting power factors in practical applications it is usual to state the phase of the current with respect to the voltage. For an inductive load the current lags the voltage and the power factor is said to be lagging. For a predominantly capacitive load the current leads the voltage and the power factor is leading.

If the power is supplied from, say, an alternator rated at 400 V and 1000 A, then these are the highest voltage and current that the machine can tolerate without overheating. The phase difference between the voltage and current is entirely dependent upon the load. Thus, if the power factor of the load is unity, then the 400 kVA alternator can supply 400 kW of power to the load. Neglecting losses, the prime mover that drives the alternator must also be capable of supplying 400 kW. If, on the other hand, the power factor of the load is 0.5, then the power supplied will only be 200 kW. This means that although the generator will be operating at its rated kVA, the prime mover which drives the generator will be operating at only half of its capacity.

An alternative way of looking at this phenomenon is to consider a load of, say, 100 kW, with a lagging power factor of 0.75. If the supply voltage is 50 V, then the required current is 2.67 A. If, however, the power factor of the load were to be increased to unity, then the required current would be reduced to 2 A. This means that the conducting cables, in supplying a reduced current, may have a correspondingly reduced cross-sectional area.

In general, the size of an electrical system including transmission lines, switchgear, and transformers is dependent upon the size of the current. It is economically viable, therefore, to ensure that the current is minimized. As a further incentive to industrial consumers, the electricity supply authorities normally operate a two-part tariff system. This consists of a fixed rate depending on the kVA rating of the maximum demand and a running charge per unit kilowatts consumed per hour.

For these reasons it is advantageous to try to increase the power factor such that it is close to (but not quite) unity. A unity power factor is in fact avoided, because it gives rise to a condition of resonance. In practice, capacitors connected in parallel are often used to improve the power factor of predominantly inductive loads such as electric motors. For large-scale power systems, a separate phase advance plant is used.

Frequency Response of Circuits

The frequency response of a circuit is usually presented as a plot of the ratio of output over input against the frequency as base. The ratio plotted could be one of voltages, currents, or powers. Since the range of frequencies involved may be quite large, a logarithmic scale is normally employed. A logarithmic scale is also usually adopted for the vertical axis and the output/input ratio quoted in decibels (dB), i.e.,

ELECTRICAL AND ELECTRONIC PRINCIPLES

$$\text{Voltage ratio in dB} = 20 \log_{10} \left[\frac{V_{\text{out}}}{V_{\text{in}}} \right] \tag{4.58}$$

Considering the series *RLC* circuit shown in Figure 4.13 and taking the voltage across the resistor as an output,

$$V_{\text{out}} = IR$$

$$V_{\text{in}} = I[R + j(\omega L - 1/\omega C)]$$

Therefore,

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{R}{R + j(\omega L - 1/(\omega C)}$$

Using the complex conjugate and calculating the modulus of the voltage ratio gives

$$\left| \frac{V_{\text{out}}}{V_{\text{in}}} \right| = \frac{R}{[R^2 + (\omega L - 1/\omega C)^2]^{1/2}} \tag{4.59}$$

The phase angle

$$\phi = -\tan^{-1} \left[\frac{(\omega L - 1/\omega C)}{R} \right] \tag{4.60}$$

The voltage ratio will have a maximum value of unity when the frequency

$$\omega = \frac{1}{\sqrt{LC}} \tag{4.61}$$

Equation (4.61) defines the resonance condition at which the inductive and capacitive reactances are equal and self-canceling. The resonant frequency is usually denoted ω_0 and is the frequency at which the power transferred through the circuit is maximum. At any other frequency above or below ω_0 the power transferred is reduced.

The impedance of the circuit is given by

$$\overline{Z} = R + j(X_L - X_C) \tag{4.62}$$

At the resonant frequency the total reactance is zero and the circuit behaves as if only the resistive element were present.

The general variation of the voltage ratio (or amplitude ratio) and phase angle with frequency is illustrated in Figure 4.18. Also shown in the figure are the two frequencies, ω_1 and ω_2, at which the amplitude ratio is -3 dB. The -3 dB amplitude ratio is chosen because it corresponds to a halving in the power transmitted.

The bandwidth is the frequency range between ω_1 and ω_2. A quality parameter, used with respect to resonant circuits, is the so-called Q factor, which is defined as the ratio of the resonant frequency to the bandwidth.

Semiconductors

The materials commonly used for semiconductors are germanium and silicon. In recent times silicon has all but replaced germanium as a semiconductor material.

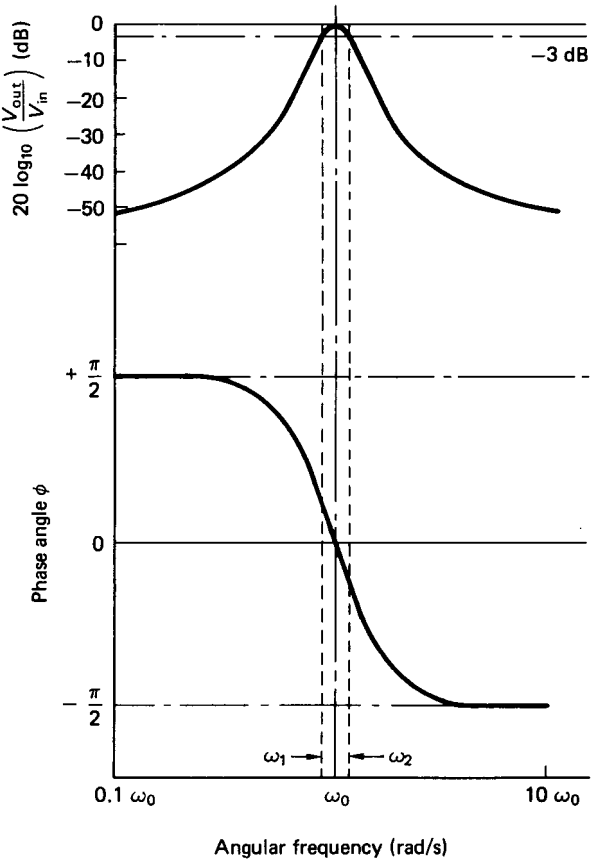


FIGURE 4.18 Voltage ratio and phase angle versus frequency (series *RLC*).

These materials have a crystalline structure such that each atom is surrounded by equally spaced neighbors. The basic structure can be visualized as a two-dimensional grid where the node points represent the central nucleus and the inner shell electrons, while the connecting lines of the grid represent the four valence electrons associated with each nucleus. This grid concept is adequate to describe an intrinsic (or "pure") semiconductor.

At absolute zero temperature the crystalline structure is perfect and the electrons are all held in valence bonds. Since there are no current carriers available, the crystal behaves as a perfect insulator. As the temperature rises above absolute zero, an increasing number of valence bonds are broken, releasing pairs of free electrons and their associated "holes". In the absence of an applied field the free electrons move randomly in all directions. When an electric field is applied, the electrons drift in a preferential direction to oppose the field and a net flow of current is established.

The covalent bond, with a missing electron, has a large affinity for electrons such that an electron from a neighboring bond may easily be captured. This will

leave the neighboring atom depleted of electrons, and the flow of electrons is generally associated with a counterflow of so-called holes. The mobile hole, to all intents and purposes, is essentially a simple positive charge.

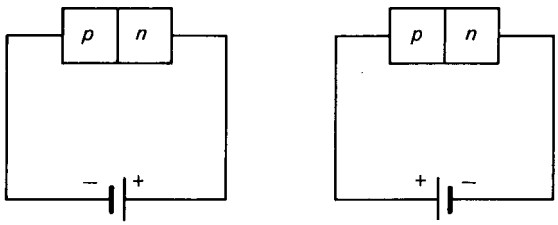
Doped Semiconductors

Doped semiconductors are those in which an impurity has been introduced into a very pure intrinsic silicon. The nature of the impurity depends on the type of semiconductor required:

1. *n-type:* Impurities with five valence electrons can be added to produce a negative type of semiconductor. These impurities are referred to as donors, since the additional electron is very easily freed within the matrix. In the *n*-type semiconductor the free electrons are the dominant current carriers.
2. *p-type:* The *p*-type semiconductor is one in which the added impurities have only three valence electrons. Such impurities are called acceptors, and they produce a positive type of semiconductor within which hole conduction is the dominant current carrier.

pn Junction Diode

A *pn* junction is formed by doping a crystal in such a way that the semiconductor changes from *p*- to *n*-type over a very short length (typically 10^{-6} in). The transition zone from *p*- to *n*-type is called the *carrier depletion layer* and, due to the high concentration of holes on one side and electrons on the other, a potential difference exists across this layer. The diffusion of holes from *p* to *n* and electrons from *n* to *p* is the majority carrier movement, called the *diffusion current*. The drift of electrons from *p* to *n* and holes from *n* to *p* is the minority carrier movement, referred to as the *drift current*. When there is no externally applied potential difference, the diffusion current and the drift current are balanced in equilibrium. If an electric field is applied across the device, then two situations can exist, as illustrated in Figure 4.19. Figure 4.19(a) shows the reverse-bias mode, in which the potential barrier is increased. The diffusion current is reduced while the drift current is barely altered. Overall, the current is negative and very small. When forward bias is applied, as in Figure 4.19(b), the potential barrier is reduced and a large diffusion current flows. Overall, the current is positive and large. These general characteristics



(a) Reverse bias (b) Forward bias

FIGURE 4.19 *pn* junction with applied potential difference.

are the basis of a semiconductor diode that displays the typical current/voltage relationship depicted in Figure 4.20.

This figure shows clearly that a very high impedance is presented by the diode to an applied voltage of reverse polarity. A low impedance is presented to a forward polarity voltage. In simple terms, the diode accommodates a forward flow of current but greatly inhibits a reverse flow. The diode may be likened therefore to a switch that is activated "on" for forward voltages and "off" for reverse voltages. The reverse saturation current, I_s, is typically of the order of a few nano-amperes and can sensibly be regarded as zero.

The general characteristic also shows that the reverse has a critical limiting value at which a breakdown occurs. Depending upon the diode construction, the breakdown (or Zener voltage) may range from as low as one volt to as much as several thousand volts. Up to the breakdown voltage, the reverse saturation current is independent of the reverse voltage.

Since the current/voltage relationship for a diode is a nonlinear exponential function, the analysis of circuits involving diodes can become complicated. A simple awareness of the diode's practical function as a rectifier is perhaps more important than a proficiency in analyzing circuits involving diode elements.

A.C. Rectification

Figure 4.21 shows an a.c. circuit with a diode in series with a load resistor. When the diode is forward biased a current will flow in the direction indicated by the

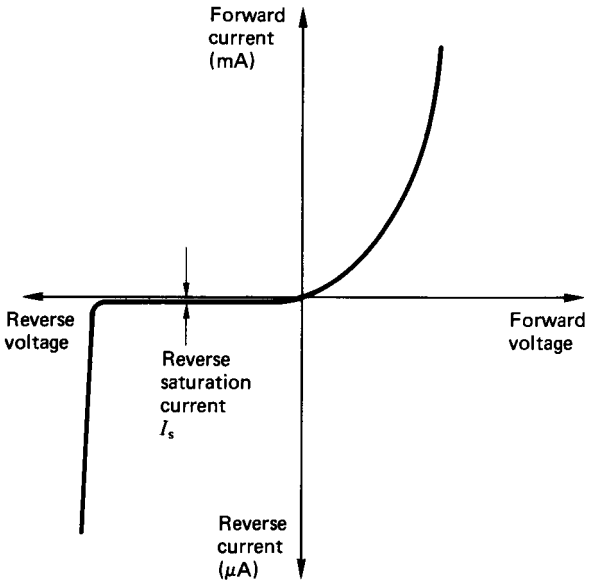


FIGURE 4.20 Current/voltage relationship for a *pn* semiconductor diode.

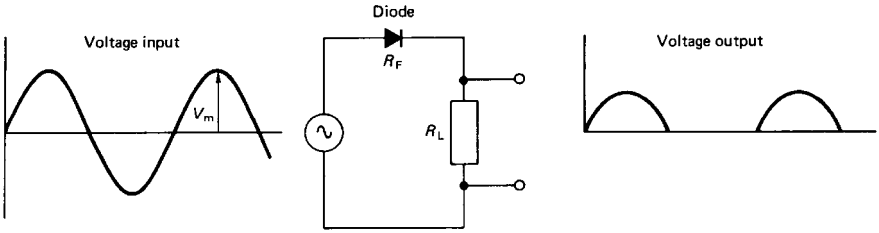


FIGURE 4.21 Half-wave rectification circuit.

arrowhead. No current can flow when the diode is reverse biased, provided that the applied voltage does not exceed the breakdown value. The resultant current waveform through the resistor, for a sinusoidal voltage input, will therefore consist of positive only half-sine waves. Since the output waveform is positive only, then it is, by definition, a d.c. voltage. It can be shown that the r.m.s. voltage across the resistor is

$$V = \frac{V_m}{2} \cdot \frac{R_L}{R_L + R_F} \tag{4.63}$$

where R_L is the load resistance, R_F is the diode forward resistance, and V_m is the peak input voltage. Determination of R_F is problematic, however, and models of varying complexity are used to simulate the diode in the circuit.

The single-diode circuit results in half-wave rectification. To obtain full-wave rectification a diode bridge circuit can be used. The diode bridge is shown in Figure 4.22. When A is positive with respect to B then diodes D_1 and D_3 are conducting. When B is positive with respect to A then diodes D_2 and D_4 are conducting. The circuit arrangement ensures that the current, which consists of a continuous series of positive half-sine waves, is always in the same direction through the load R_L.

With full-wave rectification there are twice as many half-sine pulses through the load than there are with half-wave rectification. In addition, there are always two diodes effectively in series with the load. The resultant r.m.s. voltage across the load resistor for the full-wave diode bridge rectification circuit is

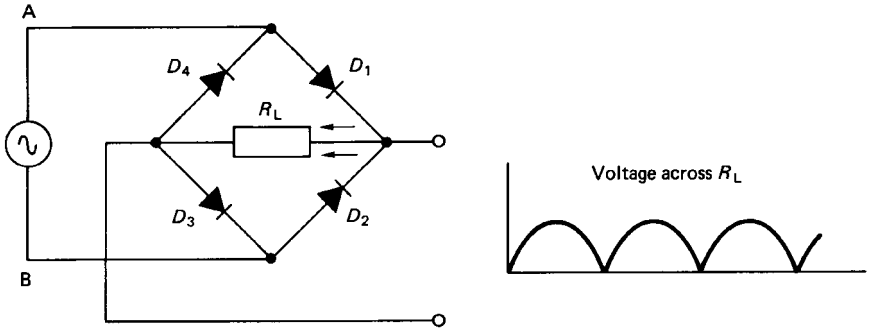


FIGURE 4.22 Full-wave rectification with a diode bridge.

$$V = \frac{V_m}{\sqrt{2}} \cdot \frac{R_L}{(R_L + 2R_F)} \tag{4.64}$$

The peak inverse voltage (PIV) is defined as the maximum reverse-biased voltage appearing across a diode. When used as a rectifier the diodes must have a sufficiently high reverse voltage rating in excess to the peak inverse voltage that the circuit can generate. For both the half- and the full-wave rectification circuits considered, the peak inverse voltage is equivalent to the maximum supply voltage, V_m. Additional manufacturers' diode specifications would normally include the maximum power rating and the maximum allowable forward current.

The Zener Diode

The diode breakdown effect is also used in a variety of circuits to provide a stabilized reference voltage. Special diodes that are designed to operate continuously in the reverse bias mode are called Zener diodes. These diodes are manufactured with a range of breakdown voltages from 3 to 20 V. Figure 4.23 shows a Zener diode being used in a circuit to give a stable voltage that is essentially independent of the current flowing through the device. The series resistor in the circuit is included to limit the reverse current through the diode to a safe value.

4.2 ELECTRICAL MACHINES

The function of a rotating electrical machine is to convert mechanical power into electrical power, or vice versa. The conversion from mechanical to electrical power is made with a a generator and the conversion of electrical to mechanical power with a motor. Electrical machines may be further subdivided into a.c. or d.c. machines. The major part of all electrical energy generated in the world today is produced by a particular type of a.c. machine called an alternator. The applications of electric motors are no less substantial, and they are used in a great variety of

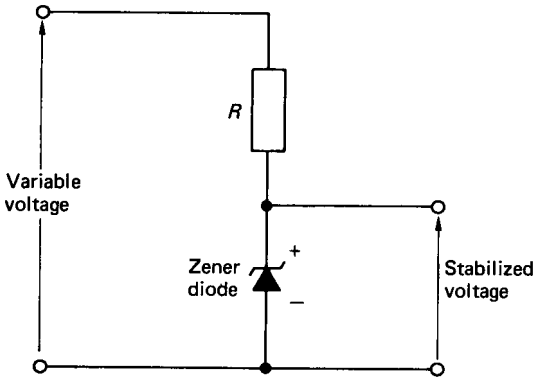


FIGURE 4.23 Zener diode as a reference voltage source.

industrial drives. It is usually the mechanical features of a particular application that determine the type of electric motor to be employed, and the torque–speed characteristics of the machine are therefore very important.

The d.c. Generator

All conventional electrical machines consist of a stationary element and a rotating element that are separated by an air gap. In d.c. machines—generator or motor—the stationary element consists of salient poles that are constructed as laminated assemblies with coils wound round them to produce a magnetic field. The function of the laminations is to reduce the losses incurred by eddy currents. The rotating element is traditionally called the armature, and this consists of a series of coils located between slots around the periphery of the armature. The armature is also fabricated in laminations that are usually keyed onto a locating shaft. A very simple form of d.c. generator is illustrated in Figure 4.24.

In the figure the single coil is rotated at constant speed between the opposite poles, north and south, of a simple magnet. From Faraday's law (equation (4.25)) the voltage generated in the coil is equal to the rate of change of flux linkages. When the coil lies in the horizontal plane there is maximum flux linking the coil but a minimum rate of change of flux linkages. On the other hand, when the coil lies in the vertical plane there is zero flux linking the coil but the rate of change of flux linkages is a maximum. The resultant variation in generated voltage in the coil, as it moves through one revolution, is shown in Figure 4.24(b). It is apparent that the generated voltage is alternating with positive and negative half-cycles. To change the a.c. output voltage into a d.c. voltage, a simple yet effective mechanical device called a commutator is used. The commutator (Figure 4.25) incorporates

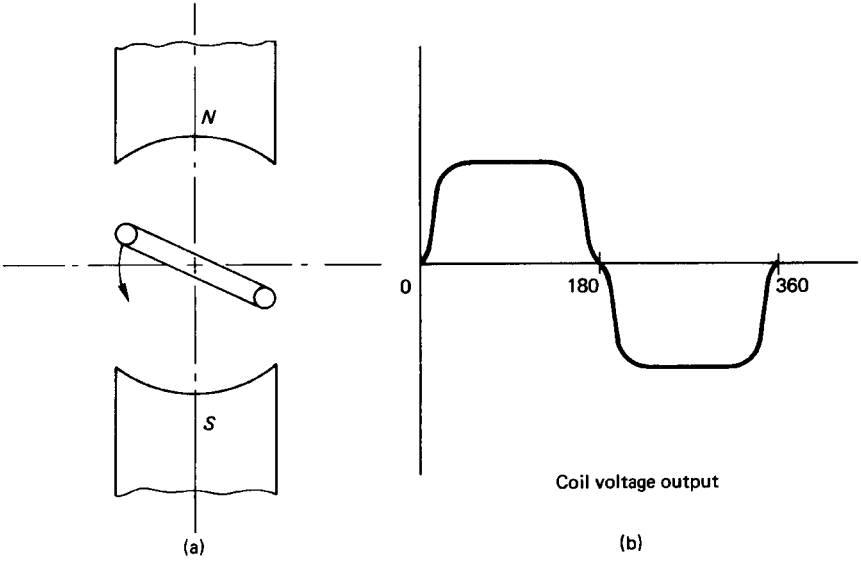


FIGURE 4.24 Single-coil, two-pole d.c. generator.

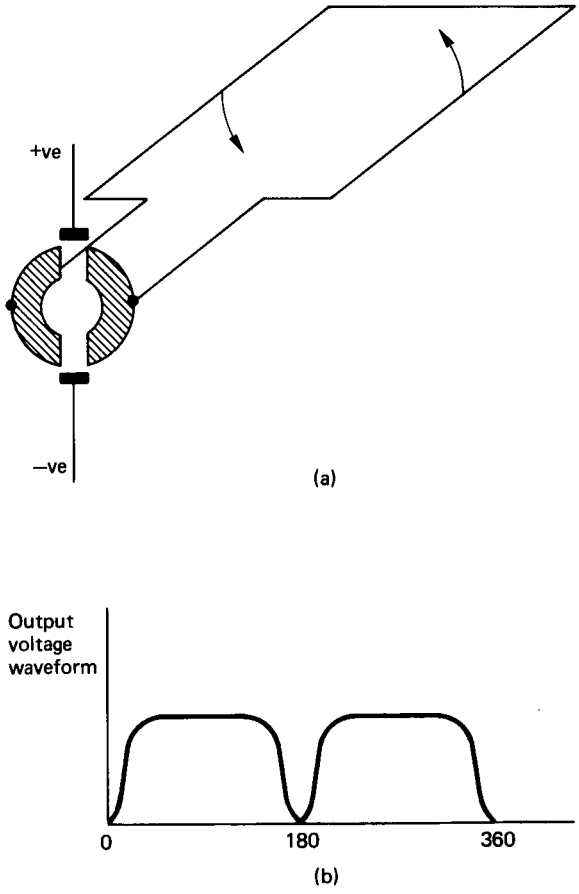


FIGURE 4.25 Commutator connections to armature.

brass segments separated by insulating mica strips. External connection to the armature coil is made by stationary carbon brushes that make sliding contact with the commutator. Referring to Figures 4.24(a) and 4.25(a), as the coil rotates from the horizontal plane through 180° the right-hand side of the coil is under the north pole and is connected via the commutator to the upper brush. Meanwhile, the left-hand side of the coil is under the south pole and is connected to the lower brush. A further 180° of rotation effectively switches the coil sides to the opposite brushes. In this manner the coil side passing the north pole is always connected to the positive upper brush, while the coil side passing the south pole is always connected to the negative lower brush. The resultant output voltage waveform is shown in Figure 4.25(b).

If two coils, physically displaced by 90°, are now used, the output brush voltage becomes virtually constant, as shown in Figure 4.26. With the introduction of a second coil, the commutator must have four separate segments. In a typical d.c.

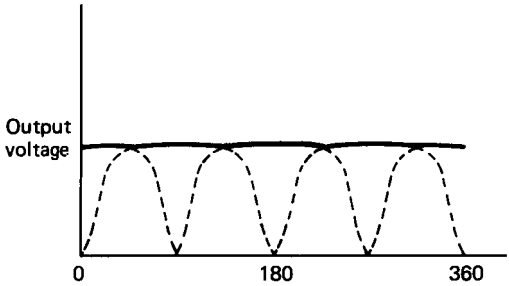


FIGURE 4.26 Two-coil, two-pole d.c. generator output voltage.

machine there may be as many as 36 coils, which would require a 72-segment commutator.

The simple d.c. generator of Figure 4.24 can be improved in perhaps three obvious ways. First, the number of coils can be increased, second, the number of turns on each coil can be increased and third, there is no reason why another pair of poles cannot be introduced. A typical d.c. machine would therefore normally incorporate four poles, wired in such a way that each consecutive pole has the opposite magnetic polarity to each of its neighboring poles. If the e.m.f.'s generated in the armature coils are to assist each other, then, while one side of the coil is moving under a north pole, the other side must be moving under a south pole. With a two-pole machine the armature coils must be wound such that one side of the coil is diametrically opposite the other. With a four-pole machine the armature coils can be wound with one side of the coil physically displaced 90° from the other. The size of the machine will generally dictate how many coils and the number of turns on each coil that can be used.

Armature e.m.f. If a conductor cuts flux, then a voltage of 1 V will be induced in the conductor if the flux is cut at the rate of 1 Wb/s. Denoting the flux per pole as Φ and the speed in revolutions per second as N, for the single-turn coil and two-pole generator of Figure 4.24(a) the e.m.f. induced in the coil is

$$E_{\text{coil}} = \frac{\text{Flux per pole}}{\text{Time for half revolution}} = \frac{\Phi}{1/(2N)} = 2N\Phi$$

For a machine having Z_s armature conductors connected in series, i.e., $Z_s/2$ turns, and $2p$ magnetic poles, the total induced e.m.f. is

$$E = 2N\Phi \frac{Z_s}{2} 2p = 2N\Phi Z_s p \text{ volts} \tag{4.65}$$

Z_s depends on the type of armature winding, and the two main types are *lap-wound* and *wave-wound*.

The lap winding is characterized by the fact that the number of parallel paths through the winding is equal to the number of poles. In the alternative wave winding the number of parallel paths through the winding is always equal to two. If Z denotes the total number of armature conductors then for the lap winding

$$Z_s = \frac{Z}{\text{Number of parallel paths}} = \frac{Z}{\text{Number of poles}} = \frac{Z}{2p} \tag{4.66}$$

and for the wave winding

$$Z_s = \frac{Z}{\text{Number of parallel paths}} = \frac{Z}{2} \tag{4.67}$$

Lap windings are generally used in low-voltage, heavy-current machines and wave winding in all other cases.

Armature Torque. The force on a current-carrying conductor is given by equation (4.27), i.e.,

$$F = BlI$$

The torque on one armature conductor is therefore

$$T = Fr = B_{\text{av}} l I_a r \tag{4.68}$$

where r = the radius of the armature conductor about the center of rotation
I_a = the current flowing in the armature conductor
l = the axial length of the conductor
B_{av} = is the average flux density under a pole

Note that

$$B_{\text{av}} = \frac{\Phi}{(2\pi r l)/2p}$$

The resultant torque per conductor is

$$T = \frac{\Phi 2 p l I_a r}{2\pi r l} = \frac{\Phi p I_a}{\pi}$$

For Z_s armature conductors connected in series the total torque on the armature is

$$T = \frac{\Phi p I_a Z_s}{\pi} \text{ newton-meters} \tag{4.69}$$

Terminal Voltage. Denoting the terminal voltage by V, the induced e.m.f. by E and the armature resistance by R_a,

$$V = E - I_a R_a \quad \text{(for a generator)} \tag{4.70}$$

$$V = E + I_a R_a \quad \text{(for a motor)} \tag{4.71}$$

For the motor, the induced e.m.f. is often called the back e.m.f.

Methods of Connection

The methods of connecting the field and armature windings may be grouped as follows:

1. *Separately excited:* where the field winding is connected to a source of supply independently of the armature supply
2. *Self-excited:*
   - ***a.*** *Shunt-wound:* where the field winding is connected across the armature terminals
   - ***b.*** *Series-wound:* where the field winding is connected in series with the armature winding
   - ***c.*** *Compound-wound:* which is a combination of shut and series windings

The four alternative methods of connection are illustrated in Figure 4.27.

The Separately Excited Generator

Consider the separately excited generator, shown in Figure 4.27(a), running at a constant rated speed with no load across the output. It is assumed that initially the

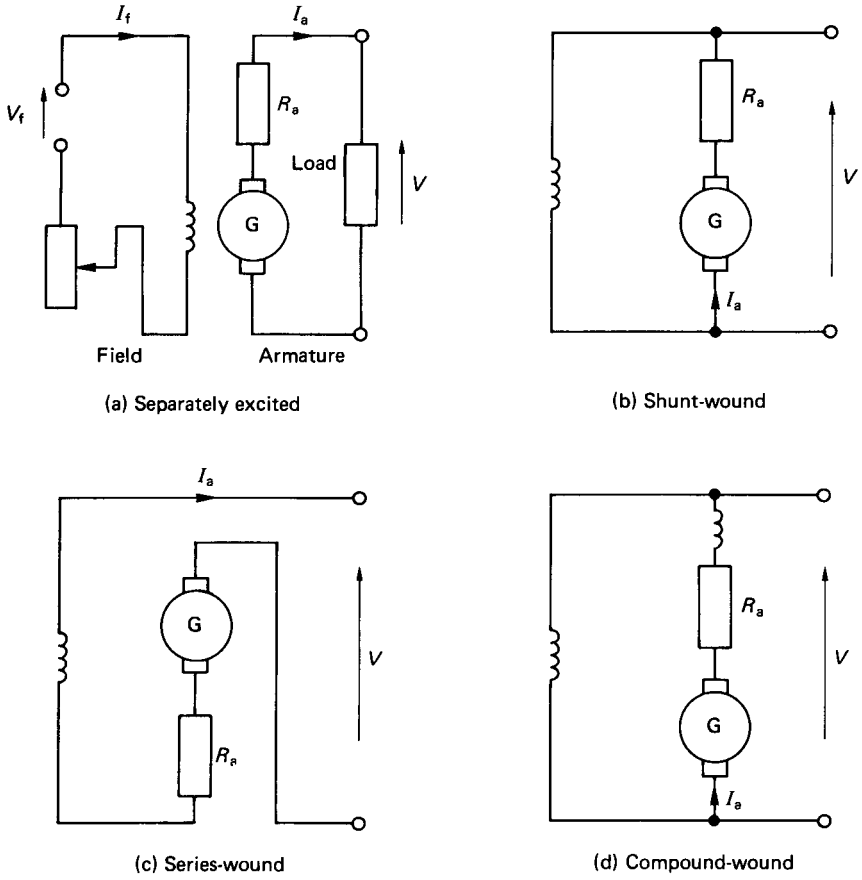


FIGURE 4.27 Methods of field connection.

poles were completely demagnetized. If the field current, and hence the magnetic field, is gradually increased, then a plot of terminal voltage against field current takes the form shown in Figure 4.28.

As the field current increases, the iron poles begin to saturate and the proportionality between the flux and the field current no longer exists. If the field current is then reduced, the magnetic hysteresis causes the terminal voltage to have a slightly greater value than that obtained when the field current was being increased. When the field current is reduced to zero, a residual voltage remains. On increasing the field current once more, the curve follows the broken line to merge with the original lower curve. These curves are termed the open-circuit characteristics of the machine.

If the generator is now connected to a variable external load and driven at constant speed with a constant field current, I_f, the terminal voltage variation with armature current is as shown in Figure 4.29. The decrease in terminal voltage with increase in load is due mainly to the voltage drop across the armature resistance, R_a. Additionally, the decrease in terminal voltage is attributed to a decrease in flux caused both by the demagnetizing ampere-turns of the armature and the magnetic saturation in the armature teeth. These effects are collectively known as armature reaction. Figure 4.29 is referred to as the load characteristic of the generator.

The separately excited generator has the disadvantage inherent with a separate source of direct current required for the field coils. They are, however, used in cases where a wide range in terminal voltage is required.

The Shunt-Wound Generator

The field winding in the shunt-wound generator is connected across the armature terminals as shown in Figure 4.27(b) and is therefore in parallel (or shunt) with the

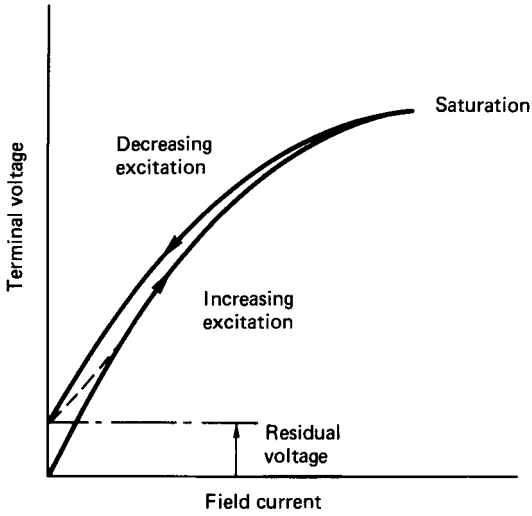


FIGURE 4.28 Open-circuit characteristics of a separately excited generator.

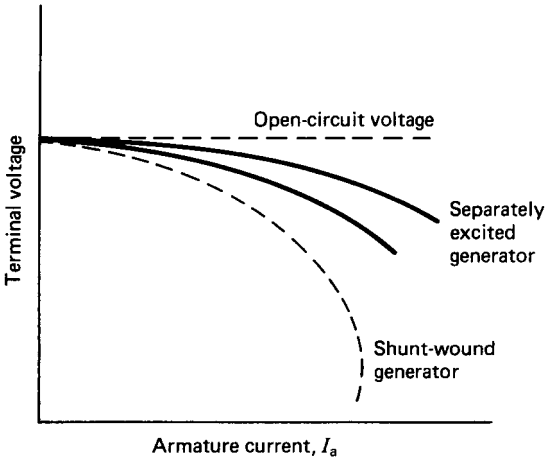


FIGURE 4.29 Load characteristics of a separately excited generator.

load. A shunt generator will excite only if the poles have some residual magnetism and the resistance of the shunt circuit is less than some critical value.

If, when running at constant speed, the field is disconnected from the armature, the voltage generated across the armature brushes is very small and entirely due to the residual magnetism in the iron. When the field is connected, the small residual voltage generates a flow of current in the field winding. The total flux in the field winding will gradually build up and the final terminal voltage will depend on the resistance of the field winding and the magnetization curve of the machine. The general characteristic is shown in Figure 4.30.

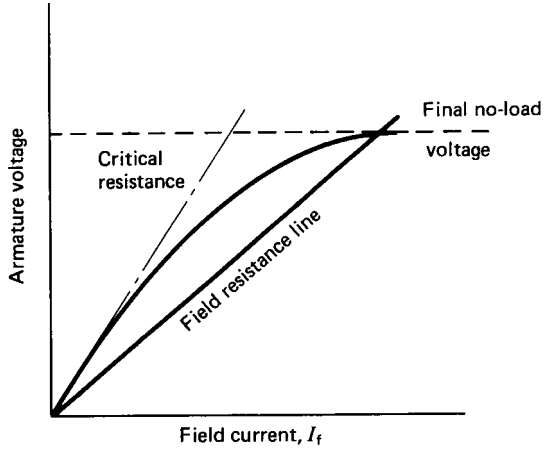


FIGURE 4.30 No-load characteristic of a shunt-wound generator.

When connected to an external load, the shunt-wound generator exhibits a drop in terminal voltage as the armature current is increased (see Figure 4.29). The drop in voltage in the shunt-wound generator is much greater than that in the separately excited generator. This stems from the fact that, as the terminal voltage drops, the field current also reduces, which causes a further drop in terminal voltage.

The shunt-wound machine is the most common type of d.c. generator employed. The load current, however, must be limited to a value well below the maximum value to avoid excessive variation in terminal voltage.

The Series-Wound Generator

For the series-wound generator the field winding is connected in series with the armature terminals as shown in Figure 4.27(c). The armature current therefore determines the flux. The constant speed load characteristic (Figure 4.31) exhibits an increase in terminal voltage as the armature (or load) current increases.

At large values of load current the armature resistance and reactance effects cause the terminal voltage to decrease. It is apparent from Figure 4.31 that the series-wound generator is totally unsuitable if the terminal voltage is required to be reasonably constant over a wide range of load current.

The Compound-Wound Generator

The compound-wound generator (Figure 4.27(d)) is a hybrid between the shunt- and the series-wound generators. Normally, a small series field is arranged to assist the main shunt field. This is termed *cumulative compounding*. The shape of the load characteristic (Figure 4.32) depends upon the number of turns on the series winding. If the series field is arranged to oppose the main shunt field (differentially compounded), a rapidly falling load characteristic is obtained. The number of turns

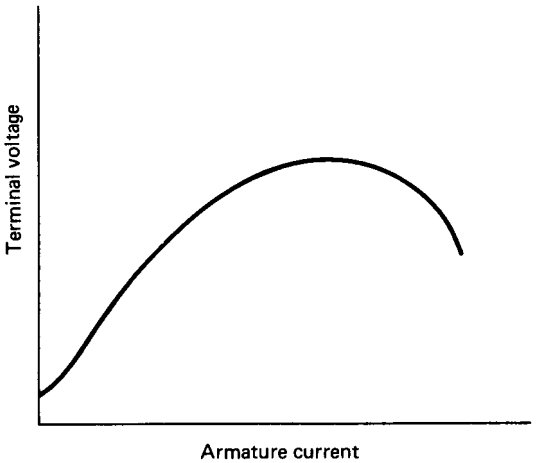


FIGURE 4.31 Constant speed load characteristic for the series-wound generator.

ELECTRICAL AND ELECTRONIC PRINCIPLES

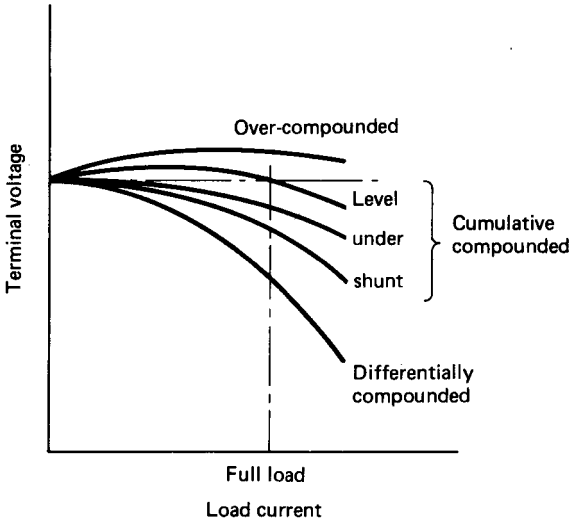


FIGURE 4.32 Load characteristic for the compound-wound generator.

on the series coil can be varied to give an overcompounded, a level-compounded, or an undercompounded characteristic, as shown in Figure 4.32.

The d.c. Motor

There is no difference in basic construction between a d.c. generator and a d.c. motor. The only significant distinction between the two machines is quantified by equations (4.70) and (4.71). These illustrate the fact that, for a d.c. generator, the generated e.m.f. is greater than the terminal voltage. For the d.c. motor, the generated e.m.f. is less than the terminal voltage.

Equation (4.65), which gives the relationship between the induced e.m.f. and the speed of a d.c. generator, applies equally well to the d.c. motor. Since the number of poles and number of armature conductors are fixed, a proportionality relationship can be derived to relate speed as a function of induced e.m.f. and flux, i.e.,

$$N = E/\phi \tag{4.72}$$

or, using equation (4.71),

$$N = (V - I_a R_a)/\Phi \tag{4.73}$$

The value of $I_a R_a$ is usually less than about *5%* of the terminal voltage such that, to a reasonable approximation,

$$N \approx V/\Phi \tag{4.74}$$

Similarly, equation (4.69), which gives the armature torque on a d.c. generator, also

applies to the d.c. motor. A proportionality relationship for the d.c. motor torque is therefore

$$T = I_a \Phi \tag{4.75}$$

Equation (4.74) shows that the speed of a d.c. motor is approximately proportional to the voltage applied to the armature and inversely proportional to the flux. All methods of controlling the speed of d.c. motors are based on these proportionality relationships. Equation (4.75) indicates that the torque of a given d.c. motor is directly proportional to the product of the armature current and the flux per pole.

The Shunt-Wound Motor

The shunt-wound motor is shown schematically in Figure 4.33. Under normal operating conditions the field current will be constant. As the armature current increases, however, the armature reaction effect will weaken the field and the speed will tend to increase. The induced voltage will decrease due to the increasing armature voltage drop, and this will tend to decrease the speed. The two effects are not self-canceling, and overall the motor speed will fall slightly as the armature current increases.

The motor torque increases approximately linearly with the armature current until the armature reaction starts to weaken the field. These general characteristics are shown in Figure 4.34, along with the derived torque–speed characteristic.

Figure 4.34(a) shows that no torque is developed until the armature current is large enough to supply the constant losses in the machine. Since the torque increases significantly for a slight decrease in speed, the shunt-wound motor is particularly suitable for driving equipment such as pumps. compressors, and machine tool elements, where the speed must remain constant over a wide range of load.

The Series-Wound Motor

The series-wound motor is shown in Figure 4.35. As the load current increases, the induced voltage, E, will decrease due to reductions in the armature and field resistance voltages. Because the field winding is connected in series with the armature, the flux is directly proportional to the armature current. Equation (4.74) therefore suggests that the speed/armature current characteristic will take the form of a rec-

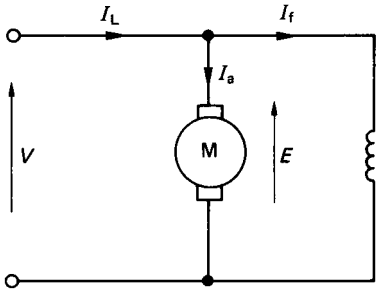


FIGURE 4.33 The shunt-wound motor.

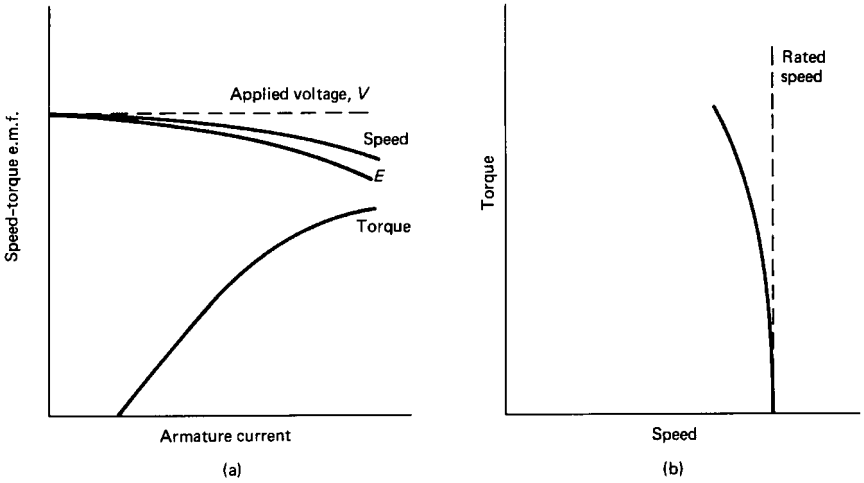


FIGURE 4.34 The shunt-wound motor load characteristics.

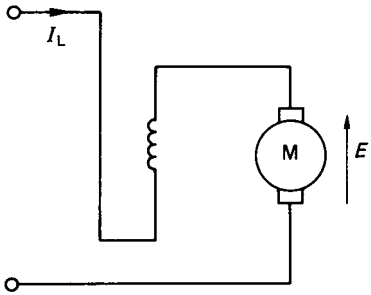


FIGURE 4.35 The series-wound motor.

tangular hyperbola. Similarly, equation (4.75) indicates that the torque/armature current characteristic will be approximately parabolic. These general characteristics are illustrated in Figure 4.36 along with the derived torque–speed characteristic.

The general characteristics indicate that if the load falls to a particularly low value then the speed may become dangerously high. A series-wound motor should therefore never be used in situations where the load is likely to be suddenly relaxed.

The main advantage of the series-wound motor is that it provides a large torque at low speeds. These motors are eminently suitable, therefore, for applications where a large starting torque is required. This includes, for example, elevators, hoists, cranes, and electric trains.

The Compound-Wound Motor

Compound-wound motors, like compound generators, are produced by including both series and shunt fields. The resulting characteristics of the compound-wound

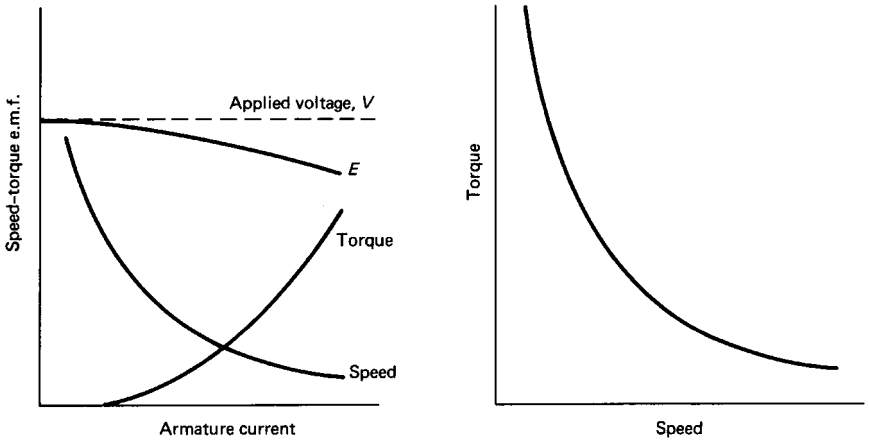


FIGURE 4.36 The series-wound motor load characteristics.

motor fall somewhere in between those of the series- and the shunt-wound machines.

Starting d.c. Motors

With the armature stationary, the induced e.m.f. is zero. If, while at rest, the full voltage is applied across the armature winding, the current drawn would be massive. This current would undoubtedly blow the fuses and thereby cut off the supply to the machine. To limit the starting current, a variable external resistance is connected in series with the armature. On start-up the full resistance is connected in series. As the machine builds up speed and increases the back e.m.f., the external resistance can be reduced until the series resistance is disconnected at rated speed.

Variable-resistance starters are also usually equipped with a return spring and an electromagnetic catch plate. The latter keeps the starter in the zero resistance position while the machine is running at its rated speed. The electromagnet is powered by the field current and, in the event of a supply failure, the electromagnet is deenergized and the return spring pulls the starter back to the full-resistance "off" position. This ensures that the full starting resistance will always be in series with the armature winding when the machine is restarted.

An overload cut-out switch is another normal feature incorporated into the starter mechanism. The overload cut-out is another electromagnetic switch, which this time is powered by the supply current. The overload switch is normally "off," but if the supply current becomes excessive, the switch is activated and it short-circuits the supply to the electromagnetic catch plate. This, in turn, deenergizes the catch plate and the return spring takes the starter back to the "off" position. Figure 4.37 illustrates the essential features of a starter device for a shunt-wound motor.

Speed Control of d.c. Motors

Equation (4.74) shows that the speed of a d.c. motor is influenced both by the applied voltage and the flux. A variation in either of these parameters will therefore effect a variation in the motor speed.

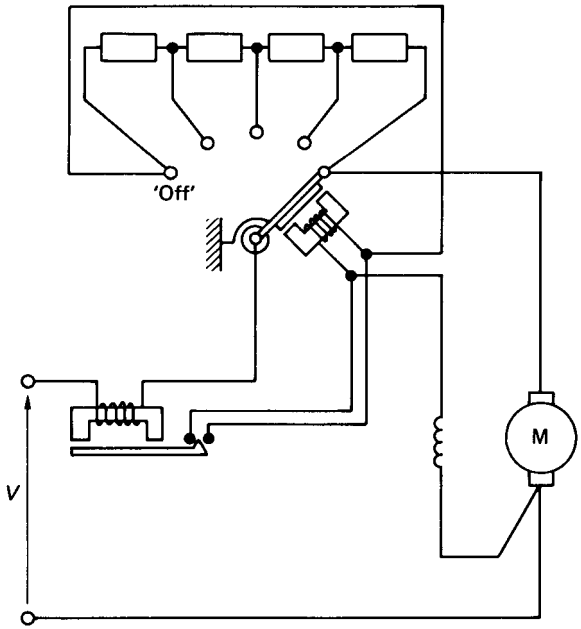


FIGURE 4.37 Starter device for d.c. motors.

Field Regulator. For shunt- and compound-wound motors a variable resistor, called a field regulator, can be incorporated in series with the field winding to reduce the flux. For the series-wound motor the variable resistor is connected in parallel with the field winding and is called a diverter. Figure 4.38 shows the various methods of weakening the field flux for shunt-, compound-, and series-wound motors.

In all the above methods of speed control the flux can only be reduced, and from equation (4.74) this implies that the speed can only be increased above the rated speed, and may, in fact, be increased to about three or four times the rated speed. The increased speed, however, is at the expense of reduced torque, since the torque is directly proportional to the flux which is reduced.

Variable Armature Voltage. Alternatively, the speed can be increased from standstill to rated speed by varying the armature voltage from zero to rated value. Figure 4.39 illustrates one method of achieving this.

The potential divider, however, carries the same current as the motor, and this limits this method of speed control to small machines. Additionally, much of the input energy is dissipated in the controller, which consequently renders the system inefficient.

Ward Leonard Drive. In this case the variable d.c. voltage for the speed-controlled motor is obtained from a separate d.c. generator that is itself driven by an induction motor (see Figure 4.40). The field coil for the d.c. generator is supplied from a center-tapped potential divider. When the wiper arm is moved from O to A, the armature voltage of the d.c. motor is increased from zero and the motor speed will

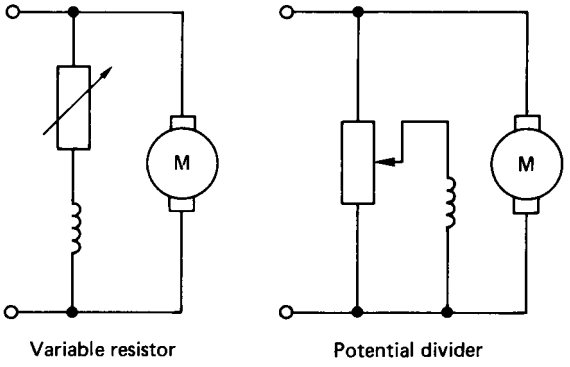


FIGURE 4.38 Speed control for flux reduction.

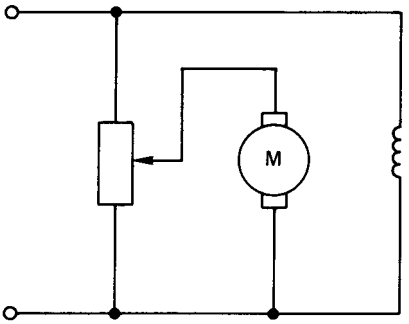


FIGURE 4.39 Speed control by varying armature voltage.

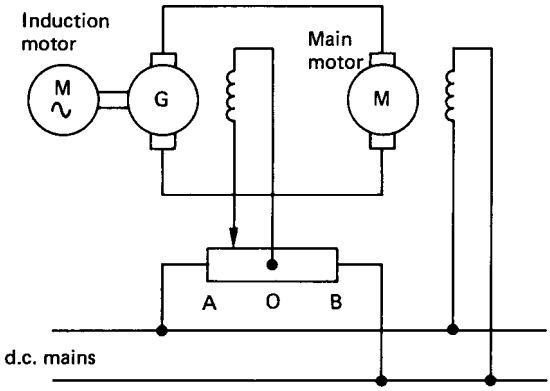


FIGURE 4.40 Ward Leonard drive.

rise. In moving the wiper from *A* to *O* and on through to *B* the motor will decelerate to a standstill and then increase in speed again, but in the opposite direction of rotation. The Ward Leonard drive is smooth and accurate in either direction and also provides for very responsive braking. Its complexity, however, makes it a very expensive system, and it is only used in high-quality applications.

Chopper Control. Figure 4.41 shows a thyristor circuit connected in series with the armature of a d.c. motor. The thyristor circuit is triggered such that it operates essentially as a high-speed on/off switch. The output waveform across the armature terminals is depicted in Figure 4.42. The ratio of time on to time off (i.e., the mark/space ratio) can be varied, with the result that the average voltage supplied to the armature is effectively between zero and fully on. The frequency of the signal may be up to about 3 kHz, and the timing circuit is necessarily complex. Speed control of d.c. motors using thyristors, is, however, effective and relatively inexpensive.

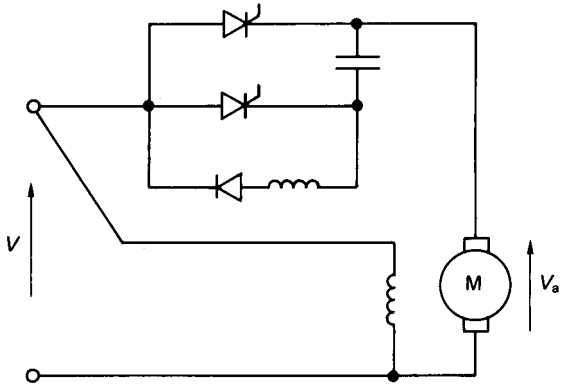


FIGURE 4.41 Speed control using thyristors.

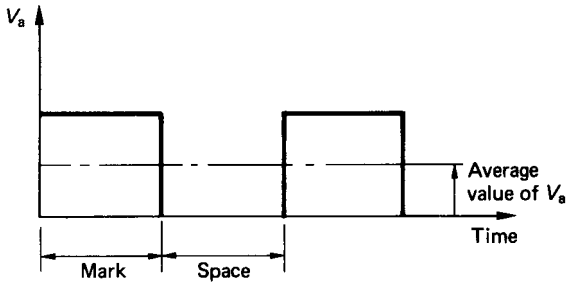


FIGURE 4.42 Voltage across armature terminals.

Efficiency of d.c. Machines

The losses in d.c. machines can be generally classified as:

Armature losses
Iron loss
Commutator losses
Excitation loss
Bearing friction and windage

Despite the variety and nature of the losses associated with d.c. machines, they have, nonetheless, a very good performance, with overall efficiencies often in excess of 90%.

Three-Phase Circuits

Since a.c. machines are generally associated with three-phase systems, it is necessary to consider some aspects of three-phase circuits before a meaningful discussion of a.c. machines can be undertaken. The limiting factor of a d.c. machine is related to the commutator which restricts the maximum voltage that can be generated. Because of their efficiency and performance, three-phase machines have emerged as the dominant type of electrical generator and motor and, on a worldwide basis, three-phase electrical distribution networks are the norm.

Generation of Three-Phase e.m.f.s

Figure 4.43 shows three similar coils displaced at 120° relative to each other. Each ioop terminates in a pair of slip-rings, and if the coils are to be isolated from one another, then six slip-rings are required in total. If the three coils are rotated in the anticlockwise direction at constant speed, then each coil will generate a sinusoidally varying e.m.f. with a phase shift of 120° between them.

ELECTRICAL AND ELECTRONIC PRINCIPLES

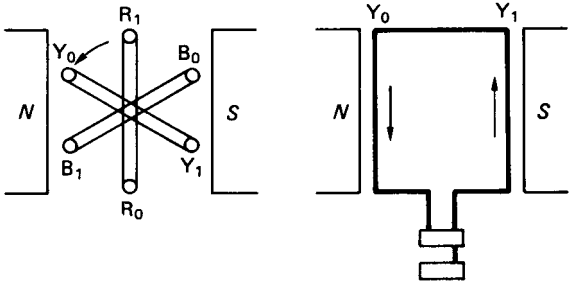


FIGURE 4.43 Generation of three-phase e.m.f.'s.

Star and Delta Connections

The three coils shown in Figure 4.43 can be connected together in either of two symmetrical patterns. These are the *star* (or *wye*) connection and the *delta* (or *mesh*) connection. The two types of connection are shown in Figure 4.44.

The star pattern is made by joining R_0, Y_0, and B_0 together. This connection point is referred to as the *neutral point*. The delta pattern is formed by connecting R_0 to Y_1, Y_0 to B_1, and B_0 to R_1.

Three-Phase Voltage and Current Relations

Figure 4.45 shows a three-phase star connected alternator supplying currents I_R, I_Y, and I_B to a balanced (or equal) resistive—inductive load. This gives the usual four-wire star-connected system. Since there are only four transmission cables involved, the alternator connected in a star pattern will only require four slip-rings.

For a balanced system the phase voltages V_{RN}, V_{YN}, and V_{BN} are all equal in magnitude and equally displaced by a phase angle of 120°. The currents I_R, I_Y, and I_B are also equal in magnitude and equally displaced in phase angle, but they all lag their respective phase voltages by some angle ϕ. Phasor addition of the currents shows that the neutral current, I_N is zero.

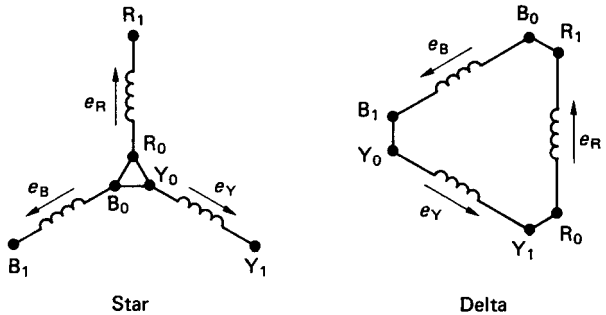


FIGURE 4.44 Star and delta connections for three-phase systems.

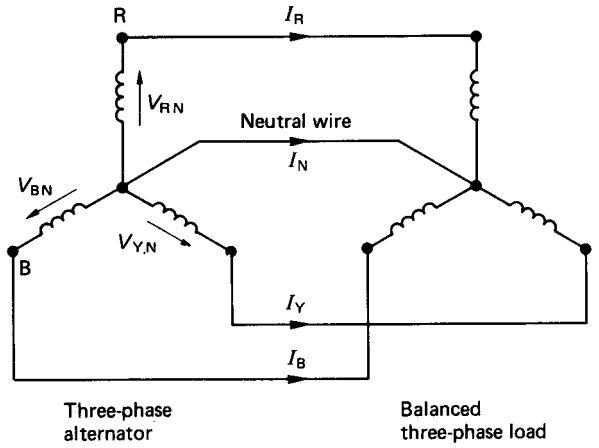


FIGURE 4.45 Three-phase supply connections.

The voltages between the transmission cables are called the line voltages. If the phase voltages are all equal then phasor addition shows that the line voltages are given by

$$V_{\text{line}} = 2V_{\text{phase}} \cos(30)$$

or

$$V_L = \sqrt{3} \times V_p \tag{4.76}$$

For the star connection, the line currents, I_L, are equal to the phase currents, I_p. Figure 4.46 shows the alternator windings connected in the delta pattern. In this pattern the line voltages are equal to the phase voltages. Phasor addition of the currents shows that if the phase currents are equal then the line currents are given by

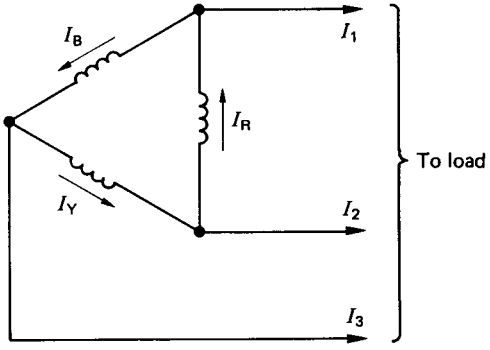


FIGURE 4.46 Alternator windings in delta connection.

$$I_L = \sqrt{3} \times I_p \tag{4.77}$$

Power in Three-Phase Circuits

The power per phase is given by

$$P_{\text{phase}} = V_p I_p \cos(\phi) \tag{4.78}$$

where V_P = the phase voltage
I_p = the phase current
ϕ = the phase angle between V_P and I_p

The total power for a three-phase circuit is simply three times the power for one of the phases, i.e., three times equation (4.78).

For a star connection:

$$P = 3 \frac{V_L}{\sqrt{3}} I_L \cos(\phi) = \sqrt{3} \times V_L \times I_L \cos(\phi) \tag{4.79}$$

For a delta connection:

$$P = 3V_L \frac{I_L}{\sqrt{3}} \cos(\phi) = \sqrt{3} \times V_L \times I_L \cos(\phi)$$

The same relation is obtained. In terms of line voltages and currents, therefore, the power in a three-phase circuit is independent of the winding connection and is given by equation (4.79). This equation does not, however, apply if the system is unbalanced. In an unbalanced system the total power can only be obtained as the summation of the powers in each of the individual phases.

Three-Phase Alternators

Alternators are constructed with a stationary a.c. winding and a rotating field system. This reduces the number of slip-rings required to two, and these have to carry only the field-exciting current as opposed to the generated current. The construction is thereby simplified and the slip-ring losses are minimized. In addition, the simpler arrangement enables heavier insulation to be used and, in consequence, much higher voltages can be generated. The robust mechanical construction of the rotor also means that higher speeds are possible and substantially higher power outputs can be generated with an alternator. A simple form of three-phase generator is depicted in Figure 4.47.

The three coils on the stator are displaced 120° and the rotor, which is a salient pole type, is supplied via the two slip-rings with a d.c. current. As the rotor is driven by some form of prime mover, a rotating magnetic field is established and the e.m.f.'s generated in the coils will be displaced with a phase shift of 120°. The magnitude of the generated voltages is dependent on the flux produced by the rotor, the number of turns on the stator coils, and the speed of rotation of the rotor. The rotor speed will also dictate the frequency of the generated voltage.

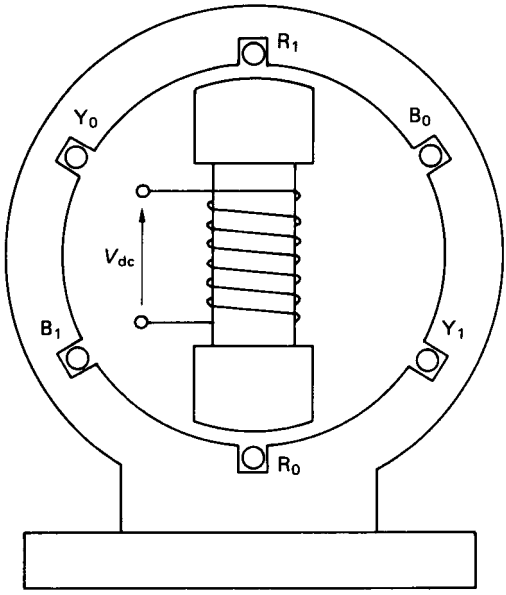


FIGURE 4.47 Simple three-phase generator.

The no-load and load characteristics of an alternator are very similar to those of the d.c. separately excited generator (Figures 4.28 and 4.29, respectively). In constant speed operation the terminal voltage exhibits a drooping characteristic, where the decrease in terminal voltage is due to armature resistance and reactance effects. For an alternator, the term *armature* is taken to imply the stator windings.

As the load on an alternator is increased, the speed of the prime mover will drop. This is an unacceptable situation because the speed controls the frequency of the generated voltage. To maintain a constant frequency, the prime mover must be governed to run at constant speed over the entire range of expected loads. This is particularly important where many alternators are to be run in parallel to supply a distribution system such as the National Grid. In such cases the prime movers are always speed controlled and the output voltage is regulated to comply with the rated values. In the United Kingdom, alternators are usually two-pole machines driven at 3000 rev/min to produce the rated frequency of 50 Hz. In the United States, a great deal of the electrical power consumed is generated from hydroelectric power stations. The water turbines used in these installations are fairly low-speed machines and the alternators, which are directly driven, are equipped with multiple poles to produce the rated frequency of 60 Hz. An alternator running at 240 rev/min, for example, must have 30 poles to give the rated output frequency.

The production of the rotating magnetic field may also be activated using three, 120° displaced, rotor coils supplied with three-phase current. The rotational speed of the field is related to the frequency of the currents, i.e.,

$$N_s = \frac{f \times 60}{\text{Number of pole pairs}} \tag{4.80}$$

where N_s is the speed of the field (rev/min) and f is the frequency of the supply

currents. The speed of the rotating field is termed the *synchronous speed* and for an equivalent single pair of poles (i.e., three coils) this is 3000 rev/min when the frequency of the supply currents is at 50 Hz.

The use of a.c. excited rotor coils to produce the rotating magnetic field simplifies the mechanical construction of the rotor and greatly facilitates the dynamic balancing of the machine. An added advantage is that the waveform of the generated voltage is improved. The a.c. method of exciting the field is used extensively in large alternators. Salient pole rotors are normally restricted to the smaller machines.

Synchronous Motors

Synchronous motors are so called because they operate at only one speed, i.e., the speed of the rotating field. The mechanical construction is exactly the same as the alternator shown in Figure 4.47. The field is supplied from a d.c. source and the stator coils with a three-phase current. The rotating magnetic field is induced by the stator coils and the rotor, which may be likened to a permanent bar magnet, aligns itself to the rotating flux produced in the stator. When a mechanical load is driven by the shaft the field produced by the rotor is pulled out of alignment with that produced by the stator. The angle of misalignment is called the *load angle*. The characteristics of synchronous motors are normally presented in terms of torque against load angle, as shown in Figure 4.48. The torque characteristic is basically sinusoidal, with

$$T = T_{\max} \sin(\delta) \tag{4.81}$$

where $T_{\max}$ is the maximum rated torque and δ is the load angle.

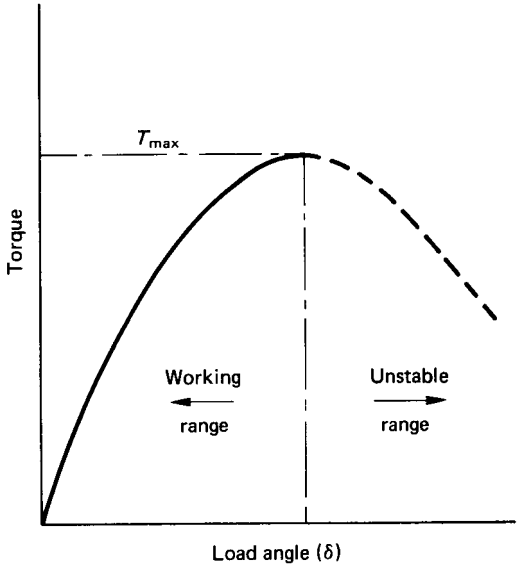


FIGURE 4.48 Torque characteristic for a synchronous motor.

It is evident from equation (4.81) that synchronous motors have no starting torque and the rotor must be run up to synchronous speed by some alternative means. One method utilizes a series of short-circuited copper bars inserted through the outer extremities of the salient poles. The rotating magnetic flux induces currents in these grids and the machine accelerates as if it were a cage-type induction motor. A second method uses a wound rotor similar to a slip-ring induction motor. The machine is run up to speed as an induction motor and is then pulled into synchronism to operate as a synchronous motor.

The advantages of the synchronous motor are the ease with which the power factor can be controlled and the constant rotational speed of the machine, irrespective of the applied load. Synchronous motors, however, are generally more expensive, and a d.c. supply is a necessary feature of the rotor excitation. These disadvantages, coupled with the requirement for an independent starting mode, make synchronous motors much less common than induction ones.

Induction Motors

The stator of an induction motor is much like that of an alternator and, in the case of a machine supplied with three-phase currents, a rotating magnetic flux is produced. The rotor may be either of two basic configurations: the squirrel-cage or the slip-ring type. In the ''squirrel-cage'' motor the rotor core is laminated and the conductors consist of uninsulated copper (or aluminum) bars driven through the rotor slots. The bars are brazed or welded at each end to rings or plates to produce a completely short-circuited set of conductors. The slip-ring machine has a laminated core and a conventional three-phase winding, similar to the stator and connected to three slip-rings on the locating shaft.

Figure 4.49 shows a schematic representation of an induction motor having three stator coils displaced by 120°. If the stator coils are supplied with three-phase

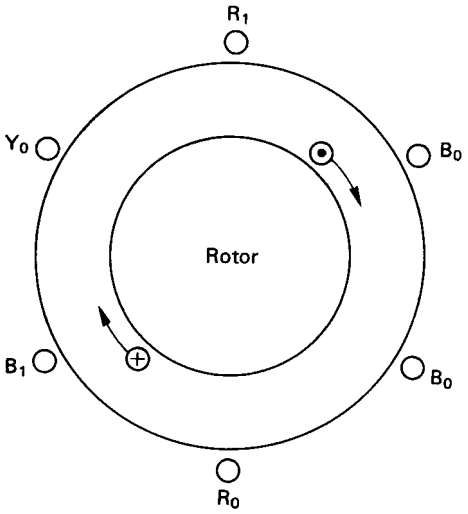


FIGURE 4.49 Schematic representation of an induction motor.

currents, a rotating magnetic field is produced in the stator. Consider the single-rotor coil shown in the figure. At standstill the rotating field will induce a voltage in the rotor coil since there is a rate of change of flux linking the coil. If the coil forms a closed circuit then the induced e.m.f. will circulate a current in the coil. The resultant force on the current-carrying conductor is a consequence of equation (4.27), and this will produce a torque that will accelerate the rotor. The rotor speed will increase until the electromagnetic torque is balanced by the mechanical load torque. The induction motor will never attain synchronous speed because, if it did, there would be no relative motion between the rotor coils and the rotating field. Under these circumstances there would be no e.m.f. induced in the rotor coils and subsequently no electromagnetic torque. Induction motors therefore always run at something less than synchronous speed. The ratio of the difference between the synchronous speed and the rotor speed to the synchronous speed is called the *slip*, s, i.e.,

$$s = \frac{N_s - N}{N_s} \tag{4.82}$$

The torque–slip characteristic is shown in Figure 4.50. With the rotor speed equal to the synchronous speed, i.e., $s = 0$, the torque is zero. As the rotor falls below the synchronous speed the torque increases almost linearly to a maximum value dictated by the total of the load torque and that required to overcome the rotor losses. The value of slip at full load varies between 0.02 and 0.06. The induction motor may be regarded therefore as a constant-speed machine. In fact, the difficulties of varying the speed constitutes one of the induction motor's main disadvantages.

On start-up the slip is equal to unity and the starting torque is sufficiently large to accelerate the rotor. As the rotor runs up to its full-load speed the torque increases in essentially inverse proportion to the slip. The start-up and running curves merge at the full-load position.

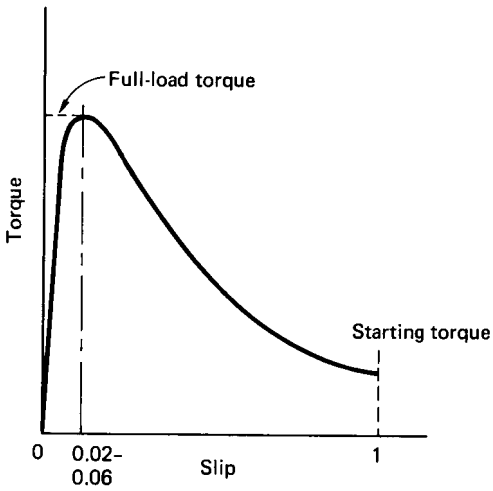


FIGURE 4.50 Torque–slip characteristic for an induction motor.

Starting Induction Motors

As with d.c. motors, the current drawn during starting of a.c. motors is very large, up to about five times full-load current. A number of devices are therefore employed to limit the starting current but they all involve the use of auxiliary equipment, which is usually quite expensive.

Star–Delta Starter. The star–delta switch (Figure 4.51) is the cheapest and most common method employed. With the machine at standstill and the starter in the "start" position, the stator coils are connected in the star pattern. As the machine accelerates up to running speed the switch is quickly moved over to the run position, which reconnects the stator windings in the delta pattern. By this simple expedient the starting supply current is reduced to one third of what it would have been had the stator windings been connected in the delta pattern on start-up.

Auto-Transformer Starter. The auto-transformer represents an alternative method of reducing the starting current drawn by an induction motor.

Rotor Resistance. With slip-ring induction motors it is possible to include additional resistance in series with the rotor circuit. The inclusion of extra resistance in the rotor provides for reduced starting current and improved starting torque.

Braking Induction Motors

Induction motors may be brought to a standstill by either plugging or dynamic braking:

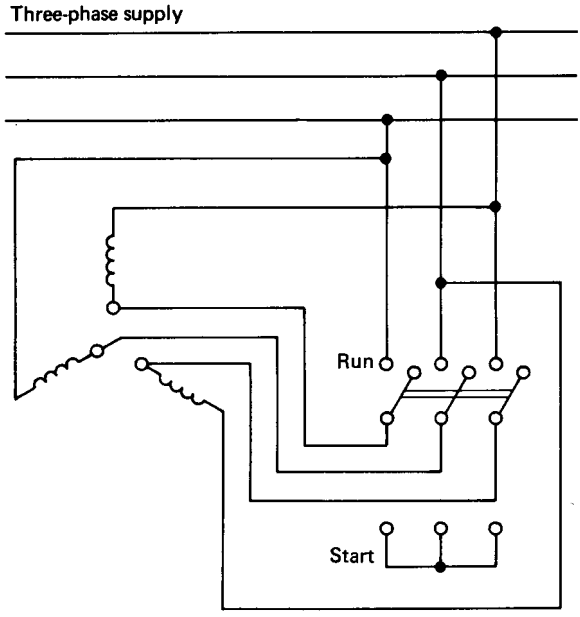


FIGURE 4.51 Star–delta starter.

1. *Plugging:* This refers to the technique where the direction of the rotating magnetic field is reversed. This is brought about by reversing any two of the supply leads to the stator. The current drawn during plugging is, however, very large, and machines that are regularly plugged must be specially rated.
2. *Dynamic braking:* In this technique the stator is disconnected from the a.c. supply and reconnected to a d.c. source. The direct current in the stator produces a stationary unidirectional field and, as the rotor will always tend to align itself with the field, it will come to a standstill.

Speed Control of Induction Motors

Under normal circumstances the running speed of an induction motor will be about 94–98% of the synchronous speed, depending on the load. With the synchronous speed given by equation (4.80) it is clear that the speed may be varied by changing either the frequency of the supply current or the number of poles.

Change of Supply Current Frequency. Solid state variable-frequency drives first began to appear in 1968. They were originally applied to the control of synchronous a.c. motors in the synthetic fiber industry and rapidly gained acceptance in that particular market. In more recent times they have been used in applications such as pumping, synchronized press lines, conveyor lines, and, to a lesser extent, in the machine-tool industry as spindle drives. Modern a.c. variable-frequency motors are available in power ratings ranging from 1 kW to 750 kW and with speed ranges from 10/1 to 100/1.

Change of Number of Poles. By bringing out the ends of the stator coils to a specially designed switch it becomes possible to change an induction motor from one pole configuration to another. To obtain three different pole numbers, and hence three different speeds, a fairly complex switching device would be required.

Changing the number of poles gives a discrete change in motor speed, with little variation in speed over the switched range. For many applications, however, two discrete speeds are all that is required, and changing the number of poles is a simple and effective method of achieving this.

Changing the Rotor Resistance. For slip-ring induction motors additional resistance can be coupled in series with the rotor circuit. It has already been stated that this is a common method used to limit the starting current of such machines. It can also be employed for marginal speed control. Figure 4.52 shows the torque characteristics of a slip-ring induction motor for a range of different resistances connected in series with the rotor windings. As the external resistance is increased from R_1 to R_3 a corresponding reduction in speed is achieved at any particular torque. The range of speeds is increased at the higher torques.

The method is simple and therefore inexpensive, but the decrease in speed is accompanied with a reduction in overall efficiency. Additionally, with a large resistance in the rotor circuit (i.e., R_3) the speed changes considerably with variations in torque.

Reduced Stator Voltage. By reducing the applied stator voltage a family of torque-speed characteristics is obtained, as shown in Figure 4.53. It is evident that as the stator voltage is reduced from V_1 to V_3, a change in speed is effected at any particular value of torque. This is provided, of course, that the torque does not

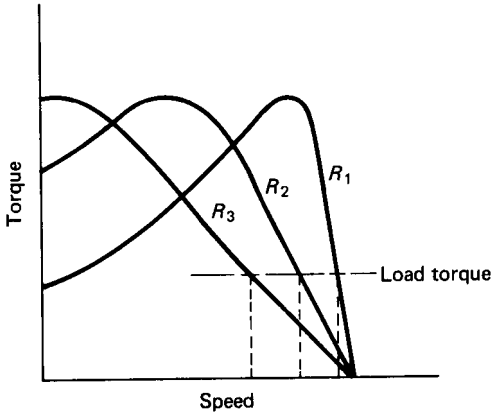


FIGURE 4.52 Torque–speed characteristics for various rotor resistances.

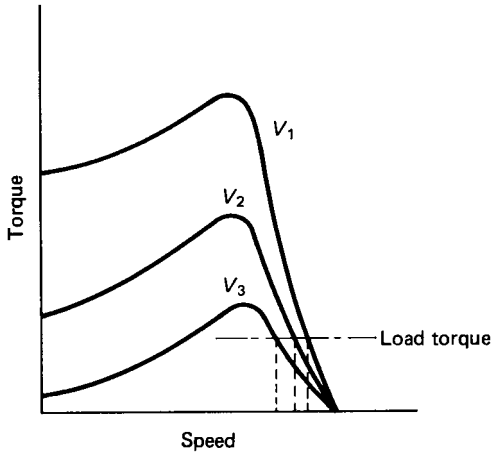


FIGURE 4.53 Torque–speed characteristics for various stator voltages.

exceed the maximum load torque available at the reduced stator voltage. This latter point is obviously a limiting factor that places a constraint on this method of speed control. Generally, only very small speed ranges can be obtained using a variable stator supply voltage.

Single-Phase Induction Motors

The operation of an induction motor depends upon the creation of a rotating magnetic field. A single stator coil cannot achieve this, and all the so-called single-phase induction motors use some or other external means of generating an

approximation to a two-phase stator supply. Two stator coils are therefore used, and these are displaced by 90°. Ideally, the currents that supply each coil should have a phase difference of 90°. This then gives the two-phase equivalent of the three-phase induction motor.

The Shaded-Pole Motor. The stator of the shaded-pole motor consists of a salient pole single-phase winding, and the rotor is of the squirrel-cage type (see Figure 4.54). When the exciting coil is supplied with alternating current the flux produced induces a current in the shading ring. The phase difference between the currents in the exciting coil and the shading ring is relatively small and the rotating field produced is far from ideal. In consequence, the shaded-pole motor has a poor performance and an equally poor efficiency due to the continuous losses in the shading rings.

Shaded-pole motors have a low starting torque and are used only in light-duty applications such as small fans and blowers or other easily started equipment. Their advantage lies in their simplicity and low cost of manufacture.

The Capacitor Motor. A schematic layout of a capacitor motor is given in Figure 4.55. The stator has two windings physically displaced by 90°. A capacitor is connected in series with the auxiliary winding such that the currents in the two wind-

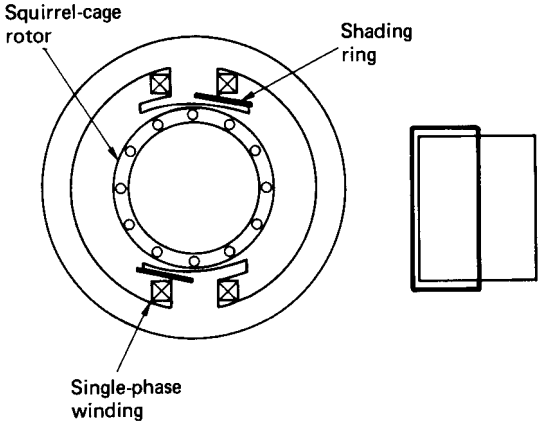


FIGURE 4.54 Shaded pole motor.

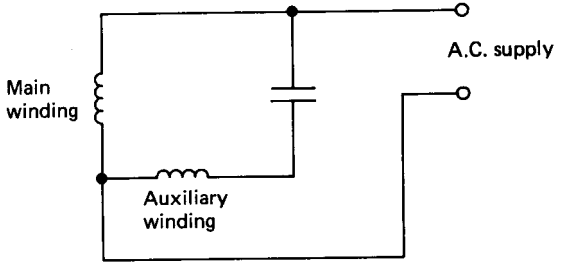


FIGURE 4.55 Capacitor motor.

ings have a large phase displacement. The current phase displacement can be made to approach the ideal 90°, and the performance of the capacitor motor closely resembles that of the three-phase induction motor.

The Universal Motor. These are small d.c. series-wound motors that operate at about the same speed and power on direct current, or on single-phase current with approximately the same root mean square voltage. The universal (or plain-series) motor is used mainly in small domestic appliances such as hair dryers, electric drills, vacuum cleaners, hedge trimmers, etc.

The d.c. Permanent Magnet (PM) Motor

The d.c. permanent magnet (PM) motor is a continuous-rotation electromagnetic actuator which can be directly coupled to its load. Figure 4.56 shows the schematic representation of a d.c. PM motor. The PM motor consists of an annular brush ring assembly, a permanent magnet stator ring and a laminated wound rotor. It is particularly suitable for servo systems where size, weight, power, and response times must be minimized and high position and rate accuracies are required.

The response times for PM motors are very fast and the torque increases directly with the input current, independently of the speed or the angular position. Multiple-pole machines maximize the output torque per watt of rotor power. Commercial PM motors are available in many sizes, from 35 million Newton-meters at about 25 mm diameter to 13.5 Newton-meters at about 3 m diameter.

Direct-drive rate and position systems using PM motors utilize d.c. tachogenerators and position sensors in various forms of closed-loop feedback paths for control purposes.

The Stepper Motor

A stepper motor is a device which converts a d.c. voltage pulse train into a proportional mechanical rotation of its shaft. The stepper motor thus functions both as an actuator and as a position transducer. The discrete motion of the stepper motor makes it ideally suited for use with a digitally based control system such as a microcomputer.

The speed of a stepper motor may be varied by altering the rate of the pulse train input. Thus, if a stepper motor requires 48 pulses to rotate through one com-

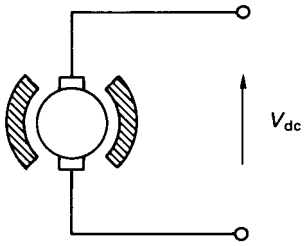


FIGURE 4.56 D.C. permanent magnet motor.

ELECTRICAL AND ELECTRONIC PRINCIPLES

plete revolution, then an input signal of 96 pulses per second will cause the motor to rotate at 120 rev/mm. The rotation is actually carried out in finite increments of time, but this is visually indiscernible at all but the lowest speeds.

Stepper motors are capable of driving a 2.2 kW load with stepping rates from 1000 to 20 000 per second in angular increments from 45° down to 0.75°. There are three basic types of stepper motor:

1. *Variable reluctance:* This type of stepper motor has a soft iron multi-toothed rotor with a wound stator. The number of teeth on the rotor and stator, together with the winding configuration and excitation, determines the step angle. This type of stepper motor provides small to medium-sized step angles and is capable of operation at high stepping rates.
2. *Permanent magnet:* The rotor used in the PM-type stepper motor consists of a circular permanent magnet mounted onto the shaft. PM stepper motors give a large step angle, ranging from 45° to 120°.
3. *Hybrid:* The hybrid stepper motor is a combination of the previous two types. Typically, the stator has eight salient poles that are energized by a two-phase winding. The rotor consists of a cylindrical magnet that is axially magnetized. The step angle depends on the method of construction and is generally in the range 0.9–5°. The most popular step angle is 1.8°.

The principle of operation of a stepper motor can be illustrated with reference to a variable-reluctance, four-phase machine. This motor usually has eight stator teeth and six rotor teeth (see Figure 4.57).

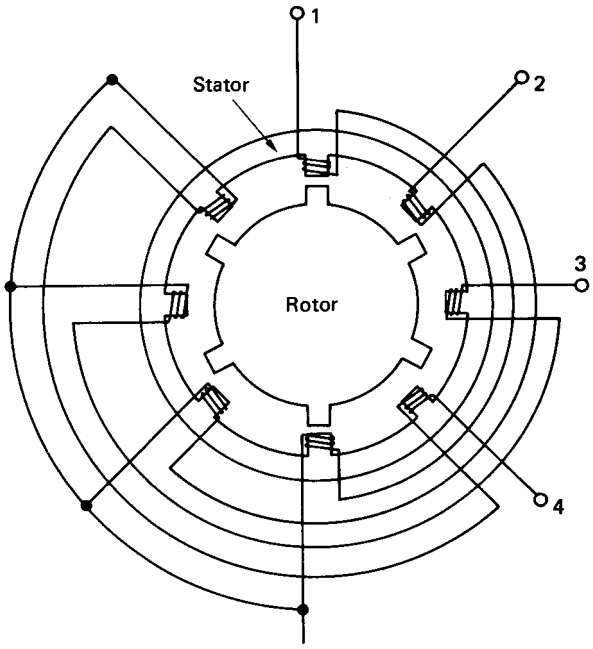


FIGURE 4.57 Variable-reluctance stepper motor.

If phase 1 of the stator is activated alone, then two diametrically opposite rotor teeth align themselves with the phase 1 teeth of the stator. The next adjacent set of rotor teeth in the clockwise direction is then 15° out of step with those of the stator. Activation of the phase 2 winding on its own would cause the rotor to rotate a further 15° in the anticlockwise direction to align the adjacent pair of diametrically opposite rotor teeth. If the stator windings are excited in the sequence 1, 2, 3, 4 the rotor will move in consecutive 15° steps in the anticlockwise direction. Reversing the excitation sequence will cause a clockwise rotation of the rotor.

Stepper Motor Terminology

Pull-out torque: The maximum torque that can be applied to a motor, running at a given stepping rate, without losing synchronism.

Pull-in torque: The maximum torque against which a motor will start, at a given pulse rate, and reach synchronism without losing a step.

Dynamic torque: The torque developed by the motor at very slow stepping speeds.

Holding torque: The maximum torque which can be applied to an energized stationary motor without causing spindle rotation.

Pull-out rate: The maximum switching rate at which a motor will remain in synchronism while the switching rate is gradually increased.

Pull-in rate: The maximum switching rate at which a loaded motor can start without losing steps.

Slew range: The range of switching rates between pull-in and pull-out in which a motor will run in synchronism but cannot start or reverse.

The general characteristics of a typical stepper motor are given in Figure 4.58. During the application of each sequential pulse the rotor of a stepper motor accelerates rapidly towards the new step position. However, on reaching the new position there will be some overshoot and oscillation unless sufficient retarding torque is

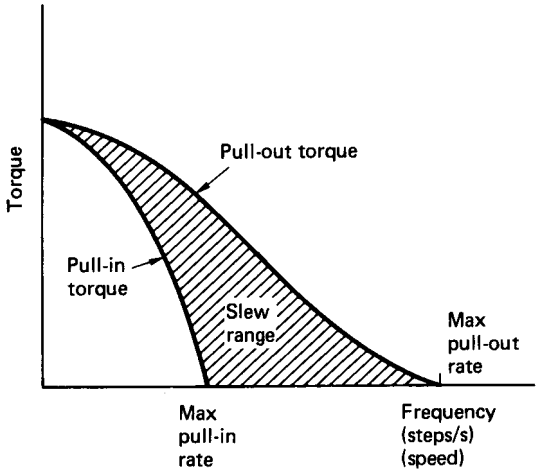


FIGURE 4.58 Stepper motor characteristics.

provided to prevent this happening. These oscillations can cause rotor resonance at certain pulse frequencies, resulting in loss of torque, or perhaps even pull-out conditions. Because variable-reluctance motors have very little inherent damping, they are more susceptible to resonances than either the permanent magnet or the hybrid types. Mechanical and electronic dampers are available that can be used to minimize the adverse effects of rotor resonance. If at all possible, however, the motor should be selected such that its resonant frequencies are not critical to the application under consideration.

Because of their unique characteristics, stepper motors are widely used in applications involving positioning, speed control, timing, and synchronized actuation. They are prevalent in X–Y plotters, punched-taped readers, floppy disk head drives, printer carriage drives, numerically controlled machine tool slide drives, and camera iris control mechanisms.

By far the most severe limitation on the purely electric stepper motor is its power-handling capability.

Brushless d.c. Motors

These motors have position feedback of some kind so that the input waveforms can be kept in the proper timing with respect to the rotor position. Solid-state switching devices are used to control the input signals, and the brushless d.c. motor can be operated at much higher speeds, with full torque available at those speeds. The brushless motor can normally be rapidly accelerated from zero to operating speed as a permanent magnet d.c. motor. On reaching operating speed the motor can then be switched over to synchronous operation.

The brushless motor system consists of a wound stator, a permanent magnet rotor, a rotor position sensor, and a solid-state switching assembly. The wound stator can be made with two or more input phases. Figure 4.59 gives the schematic representation of a two-phase brushless motor. The torque output of phase A is

$$T_A = I_A(Z\Phi/2\pi)\sin(p\theta/2) = I_A K_T \sin(p\theta/2) \tag{4.83}$$

where I_A = the current in phase A
$K_T = (Z\Phi/2\pi)$ = the torque constant of the motor
p = is the number of poles
θ = the angular position of the rotor

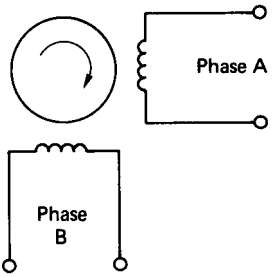


FIGURE 4.59 Two-phase brushless motor.

In the expression for the torque constant, Z is the total number of conductors and Φ is the magnetic flux.

Similarly, the torque output of phase B is

$$T_B = I_B K_T \cos(p\theta/2) \tag{4.84}$$

If the motor currents are arranged to be supplied in the following relationships:

$$I_A = I \sin(p\theta/2) \text{ and } I_B = I \cos(p\theta/2)$$

then the total torque for a two-pole motor becomes

$$\begin{aligned} T = T_A + T_B &= IK_T[\sin^2(\theta) + \cos^2(\theta)] \\ &= IK_T \end{aligned} \tag{4.85}$$

Equation (4.85) shows that if all the above conditions are satisfied then the brushless d.c. motor operates in a manner similar to the conventional d.c. motor, i.e., the torque is directly proportional to the armature current. Note that the armature current in this context refers to the stator windings.

Excitation of the phases may be implemented with sinusoidal or square-wave inputs. The sine-wave drive is the most efficient, but the output transistors in the drive electronics must be capable of dissipating more power than that dissipated in square-wave operation. Square-wave drive offers the added advantage that the drive electronics can be digitally based.

The brushless d.c. motor will duplicate the performance characteristics of a conventional d.c. motor only if it is properly commutated. Proper commutation involves exciting the stator windings in a sequence that keeps the magnetic field produced by the stator approximately 90 electrical degrees ahead of the rotor field. The brushless d.c. motor therefore relies heavily on the position feedback system for effective commutation. It might also be apparent that the brushless motor as described is not strictly a d.c. machine but a form of a.c. machine with position feedback.

The further development of the brushless d.c. motor will depend to a large extent upon future advances in semiconductor power transistor technology. It is likely, however, that within the next decade the true brushless d.c. motor, using solid-state switching, will become commercially viable and will progressively dominate the d.c. servosystem market.

This brief discussion of rotating electrical machines is in no way comprehensive. A fuller discourse on a.c. and d.c. machines is given by both Gray (1989) and Sen (1989). Orthwein (1990) presents an interesting practical discussion on the mechanical applications of a.c. and d.c. motors and Kenjo and Nagamori (1985) provide a detailed in-depth study of permanent-magnet d.c. motors.

Transformers

Transformers are used to change the level of an alternating voltage.

Basic Transformer Action

Figure 4.60 illustrates a simple single-phase transformer in which two separate coils are wound onto a ferrous core. The coil connected to the supply is called the

ELECTRICAL AND ELECTRONIC PRINCIPLES

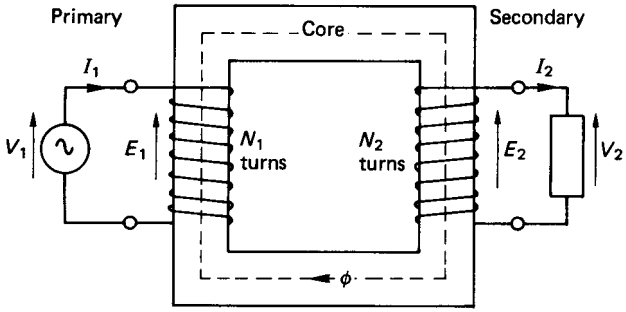


FIGURE 4.60 Single-phase transformer.

primary winding and that connected to the load is the *secondary winding*. The ferrous core is made in laminations, which are insulated from one another, to reduce eddy current losses.

If a sinusoidal voltage, V_1, is applied across the primary winding a current, I_1, in the coil will induce a magnetic flux, ϕ, in the core. From Faraday's law (equation (4.25)) the induced e.m.f. in the primary coil is

$$E_1 = N_1(d\phi/dt) \tag{4.86}$$

Since the magnetic flux is common to both coils the e.m.f. induced in the secondary winding is

$$E_2 = N_2(d\phi/dt) \tag{4.87}$$

Hence,

$$\frac{E_1}{E_2} = \frac{N_1}{N_2} \tag{4.88}$$

The ratio of primary coil turns to secondary turns, N_1/N_2, is called the *transformation ratio*. The primary and secondary winding impedances, Z_1 and Z_2, respectively, are both very small such that when the secondary winding is on open circuit, then $V_1 = E_1$ and $V_2 = E_2$. Therefore,

$$\frac{V_1}{V_2} = \frac{N_1}{N_2} \tag{4.89}$$

When a load is connected across the secondary winding a current, I_2, will flow in the secondary winding. From Lenz's law this will set up a flux which will tend to oppose the main flux, ϕ. If the main flux is reduced, then E_1 would be correspondingly decreased and the primary current, I_1, would then increase. This increased primary current would tend to produce a flux to oppose that induced by the secondary current. In this manner the main flux is generally maintained. In steady state the ampere-turns in the primary and secondary windings are balanced, i.e.,

$$I_1N_1 = I_2N_2$$

or

$$\frac{I_1}{I_2} = \frac{N_2}{N_1} \tag{4.90}$$

Transformer Voltage Equation

In normal operation the flux may be considered to be a sinusoidally varying quantity, i.e.,

$$\phi = \Phi \sin(\omega t) \tag{4.91}$$

The induced e.m.f., from Faraday's law, is

$$\text{Primary side, } e_1 = N_1(d\phi/dt) = N_1\Phi\omega \cos(\omega t)$$

The r.m.s. value of the induced e.m.f. is

$$E_1 = \frac{2\pi f N_1 \Phi}{\sqrt{2}} = 4.44\ fN\Phi \tag{4.92}$$

Similarly, for the secondary side,

$$E_2 = 4.44\ fN_2\Phi$$

Transformer Losses

Equations (4.89) and (4.90) define the ideal transformer in which there are no resistive or inductive losses. An actual transformer, of course, does involve some losses, which are:

1. *Copper losses:* These are associated with the I^2R loss in both of the coils. They may be represented therefore as a resistance in series with each coil.
2. *Iron loss:* These are associated with magnetic hysteresis effects and eddy current losses in the iron core. The iron losses are essentially constant for a particular value of supply voltage. Iron losses can be represented as a resistor in parallel with the primary coil.
3. *Flux leakage:* The useful (or main) flux is that which effectively links both coils. In practice, some of the flux will escape or otherwise fail to link both coils. The e.m.f.s produced by the leakage fluxes are proportional to (and lead the fluxes by) 90°. The effect of flux leakage may be likened therefore to having an additional inductive coil in series with the primary and secondary coils. In practice, the flux leakage loss is usually lumped together with the iron loss.

Determination of Transformer Losses

Open-Circuit Test. The secondary coil is on open-circuit and the full-rated voltage is applied to the primary winding. The transformer takes a small no-load current to supply the iron loss in the core, and the copper losses are essentially zero. Since

ELECTRICAL AND ELECTRONIC PRINCIPLES

the normal voltage and frequency are applied, a wattmeter connected to the primary side will give a measure of the iron loss. The iron loss can then be taken as a constant, irrespective of the load.

Closed-Circuit Test. With the secondary winding short-circuited the transformer requires only a small input voltage to circulate the full-load current. The wattmeter on the primary side then gives an indication of the full-load copper losses. If the load is expressed as a fraction of the full load, the copper losses at reduced loads are proportional to the load squared. At half load, for example, the copper losses are one quarter of the full-load value.

Referred Values

In dealing with transformers it is usual to base all calculations on one side of the transformer. Parameters on the neglected side are accounted for by referring them over to the side on which the calculation is to be based. The transformation ratio is used to scale the equivalent values. For example, the copper loss on the secondary side, $I_2^2 R_2$, can be referred to the primary side through the relation

$$I_2'^2 R_2' = I_2^2 R_2 \tag{4.93}$$

where the prime denotes the referred values. Using equation (4.90), the referred resistance becomes

$$R_2' = \{N_1/N_2\}^2 R_2 \tag{4.94}$$

Thus, equation (4.94) gives an equivalent resistance, R_2', in the primary side that accounts for the actual resistance, R_2, of the secondary winding. Reactances may be similarly referred to one or other side of the transformer for calculation purposes.

Transformer Efficiency

The transformer efficiency, as with any machine, is the ratio of the output power to the input power. The difference between the output and the input power is the sum of the losses, which, for the case of a transformer, is the copper and the iron losses, i.e.,

$$\eta = \frac{\text{Output}}{\text{Input}} = \frac{\text{Output}}{\text{Output} + \text{copper loss} + \text{iron loss}}$$

Therefore,

$$\eta = \frac{V_2 I_2 \cos(\theta_2)}{V_2 I_2 \cos(\theta_2) + I_2^2 \cdot R_e + F_e} \tag{4.95}$$

Note that R_e represents an equivalent resistance, which consists of the resistance of the secondary winding and that of the primary winding referred over to the secondary side, i.e.,

$$R_e = R_2 + (N_2/N_1)^2 R_1 \tag{4.96}$$

The iron loss, F_e, is assumed to be constant and $\cos(\theta_2)$ is the load power factor, also assumed constant.

By dividing the numerator and the denominator of equation (4.95) by I_2, then differentiating the denominator with respect to I_2, and equating the result to zero, it can be shown that for maximum efficiency, $I_2^2 \cdot R_e = F_e$. Maximum transformer efficiency then occurs when the copper loss is equal to the iron loss. The general efficiency characteristics for a transformer are shown in Figure 4.61.

Equation (4.95) also shows that the output will be influenced by the load power factor. At unity power factor the output (and hence also the efficiency) is maximized. As the power factor decreases, the transformer efficiency also reduces proportionally.

Voltage Regulation

As the load current drawn from a transformer is increased, the terminal voltage decreases. The difference between the no-load output voltage and the output voltage on load is called the *regulation*. The percentage regulation is defined as

$$\frac{\text{No-load voltage} - \text{load voltage}}{\text{No-load voltage}} \times 100 \tag{4.97}$$

Figure 4.62 shows the two voltages in terms of phasors referred to the primary side. In the figure V_1 is the no-load primary voltage and V_2' is the secondary-side voltage referred to the primary. R_e and X_e denote the equivalent resistance and reactance, respectively, including the referred secondary values. Since δ is very small, then, to a reasonable approximation,

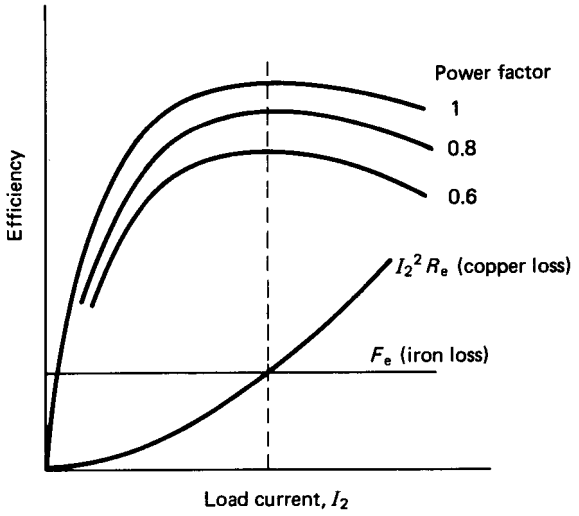


FIGURE 4.61 Transformer efficiency characteristics.

ELECTRICAL AND ELECTRONIC PRINCIPLES

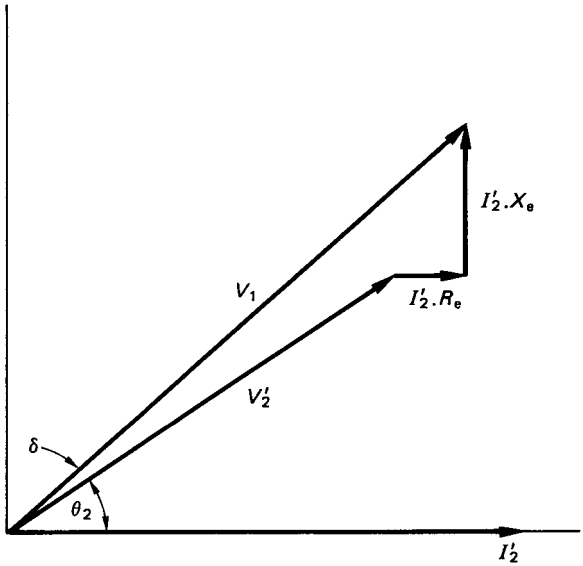


FIGURE 4.62 Phasor diagram for a transformer with a lagging power factor load current.

$$V_1 = V'_2 + I'_2 \cdot R_e \cdot \cos(\theta_2) + I'_2 \cdot X_e \cdot \sin(\theta_2) \tag{4.98}$$

The percentage regulation is therefore

$$(100/V_1)[I'_2 R_e \cos(\theta_2) + I'_2 X_e \sin(\theta_2)] \tag{4.99}$$

Equation (4.99) is based on the assumption that the load power factor is lagging, and this is the normal situation. If, however, the load power factor is leading, the plus operator within the term in square brackets must be replaced with a minus operator.

Three-Phase Transformers

Modern large three-phase transformers are usually constructed with three limbs as shown in Figure 4.63. In the figure the primary windings are star-connected and the secondary windings are delta-connected. In fact, the primary and secondary windings can be connected in any pattern, depending upon the conditions under which the transformer is to operate. It is important, however, to know how the three-phase transformer is connected, particularly when two or more transformers are to be operated in parallel. It is essential, for instance, that parallel operation transformers belong to the same main group and that their voltage ratios are perfectly compatible.

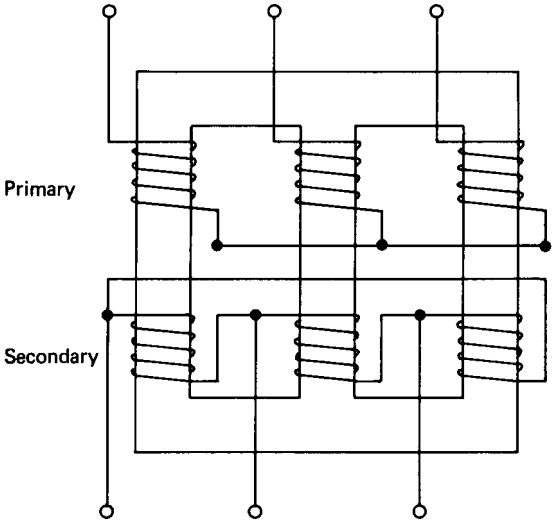


FIGURE 4.63 Three-phase transformer.

Auto-Transformers

The auto-transformer is characterized by having part of its winding common to both the primary and secondary circuits (see Figure 4.64). The main application of auto-transformers is to provide a variable voltage, and it is used, for example, to limit the starting current drawn by an induction motor.

A major disadvantage of the auto-transformer is that the primary and secondary windings are not electrically isolated from one another. This presents a serious risk of shock, and therefore auto-transformers cannot be used for interconnecting high- and low-voltage systems.

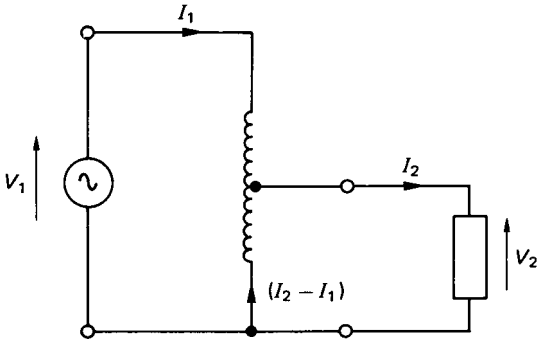


FIGURE 4.64 Auto-transformer.

4.3 ANALOG AND DIGITAL ELECTRONICS THEORY

The Bipolar (or Junction) Transistor

The term *transistor*, derived from *transfer resistor*, describes a device that can transfer a current from a low-resistance circuit to a high-resistance one with little change in current during the process. The junction transistor consists of two *pn* diodes formed together with one common section, making it a three-layer device (see Figure 4.65).

Current flow in the transistor is due to both electron and hole conduction. The common central section is referred to as the *base* and is typically of the order of 25 μm in length. Since the base can be made either an *n*-type or a *p*-type semiconductor, two basic configurations are possible. These are the *npn* and the *pnp* types, as illustrated in Figure 4.65. The two other terminals are called the *emitter* and the *collector*. An arrowhead is traditionally shown between the emitter and the base to indicate the conventional direction of the current flow in that part of the circuit.

A brief description of the physical operation of the junction transistor can be made with respect to the *npn* type. The mode of operation of the *pnp* type is the same as that of the *npn* type, except that the polarities of all applied voltages, currents, and charge carriers are reversed.

In normal use, as a linear amplifier, the transistor is operated with the emitter to base junction forward biased and the collector to base junction reversed biased. For the *npn* transistor, the emitter is therefore negative with respect to the base while the collector is positive with respect to the base (see Figure 4.66). The junction $n_1 p$ is forward biased such that the free electrons drift from n_1 to p. On the other hand, junction $n_2 p$ is reverse biased and it will collect most of the electrons from n_1. The electrons which fail to reach n_2 are responsible for the current at the base terminal, I_B. By ensuring that the thickness of the base is very small and that the concentration of impurities in the base is much lower than either that of the emitter or the collector, the resultant base current will be limited to some 2% of the emitter current. The basic transistor characteristic is therefore

$$I_C = h_{FB} \cdot I_E \tag{4.100}$$

where I_C is the collector current, I_E is the emitter current, and h_{FB} is the current

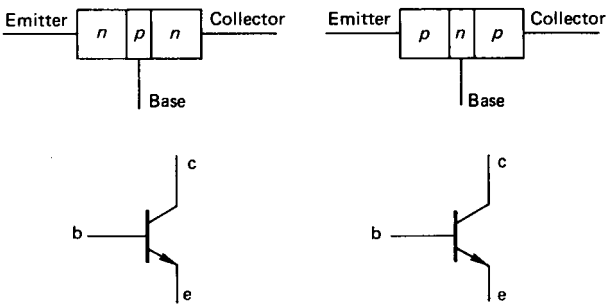


FIGURE 4.65 *npn* and *pnp* junction transistors.

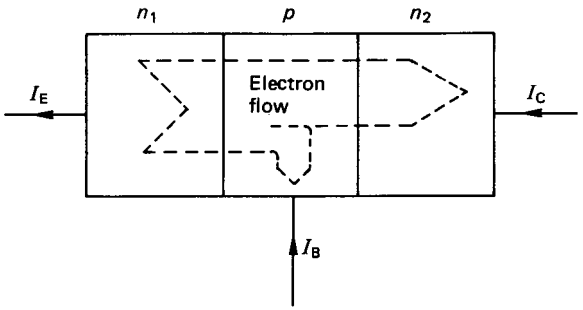


FIGURE 4.66 *npn* transistor in normal operation.

gain between the collector and the emitter. Normally, h_{FB} would range between 0.95 and 0.995 for a good-quality transistor.

Common-Base Characteristics

Figure 4.67 shows an *npn* transistor connected in a circuit to determine its static common-base characteristics. The emitter current, I_E is kept constant by varying R_1, and a range of values for I_C is imposed by varying R_2. The value of V_{CB}, the collector-base voltage, is noted. The test is repeated for another fixed value of I_E and the results are as depicted in Figure 4.68.

It is found that over a wide range of collector-base voltages the collector current is essentially independent of the collector-base voltage. This is because most of the electrons entering the *npn* junction are attracted to the collector. In effect, the collector circuit has a very high impedance and acts as a constant current source. The actual value of the collector current is determined by the emitter current and the two are related through equation (4.100), which is the common-base characteristic. The general characteristics also show that the collector-base voltage must be reversed (i.e., collector negative with respect to base) in order to reduce the collector current to zero. Finally, at a high collector-base voltage, the collector current increases rapidly in consequence of the Zener effect. The same characteristics are

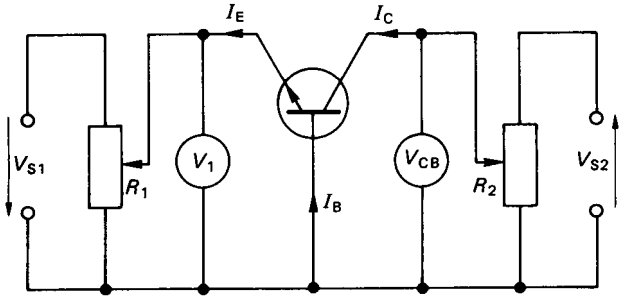


FIGURE 4.67 *npn* transistor in common-base circuit.

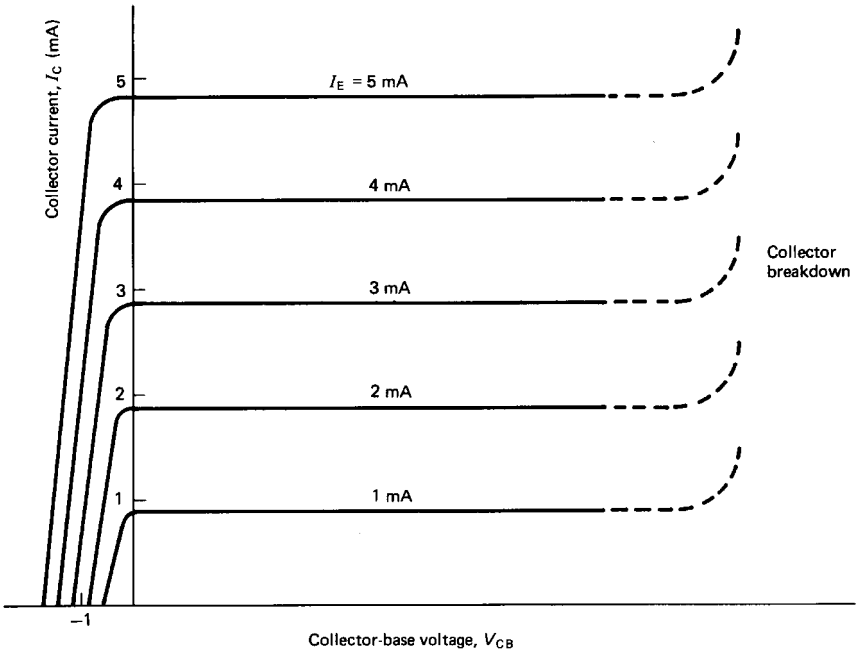


FIGURE 4.68 Common-base characteristics.

observed with the *pnp* transistor, except that the signs are in the reverse direction to that shown in Figure 4.67.

Common-Emitter Characteristics

Figure 4.69 shows the *npn* transistor with its emitter terminal connected to both the base current and the collector current circuits. Using the same test procedure as before, the resulting characteristics are as shown in Figure 4.70. The first sig-

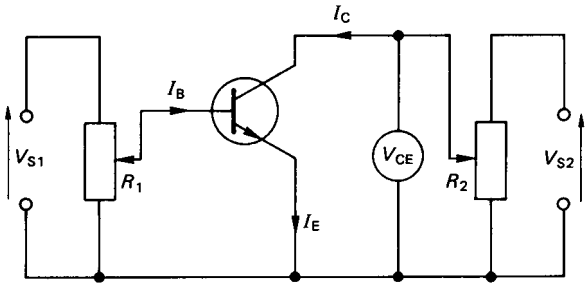


FIGURE 4.69 *npn* transistor in common emitter circuit.

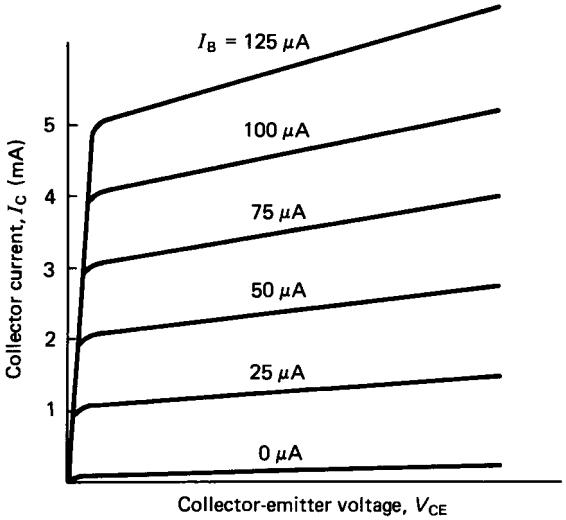


FIGURE 4.70 Common-emitter characteristics.

nificant observation is that the collector-emitter voltage, V_{CE}, must be positive to produce a positive collector current. At low values of V_{CE} the collector current I_c is also low, but when V_{CE} exceeds the so-called knee voltage the characteristic assumes a linear relationship. The gradient of the linear region is generally much higher than that for the common-base configuration and the collector impedance is therefore lower than that for the common-base circuit. When the base current is zero the collector current still has a positive finite value.

The common-emitter characteristic is generally written as

$$I_C = h_{\mathrm{FE}} \cdot I_B \tag{4.101}$$

where h_{FE} is the current gain between the collector and base. Application of Kirchhoff's first law to the common-emitter circuit gives

$$I_E = I_C + I_B$$

Using equation (4.100) and eliminating I_E, it can be shown that

$$\frac{I_C}{I_B} = \frac{h_{FB}}{1 - h_{FB}} = h_{FE} \tag{4.102}$$

For a transistor with a steady-state current gain in common base of 0.95 the common-emitter gain is

$$h_{FE} = \frac{0.95}{1 - 0.95} = 19$$

If, due to some temperature effect, h_{FB} undergoes a minor change to, say, 0.96, the new value of h_{FE} becomes 24. It is clear therefore that the common-emitter gain, h_{FE}, is much more sensitive to small-order effects than the common-base gain, h_{FB}.

ELECTRICAL AND ELECTRONIC PRINCIPLES

For a *pnp* transistor the characteristics of the common-emitter circuit are the same, except that the polarities of all voltages and currents are again in reverse order to that shown in Figure 4.69.

The Transistor in a Circuit

In most practical applications transistors are operated in the common-emitter mode where the emitter terminal forms the common connection between the input and output sections of the circuit (see Figure 4.71).

The transistor collector characteristics are shown again in Figure 4.72. The load line for the resistor, R_C, is superimposed and the operating point is given by the intersection of the load line with the collector characteristic. The operating point will therefore be dependent on the base current, since this controls the collector characteristic. Also shown in Figure 4.72 is the maximum power dissipation curve (broken line), which represents the locus of the product of collector current and collector-emitter voltage. The maximum power dissipation curve represents a physical limitation and the operating point must be constrained to lie below the curve at all times.

As the base current is reduced the operating point moves down the load line. When I_B reaches zero the collector current will be minimized and the transistor is said to be *cut off*. Alternatively, as the base current is increased, the operating point moves up the load line and eventually reaches a maximum value at which the transistor is said to be *bottomed*, or *saturated*. When saturated, the collector-emitter voltage is at a minimum of about 0.1–0.2 V and the collector current is a maximum. The two extremes between cut-off and saturation represent a very high and a very low impedance state of the transistor, respectively. These extremes have great practical application to rapid, low-power switching, and transistors operating between cut-off and saturation are frequently used in digital electronics circuitry. The low-

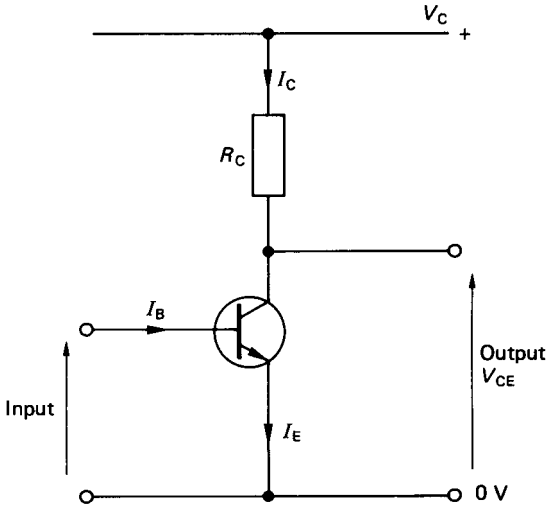


FIGURE 4.71 *npn* transistor in a practical common-emitter circuit.

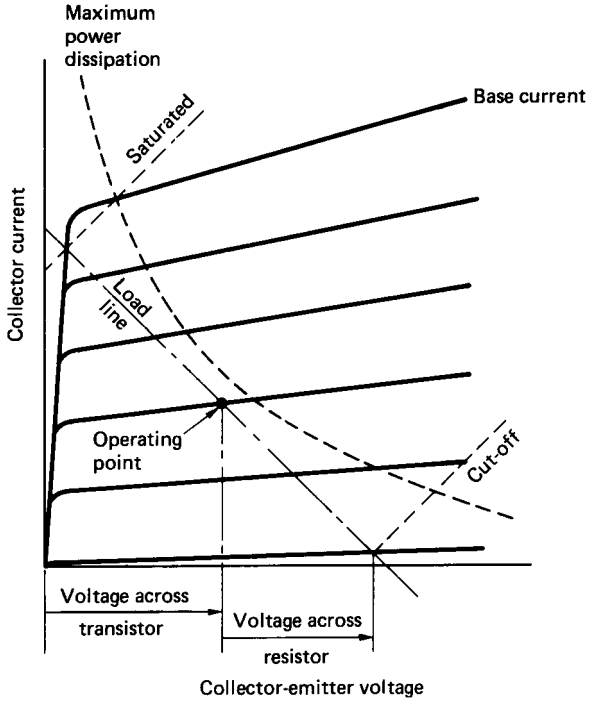


FIGURE 4.72 Common-emitter characteristics with superimposed load line.

impedance state represents a switch closed (or on) and the high-impedance state represents the switch open (or off). When operating as a linear current amplifier the operating point is ideally located in the center of the active region of the characteristic.

The analysis of circuits involving transistors is conveniently dealt with by representing the transistor in terms of an equivalent circuit and using the conventional current flow direction from positive to negative. Consideration of the charge carriers (i.e., holes or electrons) is only necessary to describe the internal physical operation of the transistor. Fully detailed worked examples are particularly informative, and these are usually provided in all standard textbooks on electrical and electronics technology.

The Field Effect Transistor (FET)

Field effect transistors (FETs) are a much more recent development than bipolar transistors and operate on a substantially different mechanism in achieving signal amplification. Operationally, FETs are voltage-controlled devices, as opposed to the bipolar transistor, which is current-operated. FETs are often described as unipolar, since conduction in the FET is the result of only one predominant charge carrier.

The junction field effect transistor (JFET) consists of a thin bar of semiconductor that forms a channel between its two end-connections, which are referred to as the *source* and the *drain*. If the semiconductor used in the construction of the FET is *n*-type, the device is called an *n*-channel. Conversely, a FET made from a *p*-type semiconductor is called a *p*-channel device.

If the channel consists of a uniformly doped semiconductor, the conductivity will be constant and the FET will function as a linear resistor. By introducing two opposite type semiconductor layers on either side of the channel, the effective thickness of the channel (and hence the current flow) can be controlled. The opposite type layers are denoted as *gates*, and in normal operation they are reverse biased by a d.c. potential, V_{GS}, referred to as the *gate source voltage*. The reverse bias ensures that no current can flow between the two gates and the gate inputs have an extremely high impedance. By using a lightly doped semiconductor for the channel the gate depletion layer, which is determined by V_{GS}, can be made to extend well into the channel width. This controls the resistance of the path between the source and the drain. The general characteristics of such a FET are shown in Figure 4.73.

For a given value of V_{GS} an increase in drain-source voltage from zero initially gives a linear rise in drain current. Further increases in drain-source voltage result in a so-called pinch-off in the drain current, which then becomes independent of the drain-source voltage. Finally, at a particular limiting value of drain-source voltage a breakdown is initiated. The similarities between Figures 4.73 and 4.70 or 4.72 are clear, and it is evident therefore that the bipolar junction transistor and the unipolar FET can perform essentially a similar function in any given application. Many other types of transistor (for example, the metal oxide semiconductor FET, or MOSFET) use alternative means to control the resistance of the source to drain channel. The general characteristics of these devices, however, are all very similar to that shown in Figure 4.73.

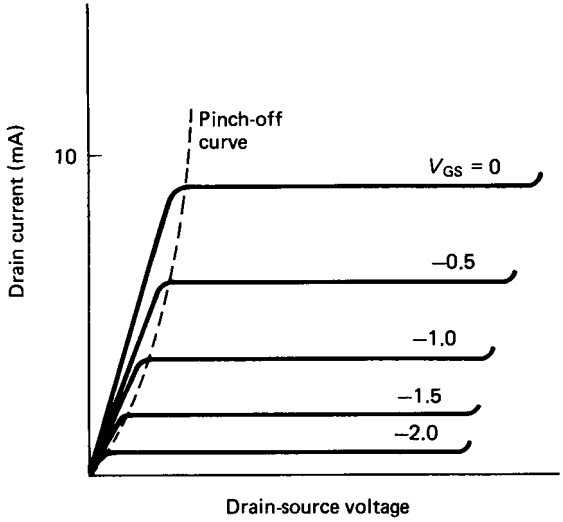


FIGURE 4.73 Characteristics of a FET.

Integrated Circuits

While transistor-based amplifiers are still found as individual elements in working circuits, the trend is towards the development of integrated circuits, where all the circuit elements are housed within a single silicon wafer. MOSFET technology is predominant in this area, since the number of components on a single silicon chip can be packed up to twenty times more densely than with bipolar technology.

The integrated circuit components include diodes and transistors, which may be either bipolar junction type or FETs. Resistors can be deposited on top of the wafer in the form of tantalum, which is a poor conductor, or built into the wafer as pinch resistors, which are partially turned-off FETs. Capacitors can also be produced within the silicon wafer. Capacitive elements may be formed when a *pn* junction diode is reverse biased. The *p*- and *n*-type layers form the plates of the capacitor and the carrier-depletion layer acts as a dielectric. The capacitance is, however, limited to a few picofarads. There is no microelectronic equivalent for an inductor, but most circuit designs can generally avoid the requirement for coiled inductive elements.

The Thyristor

Both the bipolar transistor and the FET can be utilized for switching operations. These devices, however, are usually associated with low-power switching. For switching very large currents and voltages a special device called a *thyristor* (formerly known as a silicon-controlled rectifier, SCR) is normally used. The thyristor is a four-layer, unidirectional semiconductor device with three connections referred to as the anode, cathode, and control gate (see Figure 4.74).

The current flow is from the anode to the cathode only and, with the cathode positive with respect to the anode, the device has a very high impedance. Under normal circumstances the thyristor will fail to conduct current in any direction. If a voltage is applied such that if the thyristor were a diode it would conduct in the forward-biased direction, then application of a very small current between the gate and the cathode will cause the thyristor to abruptly change from nonconducting to

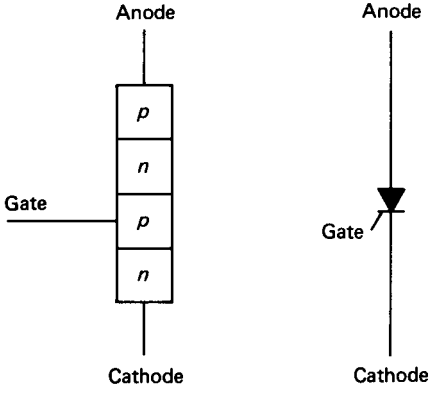


FIGURE 4.74 Thyristor device and circuit symbol.

conducting mode. The turn-on is rapid (within a few microseconds) and, once turned on, the thyristor will remain on, even if the gate current is removed.

Once triggered into conduction, the thyristor will turn off again only when the current flowing through it is reduced below a critical value. This minimum conducting current is called the *holding current* and may range between a few microamps to a few tens of milliamps. Thyristors are additionally connected in series with a resistor, which serves to limit the current to a safe value. The basic thyristor function is that of a power-control device, and thyristors are used extensively for switching mains electricity and as speed controllers for d.c. motors.

The Triac

The triac (or bidirectional thyristor) is similar in operation to the thyristor but differs in that it can be switched into conduction in either direction. In essence, the triac is equivalent to two thyristors mounted back to back. Triacs find application to switching in full-wave alternating power supplies.

Amplifiers

In general, electronic amplifiers are supplied with energy from a d.c. source. An input signal to the circuit controls the transfer of energy to the output, and the output signal should be a higher-power version of that supplied to the input. The amplifier does not, however, function as some magical source of free energy. The increased power across the amplifier is invariably drawn from the supply.

The term *amplifier* is actually a shortened form of the complete specification *voltage amplifier*. This has transpired because most amplifiers are intended to magnify voltage levels. Any other type of amplifier is normally prefixed with the name of the quantity which is amplified (e.g., current amplifier, charge amplifier or power amplifier).

Amplifiers may be broadly classified with reference to the frequency range over which they are designed to operate. In this respect there are two general categories: *wide-band* and *narrow-band* amplifiers. The names are self-explanatory in that the wide-band amplifier exhibits a constant power gain over a large range of input signal frequencies. The narrow-band (or *tuned*) amplifier, on the other hand, provides a power gain over a very small frequency range. This gain is usually expressed in decibels and is defined by equation (4.58).

The bandwidth of an amplifier is used in the same context as before to define the operating frequency range. In this respect the -3 dB amplitude ratio is used consistently to define the upper and lower input signal frequencies at which the power transferred across the amplifier is halved.

Using the system model, the amplifier can be represented as shown in Figure 4.75. In the figure the amplifier is shown enclosed within the broken lines. There is a single input, a single output, and one common connection. The amplifier also features an internal input impedance, shown as resistance R_i, and an internal output impedance, shown as resistance R_o. In fact, the input and output impedances could have both inductive and capacitive components as well as the simple resistances, as shown in the figure.

Connected to the input stage of the amplifier is a voltage source, V_s, and its associated internal resistance, R_s. This could be taken to represent some form of transducer having a low-voltage output in the millivolt range. At the output stage

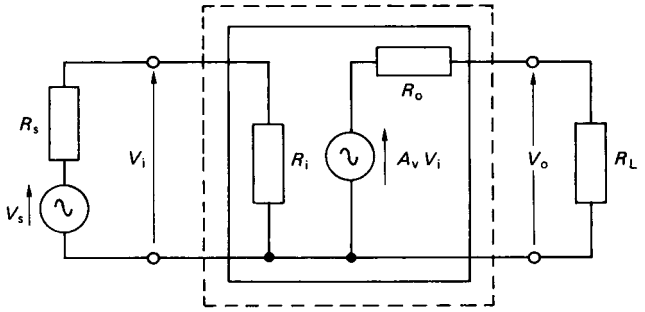


FIGURE 4.75 System representation of an amplifier.

the amplifier acts as a voltage source where A_v is the voltage gain. The output is shown connected to an external load, R_L, which might be considered to be a recording instrument such as a digital voltmeter.

Considering the input stage, it may be shown, from Ohm's law, that

$$V_i = \frac{V_s}{(1 + R_s/R_i)} \tag{4.103}$$

This equation indicates that the voltage applied to the amplifier input stage, V_i, will approach the source voltage, V_s, only when R_i tends to infinity. The amplifier should therefore ideally have a very large input impedance to prevent serious voltage attenuation at the input stage. By a similar argument, the output impedance, R_o, should be very small in comparison to the load resistance, R_L, for maximum voltage gain.

Effect of Feedback on Amplifiers

The amplifier illustrated in Figure 4.75 is specified by its input and output impedances and its open-circuit gain, A_v, this gain being obtained when the load resistance is infinite. These parameters are not fixed but will vary with ambient temperature, power supply voltage, and variation with age. The adverse effects of these variabilities can be minimized through the application of negative feedback.

One particular method of obtaining negative feedback is the so-called series voltage method (see Figure 4.76). The feedback system in Figure 4.76 is applied by connecting a potentiometer across the output terminals and tapping off a fraction, β, of the output signal. This fraction is connected in series with the input and with a polarity which will always oppose the input signal. Assuming both that the input impedance of the amplifier is very large in comparison to the internal resistance of the voltage source and that the resistance of the potentiometer is very large in comparison with the output impedance of the amplifier,

$$V_i = V_s - \beta \cdot V_o \tag{4.104}$$

Since $V_o = A_v \cdot V_i$, then $V_o = A_v \cdot V_s - \beta \cdot A_v \cdot V_o$. The overall gain of the system with feedback, A_f, is

ELECTRICAL AND ELECTRONIC PRINCIPLES

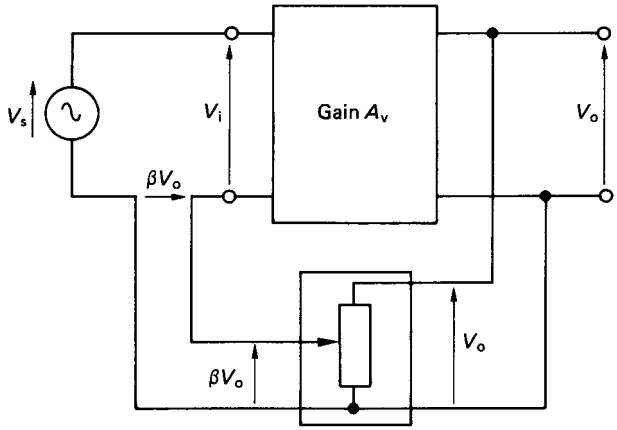


FIGURE 4.76 Series voltage method of negative feedback.

$$A_f = \frac{V_o}{V_s} = \frac{A_v}{1 + \beta A_v} \tag{4.105}$$

Equation (4.105) shows that the feedback loop has reduced the original gain by the factor $(1 + \beta \cdot A_v)$. If, in addition, the original gain A_v was in itself very large such that $\beta \cdot A_v >> 1$, then

$$A_f = A_v/(\beta \cdot A_v) = 1/\beta \tag{4.106}$$

Under the above circumstances the overall gain of the system with feedback is essentially dependent only on the feedback fraction, β. Any changes, therefore, which alter the original gain, A_v, of the amplifier will not affect the gain of the overall system with feedback.

Consideration of the system with and without the feedback loop shows that the effect of series voltage negative feedback is to increase the input resistance by the factor $1 + \beta \cdot A_v$ and to reduce the output resistance by the same factor. Both these effects are of benefit to the operation of the system. These comments refer only to a negative feedback system using the series voltage method. Other methods of obtaining negative feedback can be used, including series current feedback, shunt current, and shunt voltage feedback. These alternative methods have different effects on the overall gain and on the input and output impedances of the amplifier.

Noise and Distortion in Amplifiers

Noise is inherently present in all electronic amplifier systems. The source of the noise is due to a number of effects, which include the random charge movements within solid-state devices, thermoelectric potentials, electrostatic and electromagnetic pick-up, and interference from the power supply. The noise is fairly evenly distributed across the whole frequency spectrum and appears superimposed upon the amplified input signal. If the noise is generated at the input stage of the amplifier

then the signal-to-noise ratio is not improved by feedback. This ratio can, however, be improved if an intermediate amplifying stage, free from noise effects, can be included in the system.

Distortion is another undesirable feature which arises when the amplifier input/output (or transfer) characteristic deviates from an ideal linear relationship. If the transfer characteristic is linear then the output signal will be a faithful amplified replica of the input. A nonlinear characteristic will give a distorted output, and a nonsinusoidal output will be generated from a sinusoidal input. Distortion is usually associated with a high level of input signal, which overextends the linear operating range of the amplifier.

Amplifier Frequency Response

The frequency response of an amplifier is usually illustrated as a plot of the gain in decibels against the input signal frequency. The graph is called a Bode plot and the phase relationship between the output and input is also shown for completeness. Figure 4.77 illustrates the frequency characteristics for a typical wide-band amplifier.

In the figure the bandwidth between the −3 dB cut-off frequencies is determined either by the characteristics of the active devices used to make the amplifier or by

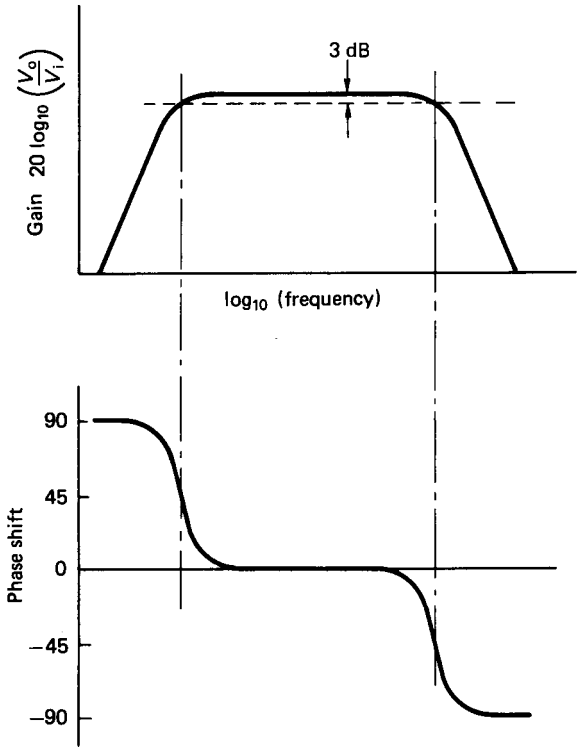


FIGURE 4.77 Frequency response for a wide-band amplifier.

other frequency-dependent elements in the amplifier circuit. The upper limiting frequency is fixed by the charge transit time through the active device. In practice, any stray capacitance, which is manifested as a parallel capacitance in the system, will considerably reduce the upper limiting frequency. In theory, the active device will respond to frequencies down to 0 Hz but, because of the variabilities due to aging effects, a lower cut-off frequency is often imposed by including series capacitors on one or both of the input connections.

Positive Feedback and Stability

In Figure 4.76 a negative feedback signal is produced by using a series voltage. If the phase of the series voltage were changed such that the feedback signal augmented the input, then the nature of the feedback loop would become positive. With this positive feedback system the overall gain would then become

$$A_f = A_v/(1 - \beta \cdot A_v) \tag{4.107}$$

Positive feedback therefore increases the overall system gain. If indeed the product $\beta \cdot A_v$ is made equal to unity then the overall gain becomes infinite. Positive feedback, however, is inherently unstable, since the output signal tends to increase indefinitely in an uncontrolled manner. Systems with positive feedback are found, nonetheless, in oscillator circuits where the amplifier produces its own input signal via a positive feedback loop.

The Operational Amplifier

Modern amplifier systems rely less on discrete active devices such as transistors and much more on the large range of integrated circuits which are readily available. One of the most prevalent operational amplifiers based on integrated circuit technology is the generic type SN72741, or, as it is often abbreviated, the 741. The 741 consists of 20 bipolar transistors, 11 resistors and one capacitor. Figure 4.78 shows the usual representation of the 741 operational amplifier (op-amp).

The internal circuitry is quite complex but is conveniently reduced to the basic schematic form shown in the figure. The operational amplifier consists of an output, an inverting input, and a noninverting input. The circuitry in addition, requires a bipolar power supply, which may range between ± 3 to ± 18 V. There is also provision for an offset null on connection pins 1 and 5. For the most part the offset pins can be ignored.

The operational amplifier has a high input impedance, a low output impedance, and a very high open-circuit gain, A. Ideally, the gain should be infinite. The bandwidth should also be infinite but the 741, for example, has an effective bandwidth limited between 0 Hz and about 1 MHz.

For operational amplifiers such as the 741 there are a number of standard circuits which are used routinely to perform specific functions.

Inverting Amplifier. Figure 4.79 shows an op-amp wired up for an inverted output. The input current, i_1, is given as V_1/R_1 and, because the amplifier input impedance is very high, the current flowing into the input terminal is approximately zero. This is equivalent to having the potential available at point E equal to zero. For this reason, E is referred to as a *virtual earth*. From Kirchhoff's first law, it is apparent that $i_1 = -i_2$. Thus, $V_1/R_1 = -V_o/R_2$, and the gain can be written as

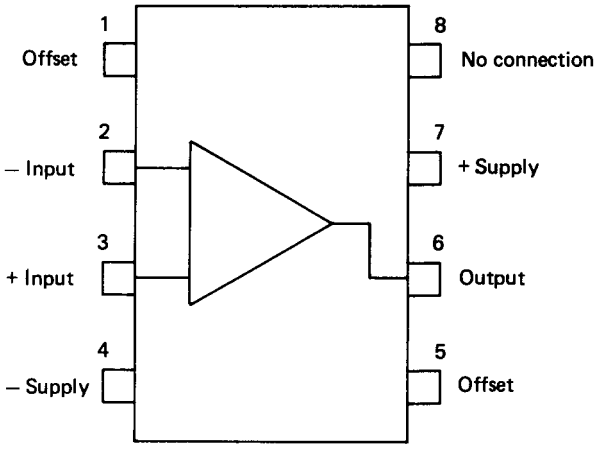


FIGURE 4.78 SN72741 operational amplifier.

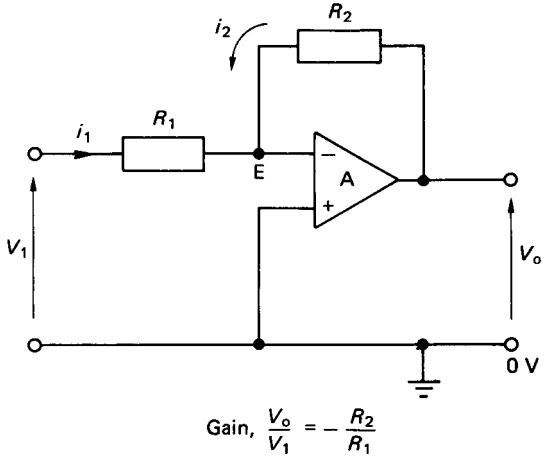


FIGURE 4.79 Inverting amplifier.

$$\frac{V_o}{V_1} = -\frac{R_2}{R_1} \tag{4.108}$$

Provided the open-circuit gain of the amplifier is very high, the overall gain with this negative feedback system is given by the ratio of the two external resistors and is independent of the open-circuit gain.

Unity Gain Amplifier. Figure 4.80 depicts a unity gain amplifier in which no external resistors are wired into the circuit. The unity gain amplifier is also known as a voltage follower or a buffer amplifier. This type of amplifier circuit is often used in instrumentation systems where the internal resistance of a voltage-

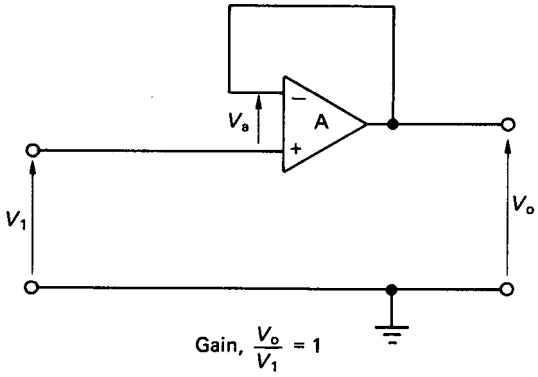


FIGURE 4.80 Unity gain amplifier.

generating transducer and that of the voltage-recording instrument are so poorly matched that the transducer voltage is seriously attenuated. This situation arises when the transducer internal resistance is large in comparison to that of the recording instrument. Since the buffer amplifier has a large input impedance and a low output impedance, it can be interfaced between the transducer and the recording instrument to provide optimum impedance matching. This gives a low source impedance and high destination impedance between both the transducer and amplifier and also between the amplifier and the instrument.

Summing the voltages round the amplifier in Figure 4.80 gives

$$V_1 + V_a = V_o$$

Since the internal impedance of the amplifier is very large then V_a is effectively zero and the gain is

$$V_o/V_1 = 1 \tag{4.109}$$

Noninverting Amplifier. Figure 4.81 shows the operational amplifier connected up for a non-inverting output. Assuming that the currents through resistors R_1 and R_2 are equal and that point E is a virtual earth,

$$\frac{V_i}{R_1} = \frac{V_o - V_i}{R_2}$$

Hence,

$$\frac{V_o}{V_i} = \frac{R_2 + R_1}{R_1}$$

Since E is a virtual earth, then $V_i = V_1$ and

$$\frac{V_o}{V_1} = \frac{R_2 + R_1}{R_1} \tag{4.110}$$

If, in addition, $R_2 >> R_1$,

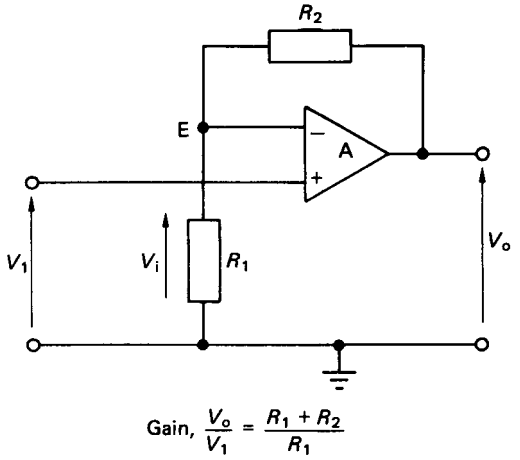


FIGURE 4.81 Noninverting amplifier.

$$\frac{V_o}{V_1} = \frac{R_2}{R_1} \tag{4.111}$$

Summing Amplifier. The summing amplifier is shown in Figure 4.82. As point E is a virtual earth, then

$$-i_4 = i_1 + i_2 + i_3$$

Therefore,

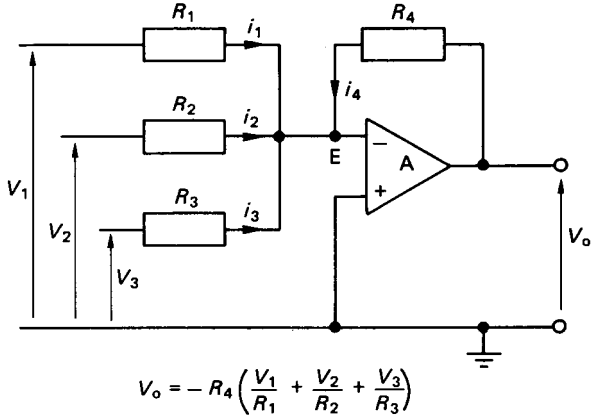


FIGURE 4.82 Summing amplifier.

$$\frac{V_o}{R_4} = -\left[\frac{V_1}{R_1} + \frac{V_2}{R_2} + \frac{V_3}{R_3}\right]$$

or

$$V_o = -R_4\left[\frac{V_1}{R_1} + \frac{V_2}{R_2} + \frac{V_3}{R_3}\right] \tag{4.112}$$

If the resistances used in the circuit are all of equal value, the output voltage will be equivalent to the summation of all the input voltages and with a reversed sign. Subtraction of any of the voltages can be performed by reversing its polarity, i.e., by first passing the voltage through a unity gain inverting amplifier before it is passed on to the summing amplifier.

Integrating Amplifier. The integrating amplifier uses a capacitor, as opposed to a resistor, in the feedback loop (see Figure 4.83). The voltage across the capacitor is

$$1/C \int_0^t i_2 dt$$

Since E is a virtual earth, $i_1 = -i_2$. Therefore, $i_2 = -(V_1/R_1)$. The voltage across the capacitor, which is, in effect, V_o, is

$$V_o = -(1/C)\int_0^t (V_1/R_1)dt = -(1/CR_1)\int_0^t V_1 dt \tag{4.113}$$

Thus, the output voltage is related to the integral of the input voltage.

Apart from various mathematical processes, operational amplifiers are also used in active filtering circuits, waveform generation and shaping, as a voltage comparator, and in analog-to-digital (A/D) and digital-to-analog (D/A) conversion ICs.

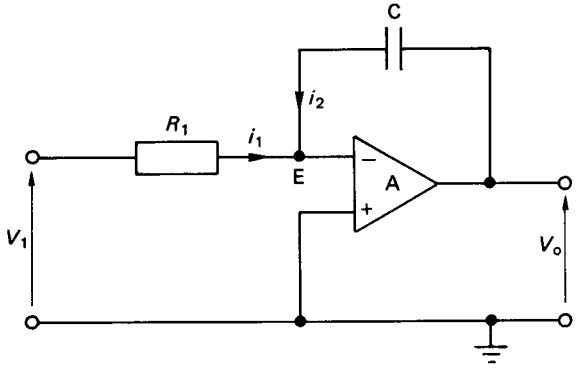


FIGURE 4.83 Integrating amplifier.

The Differential Amplifier

The differential amplifier (or subtractor) has two inputs and one output, as shown in Figure 4.84. The differential amplifier yields an output voltage which is proportional to the difference between the inverting and the noninverting input signals. By applying the superposition principle, the individual effects of each input on the output can be determined. The cumulative effect on the output voltage is then the sum of the two separate inputs. It can be shown therefore that

$$V_o = (R_2/R_1)[V_2 - V_1] \tag{4.114}$$

The input signals to a differential amplifier, in general, contain two components; the *common-mode* and *difference mode* signals. The common-mode signal is the average of the two input signals, and the difference mode is the difference between the two input signals. Ideally, the differential amplifier should affect the difference-mode signal only. However, the common-mode signal is also amplified to some extent. The common-mode rejection ratio (CMRR) is defined as the ratio of the difference signal voltage gain to the common-mode signal voltage gain. For a good-quality differential amplifier the CMRR should be very large.

Although particularly important to the differential amplifier, the common-mode rejection ratio is a fairly general quality parameter used in most amplifier specifications. The 741 op-amp has a CMRR of 90 dB, and the same signal applied to both inputs will give an output approximately 32,000 times smaller than that produced when the signal is applied to only one input line.

Instrumentation Amplifier

Instrumentation amplifiers are precision devices having a high input impedance, a low output impedance, a high common-mode rejection ratio, a low level of self-generated noise, and a low offset drift. The offset drift is attributable to temperature-

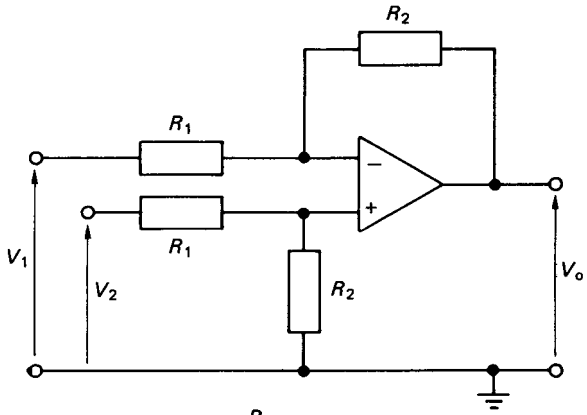


FIGURE 4.84 The differential amplifier.

dependent voltage outputs. Figure 4.85 shows the schematic representation of a precision instrumentation amplifier.

The relationship between output and input is

$$V_o = (R_4/R_3)[1 + 2(R_2/R_1)]V_1 \tag{4.115}$$

The first two amplifiers appearing in the input stage operate essentially as buffers, either with unity gain or with some finite value of gain.

A number of instrumentation amplifiers are packaged in IC form, and these are suitable for the amplification of signals from strain gauges, thermocouples, and other low-level differential signals from various bridge circuits. Kaufman and Seidman (1988) give a good practical coverage on the general use of amplifiers.

Power Supplies

Previously the use of *pn* junction diodes was illustrated as a means of a.c. voltage rectification. Both the half-wave and full-wave rectification circuits give outputs which, although varying with respect to time, are essentially d.c. in that there is no change in polarity. These rectification circuits provide a first stage in the production of a steady d.c. voltage from an a.c. power supply. Some further refinements are, however, added to the circuits to reduce the variation (or ripple) in the d.c. output voltage. The ripple factor can be greatly reduced by adding a reservoir capacitor, as shown in Figure 4.86, which is connected in parallel with the load.

A further reduction in ripple can be achieved by using a full-wave rectification circuit, since there are then twice as many voltage pulses and the capacitor discharge time is halved. The reservoir capacitor is, of necessity, quite large, and electrolytic capacitors are often used in this application. A leakage resistor is also frequently

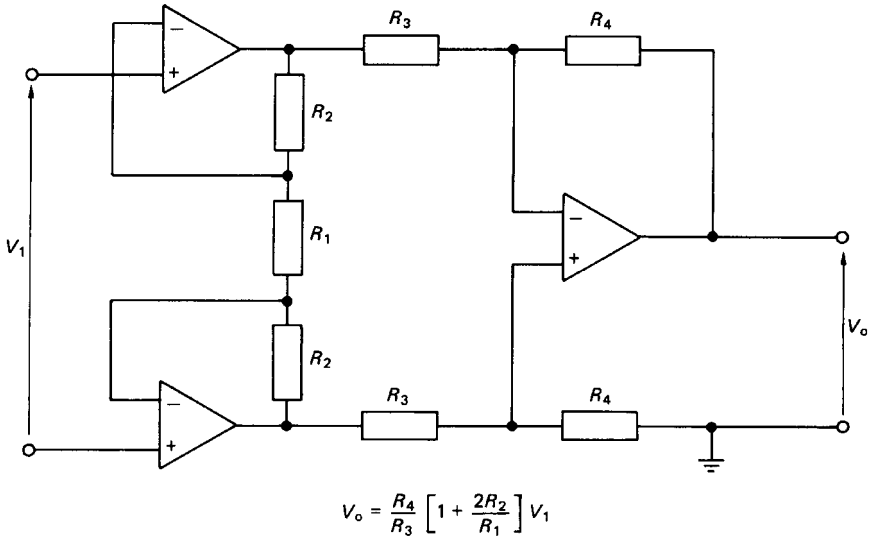


FIGURE 4.85 Precision instrumentation amplifier.

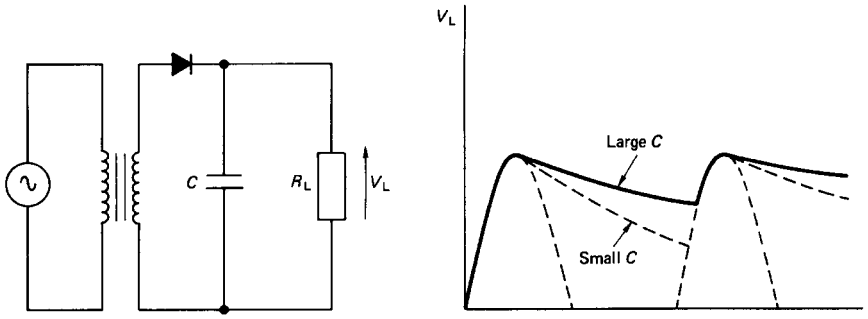


FIGURE 4.86 Half-wave rectification circuit with reservoir capacitor.

connected in parallel with the reservoir capacitor as a safety feature. In the event that the load is disconnected leaving the reservoir capacitor fully charged, the leakage resistor will dissipate the charge safely. For applications where the reservoir capacitor still cannot reduce the ripple to an acceptable level an additional ripple filtering circuit may be added.

Further enhancement might include a variable resistor either in series or in parallel with the load. The function of the variable resistor is to allow regulation of the voltage supplied to the load. The Zener diode discussed above is often used in this capacity to provide a stabilized voltage supply.

For high-power systems thyristors are used in place of diodes as the rectification element. The controlled conduction properties of thyristors allow close control to be exercised on the power supplied to the load.

Analog and Digital Systems

Thus far this chapter has been concerned with purely analog systems in which the circuit currents and voltages are infinitely variable. Digital systems, on the other hand, operate between one of two possible states, ''off'' or ''on'' (conducting or not conducting), and, as such, digital systems are essentially discrete in their operation.

Boolean Algebra

The basic rules of Boolean algebra are conveniently described with reference to simple manually switched circuits. In the binary notation a 0 denotes that the switch is off and a 1 that the switch is on. The 0 and 1 can also be taken to represent the absence or presence, respectively, of a voltage or a current.

Logical AND. Figure 4.87 shows a simple AND circuit. Obviously, the lamp will light only when both switches *A* AND *B* are closed. Writing this as a Boolean expression,

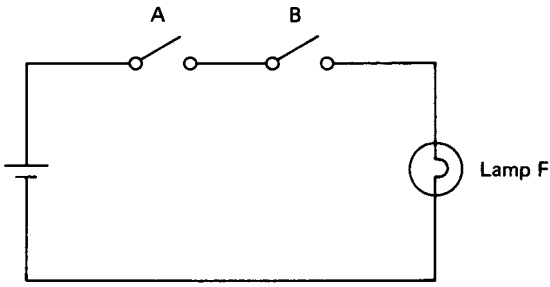


FIGURE 4.87 Simple AND circuit.

$$F = A \text{ AND } B \tag{4.116}$$

where A, B, and F are Boolean variables denoting switches A, B and the lamp, respectively. The logical operator AND is denoted by a dot, thus:

$$F = A.B$$

or

$$F = AB \tag{4.117}$$

Logical OR. Figure 4.88 shows the simple OR circuit. It is clear that the lamp will light in the OR circuit when either switch A OR switch B is closed. As a Boolean expression, the OR function is written

$$F = A \text{ OR } B$$

i.e.

$$F = A + B \tag{4.118}$$

The + sign is used to denote the logical OR and must not be confused with the arithmetical meaning.

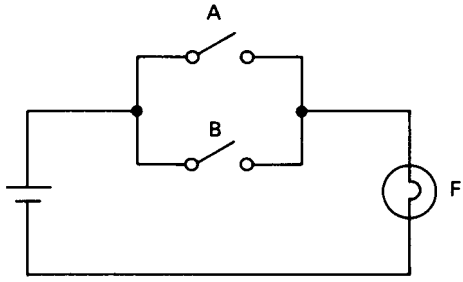


FIGURE 4.88 Simple OR circuit.

The AND and the OR are the basic logical functions, and quite complex switching circuits can be represented by them in Boolean form.

Logical NOT. The NOT function is the inverse complement, or negation of a variable. The negation of the variable A is $\overline{A}$. Thus, if $A = 1$, then $\overline{A} = 0$ and vice versa.

Logical NAND. The NAND function is the inverse of AND.

Logical NOR. Similarly, the NOR is the inverse of OR.

Exclusive OR. In Figure 4.88 it can be seen that the lamp will also light when both switches A and B are closed. The exclusive OR is a special function which does not enable an output when both switches are closed. Otherwise the exclusive OR functions as the normal OR operator.

The logical functions may also be represented in a tabular form known as a truth table. This table indicates the output generated for all possible combinations of inputs and is illustrated in Figure 4.89 for the AND and NAND operators with three inputs A, B, and C.

Using the basic logical functions, the Boolean identities are specified in Table 4.3. In this table a 0 can be taken to represent an open circuit, while a 1 represents a short circuit.

Using a truth table, it is easy to prove the validity of various logical expressions by evaluating both sides, e.g.,

A	B	C	$A.B.C$	$\overline{A.B.C}$
0	0	0	0	1
0	0	1	0	0
0	1	0	0	0
0	1	1	0	0
1	0	0	0	0
1	0	1	0	0
1	1	0	0	0
1	1	1	1	0

FIGURE 4.89 Truth table for AND and NAND operators with three imputs.

TABLE 4.3 Boolean Identities

$A + 0 = A$	$A \cdot A = A$
$A + 1 = 1$	$A + \overline{A} = 1$
$A \cdot 0 = 0$	$A \cdot \overline{A} = 0$
$A \cdot 1 = A$	$\overline{\overline{A}} = A$
$A + A = A$	$\overline{A + B} = \overline{A} \cdot \overline{B}$
	$\overline{A \cdot B} = \overline{A} + \overline{B}$

ELECTRICAL AND ELECTRONIC PRINCIPLES

$$A(B + C) = AB + AC$$

$$(A + B).(A + C) = A + BC$$

$$A + \overline{A}B = A + B \quad \text{etc.}$$

The first example shows that the parentheses may be removed by multiplying out, as in normal arithmetic. The second two examples have no arithmetic counterpart.

De Morgan's theorem states that, in any logical expression, AND can be replaced by OR and vice versa, provided that each term is also replaced with its inverse complement. The resulting expression is then the inverse of the original.

Example 1
From ABC we negate to

$$\overline{ABC} = \overline{A} + \overline{B} + \overline{C}$$

Hence,

$$ABC = \overline{\overline{A} + \overline{B} + \overline{C}}$$

Example 2
From $F = AB + CD$ we negate to

$$\overline{F} = (\overline{A} + \overline{B}) + (\overline{C} + \overline{D})$$

Applying De Morgan again,

$$F = \overline{(\overline{A} + \overline{B}).(\overline{C} + \overline{D})}$$

The equivalence of the original and the final expressions in the above two examples may be checked by using a truth table.

Digital Electronic Gates

The principles of Boolean algebra have been considered with respect to manually switched circuits. In modern digital systems the switches are formed with transistors for speed of operation, and they are generally referred to as *gates*. Over the years, various technologies have been developed in the manufacture of logic gates. The earliest forms of electronic gate were based on the unidirectional conduction properties of diodes. Diode logic gates have now been superseded by transistor—transistor logic gates (TTL) or the more recent CMOS family of logic gates.

The internal construction and operation of modern logic gates may be quite complex, but this is of little interest to the digital systems designer. Generally, all that the designer need know is the power supply voltages, the transient switching times, the ''fan out'' and the ''fan in.'' *Fan out* refers to the number of similar gates which can be driven from the output of one gate. *Fan in* denotes the number of similar gate outputs which can be safely connected to the input of one gate.

TTL. The TTL family is based on the bipolar junction transistor and was the first commonly available series of logic elements. TTL logic gates are rapid-switching devices (the SN7400, for example, takes just 15 ns to change state). The standard

power supply is 5 V with a low tolerance band of ±0.25 V. This, in turn, necessitates a reliable power supply regulation which is reasonably facilitated through the great variety of supply regulators which are now available in IC form. For the 5N74 series TTL ICs, the fan out is about 10.

A TTL-based system can draw quite large instantaneous loads on a power supply, and this can result in substantial interference spikes in the power lines. Since the spikes can upset the normal operation of the system, it is common practice to connect small capacitors directly across the power lines, as close to the TTL ICs as possible. One capacitor, 0.1–10 μF, per five ICs is sufficient in most instances.

TTL circuits are continually being improved, a major advance has been the introduction of the low-power Schottky TTL circuits. These use the same generic code numbers as the standard series, but have "LS" inserted before the type code (e.g., SN74LS00). The operating speed is about twice as high and the power consumption is about 20% that of the standard series. Schottky devices are, however, slightly more expensive.

CMOS. The problematic features of the power supply associated with the TFL family of logic devices have been largely responsible for the growth of its major competitor, CMOS. CMOS ICs are based on the field effect transistor and can operate off a range of power supply voltages between ±3 V to ±18 V. CMOS devices dissipate very little power, are very cheap, and are simple in operation. The fan out is about 50 and they have a far greater immunity to power supply noise. The noise immunity of CMOS devices means that there is no requirement for smoothing capacitors to the extent that they are generally found in TTL circuitry.

There are also some disadvantages associated with CMOS devices, the main one being that CMOS is slower than TTL, roughly about one-tenth of the equivalent TTL circuit. CMOS ICs are also very sensitive to electrostatic voltages. Manufacturers do build in some safety features to reduce the electrostatic sensitivity, but CMOS devices must still be handled with due care. Table 4.4 gives a brief comparison between TTL and CMOS devices.

Gate Symbols

Having defined a system output in terms of a Boolean expression, the actual circuit can be constructed using the required gates selected from the logic family chosen. Generally, the design will be centered round the more readily available NAND and NOR logic gates. In laying out a gate interconnection diagram, standard symbols are used to represent the individual gates. Unfortunately, no universal set of symbols

TABLE 8.4 Comparison between TTL and CMOS Devices

Property	TTL	CMOS
Power supply	5 V ± 0.25 V	3 V to 18 V d.c.
Current required	Milliamps	Microamps
Input impedance	Low	Very high
Switching speed	Fast—10 ns	Slow—300 ns
Fan out	10	50

ELECTRICAL AND ELECTRONIC PRINCIPLES

has emerged, and several systems are in current use. Figure 4.90 summarizes the most common gate symbol systems.

Logic Systems Using Simple Gates

A vending machine which dispenses either tea or coffee can serve as an illustrative example. The logic circuit may be realized using AND gates as shown in Figure 4.91.

The money input is common to both gates, and the system, although workable, has a minor fault in that if both buttons are pressed, after the money criterion is

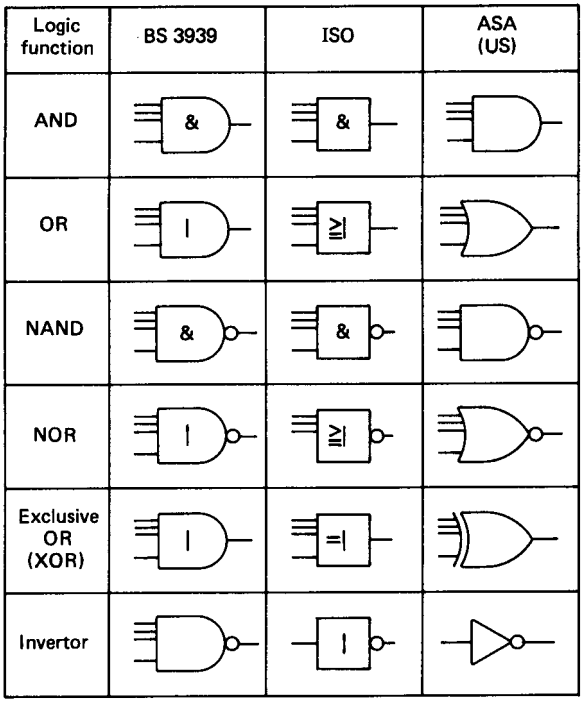


FIGURE 4.90 Gate symbol systems in current use.

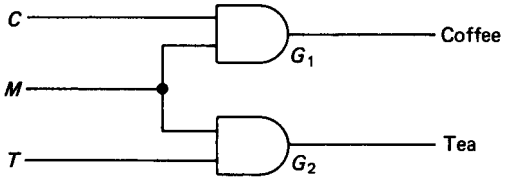


FIGURE 4.91 Logic circuit for drink-vending machine.

satisfied, then the output will be both tea and coffee. This fault can be designed out of the system by extending the logic circuit as shown in Figure 4.92.

The extended system incorporates a NAND gate and an additional AND gate. If both buttons are now pressed, then the output from G_3 will be 0. With the output 1 from G_1, the output from G_4 will be 0 and the machine will dispense tea. On either button being pressed on its own and the money input criterion being satisfied, the correct drink will be output. The operation of the extended system is verified in the truth table shown in Figure 4.93.

By inspection of Figure 4.92, the system can be represented in Boolean expressions as

$$\text{Coffee} = (CM) \,.\, (\overline{CT}) \tag{4.119}$$

$$\text{Tea} = TM \tag{4.120}$$

where C, T, and M represent the coffee button, tea button, and money input, respectively, and the overbar represents the inverse complement as usual.

Using De Morgan's theorem the system may alternatively be written as

$$\text{Coffee} = \overline{(\overline{C} + \overline{M}) + (\overline{\overline{C} + \overline{T}})} \tag{4.121}$$

$$\text{Tea} = \overline{\overline{T} + \overline{M}} \tag{4.122}$$

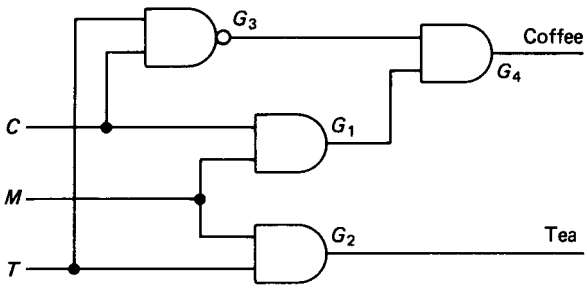


FIGURE 4.92 Extended logic circuit for drink vending machine.

Inputs			Outputs	
C	*M*	*T*	Coffee	Tea
0	0	0	0	0
0	0	1	0	0
0	1	0	0	0
0	1	1	0	1
1	0	0	0	0
1	0	1	0	0
1	1	0	1	0
1	1	1	0	1

FIGURE 4.93 Truth table for drink vending machine.

Thus, the same logic system can be implemented using one OR and three NOR gates, as shown in Figure 4.94.

The validity and equivalence of equations (4.119)–(4.122) may easily be checked using a truth table. Four logic gates are again required, but the circuit operates with inverted input signals. This means that three inverters are also required in the circuit as shown.

It is apparent that the logical function can be realized in several different ways, e.g.,

$$\text{Business} = \overline{(\overline{C} + \overline{M})} \,.\, (\overline{C} + \overline{T}) \text{ and Economy} = TM$$

Using the above realization, the circuit takes the form shown in Figure 4.95.

Logic Systems Using NAND and NOR Gates Only

Logic gates are packaged as arrays of the same type in IC form. A typical example is SN7408, which is a 14-pin DIL package containing four separate two-input AND gates. Because the logic gates are marketed in this particular form, it is advantageous to design the logic circuit using only one type of gate. This normally minimizes the number of IC packages required. Figure 4.96 shows a two-input NAND gate driving into a single-input NAND gate.

For the two-input NAND gate, the Boolean expression is

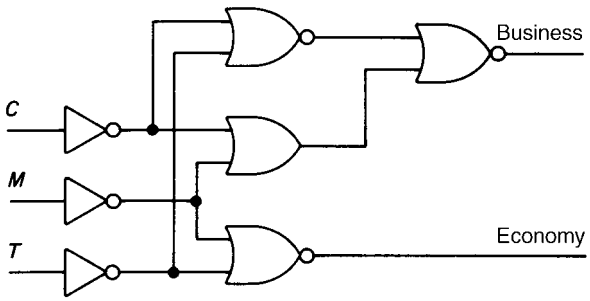


FIGURE 4.94 Logic circuit for an airline ticket-vending machine using OR and NOR gates.

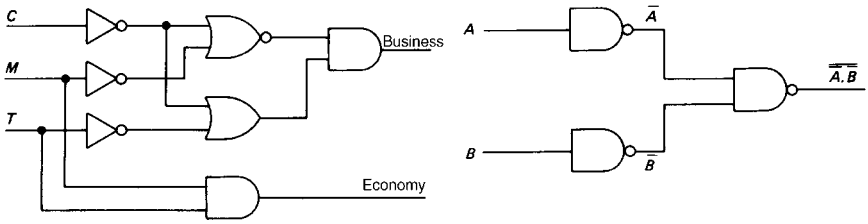
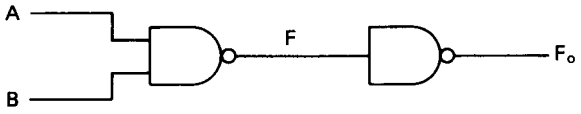


FIGURE 4.95 Alternative logic circuit for airline ticket-vending machine.



A	B	$A.B$	$\overline{A.B}$	$\overline{\overline{A.B}}$
0	0	0	1	0
0	1	0	1	0
1	0	0	1	0
1	1	1	0	1

FIGURE 4.96 AND realization using NAND gates.

$$F = \overline{A \,.\, B}$$

Since F is then fed into a single-input NAND gate, which operates as an inverter, then the final output is

$$F_o = \overline{\overline{F}} = \overline{\overline{A \,.\, B}} = A \,.\, B$$

It is apparent therefore that the circuit given in Figure 4.96, using NAND gates, performs the same function as the logical AND operator.

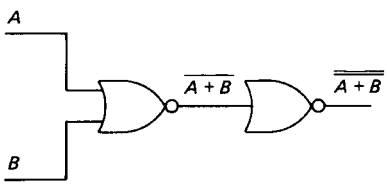
Figure 4.97 shows two single-input NAND gates with their outputs driving into a two-input NAND gate. Following through the truth table, it can be seen that the circuit performs the logical OR function. If the output F is then fed to another single-input NAND gate (not shown in the figure), then the function performed will be a logical NOR. It can be seen, therefore, that suitable combinations of NAND gates can be made to perform the logical functions AND, OR, and NOR. Similarly, it can be shown that the AND and OR functions can be realized using NOR gates only. This is illustrated in Figure 4.98. The conclusion which can be drawn is that any logic circuit can be realized using NAND gates or NOR gates alone.

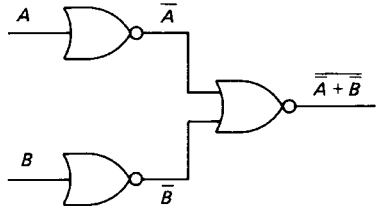
Considering again the ticket-vending machine depicted in Figure 4.94, the single OR gate may be replaced with a two-input NOR gate which then feeds directly into a single-input NOR gate. This is shown in Figure 4.99. Note that NOR gates are also used in place of invertors in the input signal lines.

By inspection of the circuit diagram the governing Boolean expressions are

A	B	$\overline{A}$	$\overline{B}$	$\overline{A}.\overline{B}$	$\overline{\overline{A.B}}$
0	0	1	1	1	0
0	1	1	0	0	1
1	0	0	1	0	1
1	1	0	0	0	1

FIGURE 4.97 OR realization using NAND gates.





A	B	$\overline{A+B}$	$\overline{\overline{A+B}} = A+B$
0	0	1	0
0	1	0	1
1	0	0	1
1	1	0	1

A	B	$\overline{A}$	$\overline{B}$	$\overline{\overline{A}+\overline{B}} = A.B$
0	0	1	1	0
0	1	1	0	0
1	0	0	1	0
1	1	0	0	1

FIGURE 4.98 OR and AND realizations using NOR gates only.

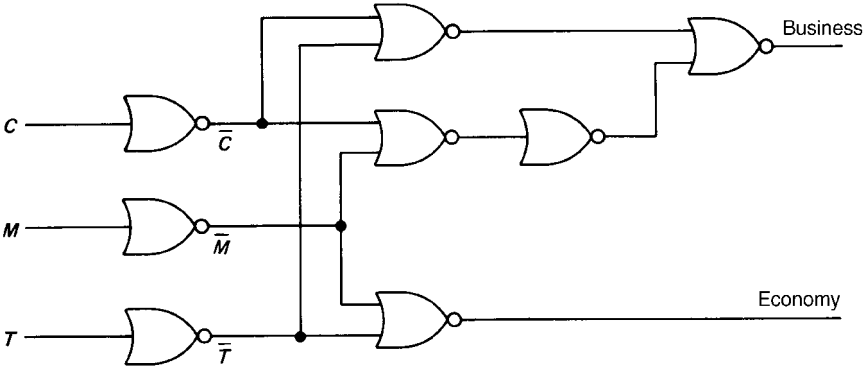


FIGURE 4.99 Logic circuit using NOR gates only for airline ticket-vending machine.

$$\text{Business} = \overline{\overline{(\overline{C} + \overline{M})} + \overline{(\overline{C} + \overline{T})}}$$

$$= (\overline{C} + \overline{M}) + (\overline{C} + \overline{T}) \tag{4.123}$$

$$\text{Economy} = \overline{\overline{T} + \overline{M}} \tag{4.124}$$

Equations (4.123) and (4.124) are identical to equations (4.121) and (4.122), respectively. This, of course, must be true, since the circuits from which the expressions were deduced perform identical logical functions.

Similarly, the circuit in Figure 4.92, involving one NAND and three AND gates, may be replaced by an equivalent circuit using only NAND gates. This equivalent circuit is shown in Figure 4.100. Inspection of the circuit gives the Boolean expressions

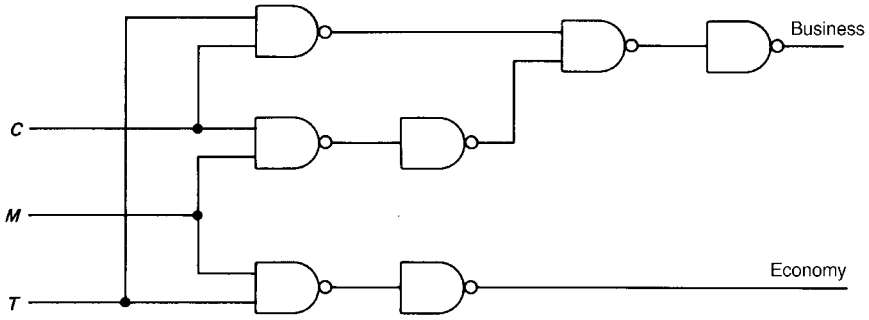


FIGURE 4.100 Logic circuit using NAND gates only for airline ticket-vending machine.

$$\text{Coffee} = \overline{\overline{(C \,.\, M)} \,.\, \overline{(C \,.\, T)}}$$

$$= (C \,.\, M) \,.\, (\overline{C \,.\, T}) \tag{4.125}$$

$$\text{Tea Q} = \overline{\overline{(T \,.\, M)}} = TM \tag{4.126}$$

Perhaps as expected, the Boolean expressions are identical to equations (4.119) and (4.120), which were deduced from the logic circuit of Figure 4.92.

The realization of Boolean expressions in either all NAND or all NOR gates can be stated in the following simple rules:

1. *NAND realization:* First, obtain the required Boolean expression in AND/OR form and construct the circuit required. The final output gate must be an OR gate. Replace all gates with NAND gates and, starting with the output gate, number each level of gates back through to the inputs. The logic level at the inputs to all "odd" level gates must be inverted.
2. *NOR realization:* Obtain the required Boolean expression in OR/AND form. The final output gate must be an AND gate. Replace all gates with NOR gates and number each level of gates from the output back through to the input. The logic level at all inputs to "odd" level gates must be inverted.

Application of these rules is best illustrated by, e.g.,

NAND realization of $F = AB + C(D + E)$

Figure 4.101 shows the realization of the function in AND/OR form. As inputs D and E appear at an odd level of gate input they must be inverted. In terms of the actual circuit this will mean that inputs D and E are inverted, using NAND gates, prior to entering the NAND gate at level 4.

A similar procedure is adopted for a NOR realization of a Boolean expression. The exclusive-OR function serves as an interesting example. Written as a Boolean expression, the exclusive-OR is

$$F = A \,.\, \overline{B} + \overline{A} \,.\, B \tag{4.127}$$

For the NOR realization, however, it is necessary that the final output gate is an AND. The exclusive-OR function must therefore be manipulated such that the final logical function in the expression is an AND. Using De Morgan's theorem,

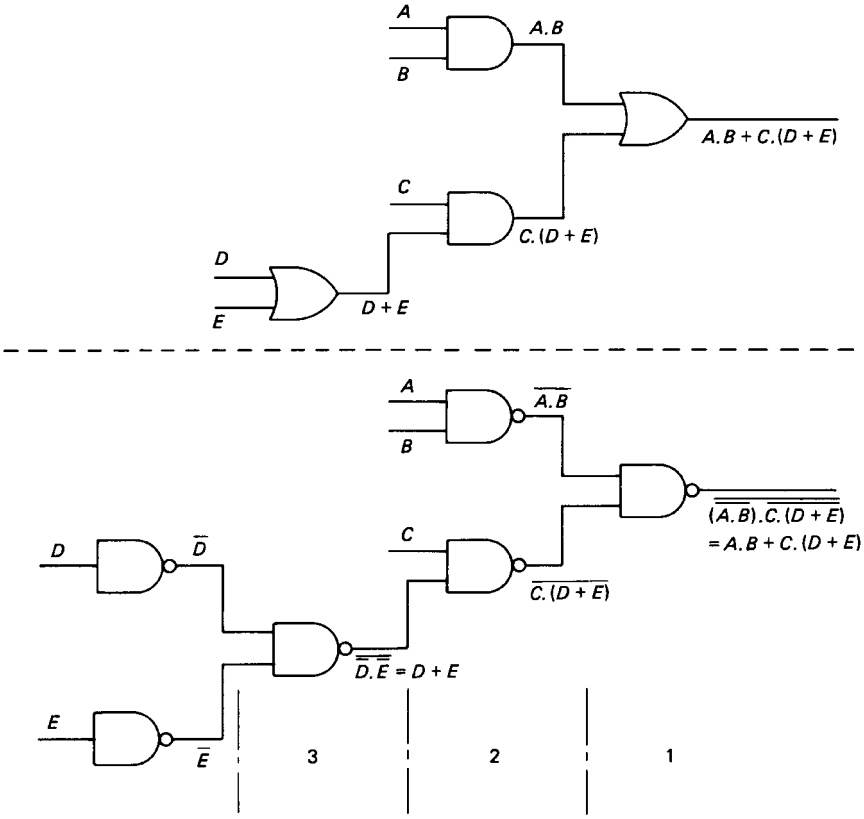


FIGURE 4.101 NAND realization of a Boolean function.

$$F = \overline{(A \,.\, \overline{B}) \,.\, (\overline{A} \,.\, B)}$$

$$= \overline{(\overline{A} + B) \,.\, (A + \overline{B})}$$

Multiplying out this expression gives

$$= \overline{\overline{A} \,.\, A + \overline{A} \,.\, \overline{B} + A \,.\, B + B \,.\, \overline{B}}$$

Since $A \,.\, \overline{A} = B \,.\, \overline{B} = 0$, the expression simplifies to

$$= \overline{\overline{A} \,.\, \overline{B} + A \,.\, B}$$

$$= (\overline{\overline{A} \,.\, \overline{B}}) \,.\, (\overline{A \,.\, B})$$

Using De Morgan again gives

$$F = (A + B) \,.\, (\overline{A} + \overline{B}) \tag{4.128}$$

The realization of this equation is shown in Figure 4.102.

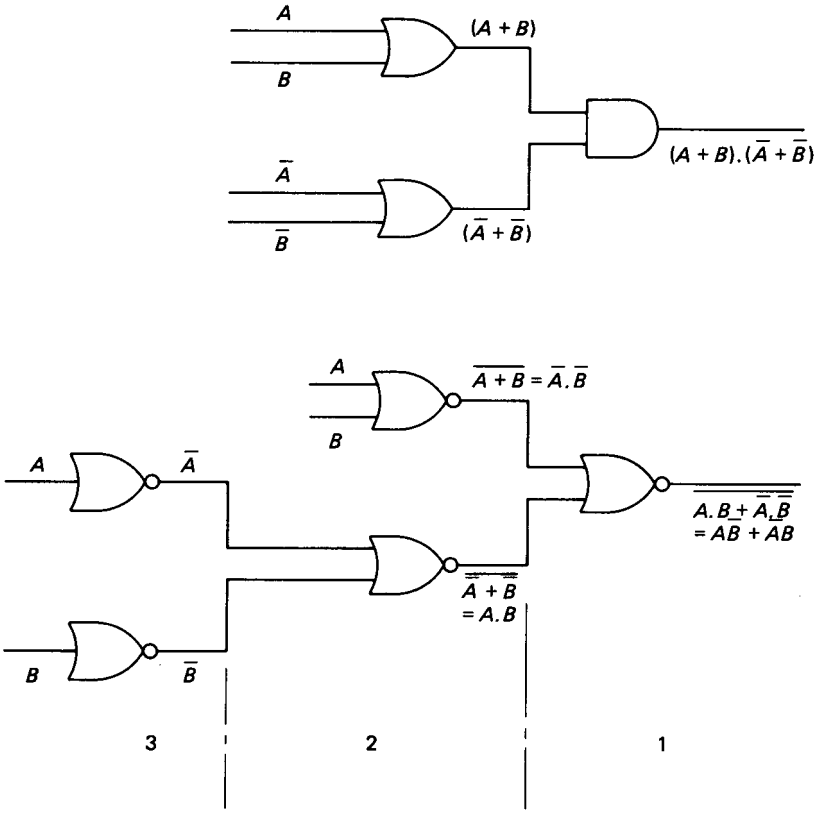


FIGURE 4.102 NOR realization of the exclusive-OR function.

Unused Inputs

Multiinput gates are also commonly available. In practical circuits, however, it is important that any unused inputs are tied, i.e., they are connected either to the positive voltage supply or to the zero voltage supply. The unused inputs are therefore set at either logic level 1 or at logic level 0, as required. In connecting an unused input to the positive supply the connection should be made through a 1 kΩ resistor. Failure to connect any unused inputs can result in intermittent malfunction of the circuit or in harmful oscillations with attendant overheating.

Latches

It is often useful to freeze a particular binary sequence, and devices called *latches* are used for this purpose. A latch has four inputs and four outputs. Normally the outputs assume the same state as the inputs. However, when a control signal (known as a *strobe* input) is taken to logic 1 the outputs are locked in whatever state they were in at the instant of the strobe input going high. This enables the binary sequence to be captured without affecting the ongoing processes, whatever they may

be. The latch therefore serves as a temporary state-recording device which may subsequently be referred to during various interrupt operations.

The Karnaugh Map

The Karnaugh map provides an alternative representation of a Boolean expression for all possible Boolean input combinations. In some respects the Karnaugh map is like a truth table in that identical logical expressions display an identical pattern on a Karnaugh map. The Karnaugh map, however, also has a great utility in simplifying Boolean expressions in a systematic manner.

The Karnaugh map consists of a set of boxes in which each box represents one possible combination of the Boolean input variables. The boxes are assigned either a 1 or a 0 to indicate the value of the Boolean expression for the particular combination of input variables that the box represents. The number of boxes required is 2^n, where n is the total number of input variables. Although any number of input variables can be represented, a practical limitation is about seven. Figure 4.103 shows the Karnaugh map for a four-input system. Within each box the unique Boolean input combination is represented by assigning each variable the logic values indicated along the horizontal and vertical axes. These values conform to the binary Gray code in which adjacent consecutive characters differ only in one variable. This imparts a property to the Karnaugh map in that the adjacent squares (vertically or horizontally) differ only in one variable.

As an example, the Boolean expression, $F = \overline{A}BCD + A\overline{B}\overline{C}\overline{D} + \overline{A}B\overline{C}D$ is represented by the Karnaugh map given in Figure 4.104. The maps are drawn up by placing a 1 in each box for which the combination of input variables makes the logical expression have a value of 1. All the other boxes represent the combination of input variables which make the expression have a logical value of 0. Usually the 0 is not entered in the box.

A second example for consideration is

$$F = \overline{A}\overline{B}CD + AC + \overline{C}\overline{D}$$

The Karnaugh map for this expression is shown in Figure 4.105. It can be seen

CD \ AB	00	01	11	10
00	$\overline{A}\overline{B}\overline{C}\overline{D}$	$\overline{A}B\overline{C}\overline{D}$	$AB\overline{C}\overline{D}$	$A\overline{B}\overline{C}\overline{D}$
01	$\overline{A}\overline{B}\overline{C}D$	$\overline{A}B\overline{C}D$	$AB\overline{C}D$	$A\overline{B}\overline{C}D$
11	$\overline{A}\overline{B}CD$	$\overline{A}BCD$	$ABCD$	$A\overline{B}CD$
10	$\overline{A}\overline{B}C\overline{D}$	$\overline{A}BC\overline{D}$	$ABC\overline{D}$	$A\overline{B}C\overline{D}$

FIGURE 4.103 Karnaugh map for a four-input system.

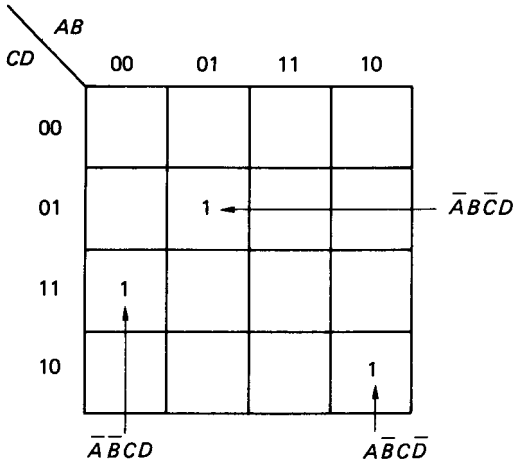


FIGURE 4.104 Karnaugh map for a Boolean expression (1), see text for definition.

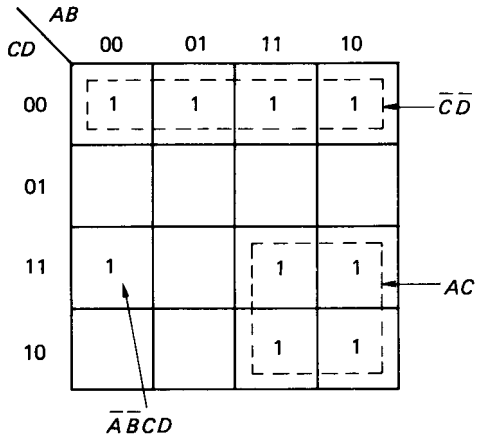


FIGURE 4.105 Karnaugh map for a Boolean expression (2), see text for definition.

that the term AC includes all four squares in which both A and C are included. Similarly, the term $\bar{C}\bar{D}$ also encompasses four squares on the map. It may be concluded that in a four-variable expression any term which contains the four variables will occupy one square on the Karnaugh map. Any term which contains only three of the variables will occupy two squares, and any which contains only two of the variables will occupy four squares. A term containing only one of the variables will occupy eight squares in the Karnaugh map. The Karnaugh map may be used in a reverse mode to deduce the Boolean expression.

ELECTRICAL AND ELECTRONIC PRINCIPLES

Minimization of Boolean Expressions

The principle of minimization is based on the Boolean identity $A + \overline{A} = 1$. Thus,

$$F = ABCD + ABC\overline{D} = ABC(D + \overline{D}) = ABC \tag{4.129}$$

The grouping of squares along any axis therefore enables the minimization, which is typified by equation (4.129). An extension of this principle is shown in Figure 4.106. The Boolean expression depicted in this figure can be written as

$$\begin{aligned} F &= AB\overline{C}\,\overline{D} + ABC\overline{D} + \overline{A}\,\overline{B}CD + A\overline{B}CD \\ &= AB\overline{D}(C + \overline{C}) + \overline{B}CD(\overline{A} + A) \\ &= AB\overline{D} + \overline{B}CD \end{aligned}$$

Minimization in the above example reduces the four terms in the expression to two, each involving three variables. The groupings in the example are akin to the idea of rolling the map into a cylinder about either axis to complete the two groupings as shown.

In extending the minimization principle to five variables the number of squares required is $2^5 = 32$. This is best handled as two sets of 16 squares in a top and bottom arrangement. The 16-square layers represent the first four-input variables and each layer accommodates the two possible input combinations for the fifth variable. Higher numbers of input variables can be dealt with, but the map becomes increasingly more difficult to handle.

In certain situations involving a number of input variables, particular combinations of the variables never actually occur in practice. Under these circumstances the output which would occur with these combination of variables is irrelevant. The output can therefore have any value, since it is a situation that never occurs. Such

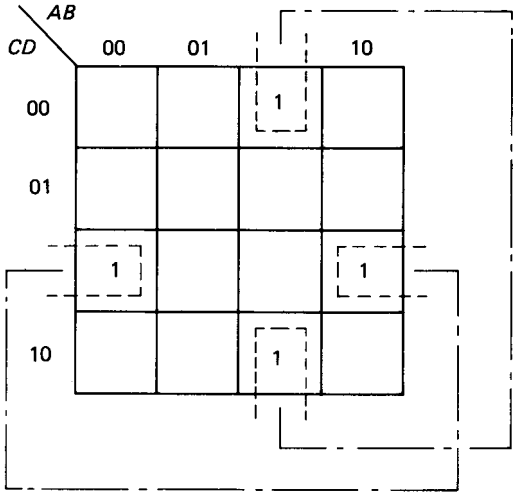


FIGURE 4.106 Extended minimization principle.

input combinations are called "don't care" conditions, and they can be incorporated into a system to allow a simpler circuit realization.

The principle can be illustrated by means of an example:

$$F = \overline{A}B\overline{C}D + AB\overline{C}D + \overline{A}BCD$$

It is stated that the combination $ABCD$ will never occur. Including the don't care condition in the expression gives

$$F = \overline{A}B\overline{C}D + AB\overline{C}D + \overline{A}BCD + \{ABCD\}_x \tag{4.130}$$

The don't care combination is usually enclosed within parentheses and subscripted with either x or 0.

The Karnaugh representation for the expression is shown in Figure 4.107 and the don't care condition is clearly indicated in the figure. By ignoring the don't care condition, minimization of the expression results in

$$F = B\overline{C}D + \overline{A}BD$$

If the network output is allowed to be 1 for the don't care condition, the minimization yields

$$F = BD$$

The example evidently shows that considerable savings in the realization of an expression can be made by including a relevant don't care condition.

It is also worth bearing in mind that although the Karnaugh map can yield a minimum gate solution to a given problem, it might not be an optimum solution. In the real world other considerations may well dictate in terms of parts, design, assembly costs, and the number of IC packages required.

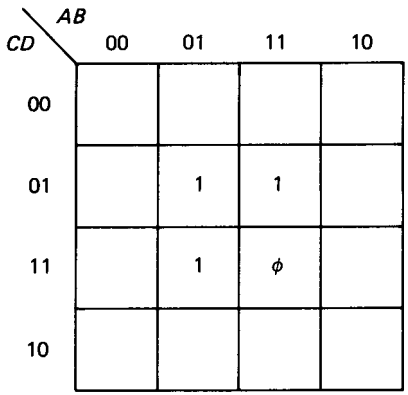


FIGURE 4.107 Karnaugh representation for equation (4.130).

Positive and Negative Logic

In considering the digital logic systems so far, no mention has been made of the significance of the logic levels in terms of the actual voltages applied. Two possibilities exist to differentiate between logic 1 and 0. In a positive logic system, logic level 1 is represented by a more positive voltage level than logic level 0. Both logic voltage levels could actually be negative, but many digital systems operate with a voltage between 0 V and 0.8 V, denoting logic level 0, and a voltage between 2.4 V and 5 V, denoting logical level 1. This standard is used in the TTL and CMOS series of logic devices.

In a negative logic system, logic level 1 is represented by a less positive voltage than logic level 0. This standard applies to data-transmission interfaces where a voltage in the range −3 V to −15 V denotes logic 1 and a voltage in the range +3 V to +15 V logic 0. The large differentiation between 0 and 1 ensures good immunity to electrical noise. These voltages, however, are not compatible with TTL and CMOS devices, and interconversion ICs are required within the data-transmission interface.

As an alternative to using the terms logic 1 and logic 0, "high" and "low" are often substituted. In a positive logic system a transition from logic 0 to logic 1 can be termed a transition from low to high.

The logic level definitions also influence the function of the logic device. Figure 4.108 shows two types of two input NOR gates. In Figure 4.108(a) the inputs are negative logic and the output is positive logic. The NOR gate therefore performs the logical AND function. In Figure 4.108(b) the inputs are positive logic while the output is negative logic. This NOR gate therefore performs the logical OR function.

Tri-state Logic

Tri-state logic does not represent three logic levels but denotes three states which may be logic 1, logic 0, or "unconnected." A separate *enable* input determines whether the output behaves as a normal output or goes into the third (open-circuit) state. Tri-state devices are used in applications where different logic devices are required to be connected into output lines which are common to other logic devices (for example, computer data buses). While one set of logic devices is transmitting signals the other set is temporarily disconnected or disabled.

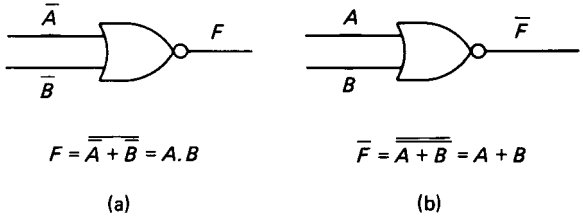


FIGURE 4.108 NOR gates using positive and negative logic input/output systems.

Sequential Logic Circuits

The logic circuits considered so far are all examples of combinational logic systems where the output is determined by the combination of input variables present at that time. Sequential logic circuits are those in which the outputs depend upon the sequence of prior inputs. The main difference between sequential and combinational logic systems is that the former circuits must possess some semblance of memory. The basic memory element in sequential logic systems is provided by one of several bistable gates, so called because of the two different but stable outputs which the gates produce.

The SR Bistable (Flip-Flop). The term *flip-flop* is traditionally used with respect to basic memory elements, and in the SR flip-flop "S" denotes Set and "R" denotes Reset. The SR flip-flop was an early development, commonly constructed using discrete transistors. The internal operation, in which two transistors alternate between the cut-off and saturated states, is of less importance than the external function which the device performs.

Using the systems approach, the SR flip-flop can be represented as shown in Figure 4.109. The system shows the two inputs S and R and the two output lines traditionally denoted as Q and $\overline{Q}$. For sequential circuits the truth table is more usually called a state table. The state table for the SR flip-flop is given in Figure 4.110. Each set of input variable values is considered for both possible states of the output. This is necessary because the output values depend not only on the input variable values but also on the current values of the outputs themselves.

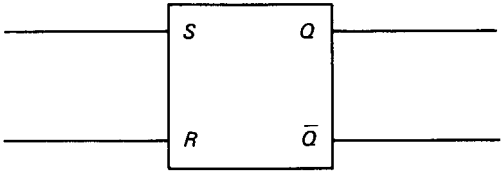


FIGURE 4.109 The SR flip-flop.

Inputs		Output changes	
S	R	$Q_n \rightarrow Q_{n+1}$	$\overline{Q}_n \rightarrow \overline{Q}_{n+1}$
0	0	0 → 0	1 → 1
0	0	1 → 1	0 → 0
0	1	0 → 0	1 → 1
0	1	1 → 0	0 → 1
1	0	0 → 1	1 → 0
1	0	1 → 1	0 → 0
1	1	Not available	
1	1		

FIGURE 4.110 State table for the SR flip-flop.

The operation of the SR flip-flop may be summarized as follows:

1. With $S = 0$ and $R = 0$, the output is not affected and remains as it was.
2. With $S = 1$ and $R = 0$, the output will change to $Q = 1$ if previously Q was 0. Q will remain at 1 if previously Q was 1.
3. With $S = 0$ and $R = 1$, the output will change to $Q = 0$ if previously Q was 1. Q will remain at 0 if previously Q was 0.
4. In all cases considered, the output $\overline{Q}$ will be the inverse complement of Q.

The SR flip-flop may be constructed using cross-coupled NOR or NAND gates, as shown in Figure 4.111.

The T (Trigger) Flip-Flop. The T flip-flop is another bistable circuit having two outputs, Q and $\overline{Q}$, but only one input, T. The T flip-flop changes state on every T input signal and then remains in that state while the T input remains low.

The JK Flip-Flop. The JK flip-flop uses integrated-circuit technology and, since it can perform both the SR and T flip-flop functions, it has become the most common flip-flop in current use. Figure 4.112 gives the state table and logic symbol for the JK flip-flop. The state table is identical to the SR flip-flop with the exception that the input condition $J = 1$, $K = 1$ is allowed. For these latter inputs the JK flip-flop functions as a T flip-flop using an input clock signal, in the form of a pulse train, as the trigger.

The JK flip-flop operates in a clocked or synchronous mode. In synchronous mode, the J and K inputs do not in themselves initiate a change in the logic outputs but are used to control inputs to determine the change of state which is to occur. A pulsed input to the clock terminal (CK) then determines the timing of the state changes. The clocked mode allows for precise timing of the state changes in a sequential circuit.

JK flip-flops may also be provided with additional Set, S, and Clear, C, inputs, which can be used to set output Q to 1 or clear output Q to 0 at any time. Multiple J and K inputs are also commonly available to enable logical ANDing of multiple-input signals.

A slightly more complicated flip-flop arrangement is the JK master–slave flip-flop. This consists of a pair of SR flip-flops connected together by various logic gates as shown in Figure 4.113. The JK master–slave flip-flop differs from the simpler arrangement in that if the clock pulse is at logic 1, a logic 1 applied to either J or K will not set the outputs. The new data, however, are accepted by the

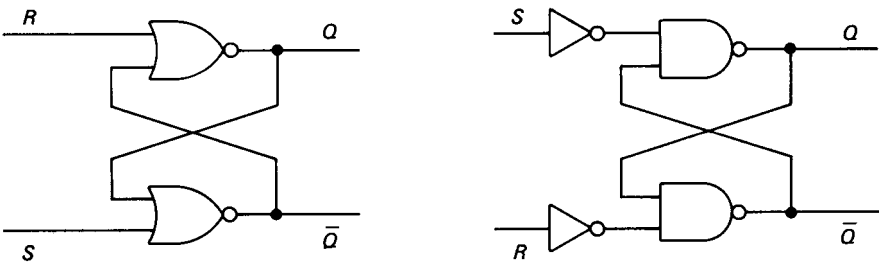
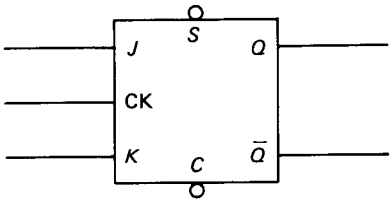


FIGURE 4.111 SR flip-flops using cross-coupled gates.



Inputs		Outputs
J	K	$Q_n - Q_{n+1}$
0	0	0 → 0
0	0	1 → 1
0	1	0 → 0
0	1	1 → 0
1	0	0 → 1
1	0	1 → 1
1	1	0 → 1
1	1	1 → 0

FIGURE 4.112 JK flip-flop and corresponding state table.

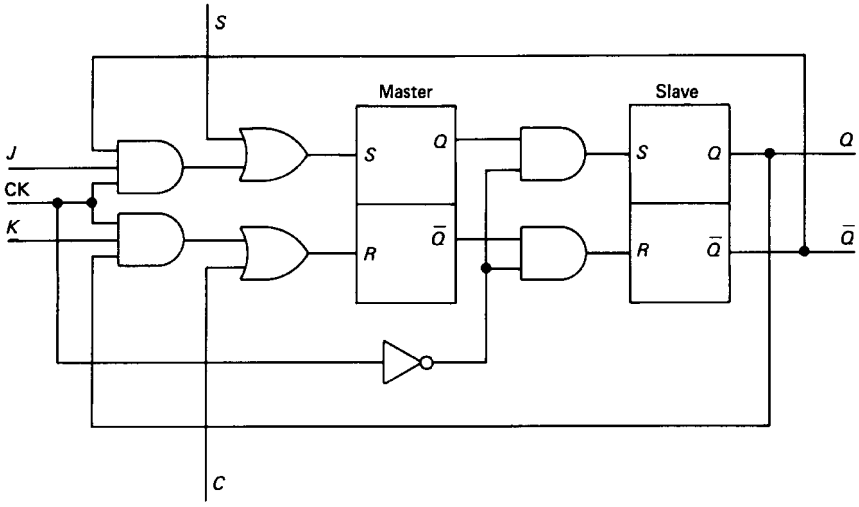


FIGURE 4.113 The JK master–slave flip-flop.

"master." When the clock pulse returns to 0, the master is isolated from the inputs but its data are transferred to the slave, with the result that Q and $\overline{Q}$ can then change state. In a circuit involving many such flip-flops, the advantage of the master–slave arrangement is that it can allow for synchronization of all the output state changes.

Registers and Counters

In the previous section it was shown that a logic level of 1 on a particular input line to a flip-flop can set an output line to 1. In this way the flip-flop can perform an elementary memory function. For a binary signal of length n bits, n flip-flops are required to construct a memory device for the n-bit input signal. A group of flip-flops used together in this manner constitutes a *register*.

Data may be entered into the register in a serial or a parallel manner. In the parallel method the n-bit binary "word" is available on n input lines. Each line is connected to its own flip-flop and the n data bits are entered simultaneously into the register. In the serial entry method the data are available on only one input line in a time sequence. They are entered consecutively and are timed into the register by a system clock. Serial entry registers are also called shift registers, since the data bits are entered into the first flip-flop and moved consecutively along into the next flip-flop as the next data bit arrives at the first flip-flop, and so on. The serial method of data entry requires as many shift and store operations as the number of bits in the binary word. This means that the serial entry method is much slower than the parallel method. Serial entry, however, is also much less expensive than parallel entry.

Yet another type of register is the counting register. This consists of a number of flip-flops arranged to store a binary word which is representative of the number of input pulses applied at the input terminal. Using n flip-flops, a total count of 2^n can be made.

Counting registers (or counters) may be synchronous in which the state changes in all the flip-flops occur simultaneously, or asynchronous in which the state changes in various flip-flops do not occur at the same time. Figure 4.114 illustrates an asynchronous, 3-bit binary counter composed of JK flip-flops.

All J and K inputs are held at logic level 1 and the input signal consists of a pulse train fed to the clock input of the first flip-flop. In this mode the JK flip-flop is operating as a T flip-flop. The output Q from the first flip-flop provides an input for the clock of the second flip-flop, and so on through the network. The outputs Q also form the binary representation of the counter, where A is the least significant

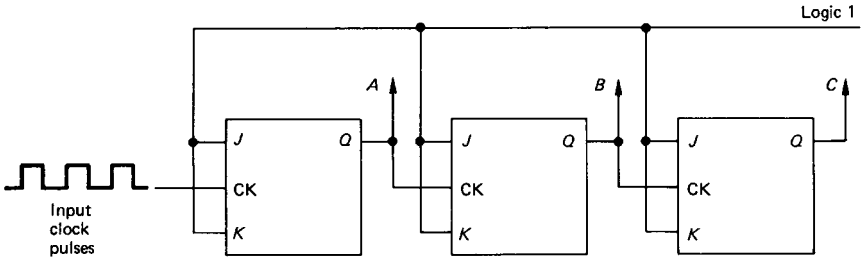
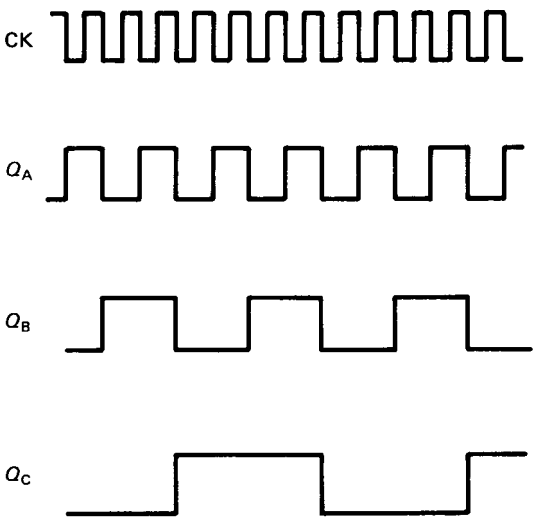


FIGURE 4.114 Asynchronous, 3-bit binary counter.

bit, increasing through to *C*, which represents the most significant bit. The state table and timing diagram for the counter are shown in Figure 4.115.

The state table shows that each flip-flop changes state when the next-less significant flip-flop output changes from 1 to 0. The output signal from each flip-flop, moving through the network, is at half the frequency of that of the previous flip-flop. These output signals thus provide the correct binary count of the number of input pulses applied to the input. The 3-bit binary counter can count up to a maximum of 8 decimal. If a ninth pulse is applied at the input, the count reverts back to the initial zero setting and the count continues again as normal for further input pulses.

Asynchronous counters are also referred to as *ripple* counters because of the way that the changes of state ripple through the network of flip-flops.



Clock pulses	Flip-flop		
	C	*B*	*A*
0	0	0	0
1	0	0	1
2	0	1	0
3	0	1	1
4	1	0	0
5	1	0	1
6	1	1	0
7	1	1	1
0	0	0	0

FIGURE 4.115 State table and timing diagram for a 3-bit binary counter.

A synchronous version of the counter can also be realized using a network of JK flip-flops. The synchronous counter additionally uses the outputs Q and $\overline{Q}$ of each flip-flop, in logic gate networks, to produce the necessary control signals for the J and K inputs. This ensures that all flip-flops change state correctly to the desired state table for each clock pulse. The synchronous counter alleviates the problems associated with transient operation inherent in the asynchronous counter.

There are, of course, many other types of flip-flop available, but the only one of significant practical importance is the D flip-flop (see Figure 4.116), where the "D" refers to Data.

In the D-type flip-flop the D input is fed directly into the J input line and the inverse complement of D is fed to the K input line. This ensures that J and K are always the inverse complement of one another. A logic 0 or 1 on the data input will then flip (or flop) the outputs when the clock pulse is at logic 0.

Timers and Pulse Circuits

An essential feature of the flip-flop circuits described in the previous two sections was the provision of a pulsed clock signal. Although timers can be designed using discrete components, it is normal to design around the commonly available timers which are already available in IC form.

Monostable. Figure 4.117 shows the 555 wired up for monostable operation. When the trigger is taken from +5 V to 0 V (i.e., high to low), the output will go high for a period determined by the values selected for R and C. The length of the output pulse is given by $1.1RC$.

The timer can deliver currents of more than 100 mA, and it can therefore be used to drive a DIL reed relay directly. When such a relay is switched off, however, the back e.m.f. generated by the relay coil could damage the timer. As a precaution, a diode is normally connected in parallel with the relay coil, in the opposite direction to the current flow, to absorb the high induced voltage.

Astable. Figure 4.118 depicts the 555 wired up for astable operation. The 100 nF capacitor is only required for TTL-based timers. In astable operation the output is a continuous pulse train. The ON and OFF times can be controlled independently within certain limitations with

$$\text{ON time} = 0.693(R_1 + R_2)C \tag{4.131}$$

$$\text{OFF time} = 0.693(R_2)C \tag{4.132}$$

Obviously, the ON time can only be equal to or greater than the OFF time. The

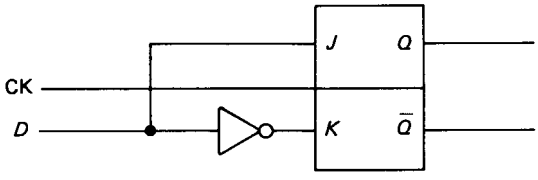


FIGURE 4.116 D-type flip-flop.

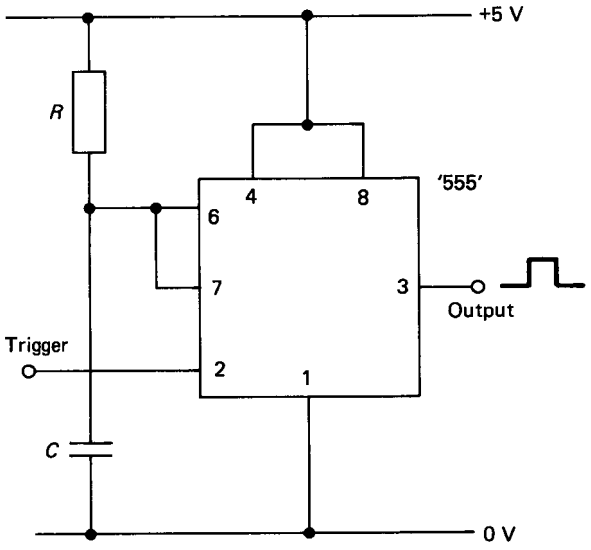


FIGURE 4.117 555 timer in monostable operation.

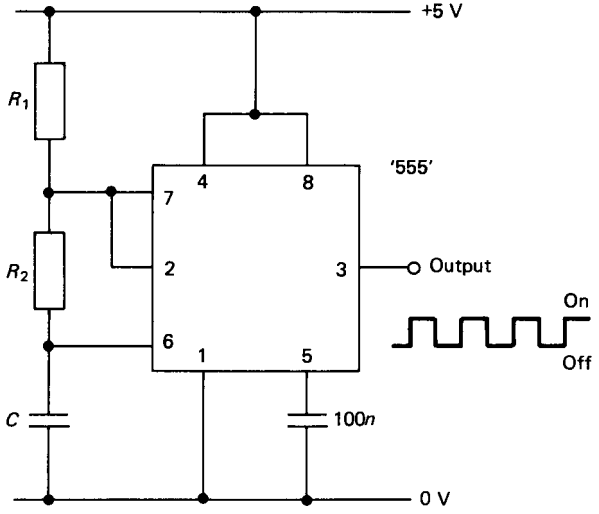


FIGURE 4.118 555 timer in astable operation.

output signal, however, can always be inverted if, in a particular application, short-duration positive pulses are required.

The maximum operating frequency for the 555 timer is about 500 kHz and the minimum frequency, limited by the leakage of the capacitor, is about one cycle per several hours.

Alternative pulsed output circuits can be constructed using TTL or CMOS gates (see Kaufman and Seidman 1988 and Watson 1983).

Digital Computers and Microprocessors

No coverage of digital electronics, however brief, can fail to give some cognizance to the impact of the digital computer and its associated microprocessor. The modern digital computer, although a complex digital system, consists of no more than the basic logical sub-systems previously discussed. This includes AND, OR, NAND, and NOR gates, registers, counters, and communication interfaces.

The main advantages of the microprocessor-based system are that the logical functions for a particular application can be developed and implemented in software, as opposed to electronic hardware. In many instances the microprocessor-based system may actually be the cheaper alternative to a hardwired logic gate circuit. The software is easy to alter in the event of incorrect system operation, and the complete system can be tested as a simulation before being committed.

For relatively small logical switching applications (up to, say, 32 inputs) the single-card (or single-chip) microcomputer represents an ideal low-cost solution (see Milne and Fraser 1990). These microsystems can be used as dedicated devices where all the system components reside on a single card or a single chip, respectively.

Application-Specific Integrated Circuits (ASICs)

Application-specific integrated circuits are programmable logic devices (PLDs) which have their internal logic configuration determined by the user, as opposed to the manufacturer. The systems design engineer therefore customizes the actual silicon building blocks to meet the requirements of the system. Such customization provides for performance, reliability, compactness, low cost, and design security. PLDs are available in both TTL and CMOS technology. The latter are erasable and can be reprogrammed almost indefinitely. PLDs represent the fastest-growing segment of the semiconductor industry in recent times, and it can be expected that they will play an increasingly important role in the design of digital logic systems in the future.

Internally, PLDs consist of an array of AND gates connected to an array of OR gates, with input and output blocks containing registers, latches, and feedback options. Figure 4.119 shows the general architecture of a programmable logic device.

In customizing the PLD the user essentially determines which of the interconnections between the gate arrays will remain open and which will be closed. The customization procedure, however, requires additional development tools, which consists of:

1. A word processor to generate the source code.
2. Development software to transform the high-level language source code into a fuse pattern for the PLD. The code which is generated is referred to as a JEDEC file.
3. A PLD programmer to implement the program within the device.

The PLD programmer programs the PLD by ‘‘burning’’ the fuse pattern in the memory array of the device. When returned to its normal operating mode, the PLD

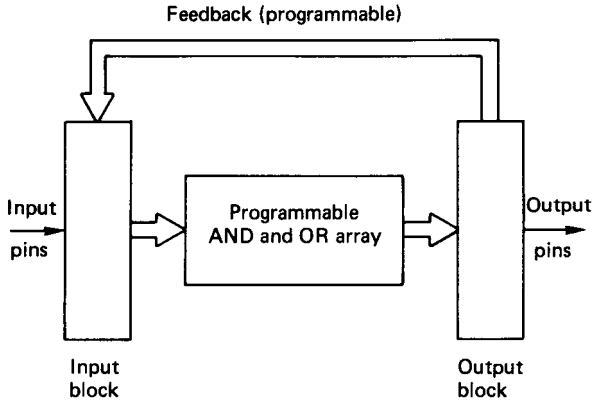


FIGURE 4.119 General architecture of a PLD.

then performs the customized logic function. Horowitz and Hill (1989) provide a reasonably detailed coverage on applications of programmable logic devices.

4.4 ELECTRICAL SAFETY

Electric Shock

Most serious injuries and deaths from electric shock occur from contact with the mains electricity supply. In the United Kingdom, the mains supply is about 240 V a.c., and in the United States it is about 110 V a.c. The live wire is at a higher potential with respect to the earth. If a person is in contact with the ground, an electric shock can be sustained by touching the live wire only. When a shock is received the passage of the electric current through the body may cause muscular contractions, respiratory failure, fibrillation of the heart, cardiac arrest, or injury from internal burns. Any one of these can be fatal. The greatest danger occurs when the current flows across the chest. This can happen either when the current flows from one arm to the other or when it flows from one arm to the opposite leg.

The magnitude of an electric shock depends on the strength of the current, which in turn depends on the voltage and the ohmic resistance of the body. The resistance of the human body varies in different persons and is primarily dependent on the resistance of the skin. This variability in skin resistance means that a "safe" voltage cannot be readily specified, and all voltages in excess of about 50 V must therefore be regarded as being potentially lethal. If the skin is damp, either from water or from perspiration, then the skin resistance is considerably reduced. Under such circumstances the chances of an electric shock proving to be fatal are greatly increased.

Injury can also be caused by a minor shock, not serious in itself but which has the effect of contracting the muscles sufficiently to result in a fall.

ELECTRICAL AND ELECTRONIC PRINCIPLES

Electric Burn

Burns can be caused either by the passage of heavy current through the body or by direct contact with an electrically heated surface. They may also be caused by the arcing across a short circuit or as a result of an electrically originated fire or explosion.

Any circuit brought near an induction heater will receive energy and heat will be generated extremely rapidly. No rings or other metal objects should therefore be worn when in the vicinity of an induction heater, nor should any metal be held in the hands. In general, no part of the body should come within close proximity (about a meter) to an induction heater.

Rescue

To render assistance to a person undergoing an electric shock the rescuer should first attempt to isolate the circuit by switching off the supply. If the rescuer cannot isolate the supply, he or she should try to break the victim's contact with the live apparatus by using insulating material. This material is essential to prevent the rescuer from becoming a second victim.

If the victim is unconscious the rescuer should send for medical assistance and start artificial resuscitation. This assumes, of course, that the rescuer is trained in modern methods of such resuscitation.

Protection

Insulation. Electrical cables consist of one or more metal cores, which may be single wires but are more usually stranded wire and surrounded by insulation. The insulation serves to contain the flow of current and prevent a person from touching the live metal and thereby receiving a shock.

Fuses. A fuse is a device which will melt when the current exceeds a predetermined value. In operation, fuses serve as current-limiting devices, and they are used for overload protection of electrical equipment. Two types of fuses are generally available: rewirable and cartridge.

Fuses are specified by the maximum current that they can transmit and the correctly rated fuse must be used at all times. A high-rating fuse must never be substituted for a low-rating one.

Circuit Breakers. A circuit breaker is a mechanical device in the form of a switch which opens automatically if the circuit which it controls becomes overloaded. Circuit breakers may be operated magnetically or thermally, and they can also be manually reset and adjusted. Plug-in circuit breakers are available and are recommended for use with small electric power tools.

Earth-Leakage Protection. Normally there is no net flow of electricity to an electrical device. The flow in from the live wire is exactly balanced by the return flow in the neutral cable. If an earth fault develops, however, a leakage current will result, and this can be detected by the earth-leakage apparatus. Modern earth-

leakage devices are so sensitive that the supply is immediately disconnected before a lethal current can be drawn from the mains. It should be noted that earth leakage protection will operate only when the fault occurs between line and ground.

Isolation. Effective means of disconnecting cables or apparatus from the source of supply must be provided so that maintenance, repair or alteration may be carried out safely. This is achieved by isolating switches which have no automatic features. Various circuits and motors should not share an isolation switch unless it is clear that under no circumstances will it be necessary, or convenient, to use one circuit while the other is being serviced.

Isolation switches should be capable of being locked in the OFF position but not in the ON one. If the isolation switch cannot be locked it should be possible to remove the fuse on the power line so that the line cannot be energized by inadvertent closing of the isolation switch.

Earthing

The external metal casing of electrical apparatus and cables must be earthed for three reasons:

1. To prevent the casing rising to a dangerous voltage if there is a fault such as a short circuit between the conductor and the casing
2. To conduct any current away by a safe path
3. To ensure that the faulty circuit is automatically disconnected from the supply by drawing sufficient current to blow the protective fuse or operate the circuit breaker

Earthing consists of connecting the metal casing by means of a conductor to an earth electrode. The earth electrode may be a buried pipe or other such conductor which is known to be making an effective connection to the general mass of the earth. Where the earth connection to a casing is made with a nut and bolt, a spring washer or other similar locking device must be used. Earthing is a legal requirement and must be effective at all times.

Double Insulation

Although the electricity regulations require all portable apparatus used at normal mains voltage to have an earthing conductor, these can introduce their own hazards. As a result, double-insulated or all-insulated apparatus is made which does not require earthing. Double insulation means what its name says, and all live conductors are separated from the outside world by two separate and distinctive layers of insulation. Each layer of insulation would adequately insulate the conductor on its own, but together they virtually negate the probability of danger arising from insulation failure. Double insulation avoids the requirement for any external metalwork of the equipment to be protected by an earth conductor.

Low-Voltage Supplies

Portable tools (particularly hand inspection lamps) can be a source of danger because they are subject to severe wear and tear and are likely to be used in confined spaces where the skin resistance could easily be reduced by damp conditions. In cases where work is carried out within confined metal enclosures, mains voltage equipment must not be used. A double-wound transformer with a secondary center tap to earth is allowable in these cases. This transformer gives 50 V for lighting and 100 V for portable tools.

These few notes on general electricity safety are by no means extensive or authoritative. Reference should always be made to the full guide to the regulations (Health and Safety Executive 1989). Further recommended reading on electrical safety guidelines can be obtained from Imperial College of Science and Technology (1976) and Reeves (1984).

REFERENCES

Bell, D. A. 1984. *Fundamentals of Electric Circuits,* 3d ed., Reston 1989, New York.

Bell, E. C. and Whitehead, R. W. 1987. *Basic Electrical Engineering & Instrumentation for Engineers,* 3d ed., Granada, St. Albans.

Gray, C. B. 1989. *Electrical Machines and Drive Systems,* Longman Scientific & Technical, Harlow.

Health and Safety Executive. 1989. Memorandum of Guidance on the Electricity at Work Regulations, 1989, Health & Safety series booklet HS(R)25, HMSO.

Horowitz, P. and Hill, W. 1989. *The Art of Electronics,* 2d ed., Cambridge University Press, Cambridge.

Hughes, W. 1987. *Electrical Technology,* 6th ed., rev. I. McKenzie Smith, Longman, Harlow.

Imperial College of Science and Technology. 1976. *Safety Precautions in the Use of Electrical Equipment,* 3d ed., Imperial College Safety Booklets.

Kaufman, M. and Seidman, A. H. 1988. *Handbook of Electronics Calculations for Engineers and Technicians,* 2d ed., McGraw-Hill, New York.

Kenjo, T. and Nagamori, S. 1985. *Permanent-Magnet and Brushless dc Motors,* Monographs in Electrical and Electronic Engineering, Clarendon Press, Oxford.

Milne, J. S. and Fraser, C. J. 1990. "Development of a Mechatronics Learning Facility," in *Mechatronic Systems Engineering,* vol. 1, pp. 31–40.

Orthwein, W. 1990. *Machine Component Design,* West, New York.

Reeves, E. A., ed. 1984. *Handbook of Electrical Installation Practice,* vol. 1, *Systems, Standards and Safety,* Granada, St. Albans.

Sen, P. C. 1989. *Principles of Electric Machines and Power Electronics,* John Wiley, Chichester.

Watson, J. 1983. *Mastering Electronics,* Macmillan, London.

FURTHER READING

Fitzgerald, A. E., D. Higginbottom, and A. Grabel, *Basic Electrical Engineering,* 5th ed., McGraw-Hill, New York (1981).

Nasar, S. A., *Handbook of Electric Machines,* McGraw-Hill, New York (1987).

Say, M. G., *Alternating Current Machines,* Pitman, London (1983).

SECTION 5
COMPUTING

Section Editor: Gordon Young
Author: Ian Robertson

5.1 INTRODUCTION

Although the advent of computers in our everyday lives may seem very recent, the principles of the modern computer were established before the existence of any electronic or electromechanical technologies as we know them today, and electronic computers were beginning to take shape in laboratories in 1945.

The work of Charles Babbage, a Cambridge mathematician of the nineteenth century, in attempting to build an ''analytical engine'' from mechanical parts, remarkably anticipated several of the common features of today's electronic computers. His proposed design, had he been able to complete it and overcome mechanical engineering limitations of the day, would have had the equivalent of punched-card input and storage registers, the ability to branch according to results of intermediate calculations, and a form of output able to set numeric bits in type.

Many purely mechanical forms of analog computer have existed over the last few centuries. The most common of these is the slide rule, and other examples include mechanical integrators and even devices for solving simultaneous equations. Much of the development leading to modern electronic computers, both analog and digital, began during World War II with the intensified need to perform ballistics calculations. The development of radar at this time also provided the stimulus for new forms of electronic circuits that were to be adopted by the designers of computers. A further development of momentous importance to the technology of computers, as it was for so many branches of electronics, was that of the transistor in 1949. Continued rapid strides in the field of semiconductors have brought us the integrated circuit, which allows a complete digital computer to be implemented in a single chip.

5.2 GENERATIONS OF DIGITAL COMPUTERS

Beginning with circuits consisting of relays, the history of the digital computer can be seen as having fallen into four generations between the 1940s and today.

First Generation

These computers were built with valve circuits and delay-line storage, physically very massive, taking up complete rooms, requiring very large amounts of electricity with corresponding high heat dissipation and low overall reliability, requiring extensive maintenance often resulting in engineers being on-site 24 hours per day. Input/output was rudimentary (teleprinters, punched cards) and programming very laborious, usually in a binary form that the machine could understand without further interpretation.

Second Generation

Developed during the 1950s with transistorized circuits, these computers were faster, smaller, and more reliable than the first generation, but still large by today's standards. Magnetic core main stores with magnetic drums were used with tapes as back-up, and line printers were employed for faster printed output. Programming language translators emerged, resulting in the widespread use of Assembler-type languages.

Third Generation

Developed during the mid-1960s, these computers were heralded by the integrated circuit, allowing more compact construction and steadily improving speed, reliability and capability. The range and capabilities of input/output and mass storage devices increased remarkably. In the software area, high-level languages (e.g., FORTRAN, COBOL, BASIC) became common and manufacturers offered operating system software developed, for example, to manage time-sharing for a large number of computer users or real-time process control.

Most significantly, a trend of downward cost for given levels of performance was established. The minicomputer, aimed at providing a few users or even one single user with direct access to and control over his own computing facility, began to gain in numbers over the large, centrally managed computer system.

Fourth Generation

The semiconductor technology of large-scale integration (LSI) of the 1970s brought complete computers on a chip, known as microprocessors, allowing further refinement and enhancement of third-generation equipment.

Semiconductor memory has completely replaced core memory, and the continuing reduction in size and cost has brought the personal computer, numbered in millions of units supplied, truly within the reach of individuals in their own homes or offices. The 1980s saw the arrival of very large scale integration (VLSI) and ultra large scale integration (ULSI) applied to semiconductor memory and processor circuit design. This has resulted in a scalar increase in the density of memory capacity that can be packed onto a single chip, as well as the proliferation of multiprocessors on a single board. Many computers of power and memory capacity equal to or greater than that of the earliest minis can now be found on a single board.

The turn of the decade also saw the introduction of surface mount technology, which broke through the previous barrier of physical limitation of the number of connections that could be built between the memory or processor chip and the outside world with which they communicated. This has also permitted closer packing of chips onto each board.

Fifth Generation

People disagree as to what will be the fifth generation of computer development but some feel it will be the development of a ''thinking computer'' using techniques developed from artificial intelligence research. Artificial intelligence techniques have been used in specialized applications to develop what is known as expert systems, but these cannot be regarded yet as thinking computers.

5.3 DIGITAL COMPUTER SYSTEMS

Digital computers in various forms are now used universally in almost every walk of life. In many cases unseen, computers nonetheless influence people in activities such as travel, banking, education and medicine. There is also a very wide range of computers designed for use by people without technical skills or training in their own homes. These are generally used for some form of entertainment or for introduction to the wider subject of computing, but some are large enough and powerful enough to support small businesses. There is also a growing range of truly portable computers now available that will fit into briefcases or even pockets. As can be imagined from the variety of applications, computers exist in many different forms, spanning a range of price (from the smallest personal system to the largest supercomputer). Yet there are certain features that are common to all digital computers:

1. Construction from circuits that have two stable states, forming binary logic elements.
2. Some form of binary storage of data.
3. Capability to receive and act on data from the outside world (input) and to transmit data to the outside world (output).
4. Operation by executing a set of discrete steps or instructions, the sequence of which can be created and modified at any time to carry out a particular series of tasks. This ability to be programmed, with a program stored in the system itself, is what gives great flexibility to the digital computer. Recent advances have enabled changes to be made to a program dynamically while it is in operation. In addition, modern techniques permit processing to be carried out on the same set of discrete steps on more than one CPU at the same time within the same logical computer. This technique is referred to as *parallel processing*.

The computer has also given rise to a whole new series of professions—those of the computer programmer, systems analyst, data specialist, business analyst, and computer operator—as well as the industry of designing, building and maintaining computers themselves.

Central Processor Unit

The CPU is where instructions forming the stored program are examined and executed and is therefore in control of the operation of the system. Instructions and data for immediate processing by the CPU are held in main memory, which is linked directly to the CPU. This general term covers the units that perform logic decisions and arithmetic, collectively known as the arithmetic and logic unit (ALU). See Figure 5.1.

In recent years considerable progress has been made on computers that contain more than one discrete CPU. These are referred to as multiple processor computers. The main area of difficulty in the development of these machines has not been in the physical co-location of processors but rather in the design of the operating system. It has to know which instructions can be operated upon simultaneously and which require prior instructions to be completed. It should also be noted that multiple processor computers could operate in two distinct modes. In *parallel* processing each processor is operating on the same instruction at the same time. This is particularly useful in fail-safe or nonstop critical applications. In *multiprocessing,* each processor is operating on a different instruction and each instruction is processed once only. This does not give the fail-safe advantages but does provide large gains in speed of processing overall.

Input/Output

This is the structure, which provides optimum communication between the CPU and other parts of the system.

Peripherals

These are the devices external to the CPU and memory, which provide bulk storage, human/machine interaction, and communication with other electronic systems.

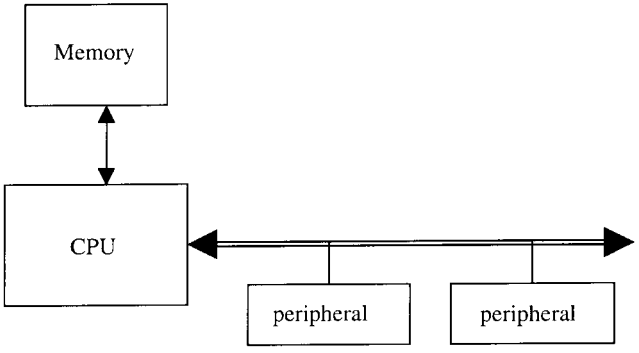


FIGURE 5.1 Components of a computer system.

BUS Paths

These act as the vehicle for the passing of data and program instructions between all other parts.

5.4 CATEGORIES OF COMPUTER SYSTEMS

Several distinct categories of computer system can be identified. Going from least to most comprehensive and powerful, these are described in subsequent sections.

Microcomputers

Microcomputers are the most widely available computers, with today's microcomputer being much more powerful than mainframes of even 5 years ago. Very few people in Western society have not come into contact with microcomputers. They are not only in use in direct applications but can also be found in domestic appliances, automobiles, etc.

The first microcomputers were appearing in the mid-1970s and were mainly intended for use by specialist engineers requiring computing power remote from central resources for large and intricate calculations. However, they were still large and expensive by modern standards and so did not come into general use.

The second half of the 1970s saw the introduction of mass-produced (and hence cheap) microcomputers, often with an integral screen, such as Apple, Nascom, and the Commodore PET. These were aimed initially at the personal market and great numbers were sold on the basis of video games provided for them. This quickly led to microcomputers being developed for the business sector. These possess memories and facilities far exceeding those of early minicomputers and, in some cases, on a par with smaller mainframes.

The late 1990s saw the rapid adoption of microcomputers by society in general with considerable home ownership of computers.

Microcomputers fall into two basic categories—desktop and portable.

Desktop computers (also called personal computers) are small enough to fit onto the desk. They usually consist of at least a keyboard, mouse, display screen, memory, processor, and disk drive although more and more home computers also have multimedia peripherals to enable sound and video output. Mostly they are used to run a set of purchased easy-to-use application software. Some desktop computers are referred to as workstations, but the distinction between a workstation and a basic desktop computer has become blurred. Workstations were designed to run more advanced applications (such as engineering design software) and consequently were much more powerful versions of desktop computers. However the development of the desktop computer, particularly to meet the demands of the home video game market, has meant that there is little to distinguish between the desktop computers and workstations.

Portable computers are microcomputers that are both small enough and light enough to carry around. Although names such as notebook computer and subnotebook computer are still used, there are really only two types of portable computers—the notebook and the personal digital assistants (PDA).

Notebook computers weigh between 1 and 4 kgs and usually consist of the same configuration as the desktop computer but are made with components that are much lighter in weight (and generally more expensive). They can run the same applications software as the desktop computers and use the same operating system software. The price of notebooks is such that a number of people are purchasing notebook computers in preference to desktop computers to save space on their desks or because they want to use the same computer at work and away from the office.

Personal digital assistants (PDAs) also known as palmtop computers, handheld PCs, or mass access devices (MADS), are very small computers that usually fit into the pocket and run a restricted set of applications—usually diary management, task list management, note taking and e-mail.

Minicomputers

Since its introduction as a recognizable category of system in the mid-1960s, with machines such as the Digital Equipment Corporation PDP8, the minicomputer has evolved rapidly. It has been the development that has brought computers out of the realm of specialists and large companies into common and widespread use by non-specialists.

The first such systems were built from early integrated-circuit logic families, with core memory. Characteristics were low cost, ability to be used in offices, laboratories, and even factories, and simplicity of operation allowing them to be used, and in many cases programmed, by the people who actually had a job to be done, rather than by specialist staff remote from the user. These were also the first items of computer equipment to be incorporated by original equipment manufacturers (OEMs) into other products and systems, a sector of the market that has contributed strongly to the rapid growth of the minicomputer industry. They led to the development of the microcomputer with early microcomputers being formally minicomputers but now made onto a single chip.

Applications of minicomputers are almost unlimited in areas such as laboratories, education, commerce, industrial control, medicine, engineering, government, banking, networking, CAD/CAM, CAE, and CIM.

There is now no real distinction between a minicomputer and a microcomputer. A better definition is to use the term *midrange computer.* These are usually used to support a number of people simultaneously for applications such as data processing or research.

Mainframes

The mainframe is the class of system typically associated with commercial data processing in large companies where a centralized operation is feasible and desired and very large volumes of data are required to be processed at high process speeds, or where a large user base (often in excess of 100 simultaneous users) requires immediate responses during interactive sessions. Today's mainframes, all products of large, established companies in the computer business (except for systems that are software-compatible emulators of the most popular mainframe series) are the successors to the first and second generation. They inherit the central control and location, emphasis on batch processing and line printers, third- and fourth-generation programming, and the need for specialized operating staff.

Mainframes are capable of supporting very large amounts of on-line disk and magnetic tape storage as well as large main memory capacity, and data communications capabilities supporting remote terminals of various kinds. Although some of the scientific mainframes have extremely high operating rates, most commercial mainframes are distinguished more by their size, mode of operation, and support than by particularly high performance.

Supercomputer

The most powerful type of computer is referred to as a supercomputer. These computers are generally specially built to perform large numbers of calculations very fast and are used for such applications as weather forecasting, mathematical modelling, etc.

5.5 CENTRAL PROCESSOR UNIT

This part of the system controls the sequence of individual steps required for the execution of instructions forming a program. These instructions are held in storage and, when executed in the appropriate order, carry out a task or series of tasks intended by the programmer.

Within any particular computer system, the word length is the fixed number of binary digits of which most instructions are made up. Arithmetic operations within the CPU are also performed on binary numbers of this fixed word length, normally 8,16, 24, 32, 36, 64 binary digits or bits. The CPU is connected via a memory bus, as in Figure 5.2, to a section of memory organized as a number of randomly accessible words, each of which can be written to or read from individually. The time for reading one word from or writing one word into main memory is typically in the range of nanoseconds or microseconds, depending upon CPU and memory BUS speed. Each word or location of memory can contain either an instruction or data. Apart from simple systems, some form of magnetic tape or disk memory peripheral is present on a system as file storage and back-up to main memory.

Control and timing circuits in the CPU enable instructions and data to be fetched from memory for processing and the results of any instructions required to be stored for further processing to be written into memory. The program counter holds the memory address of the next instruction to be fetched after each instruction has been processed. Frequently, the next instruction is held in the next location in memory and the counter need simply be incremented by one. However, some systems work by placing one or more parameters for the instruction immediately after it in memory, thereby causing the next instruction to be displaced further down memory. At other times the sequence of the program dictates that a new value be written into the program counter. Instructions, which alter the sequence of a program, calculate and insert a new value into the program counter for the next instruction are called branch instructions.

In order to start the CPU when no programs are already in memory, the program counter is loaded with a predetermined address, usually by the action of switching power onto the system, and a simple loader program held in ROM is loaded into memory. Its function is to load a comprehensive general-purpose loader, which automatically loads user or system programs. This process is known as bootstrap-

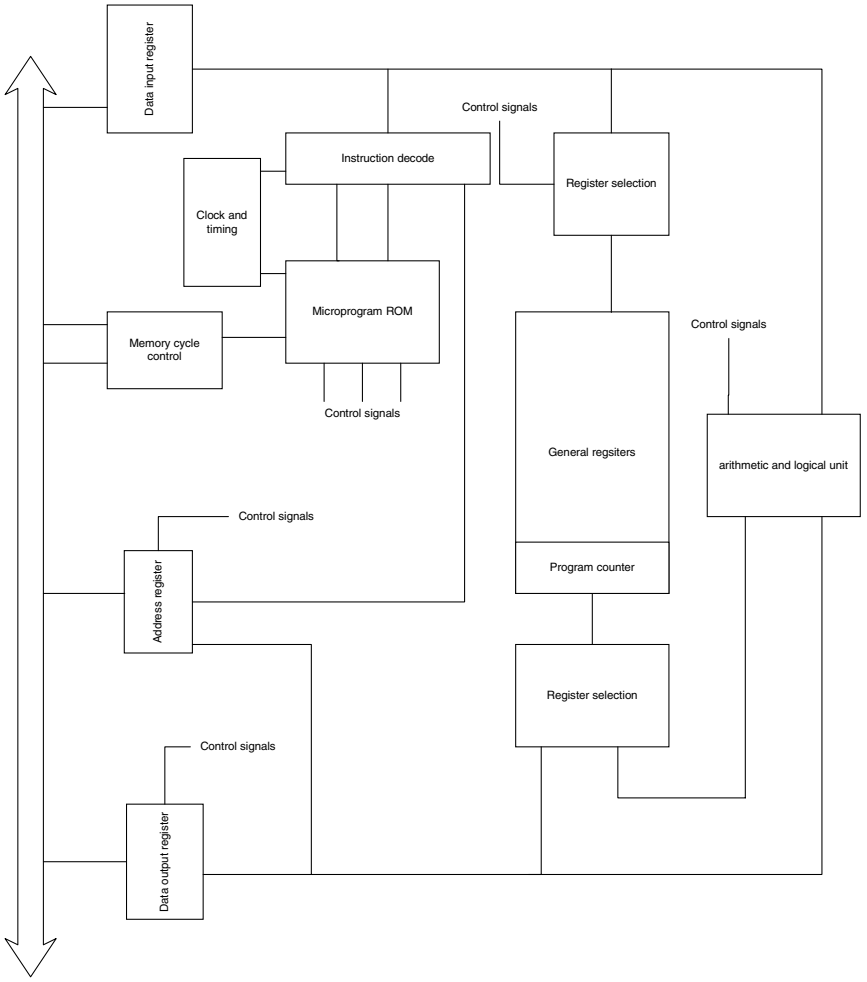


FIGURE 5.2 CPU block diagram.

ping, or booting the system, and the initial ROM program is known as the bootstrap loader. However, today most systems perform all these functions as a result of switching on the power automatically and are ready for use almost immediately, requesting date and time only if they are not held in any form of battery backed-up memory.

Instruction Set

The number and complexity of instructions in the instruction set or repertoire of different CPUs varies considerably. The longer the word length, the greater is the

variety of instructions that can be coded within it. This means, generally, that for a shorter word-length CPU a larger number of instructions will have to be used to achieve the same result, or that a longer word-length machine with its more powerful set of instructions needs fewer of them and hence should be able to perform a given task more quickly.

Instructions are coded according to a fixed format, allowing the instruction decoder to determine readily the type and detailed function of each instruction presented to it. Digits forming the operation code in the first byte (sometimes also the second) are first decoded to determine the category of instruction, and the remaining bytes interpreted in a different way, depending into which category the instruction falls. There are variations to the theme outlined above for CPUs from differing manufacturers, but generally they all employ the principle of decoding a certain group of digits in the instruction word to determine the class of instruction, and hence how the remaining digits are to be interpreted.

The contents of a memory location containing data rather than an instruction are not applied to the instruction decoder. Correct initial setting of the program counter (and subsequent automatic setting by any branch instruction to follow the sequence intended by the programmer) ensures that only valid instructions are decoded for execution. In the cases where operands follow the instruction in memory, the decoder will know how many bytes or words to skip in order to arrive at the next instruction in sequence.

Logic and arithmetic instructions perform an operation on data (normally one or two words for any particular instruction) held in either the memory or registers in the CPU. The addressing modes available to the programmer define the range of possible ways of accessing the data to be operated on. This ranges from the simple single-operand type of CPU (where the accumulator is always understood to contain one operand while the other is a location in memory specified by the addressing bits of the instruction) to a multiple-operand CPU with a wide choice of how individual operands are addressed.

In some systems instructions to input data from (and output data to) peripheral devices are the same as those used for manipulating data in memory. This is achieved by implementing a portion of the memory addresses at the high end as data and control registers in peripheral device controllers,

Certain basic data transfer, logical, arithmetic and controlling functions, must be provided in the instruction sets of all CPUs. This minimum set allows the CPU to be programmed to carry out any task that can be broken down and expressed in these basic instructions. However, it may be that a program written in this way will not execute quickly enough to perform a time-critical application such as control of an industrial plant or receiving data on a high-speed communications line. Equally, the number of steps or instructions required may not fit into the available size of memory. In order to cope more efficiently with this situation (i.e., to increase the power of the CPU), all but the very simplest CPUs have considerable enhancements and variations to the basic instruction set. The more comprehensive the instruction set, the fewer the steps required to program a given task and the shorter and faster in execution the resulting programs.

Basic types of instruction, with the examples of the variations to these, are described in the following sections.

Data Transfer. This loads an accumulator from a specified memory location and writes the contents of the accumulator into a specific memory location. Most CPUs have variations such as adding contents of memory location to the accumulator and exchanging the contents of the accumulator and memory locations.

CPUs with multiple registers also have some instructions that can move data to and from these registers, as well as the accumulator. Those with 16-bit or greater word lengths may have versions of these and other instruction types that operate on bytes as well as words.

Boolean Logical Function. This is a logical "AND" function on a bit-by-bit basis between the contents of a memory location and a bit pattern in the accumulator. It leaves ones in accumulator bit positions that are also one in the memory word. Appropriate bit patterns in the accumulator allow individual bits of the chosen word to be tested.

Many more logical operations and tests are available to more powerful CPUs, such as "OR," exclusive "OR," complement, branch if greater than or equal to zero, branch if less than or equal to zero, branch if lower or the same. The branch instructions are performed on the contents of the accumulator following a subtraction of comparison of two words, or some other operation that leaves data in the accumulator. The address for branching to is specified in the address part of the instruction. With a skip, the instruction in the next location should be an unconditional branch to the code that is to be followed if the test fails, while for a positive result, the code to be followed starts in the next but one location.

Branch or skip tests on other status bits in the CPU are often provided (e.g., on arithmetic carry and overflow).

Input/Output. CPUs with memory-mapped input/output do not require separate instructions for transferring data and status information between CPU and peripheral controllers. For this function, as well as performing tests on status information and input data, the normal data transfer and logical instructions are used.

Otherwise, separate input/output instructions provide these functions. Their general format is a transfer of data between the accumulator or other registers and addressable data, control or status registers in peripheral controllers. Some CPUs also implement special input/output instructions, such as:

1. *Skip if "ready" flag set.* For the particular peripheral being addressed, this instruction tests whether it has data awaiting input or whether it is free to receive new output data. Using a simple program loop, this instruction will synchronize the program with the transfer rate of the peripheral.
2. *Set interrupt mask.* This instruction outputs the state of each accumulator bit to an interrupt control circuit of a particular peripheral controller, so that by putting the appropriate bit pattern in the accumulator with a single instruction, interrupts can be selectively inhibited or enabled in each peripheral device.

Arithmetic

1. *Add contents of memory location to contents of accumulator, leaving result in accumulator.* This instruction, together with instructions for handling a carry bit from the addition and for complementing a binary number, can be used to carry out all the four arithmetic functions by software subroutines.
2. *Shift.* This is also valuable in performing other arithmetic functions, or for sequentially testing bits in the accumulator contents. With simpler instruction sets, only one bit position is shifted for each execution of the instruction. There is usually a choice of left and right shift, and arithmetic shift (preserving the sign of the word and setting the carry bit) or logical rotate.

Extended arithmetic capability, either as standard equipment or a plug-in option, provides multiply and divide instructions and often multiple-bit shift instructions.

Control. Halt, no operation, branch, jump to subroutine, interrupts on, interrupts off, are the typical operations provided as a minimum. A variety of other instructions will be found, specific to individual CPUs.

CPU Implementation

The considerable amount of control logic required to execute all the possible CPU instructions and other functions is implemented in one of two ways.

Random Logic. Random logic uses the available logic elements of gates, flip-flops, etc., combined in a suitable way to implement all the steps for each instruction, using as much commonality between instructions as possible. The various logic combinations are invoked by outputs from the instruction decoder.

Microcode. This is a series of internally programmed steps making up each instruction. These steps or microinstructions are loaded into ROM using patterns determined at design time, and for each instruction decoded, the microprogram ROM is entered at the appropriate point for that instruction. Under internal clock control, the microinstructions cause appropriate control lines to be operated to effect the same steps as would be the case if the CPU were designed using random logic.

The great advantage of microcoded instruction sets is that they can readily be modified or completely changed by using an alternative ROM, which may simply be a single chip in a socket. In this way a different CPU instruction set may be effected. In conjunction with microcode, bit-slice microprocessors may be used to implement a CPU. The bit-slice microprocessor contains a slice or section of a complete CPU, i.e., registers, arithmetic, and logic, with suitable paths between these elements. The slice may be 1, 2, or 4 bits in length, and by cascading a number of these together, any desired word length can be achieved. The required instruction set is implemented by suitable programming of the bit-slice microprocessors using their external inputs controlled by microcode.

The combination of microcode held in ROM and bit-slice microprocessors is used in the implementation of many CPU models, each using the same bit-slice device.

CPU Enhancements

There are several areas in which the operating speed of the CPU can be improved with added hardware, either designed in as an original feature or available as an upgrade to be added in-field. Some of the more common areas are described below.

Cache Memory. An analysis of a typical computer program shows that there is a strong tendency to access repetitively instructions and data held in fairly small contiguous areas of memory. This is due to the fact that loops (short sections of program reused many times in succession) are very frequently used, and data held in arrays of successive memory locations may be repetitively accessed in the course of a particular calculation. This leads to the idea of having a small buffer memory,

of higher access speed than the lower-cost technology employed in main memory, between CPU and memory. This is known as cache memory. Various techniques are used to match the addresses of locations in cache with those in main memory, so that for memory addresses generated by the CPU, if the contents of that memory location are in cache, the instruction or data are accessed from the fast cache instead of slower main memory. The contents of a given memory location are initially fetched into cache by being addressed by the CPU. Precautions are taken to ensure that the contents of any location in cache that is altered by a write operation are rewritten back into main memory so that the contents of the location, whether in cache or main memory, are identical at all times.

A constant process of bringing memory contents into cache (thus overwriting previously used information with more currently used words) takes place completely transparently to the user. The only effect to be observed is an increase in execution speed. This speeding up depends on two factors: hit rate (i.e., percentage of times when the contents of a required location are already in cache) and the relative access times of main and cache memory. The hit rate, itself determined by the size of cache memory and algorithms for its filling, is normally better than 90%. This is dependent, of course, on the repetitiveness of the particular program being executed. The increased speed is achieved by using faster, more expensive memory. The additional expense for the relatively small amount of memory being used is more than offset by the speed advantage obtained.

RISC Computers. Most computers require an instruction set of considerable size and complexity in order to provide all the facilities contained in the operating systems that support and manage them. This arrangement has many advantages, especially for commercial organizations, but also suffers from a distinct disadvantage—the more complex the instruction set, the more processor time and effort is required to decode and carry out each instruction. This can (and does) lead to significant reductions in overall processor performance for very large and complex operating.

Research into ways of solving this problem began in the late 1970s, principally in the United States, but it was not until 1984 that the first commercially available computer with a *reduced* instruction set was sold by Pyramid Technology. This design gave rise to the term *reduced instruction set computer,* or RISC, as it is more commonly referred to today. In order to distinguish between these processors and the normal complex instruction set computers that preceded them, the term *complex instruction set computer* (CISC) was also brought into general use.

Within a RISC processor all superfluous or little-used instructions are removed from the operating system. All instructions will generally be of the same length and take the same amount of time to process. Both of these characteristics enable pipelining and other techniques to be used to effect savings in the time taken to execute each instruction. Typically, all instructions are hardwired in the processor chip (also faster than resorting to microcode). Much use is made of a higher number of registers than normal, thus there are many more instructions address registers as opposed to main memory. Where memory is addressed, it is often within the very large cache memories that are another feature of RISC processors. All these characteristics contribute to the faster processing speed per instruction with a RISC architecture. However, since the instructions are simpler and microcode is not used, some functionality requires many more instructions on RISC than on CISC processors.

Most RISC processors run under the UNIX operating system (or one of its clones) since this system is simpler and easier to gain entry to than most proprietary operating systems. An important player in the RISC arena is Sun Microsystems Inc. with its open SPARC (Scalable Processor ARChitecture) RISC architecture.

Fixed and Floating-Point Arithmetic Hardware

As far as arithmetic instructions go, simpler CPUs only contain add and subtract instructions, operating on single-word operands. Multiplication, of both fixed and floating point numbers, is then accomplished by software subroutines, i.e., standard programs that perform multiplication or division by repetitive use of the add or subtract instructions, which can be invoked by a programmer who must perform a multiplication or division operation.

By providing extra hardware to perform fixed-point multiply and divide, which also usually implements multiple place-shift operations, a very substantial improvement in the speed of multiply and divide operations is obtained. With the hardware techniques used to implement most modern CPUs, however, these instructions are wired in as part of the standard set.

Floating-point format provides greater range and precision than single-word fixed-point format. In floating-point representation, numbers are stored as a fraction times 2^n, where n can be positive or negative. The fraction (or mantissa) and exponent are what is stored, usually in two words for single-precision floating-point format or four words for double precision.

Hardware to perform add, subtract, multiply, and divide operations is sometimes implemented as a floating-point processor, an independent unit with its own registers to which floating-point instructions are passed. The floating-point processor (sometimes called co-processor) can then access the operands, perform the required arithmetic operation, and signal the CPU, which has meanwhile been free to continue with its own processing until the result is available.

An independent floating-point processor clearly provides the fastest execution of these instructions, but even without that, implementing them within the normal instruction set of the CPU, using its addressing techniques to access operands in memory provides a significant improvement over software subroutines. The inclusion of the FTP into standard CPUs has become standard.

Array Processors

Similar to an independent floating-point processor described above, an optional hardware unit that can perform complete computations on data held in the form of arrays of data in memory, independent from the CPU and at high speed, is known as an array processor. These are used in specialized technical applications such as simulation, modeling, and seismic work. An example of the type of mathematical operation that would be carried out by such a unit is matrix inversion. The ability of these units to perform very high-speed searches upon text keys has also led to a growing use of them for the rapid retrieval of data from large data banks, particularly in areas such as banking, where real-time ATM terminals require fast response to account enquiries from very large data sets.

Timers and Counters

For systems that are used in control applications, or where elapsed time needs to be measured for accounting purposes (as, for example, in a time-sharing system where users are to be charged according to the amount of CPU time they use), it is important to be able to measure intervals of time precisely and accurately. This measurement must continue while the system is executing programs, and must be real-time, i.e., related to events and time intervals in the outside world.

Some CPUs are equipped with a simple real-time clock that derives its reference timing from 50–60-Hz mains. These allow a predetermined interval to be timed by setting a count value in a counter that is decremented at the main's cycle rate until it interrupts the CPU on reaching zero.

More elaborate timers are available on some CPUs. These are driven from high-resolution crystal oscillators and offer such features as:

1. More than one timer simultaneously
2. Timing random external events
3. Program selection of different time bases
4. External clock input

The system supervisory software normally keeps the date and time of day up to date by means of a program running in the background all the time the system is switched on and running. Any reports, logs, or printouts generated by the systems can then be labeled with the date and time they were initiated. To overcome having to reset the data and time every time the system is stopped or switched off, CPUs have a permanent battery-driven date and time clock which keeps running despite stoppages and never needs reloading once loaded initially (with the exception of change to and from Daylight Savings Time).

Counters are also useful in control applications to count external events or to generate a set number of pulses (for example, to drive a stepping motor). Counters are frequently implemented as external peripheral devices, forming part of the digital section of a process input/output interface.

5.6 MEMORY

Computers recognize two states—on or off, electrically charged or not. These two states are represented by the binary digits 0 and 1. Various combinations of 1's and 0's can represent numbers, letters, and symbols and form the memory of the computer. Each individual 0 or 1 is called a binary digit (bit). Grouping of bits that represent characters are called bytes. For example, the character A is represented by the binary byte 1000001. The size of a computer's memory is expressed as the number of bytes of storage that it can hold, usually expressed as kilobytes or megabytes—a 34-megabyte memory chip holds 34,000,000 bytes of information.

The computer handles sequences of bytes as words, with a word being defined as the maximum number of bits that can be handled as one unit by the CPU. Most data processing computers will have a word size of 32 or 64 bits, but some signal processing computers will still have small word sizes of only 8 bits.

In order to provide storage for program instructions and data in a form in which they can be directly accessed and operated upon, all computers have main memory that can be implemented in a variety of technologies and methods of organization.

Memory Organization

Memory is organized into individually addressable words into which binary information can be loaded for storage and from which the storage data pattern can be read. On some systems, memory is arranged in such a way that more than one word at a time is accessed. This is done to improve effective memory access rates, on the basis that by accessing, say, two consecutive words, on most occasions the second word will be the one that the CPU requires next. This is generally referred to as interleaved memory.

A memory controller is required between the memory arrays and the CPU to decode the CPU requests for memory access and initiate the appropriate read or write cycle. A controller can only handle up to a certain maximum amount of memory, but multiple controllers can be implemented on a single system. This can be used to speed up effective memory access by arranging that sequentially addressed locations are physically in different blocks of memory with different controllers. With this interleaved memory organization, in accessing sequential memory locations the operation of the controllers is overlapped, i.e., the second controller begins its cycle before the first has completed. Aggregate memory throughput is thus speeded up. In some more complex computer systems all or part of the memory can be shared between different CPUs in the multiprocessor configuration. Shareable memory has a special form of controller with multiple ports allowing more than one CPU access to the memory.

It is sometimes appropriate to implement two types of memory in one system: random access or read/write memory (RAM) and read-only memory (ROM). Programs have to be segregated into two areas:

1. Pure instructions, which will not change, can be entered into ROM.
2. Areas with locations that require to be written into (i.e., those containing variable data or modifiable instructions) must occupy RAM.

Read-only memory is used where absolute security from corruption of programs, such as operating system software or a program performing a fixed control task, is important. It is normally found on microprocessor-based systems and might be used, for example, to control the operation of a bank's cash dispenser.

Use of ROM also provides a low-cost way of manufacturing in quantity a standard system that uses proven programs that never need to be changed. Such systems can be delivered with the programs already loaded and secure, and do not need any form of program-loading device.

Memory Technology

The most common technologies for implementing main memory in a CPU are described in the following sections.

MOS Random Access Memory (RAM). MOS (metal oxide semiconductor) technology is very widely used, with abundant availability from the major semiconductor suppliers. Very high density has been achieved, with up to 256 megabytes of memory available on a single chip. Dynamic MOS RAMS require refresh circuitry that, at intervals, automatically rewrites the data in each memory cell. Static RAM, which does not require refreshing, can also be used. This is generally faster, but also more expensive, than dynamic RAM.

Semiconductor RAMs are volatile, i.e., they lose contents on powering down. This is catered for in systems with back-up storage (e.g., disks) by reloading programs from the back-up device when the system is switched on, or by having battery back-up for all or part of the memory.

In specialized applications requiring memory retention without mains for long periods, CMOS (complementary metal-oxide semiconductor) memory can be used. CMOS memory is powered by a battery and does not lose its contents when the power is turned off. Although it has a very low current drain, it has the disadvantage of being more expensive than normal MOS memory. Where it is essential to use CMOS, circuit boards with on-board battery and trickle charger are now available.

ROM. Read-only memories, used as described above, can be either erasable ROMs or a permanently loaded ROM.

ROM BIOS. The ROM Basic Input and Output System (BIOS) is a special chip that holds special software for the computer. This software at switch on first checks that the computer is working correctly and then loads the operating system from the disk into the RAM.

Flash BIOS. Most modern computers use flash BIOS rather than ROM BIOS; they both are used to hold the same type of software, but flash BIOS has the advantage that the software on the chip may be upgraded. This upgrade is achieved by running a program supplied by the manufacturer of the computer system.

Registers. The CPU contains a number of registers accessible by instructions, together with more that are not accessible but are necessary part of its implementation. Other than single-digit status information, the accessible registers are normally of the same number of bits as the word length of the CPU.

Registers are fast-access temporary storage locations within the CPU and implemented in the circuit technology of the CPU. They are used, for example, for the temporary storage of intermediate results or as one of the operands in an arithmetic instruction. A simple CPU may have only one register, often known as the accumulator, plus perhaps an auxiliary accumulator or quotient register used to hold part of the double length result of a binary multiplication.

More sophisticated CPUs typically have many more general-purpose registers that can be selected as operands by instructions. Some systems use one of the general-purpose registers as the program count and can use any register as a stack pointer. A stack in this context is a temporary array of data held in memory on last-in, first-out basis. It is used in certain types of memory reference instructions and for internal housekeeping in interrupt and subroutine handling. The stack pointer register is used to hold the address of the top element of the stack. The address, and hence the stack pointer contents, are incremented or decremented one at a time as data are added to or removed from the stack.

Memory Addressing. Certain instructions perform an operation in which one or more of the operands is the contents of a memory location (for example, arithmetic, logic, and data-movement instructions). In most sophisticated CPUs various addressing modes are available to give, for example, the capacity of adding together the contents of two different memory locations and deposit the result in a third.

In such CPUs instructions are double operand, i.e., the programmer is not restricted to always using one fixed register as an operand. In this case, any two of

the general-purpose registers can be designated either as each containing operand or through a variety of addressing modes, where each of the general-purpose registers selected will contain one of the following:

1. The memory address of an operand
2. The memory address of an operand, and the register contents are then incremented following execution
3. The memory address of an operand, and the register contents are then decremented following execution
4. A value to which is added the contents of a designated memory location (this is known as indexed addressing)
5. All of the above, but where the resultant operand is in the address of the final operand (known as indirect deferred addressing)

This richness of addressing modes is one of the benefits of more advanced CPUs, as, for example, it provides an easy way of processing arrays of data in memory, or of calculating the address portion of an instruction when the program is executed.

Further flexibility is provided by the ability on many processors for many instructions to operate on multiples of bits (known as a byte), on single bits within a word and some more comprehensive CPUs, on double- and quadruple-length words and also arrays of data in memory.

Memory Management. Two further attributes may be required of memory addressing. Together they are often known as memory management. This is the ability, particularly for a short word-length system (16 bits or less), for a program to use addresses greater than those implied by the word length. For example, with the 16-bit word length of older minicomputers the maximum address that can be handled in the CPU is 65,536. As applications grew larger this became a limitation. Extended addressing operates by considering memory as a number of pages. Associated with each page at any given time is a relocation constant that is combined with relative addresses within its page to form a longer address. For example, with extension to 18 bits, memory addresses up to 262,144 can be generated in his way. Each program is still limited at any given time to 65,536 words of address space, but these are physically divided into a number of pages that can be located anywhere within the larger memory. Each page is assigned a relocation constant, and as a particular program is run, dedicated registers in the CPU memory management unit are loaded with the constant for each page.

Thus, many logically separate programs and data arrays can be resident in memory at the same time, and the process of setting the relocation registers, which is performed by the supervisory program, allows rapid switching between them, in accordance with a time-scheduling scheme that is usually based upon resource usage quota, time allocation, or a combination of both. This is known as multiprogramming. Examples of where this is used are a time-sharing system for a number of users with terminals served by the system, or a real-time control system where programs of differing priority need to be executed rapidly in response to external events.

Memory Protection. As an adjunct to the hardware for memory paging or segmentation described above, a memory-protection scheme is readily implemented. As well as a relocation constant, each page can be given a protection code to prevent its being illegally accessed. This would be desirable, for example, for a page hold-

ing data that are to be used as common data among a number of programs. Protection can also prevent a program from accessing a page outside of its own address space.

Multiprogramming. Memory addressing and memory management are desirable for systems performing multiprogramming. In such systems the most important area to be protected is that containing the supervisory program or operating system, which controls the running and allocation of resources for users' programs.

Virtual Memory. Programmers frequently have a need for a very large address space within a single program for instructions and data. This allows them to handle large arrays and to write very large programs without the need to break them down to fit a limited memory size.

One solution is known as virtual memory, a technique of memory management by hardware and operating systems software whereby programs can be written using the full addressing range implied by the word length of the CPU, without regard to the amount of main memory installed in the system. From the hardware point of view, memory is divided into fixed-length pages and the memory management hardware attempts to ensure that pages in most active use at any given time are kept in main memory. All the current programs are stored in a disk-backing store, and an attempt to access a page that is not currently in main memory causes paging to occur. This simply means that the page concerned is read into main memory into the area occupied by an inactive page, and if any changes have been made to the inactive page since it was read into memory, then it is written out to disk in its updated form to preserve its integrity.

A table of address translations holds the virtual physical memory translations for all the pages of each program. The operating system generates this information when programs are loaded onto the system, and subsequently keeps it updated. Memory protection on a per-page basis is normally provided, and a page can be locked into memory as required to prevent it being swapped out if is essential for it to be immediately executed without the time overhead of paging.

When a program is scheduled to be run by the operating system, its address translation table becomes the one in current use. A set of hardware registers to hold a number of the most frequent translations in current use speeds up the translation process when pages are being repeatedly accessed.

Input/Output. In order to perform any useful role, a computer system must be able to communicate with the outside world, either with human users via keyboards, display screens, printed output, etc. or with some external hardware or process being controlled or monitored. In the latter case, where connection to other electronic systems is involved, the communication is via electrical signals.

All modern computer systems have a unified means of supporting the variable number of such human or process input/output devices required for a particular application, and indeed for adding such equipment to enhance a system at the user's location. As well as all input/output peripherals and external mass storage in the form of magnetic tape, compact disk, and disk units, some systems also communicate with main memory in this common, unified structure. In such a system there is no difference between instructions that reference memory and those that read from and write to peripheral devices. The benefits of a standard input/output bus to the manufacturer are:

1. It provides a design standard allowing easy development of new input/output devices and other system enhancements.
2. Devices of widely different data transfer rates can be accommodated without adaptation of the CPU.
3. It permits development of a family concept.

Many manufacturers have maintained a standard input/output highway and CPU instruction set for as long as a decade or more. This has enabled them to provide constantly improving system performance and decreasing cost by taking advantage of developing technology while protecting very substantial investments in peripheral equipment and software. For the user of a system in such a family, the benefits are:

1. The ability to upgrade to a more powerful CPU while retaining existing peripherals
2. Retention of programs in moving up- or down-range within the family
3. In many cases the ability to retain the usefulness of an older system by adding more recently developed peripherals, and in some cases even additional CPU capacity of newer design and technology

Input/Output Bus. The common structure for any given model of computer system is implemented in the form of an electrical bus or highway. This specifies the number, levels, and significance of electrical signals and the mechanical mounting of the electrical controller or interface that transforms the standard signals on the highway to ones suitable for the particular input/output or storage device concerned. A data highway or input/output bus needs to provide the following functions.

Addressing. A number of address lines are provided, determining the number of devices that can be accommodated on the system. For example, six lines would allow 63 devices. Each interface on the bus decodes the address lines to detect input/output instructions intended for it.

Data. The number of data lines on the bus is usually equal to the word length of the CPU, although it may alternatively be a submultiple of the word length, in which case input/output data are packed into or unpacked from complete words in the CPU. In some cases data lines are N-directional, providing a simpler bus at the expense of more complex drivers and receivers.

Control. Control signals are required to synchronize transactions between the CPU and interfaces and to get address and data signals to and from the bus. Although all the bits of an address or data word are transmitted at the same instant, in transmission down the bus, because of slightly different electrical characteristics of each individual line, they will arrive at slightly different times. Control signals are provided to get these skewed signals at a time when they are guaranteed to have reached their correct state.

Types of Input/Output Transactions. Three types of transaction via the input/output bus between CPU and peripheral device are required, as described below.

Control and Status. This type of transfer is initiated by a program instruction to command a peripheral device to perform a certain action in readiness for transferring data or to interrogate the status of a peripheral. For example, a magnetic tape unit can be issued with a command to rewind, the read/write head in a disk

unit to be positioned above a certain track on the disk, the completion of a conversion by an analog-to-digital converter verified, or a printer-out-of-paper condition may be sensed.

Normally, a single word of control or status information is output or input as a result of one instruction with each bit in the word having a particular significance. Thus multiple actions can be initiated by a single control instruction and several conditions monitored by a single status instruction. For the more complex peripheral devices, more than one word of control or status information may be required.

Programmed Data Transfer. For slow and medium-speed devices (for example, floppy disk units or line printers) data are input or output one word at a time, with a series of program instructions required for every word transferred. The word or data are transferred to or from one of the CPU registers, normally the accumulator. In order to effect a transfer of a series of words forming a related block of data (as is normally required in any practical situation) a number of CPU instructions per word transferred are required. This is because it is necessary to take the data from (or store them into) memory locations. As a minimum, in a simple case at least six CPU instructions are required per word of data transferred.

In some systems where instructions can reference equally memory locations, peripheral device registers, and CPU registers, the operation is simplified since a MOVE instruction can transfer a word of data directly from a peripheral to memory without going through a CPU register. This applies equally to control and status instructions, with a further advantage that the state of bits in a peripheral device status register can be tested without transferring the register contents into the CPU.

The rate of execution of the necessary instructions must match the data transfer rate of the peripheral concerned. Since it is usually desired that the CPU continue with the execution of other parts of the user's program while data transfer is going on, some form of synchronization is necessary between CPU and peripheral to ensure that no data are lost. In the simplest type of system, the CPU simply suspends any other instructions and constantly monitors the device status word, awaiting an indication that the peripheral has data ready for input to the CPU or is ready to receive an output from it. This is wasteful of CPU time where the data transfer rate is slow relative to CPU instruction speeds, and in this case the use of interrupt facilities provides this synchronization.

Direct Memory Access. For devices that transfer data at a higher rate (in excess of around 20,000 words per second) a different solution is required. At these speeds, efficiency is achieved by giving the peripheral device controller the ability to access memory autonomously without using CPU instructions. With very fast tape or disk units that can transfer data at rates in excess of 6 million bytes per second, direct memory access (DMA) is the only technique that will allow these rates to be sustained.

The peripheral controller has two registers that are loaded by control instructions before data transfer can begin. These contain:

1. The address in memory of the start of the block of data
2. The number of words that it is desired to transfer in the operation

When the block transfer is started, the peripheral controller, using certain control lines in the input/output bus, sequentially accesses the required memory locations until the specified number of words has been transferred. The memory addresses are placed on address lines of the input/output bus, together with the appropriate control and timing signals, for each word transferred. On completion of the number

of words specified in the word-count register, the peripheral signals to the CPU that the transfer of the block of data is completed.

Other than the instructions required initially to set the start address and word-count registers and start the transfer, a DMA transfer is accomplished without any intervention from the CPU. Normal processing of instructions therefore continues. Direct memory access (more than one peripheral at a time can be engaged in such an operation) is, of course, competing with the CPU for memory cycles, and the processing of instructions is slowed down in proportion to the percentage of memory cycles required by peripherals. In the limit, it may be necessary for a very high-speed peripheral to completely dominate memory usage in a burst mode of operation, to ensure that no data are lost during the transfer through conflicting requests for memory cycles.

Interrupts. The handling of input/output is made much more efficient through the use of a feature found in varying degrees of sophistication on all modern systems. Known as automatic priority interrupt, it is a way of allowing peripheral devices to signal an event of significance to the CPU (e.g., in some systems a keyboard having a character ready for transmission, or completion of DMA transfer) in such a way that the CPU is made to suspend temporarily its current work to respond to the condition causing the interrupt. Interrupts are also used to force the CPU to recognize and take action on alarm or error conditions in a peripheral (e.g., printer out of paper, error detected on writing to a magnetic tape unit).

Information to allow the CPU to resume where it was interrupted (e.g., the value of the program counter) is stored when an interrupt is accepted. It is necessary also for the device causing the interrupt to be identified and for the program to branch to a section to deal with the condition that caused the interrupt.

Examples of two types of interrupt structure are given below, one typical of a simpler system such as an 8-bit microprocessor or an older architecture minicomputer, the other representing a more sophisticated architecture such as the Pentium. In the simpler system a single interrupt line is provided in the input/output bus, onto which the interrupt signal for each peripheral is connected. Within each peripheral controller, access to the interrupt line can be enabled or disabled, either by means of a control input/output instruction to each device separately or by a ''mask'' instruction that, with a single 16-bit word output, sets the interrupt enabled/disabled state for each of up to 16 devices on the input/output bus. When a condition that is defined as able to cause an interrupt occurs in a peripheral, and interrupts are enabled in that device, a signal on the interrupt line will be sent to the CPU. At the end of the instruction currently being executed this signal will be recognized.

In this simple form of interrupt handing the interrupt servicing routine always begins at a fixed memory location. The interrupt forces the contents of the program counter (which is the address of the next instruction that would have been executed had the interrupt not occurred) to be stored in this first location and the program to start executing at the next instruction. Further interrupts are automatically inhibited within the CPU, and the first action of the interrupt routine must be to store the contents of the accumulator and other registers so that on return to the main stream of the program these registers can be restored to their previous state.

Identification of the interrupting device is done via a series of conditional instructions on each in turn until an interrupting device is found. Having established which device is interrupting, the interrupt-handling routine will then branch to a section of program specific to that device. At this point or later within the interrupt

routine an instruction to re-enable the CPU interrupt system may be issued, allowing a further interrupt to be received by the CPU before the existing interrupt-handling program has completed. If this "nesting" of interrupts is to be allowed, each interruptible section of the interrupt routine must store the return value of the program counter elsewhere in the memory, so that as each section of the interrupt routine is completed, control can be returned to the point where the last interrupt occurred.

A more comprehensive interrupt system differs in the following ways from that described above:

1. Multiple interrupt lines are provided, and any number of devices can be on each line or level.
2. The CPU status can be set to different priority levels, corresponding to different interrupt lines. Only interrupts on a level higher than the current priority are immediately serviced by the CPU. This provides a more adaptable way of dealing with a wide range of devices of different speeds and with different degrees of urgency.
3. When an interrupt is accepted by the CPU the interrupting device sends a vector or pointer to the CPU on the input/output bus address lines. This points to a fixed memory address for each device, which holds the start address of its interrupt routine, and in the following memory word, a new status word for the CPU, defining its priority level and hence its ability to respond to other levels of interrupt during this interrupt routine. Avoiding the need for the CPU to test each device until it finds the interrupting one means that response to interrupts is much faster.
4. The current value of the program counter and processor status word are automatically placed on a push-down stack when an interrupt occurs. A further interrupt accepted within the current interrupt routine will cause the program counter and status word to be stored on the top of the stack and the existing contents to be pushed down into the stack. On return from an interrupt routine, the program counter and status word stored when that interrupt occurred are taken from the top of the stack and used by the CPU, allowing whatever was interrupted to continue as before. This can take place for any number of interrupts, subject only to she capacity of the stack. Thus, nesting to any level is handled automatically without the need for the programmer to store the program counter at any stage.

5.7 PERIPHERALS

Peripheral devices fall into the following three categories.

Interactive

These are designed to allow humans to interact with the system by outputting information in the form of voice, readable alphanumeric text, or graphics, either on paper or on a display screen, and accepting information from humans through manual devices such as keyboards or voice-recognition devices or by scanning

printed text or images. The general function performed by devices in this class is sometimes referred to as human–machine interaction.

Storage

These act as a back-up form of storage to supplement the main memory of the system. The most simple of these (now superseded) was punched paper tape or cards, and the most complex range up to very large disks and magnetic tapes, each capable of holding up to millions of millions of characters of information. Peripherals of this type are generally known as mass-storage devices.

Developments in mass storage include compact disk (CD) and digital versatile disk (DVD). The CD used is essentially the same material and technique employed for the commercial recording of music. Most personal computers are now equipped with both readable and writable CD drives. The advantage is the considerable increase in data-storage density over conventional magnetic tape or disc.

Communication

These are interfaces between the computer system and other systems or electronic devices. Analog/digital converters, digital input/output, and communication line interfaces are good examples.

5.8 OUTPUT DEVICES

Either interactive or communications peripherals (or both) are required in every system. The existence of storage depends on the need for additional storage over and above main memory. All peripheral devices in a system are connected via the input/output structure to the CPU, memory, and in some systems a separate input /output processor.

The throughput rates and flexibility of the input/output structure determine the number and variety of peripheral devices which can be handled in a system before the input/ output requirements begin to saturate the system and prevent any processing of instructions being done by the CPU. In deciding on the configuration of a particular system it is important to analyze the throughput requirement dictated by peripheral devices, to ensure the system does not become input/output bound and that data from any peripheral devices are not lost due to other devices taking too many of the input/output resources.

Historically, in the computer industry independent manufacturers as well as the large computer systems companies have developed and manufactured peripherals. The products of the independent manufacturers are either bought by system manufacturers for design into their systems or sold by the independent manufacturers directly to users of the more popular computers, with an interface providing compatibility with the input/output bus of the system. This has fostered the development of many of the widely used, cost-effective peripherals available today, such as floppy disks and printers.

Certain storage devices with removable storage media, where the format recording data on the media have been standardized, can be used for exchanging data between systems from different suppliers. This is important where data may be gathered on one system and need to be analyzed on a different one more suitable for that purpose. However, due to the very large growth of networking in the 1980s, even between equipment from different manufacturers, the moving of data from one machine to another is most commonly achieved by file transmission, though for massively large files magnetic tapes are still sometimes used for this purpose.

5.9 TERMINALS

A data terminal is essentially an input/output device at which a computer user sits, either to receive data in alphanumeric or graphic form or to input data through a keyboard or other form of manual input, or both, to a computer system (sometimes called a host computer or a server). There are four types of terminals in common usage:

1. A *dumb terminal* is used to input and receive data only. It cannot process any data independently of a host computer system. Such terminals are used at airline reservation desks and banking applications.
2. An *intelligent terminal* includes a local processing unit, memory, and some form of secondary storage. Such a device is normally a microcomputer with some communications software to enable access to a host computer and a communications link such as a modem or network connection.
3. A *network terminal/computer* (also known as a *thin client*) is similar to an intelligent terminal but does not normally have any secondary storage. Any data used by the network computer is kept on a host or server computer. In large institutions this reduces the management problem of maintaining many independent disks of data on personal computers.
4. An *Internet terminal* (also known as a *Web terminal*) provides access to the Internet for email and Web browsing. These devices generally use a TV screen for the display and are directed at the home market.

Most applications in industry that still require the use of a terminal generally use a personal computer emulating the characteristics of a terminal.

Specialized terminals are common in areas such as:

1. Stores, as point-of-sale terminals, i.e., sophisticated cash registers linked to an in-store computer that is often adjusting a stock control system at the same time as registering the sale and debiting the customer's bank or credit card account. These are referred to as EFTPOS (Electronic Funds Transfer at Point of Sale) terminals.
2. Banking for customer cash dispensing, enquiries, and other transactions, or for teller use, including the ability to print entries in passbooks or to read a card's magnetic strip containing details of the customer's account and so eliminate the need for completion of a check or deposit slip.
3. Manufacturing for shop floor data collection and display.

These typically use features found in the terminals described above and, in addition, may have the capability to read magnetic stripes on credit cards, punched plastic cards, or identity badges, or bar codes on supermarket goods or parts in a factory.

5.10 PRINTERS AND PLOTTERS

Printers are categorized by how they produce characters on a page. There are two methods of impacting characters onto a page—direct impact and non-impact. An impact printer produces characters when a hammer, pin, or complete character strikes an ink ribbon, which in turn presses against a sheet of paper and then leaves an impression of the character on the paper. An example of such a device is a typewriter. Large computer installations still use line printers to produce large amounts of output, but the use of direct impact devices for personal computers has now largely been replaced by non-impact devices, which are generally much quieter.

Impact Printers

Dot Matrix Printer. The print head of a dot matrix printer consists of seven or more needles held in a vertical plane in the head assembly, which is positioned with the needles perpendicular to the paper and spaced a short distance from it, with carbon ribbon interposed between. Each needle can be individually driven by a solenoid to contact the paper through the ribbon, thus printing a dot. A complete character is formed stepping the head through five or more positions horizontally and at each position energizing the appropriate solenoids. The head is then stepped onto the position for the next character.

When the end of a line is reached the paper is advanced one line and the print head either returns to the left margin position or, in some faster printers, prints the next line from right to left. This is possible where the printer is provided with storage for a line or more of text and the characters can be extracted from this store in the reverse order. Throughput speed is improved where this technique is used, by saving redundant head movement. The 7×5 dot matrix within which this type of printer forms each character allows an acceptable representation of alpha and numeric characters.

Better legibility, particularly of lower-case characters with descenders, can be achieved by using a larger matrix such as 9×7, i.e., a head with nine needles stepping through seven positions for each character. Manufacturers can also supply 24-pin dot matrix printers in order to provide higher-quality print and a larger character set.

Character codes are received for printing on a line driven by a data-communications interface in the computer, whose transmission speed determines the overall printing throughput. Buffer storage of many millions of characters is provided in printers that use serial data communications, to make the most efficient use of communication lines, and in printers with built-in intelligence, to allow look-ahead so that the print head can skip blanks and take the shortest route to the next printable character position.

Character sets can be readily changed by replacing the ROM chip that contains the dot patterns corresponding to each character code or, more usually, by sending

the character patterns over the network to the printer from the host CPU, often referred to as *downline loading*. The latter method has the advantage of character selection at any time under program control and without any human intervention. Some printers already contain inbuilt multiple character sets.

Line Printer. For greater volume of printed output than can be achieved with serial printers, line printers that can produce a whole line of characters almost simultaneously are available. Using impact techniques, speeds up to 3500 full lines (usually 13 characters each) per minute are possible. Continuous paper fan-fold form, which may be multipart to produce copies, is fed through the printer a line at a time by a transport system consisting of tractors that engage sprocket holes at the edges of the paper to move it upwards and through the printer from front to rear. A paper tray at the rear allows the paper to fold up again on exit from the printer.

As well as advancing a line at a time, commands can be given to advance the paper to the top of the next page or to advance a whole page or line. This is important, for example, where preprinted forms are being used.

Two types of line printer are in common use: drum printers and band printers. Both use a horizontal row of hammers, one per character position or, in some eases, shared between two positions. These are actuated by solenoids to strike the paper through a carbon film against an engraved representation to print the desired character. In a drum printer, a print drum the length of the desired print line rotates once per print line. In each character position the full character set is engraved around the circumference of the drum. A band printer has a horizontal revolving band or chain of print elements, each with a character embossed on it. The full character set is represented on the band in this way. To implement different character fonts involves specifying different barrels in the case of a drum printer, whereas an operator changing bands can make a change readily on a band printer, or individual print elements in the band can be replaced.

The printer has a memory buffer to hold a full line of character codes. When the buffer is full (or terminated if a short print line is required) a print cycle is initiated automatically. During this print cycle the stored characters are scanned and compared in synchronism with the rotating characters on the drum or band. The printer activates the hammer as the desired character on the drum or band approaches in each print position.

Nonimpact Printers

Laser Printers. These are available to meet three different types of printing requirements:

1. Very high volumes of output at speeds exceeding 200 pages per minute. Those requiring a constant high-volume printing service, since this equipment is expensive to buy, run, and service, normally use them.
2. Departmental printing requirements, usually consisting of medium to high volumes on an *ad hoc* basis. This equipment would normally be networked to many CPUs and shared by a group of common users. They print at speeds of up to 40 pages per minute.
3. Desktop printing uses laser printers small enough to fit on an individual's desk, designed for intermittent low-volume personal printing requirements, usually at 10 pages per minute.

However, the technology used is common to all three. The principle used is that of the everyday photocopier, the difference being that the image to be copied is set up according to digital signals received from the host CPU instead of from a photoscan of the document to be copied. The main advantage of this form of output is the clarity and quality of the image printed. It is so good that it is possible not only to print data but also to print the form or letterhead of the paper at the same time, thus avoiding the cost of preprinted stationery. The disadvantage is that it is currently not possible to print multiple copies simultaneously.

Inkjet Printers. Inkjet printing, like laser printing, is a non-impact method. Ink is emitted from nozzles as they pass over a variety of possible media. Liquid ink in various colors is squirted at the paper to build up an image. A print head scans the page in horizontal strips, using a motor assembly to move it from left to right and back as another motor assembly rolls the paper in vertical steps. A strip of the image is printed, then the paper moves on, ready for the next strip. To speed things up, the print head prints not just a single row of dots in each pass, but a vertical row of dots at the same time.

Although you can now purchase color laser printers, they are far more expensive to buy than inkjet printers. Inkjets can produce much higher definition and hence are important for the digital photography market. They can produce approximately 10 pages per minute. However, they are expensive to maintain (but cheap to buy), their operating costs being ten times those of laser printers. Most home printers are of the inkjet variety.

Plotters

Pen Plotter. Pen plotters provide another form of hard copy output. These are devices aimed primarily at high-complexity graphics with a limited amount of text. Their uses range from plotting graphs of scientific data to producing complex engineering drawings in computer-aided design applications such as drawings used in integrated-circuit chip design.

The plotter has one or more pens held vertically above a table on which the paper lies. These can be of different colors, and as well as being raised or lowered on to the paper individually by program commands, they can be moved in small steps, driven by stepping motors. They plot in the X and Y directions or achieve control in one axis by moving the paper back and forth between supply and take-up rolls under stepping motor control, and in the other axis by pen movement. Combinations of movements in both axes produce diagonal lines.

With step sizes as small as 0.01 mm, high-accuracy plots can be produced (in multiple colors where more than one pen is used) and annotated with text in a variety of sizes and character sets. Supporting software is usually provided with a plotter. This will, for example, scale drawings and text and generate alphanumeric characters.

Electrostatic Plotter. The objectives of the electrostatic plotter are the same as those of the pen plotter: the production of high-quality graphics in hardcopy form. However, electrostatic plotters achieve their output by setting an electrostatic charge on the paper in the same pattern as the required output image and then attracting and retaining ink particles according to that pattern. This ink is then fused onto the paper in order to make a permanent image. This can even be accomplished in color with almost unlimited ability to recreate the spectrum. The advantage of this ap-

proach over conventional pen plotters is speed, with electrostatic plotters achieving speeds up to 50 times faster. There is also considerably less movement of paper and equipment parts. However, electrostatic plots incur heavy production costs when compared to pen plotters, and therefore have not come into large-scale general use.

5.11 DIRECT INPUT

Direct-input devices are those that use pointing, scanning, or voice input, thus eliminating the need for typing on the keyboard of a computer.

Pointing Devices

Pointing devices allow users to use their hands directly to input data. The most common device in-use is the mouse. Other devices would include:

- Touch screens—the user physically touches the screen to control the device, usually with their finger. Commonly used for public information systems.
- Light pen—the user points a light pen at the point on the screen. Commonly used in graphic design systems.
- Digital camera—the user records a digital image on disk or memory of the camera and then later transfers the image to the computer for further processing or for record purposes.

Scanning and Character Recognition

This technique offers a high-speed method for the capture of source data as an alternative to keyboard input and for processing documents such as checks. Several types of device exist, with varying capabilities and functions:

1. Page and document readers, with the capability to read several special fonts, plus, in some cases, lower-quality print, including hand printing and hand-marked forms as opposed to written or printed documents. Most character readers have some form of error handling, allowing questionable characters to be displayed to an operator for manual input of the correct character. A wide range of capabilities and hence prices are found, from simple, low-speed (several pages per minute) devices handling pages only to high-speed readers for pages and comments, the former at up to two pages per second and the latter several times faster.
2. Document readers/sorters that read and optionally sort simple documents such as checks and payment slips with characters either in magnetic ink or special font. These are geared to higher throughputs (up to 3,000 documents per minute) of standard documents,
3. Transaction devices, which may use both document reading and keyboard data entry, and where single documents at a time are handled.

Scanners are also to scan images into the computer, including photographs and slides.

Writing Tablets

Devices using a variety of techniques exist for the conversion of hand-printed characters into codes for direct input to a CPU. The overall function of these is the same—the provision of a surface on which normal forms (typically up to A4 size) can be filled in with hand-written alphanumeric characters, using either a normal writing instrument in the case of pressure-sensitive techniques or a special pen. Some devices recognize the handwriting of an individual, others use a special character set that simplifies the recognition of writing (such as Graffiti on the Palm computer). The benefits of this type of device include:

1. Immediate capture of data at source, avoiding time-consuming and error-prone transcription of data.
2. By detecting the movements involved in writing a character, additional information is gained compared with optical recognition, allowing characters that are easily confused by Optical Character Recognition (OCR) to be correctly distinguished.

These devices are used extensively on hand-held computers where the size of a keyboard would make the device too large.

5.12 SECONDARY STORAGE

The contents of the computer's main memory are lost every time the power to the computer is turned off, and therefore another permanent means of storage is required to avoid having to re-enter programs and data every time the computer is switched on. Secondary storage is used to provide this. It provides a permanent means for keeping data and programs. The contents of secondary storage may be loaded into the computer's memory when required. Secondary storage may also be copied and stored away from the CPU for security and safekeeping.

There are two types of secondary storage—those that support direct access to the data and those that provide sequential access. With direct access system you can retrieve the particular data directly but with sequential access all the data recorded before the item you want has to be read, first making access very slow. Disk media are direct access and magnetic tapes are sequential.

Magnetic Disk Storage

Disk storage is price-effective, reliable, and easy to exchange, and consequently its usage has grown considerably over the last ten years.

Disks are connected to CPU by a controller, which is normally a DMA device attached to the input/output bus or to a high-speed data channel, except in the case of the slowest of disks, which may be treated as a programmed transfer device.

The controller is generally capable of handling a number of drives. Having multiple-disk drives on a system, as well as providing more on-line storage, allows copying of information from one disk to another and affords a degree of redundancy since, depending on application, the system may continue to function usefully with one less drive in the event of a failure. A disk controller is relatively complex since it has to deal with high rates of data transfer, usually with error code generation and error detection, a number of different commands, and a large amount of status information. Four types of disk drive will be described in the following sections: floppy disk, internal hard disk, cartridge disk, and removable pack disk. The following elements are the major functional parts common to all the above types of drive, with differences in implementation between the different types.

Drive Motor. This drives a spindle on which the disk itself is placed, rotating at a nominally fixed speed. The motor is powered up when a disk is placed in the drive, and powered down (normally with a safety interlock to prevent operator access to rotating parts) until it has stopped spinning, when the disk must be removed from the system.

Disk Medium. The actual recording and storage medium is the item that rotates. It is coated with a magnetic oxide material and can vary from a flexible diskette of less than one-megabyte capacity recording on one surface only (single-sided floppy disk) to an assembly of multiple disks stacked one above the other on a single axle (disk pack) holding thousands of megabytes of data.

Head Mechanism. This carries read/write heads, one for each recording surface. The number of recording surfaces ranges from only 1 on a single-sided floppy disk to 10 or more for a multi-surface disk pack. In the latter case, the heads are mounted on a comb-line assembly, where the teeth of the comb move together in a radial direction between the disk surfaces.

During operation, the recording heads fly aerodynamically extremely close to the disk surface—except in the case of floppy disks, where the head is in contact with the surface. When rotation stops, the heads either retract from the surface or come to rest upon it, depending on the technology involved.

The time taken for the read/write head to be positioned above a particular area on the disk surface for the desired transfer of data is known as the access time. It is a function partly of the rotational speed of the disk, which gives rise to what is known as the average rotational latency (i.e., one half of the complete revolution time of the disk). Out of a number of accesses, the average length of time it is necessary to wait for the desired point to come below the head approaches this figure. The second component of access time is the head-positioning time. This is dependent upon the number of tracks to be traversed in moving from the current head position to the desired one. Again, an average figure emerges from a large number of accesses. The average access time is the sum of these two components. In planning the throughput possible with a given disk system the worst-case figures may also need to be considered.

Electronics. The drive must accept commands to seek (i.e., position the head assembly above a particular track) and must be able to recover signals from the read heads and convert these to binary digits in parallel form for transmission to the disk controller. Conversely, data transmitted in this way from the controller to

the disk drive must be translated into appropriate analog signals to be applied to the head for writing the desired data onto the disk.

Drive Control. Various other functions concerned with control of the drive and sensing of switches on the control panel are performed. On some more advanced drives, much of the operation of the drive and electronics can be tested off-line from the system, allowing fault diagnosis to be performed without affecting the rest of the system.

Information is recorded in a number of concentric, closely spaced tracks on the disk surfaces, and in order to write and thereafter read successfully on the same or a different drive it must be possible to position the head to a high degree of accuracy and precision above any given track. Data are recorded and read serially on one surface at a time, hence transfer of data between the disk controller and disk surface involves conversion in both directions between serial analog and parallel digital signals. A phase-locked loop clock system is normally used to ensure reliable reading by compensating for variations in the rotational speed of the disk.

Data are formatted in blocks or sections on all disk systems, generally in fixed block lengths preformatted on the disk medium at the time of manufacture. Alternatively, *soft sectoring* allows formatting into blocks of differing length by program**.** The drive electronics are required to read sector leaders, which contain control information to condition the read circuitry of the drive and sector address information and to calculate, write, and check an error-correcting code—normally a cyclic redundancy check—for each block.

Finding the correct track in a seek operation, where the separation between adjacent tracks may be as little as 0.01 mm, requires servo-controlled positioning of the head to ensure accurate registration with the track. All rigid-disk systems have servo-controlled head positioning, either using a separate surface prewritten with position information and a read head only or with servo information interspersed with data on the normal read/write tracks being sampled by the normal read/write head. Floppy disk systems, where the tolerances are not so fine, have a simpler stepping motor mechanism for head positioning.

Floppy Disk. The floppy disk, while having the four elements described above, was conceived as a simple, low-cost device providing a moderate amount of random access back-up storage to microcomputers. As the name implies, the magnetic medium used is a flexible, magnetic oxide-coated diskette, which is contained in a square envelope with apertures for the drive spindle to engage a hole in the center of the disk and for the read/write head to make contact with the disk. Diskettes are now nearly all 59 mm (3.5 in.). The compactness and flexibility of the disk makes it very simple to handle and store, possible for it to be sent by post. Floppy disks are now mainly used for the backing up of small amounts of data and for interchanging data with other computers.

One major simplification in the design of the floppy disk system is the arrangement of the read/write head. This runs in contact with the disk surface during read/write operations and is retracted otherwise. This feature and the choice of disk coating and the pressure loading of the head are such that, at the rotational speed of 360 rev/mm, the wear on the recording surface is minimal. Eventually, however, wear and therefore error rate are such that the diskette may have to be replaced and the information copied onto a new diskette.

Capacities vary from the 256 kilobytes of the earliest drives, which record on one surface of the diskette only, to a figure of over 100 megabytes on more recent

units, most of which use both surfaces of the diskette. Access times, imposed by the rather slow head-positioning mechanism using a stepping motor, are in the range of 100–500 ms. Transfer rates are normally about 500 kilobytes per second, although there are some devices on the market, but not in general use, that have much greater transfer rates. Another simplification is in the area of operator controls. There are generally no switches or status indicators, the simple action of moving a flap on the front of the drive to load or removing the diskette being the only operator action. The disk motor spins all the time that a disk is present.

Internal Hard Drive. An internal hard disk (also known as an affixed disk because it is located inside the system unit) consists of one or more metallic platters sealed inside a container that includes the motor for rotating the disk, the access arms and read/write heads. Typically inside are four 59 mm metallic platters.

The principal feature of the drive—the fixed unit—is known as a head disk assembly (HDA). By its being fixed and totally sealed, with the read/write heads and arm assembly within the enclosure, the following benefits are realized:

- Contaminant-free environment for the medium allows better data integrity and reliability, at the same time as less stringent environmental requirements. Simpler maintenance requirements follow from this.
- Lighter-weight heads, flying to tighter tolerances closer to the recording surface, allow higher recording densities. Since the disk itself is never removed, instead of retracting, the heads actually rest on special zones of the disk surface when power is removed,
- The arrangement of read/write heads is two per surface, providing lower average seek times by requiring less head movement to span the whole recording area.

The head-positioning arrangement differs mechanically from that of the drives previously described by being pivoted about an axis outside the disk circumference. Three general types of internal hard drive exist, with approximate disk diameters of 133, 203, and 355 mm, providing capacities from 25 megabytes to many gigabytes. Performance, for the reason described above, can exceed that for disk cartridge or pack drives of corresponding capacity.

Operationally, the fact that the disks are not removable from the drive means that a separate form of storage medium that is removable must be present on a system using an internal hard disk drive. Back-up and making portable copies is done using this separate medium, which is usually another type of disk drive or a magnetic tape system matched to the disk speed and capacity.

Cartridge Disk. This type of disk system is so called because the medium—one or two rigid disks on a single spindle, of aluminum coated with magnetic oxide and approximately 350 mm in diameter—is housed permanently in a strong plastic casing or cartridge. When the complete cartridge assembly is loaded into the drive, a slot opens to allow the read/write heads access to the recording surfaces. As well as providing mechanical mounting, the cartridge provides protection for the disk medium when it is removed from the drive.

Drives are designed either for loading from the top when a lid is raised or from the front when a small door is opened allowing the cartridge to be slotted in. Power to the drive motor is removed during loading and unloading, and the door is locked until the motor has slowed down to a safe speed. On loading and starting up, the controller cannot access the drive until the motor has reached full speed. Operator

controls are normally provided for unload, write protection, and some form of unit select switch allowing drive numbers to be reassigned on a multiple-drive system. Indicators typically show drive on-line, error, and data transfer in progress.

Access times are normally in the region of 30–75 ms, aided by a fast servo-controlled head-positioning mechanism actuated by a coil or linear motor, the heads being moved in and out over the recording surface by an arm which operates radially. Heads are lightweight, sprint-loaded to fly aerodynamically in the region of 0.001 mm from the surface of the disk when it is rotating at its full speed (usually 2400 or 2600 rev/mm). Because of the extremely small gap, cleanliness of the oxide surface is vital, as any particle of debris or even smoke will break the thin air gap, causing the head to crash into the disk surface. In this rare event, permanent damage to the heads and disk cartridge will occur. Positive air pressure is maintained in the area around the cartridge in order to minimize the ingress of fine particles. Care should be taken to ensure cleanliness in the handling and storage of cartridges when not mounted in the drive.

The capacity of cartridges is in a range of several gigabytes of data, with data transfer rates in the region of 9 megabytes per second. They are now very commonly used for back-up purposes and the exchange of data, and since they are easily removed from the computer, they are also used to hold sensitive information.

Disk Pack. The medium used in this type of drive has multiple platters (five or more) on a single spindle and is protected by a plastic casing when removed from the drive. When loaded on the drive, however, the casing is withdrawn. The drives are top loading and unlike cartridge disks, which can generally be rack mounted in the cabinet housing the CPU, are freestanding units.

Other than this difference, most of the design features of disk pack drives follow those of cartridge units. The significant difference is the larger capabilities and generally high performance in terms of access times and transfer rates.

Magnetic Tape

Reliable devices for outputting digital data to and reading from magnetic tape have been available for a considerable time. The use of this medium, with agreed standards for the format of recorded data, has become an industry standard for the interchange of data between systems from different manufacturers. In addition, low-cost magnetic tape cartridge systems exist providing useful minimal-cost large-scale backup storage plus a convenient medium for small-volume removable data and the distribution of software releases and updates.

Industry-Standard Tape Drives. These allow reels of 12.7-mm-wide oxide-coated magnetic tape, which are normally 731 m in length on a 267-mm-diameter reel (or 365 m on a 178-m reel) to be driven past write-and-read head assemblies for writing, and subsequent reading, at linear densities from 800 to 6250 bits per inch. Tapes are written with variable-length blocks or records with inter-record gaps in the region of 12.7 mm. Each block has lateral and longitudinal parity information inserted and checked, and a cyclic redundancy code is written and checked for each block. The latter provides a high degree of error-correction capability. The tape motion and stop-start characteristics are held within precise limits by a servo-controlled capstan around which the tape wraps more than 180° for sufficient grip.

Correct tape tension and low inertia are maintained by motors driving the hubs of the two tape reels in response to information on the amount of tape in the path between the two reels at any time. One of the following forms of mechanical buffering for the tape between the capstan and reels is used:

1. *Tension arm:* This uses a spring-loaded arm with pulleys over which the tape passes, alternating with fixed pulleys such that, when loaded, the tape follows a W-shaped path. The position of the arm is sensed and the information used to control the release and take-up of tape by the reel motors. This technique has almost universally been overtaken by the vacuum method described below.
2. *Vacuum chamber:* This technique, used in modern higher-performance tape drives, has between each reel and the capstan a chamber of the same width as the tape, into which a U-shaped 1oop of tape of around 1 to 2 m is drawn by vacuum in the chamber. The size of the tape loops is sensed photoelectrically to control the reel motors.

To prevent the tape from being pulled clear of the reel when it has been read or written to the end or rewound to the beginning, reflective tape markers are applied near each end of the reel. These are sensed photoelectrically and the resulting signal used to stop the tape on rewind or to indicate that forward motion should stop on reading or writing.

Three different forms of encoding the data on the tape are encountered, dependent upon which of the standard tape speeds is being used. Up to 800 bits per inch, the technique is called *non-return to zero* (NRZ), while at 1600 bits per inch *phase encoding* (PE) and at 6250 bits per inch *group code recording* (GCR) are used. Some drivers can be switched between 800 bits per inch NRZ and 1600 bits per inch PE. Very few systems below 1600 bits per inch are now manufactured.

Block format on the tape is variable under program control between certain defined limits, and as part of the standard, tape marks and labels are recorded on the tape and the interblock gap is precisely defined. Spacing between write and read heads allows a read-after-write check to be done dynamically to verify written data. Writing and reading can only be carried out sequentially. These tape units do not perform random access to blocks of data, though those units that permit selective reverse under program control do make it possible for the application to access data other than by sequential read of the tape. However, this requires prior knowledge by the application of the layout and contents of the tape and is particularly slow and cumbersome, such applications being far better serviced by a disk-storage device.

For PE and GCR, a formatter is required between the controller and drive to convert between normal data representation and that required for these forms of encoding. Tape drives can vary in physical form from a rack-mountable unit that is positioned horizontally to a floor-standing unit around 1.75 m in height.

Operator controls for on-line/off-line, manually controlled forward/reverse and rewind motion, unit select, and load are normally provided. To prevent accidental erasure of a tape containing vital data by accidental write commands in a program, a write-protect ring must be present on a reel when it is to be written to. Its presence or absence is detected by the drive electronics. This is a further part of the standard for interchange of data on magnetic tapes.

Cartridge Tape. Low-cost tape units storing many gigabytes of data on a tape cartridge are now commonly used for back-up storage, especially on desktop computers and small servers.

Tape Streamer Unit. The emergence of large-capacity, nonremovable disk storage has posed the problem of how to make up copies of complete disk contents for security or distribution to another similarly equipped system. An alternative to tape cartridges is a tape drive very similar to the industry-standard units described above but with the simplification of writing in a continuous stream, rather than in blocks. The tape controller and tape motion controls can, therefore, be simpler than those for the industry-standard drive. Many modern tape units are able to operate in both *block* and *streamer* mode, according to operator or program selection, but not on the same tape. A streamer unit can normally accept the full disk contents on to a single reel of tape. The digital audiotape (DAT) is now the most common type of back-up device in use.

Optical Disk

In optical disks a laser beam alters the surface of the disk to represent data. Optical disks use reflected light to represent 1's and 0's. The bits are represented by flat areas called *lands* and bumpy areas called *pits* on the disk. The disk is read by a laser beam projected onto these areas and the amount of light reflected determines whether the area represents a 1 or 0. The most common forms of optical disk are the CD (compact disk) and DVD (digital versatile disk).

Compact Disk. There are three types of CDs in common use as secondary storage: CD-ROM, CD-R, and CD-RW.

1. CD-ROM (compact disk-read only memory) is very similar to the disk used in the music industry. It cannot be written to or erased by the computer user. The user can only access the data written to the disk by the publisher of the disk. They are used to distribute databases, references, and software application packages.
2. CD-R (compact disk recordable) can be written to once by a user with a writable CD drive. After that they can be read many times but cannot be written to or erased. They are normally used to archive or permanently store large amounts of data.
3. CD-RW (compact disk rewritable) is similar to CD-R but since the disk surface is not permanently altered when data is recorded the disks can be rewritten to.

Digital Versatile Disk. These disks are very similar to CD disks except that more data can be written to them. At present seven times more information can be written to a DVD than to a CD (approximately 4.7 gigabytes of data of one side of a DVD disk). As for CD disks, there are DVD-ROM, DVD-R, and DVD-RAM (random access memory allowing the disk to be reused over and over again). It is expected that DVD disks and drives will replace CD disks and drives in the very near future.

5.13 DIGITAL AND ANALOG INPUT/OUTPUT

One of the major application areas for minicomputers and microcomputers is direct control of and collection of data from other systems by means of interfaces that provide electrical connections directly or via transducers to such systems. Both continuously varying voltages (analog signals) and signals which have discretion

or off states (digital signals) can be sensed by suitable interfaces and converted into binary form for analysis by programs in the CPU. For control purposes, binary values can also be converted to analog or digital form by interfaces for output from the computer system.

In process and/or machine control and monitoring, data acquisition from laboratory instruments, radar, and communications (to take some common examples) employs computer systems equipped with a range of suitable interfaces. They may be measuring other physical quantities such as temperature, pressure, and flow converted by transducers into electrical signals.

Digital Input/Output

Relatively simple interfaces are required to convert the 1's and 0's in a word output from the CPU into corresponding on or off states of output drivers. These output signals are brought out from the computer on appropriate connectors and cables. The output levels available range from TTL (± 5 V) for connection to nearby equipment that can receive logic levels, to over 100 V d.c. or a.c. levels for industrial environments. In the former case, signals may come straight from a printed circuit board inside the computer enclosure, while in the latter they are required to go through power drivers and be brought out to terminal strips capable of taking plant wiring. The latter type of equipment may need to be housed in separate cabinets.

Similarly, for input of information to the computer system, interfaces are available to convert a range of signal levels to logic levels within the interface, which are held in a register and can be input by the CPU. In some cases, input and output are performed on the same interface module.

Most mini and micro systems offer a range of logic level input/output interfaces, while the industrial type of input and output equipment is supplied by manufacturers specializing in process control. Optical isolators are sometimes included in each signal line to isolate the computer electrically from other systems. Protection of input interfaces by diode networks or fusible links is sometimes provided to prevent damage by over voltages. In industrial control, where thousands of digital points need to be scanned or controlled, interfaces with many separately addressable input and output words are used.

Although most digital input and output rates of change are fairly slow (less than l,000 words per second), high-speed interfaces at logic levels using direct memory access are available. These can, in some cases, transfer in burst mode at speeds up to 3 million words per second. High transfer rates are required in areas such as radar data handling and display driving.

Analog Input

Analog-to-digital converters, in many cases with programmable multiplexers for high- or low-level signals and programmable gain pre-amplifiers covering a wide range of signals (from microvolts to 10 V), allow conversion commands to be issued and the digital results to be transferred to the CPU by the interface. Industrial-grade analog input subsystems typically have a capacity of hundreds of multiplexer channels, low-level capability for sources such as thermocouples and strain gauges, and high common-mode signal rejection and protection. As with digital input/output,

this type of equipment is usually housed in separate cabinets with terminal strips and is supplied by specialized process control equipment or data logger manufacturers. For laboratory use, converters normally have higher throughput speed, lower multiplexer capacity, and often direct cable connection of the analog signals to a converter board housed within the CPU enclosure. Where converters with very high sampling rates (in the region of 100,000 samples per second) are used, input of data to the CPU may be by direct memory access. Resolution of analog-to-digital converters used with computer systems is usually in the range 10–12 bits, i.e., a resolution of 1 part in 1024 to 1 part in 4096. Resolutions of anything from 8 to 32 bits are, however, available. Where a programmable or auto-ranging preamplifier is used before the analog-to-digital converter, dynamic signal ranges of 1 million: 1 can be handled.

Analog Output

Where variable output voltages are required (for example, to drive display or plotting devices or as set points to analogue controllers in industrial process control applications), one or more addressable output words is provided, each with a digital-to-analogue converter continuously outputting the voltage represented by the contents of its register. Resolution is normally no more than 12 bits, with a usual signal range of ± 1 V or ± 10 V. Current outputs are also available.

Input/Output Subsystems

Some manufacturers provide a complete subsystem with its own data highway separate from the computer system input/output bus, with a number of module positions into which a range of compatible analog and digital input/output modules can be plugged. Any module type can be plugged into any position to make up the required number of analog and digital points.

5.14 DATA COMMUNICATIONS

Introduction

Since 1980 there has been a large growth in the use of data communications between different types and makes of equipment both within a physical location or building and between different buildings situated anywhere in the world. Even when this communication appears to take place between two points on earth, it has very often done so by means of a geo-stationary satellite positioned in orbit. The creation and maintenance of such networks is now nearly always the role of network managers and their staff, a function that is separate from (though working closely with) the traditional computer departments.

The requirement for the communication of data is not, of course, new, but what has changed is the basis for that requirement. Previously, the only other means available for the transfer of data between machines was a copy by magnetic media (such as tape or disk) or to key in the data again, with the consequent high risk of error and increased time taken. It was seen that data transmission would be faster

and more accurate than both of these methods. Interestingly, data communication was not regarded as a replacement for the data in hardcopy form. Today, more emphasis is being placed on eliminating hardcopy transactions, such as the growing use of EDI (Electronic Data Interchange) to replace paper as the medium for moving order information between companies.

Large-scale integration and consequent lower costs have made very powerful computers much more readily available. These can contain the sophisticated software required to handle complex networks and overcome complex problems such as finding alternative routes for messages when a transmission line is broken. Interface devices between the computer and the data network are very intelligent and powerful and are usually computers themselves. Thus, they relieve the main computer of much of the previous load that it historically handled for data communications.

Computers have always been able to communicate with their peripheral devices such as card readers, mass-storage devices, and printers, but in the 1960s it was not typical for the communications to extend beyond this. Data were transcribed onto punching documents by functional departments within an organization. Now the widespread use of personal computers by users at their desks has eliminated almost all these departments. Even the very large traditional data-entry organizations such as the utility companies have introduced data capture at source using hand-held terminals or OCR techniques.

However, in the late 1960s and 1970s the development of both hardware and software technology made it increasingly attractive to replace terminals with more intelligent remote systems. These systems varied in their sophistication. At one end of the spectrum were interactive screen-based terminals that could interrogate files held on the central computer. Greater sophistication was found in data-validation systems that held sufficient data locally to check that, for example, part numbers on a customer order really existed before sending the order to the computer for processing. More sophisticated still were complete minicomputers carrying out a considerable amount of local data processing before updating central files to be used in large number-crunching applications such as production scheduling and materials planning. From these systems have grown a whole range of requirements for data communications.

Data Communications Concepts

Computers communicate data in binary format, the bits being represented by changes in current or voltage on a wire, or, more recently, by patterns of light through an optic-fiber cable. There are various ways that characters are represented in binary format. One of the earliest of these was the 5-bit Baudot code, invented towards the end of the 19th century by Emile Baudot for use on telegraphic circuits. Five bits can be used to represent 32 different characters, and while this was adequate for its purpose, it cannot represent enough characters for modern data communications. Nonetheless, Baudot gave his name to *baud,* the commonly used unit of speed, which, although strictly meaning signal events per second, is frequently used to denote bits per second.

Nowadays, one of the most commonly used codes is the ASCII (American Standard Code for Information Interchange) code (Table 5.1). This consists of seven information bits plus one parity (error-checking) bit. Another is EBCDIC (Extended Binary Coded Decimal lnterchange Code), an 8-bit character code used primarily on IBM equipment.

TABLE 5.1 ASCII Code Table

Char	Octal	Binary
Nul	000	0000000
SOH	001	0000001
STX	002	0000010
ETX	003	0000011
EOI	004	0000100
ENQ	005	0000101
ACK	006	0000110
BEL	007	0000111
BS	010	0001000
HT	011	0001001
LF	012	0001010
VT	013	0001011
FF	014	0001100
CR	015	0001101
SO	016	0001110
SI	017	0001111
DLE	020	0010000
DC1	021	0010001
DC2	022	0010010
DC3	023	0010011
DC4	024	0010100
NAK	025	0010101
SYN	026	0010110
ETB	027	0010111
CAN	030	0011000
EM	031	0011001
SUB	032	0011010
ESC	033	0011011
FS	034	0011100
GS	035	0011101
RS	036	0011110
US	037	0011111
DEL	177	1111111

TABLE 5.1 ASCII Code Table (*Continued*)

Char	Octal	Binary
A	101	1000001
B	102	1000010
C	103	1000011
D	104	1000100
E	105	1000101
F	106	1000110
G	107	1000111
H	110	1001000
I	111	1001001
J	112	1001010
K	113	1001011
L	114	1001100
M	115	1001101
N	116	1001110
O	117	1001111
P	120	1010000
R	121	1010001
S	122	1010011
T	123	1010100
U	124	1010101
V	125	1010110
W	126	1010111
X	127	1011000
Y	130	1011001
Z	131	1011010
a	141	1100001
b	142	1100010
c	143	1100011
d	144	1100100
e	145	1100101
f	146	1100110
g	147	1100111
h	150	1101000

TABLE 5.1 ASCII Code Table (*Continued*)

Char	Octal	Binary
i	151	1101001
j	152	1101010
k	153	1101011
l	154	1101100
m	155	1101101
n	156	1101110
o	157	1101111
p	160	1110000
q	161	1110001
r	162	1110010
s	163	1110011
t	164	1110100
u	165	1110101
v	166	1110110
w	167	1110111
x	170	1111000
y	171	1111001
z	172	1111010
sp	040	0100000
!	041	0100001
“	042	0100010
#	043	0100011
$	044	0100100
%	045	0100101
&	046	0100110
‘	047	0100111
(	050	0101000
)	051	0101001
*	052	0101010
+	053	0101011
,	054	0101100
-	055	0101101
.	056	0101110

TABLE 5.1 ASCII Code Table (*Continued*)

Char	Octal	Binary
/	057	0101111
0	060	0110000
1	061	0110001
2	062	0110010
3	063	0110011
4	064	0110100
5	065	0110101
6	066	0110110
7	067	0110111
8	070	0111000
9	071	0111001
:	072	0111010
;	073	0111011
<	074	0111100
=	075	0111101
>	076	0111110
?	077	0111111
@	100	1000000

Within the computer and between the computer and peripheral devices such as mass-storage devices and its printer, data are usually transferred in parallel format. In parallel transmission a separate wire is used to carry each bit, with an extra wire carrying a clock signal. This clock signal indicates to the receiving device that a character present on the information wires. The advantage of parallel transmission is, of course, speed, since an entire character can be transmitted in the time it takes to send one bit. However, the cost would prove prohibitive where the transmitter and receiver are at some distance apart. Consequently, for sending data between computers and terminal devices and between computers that are not closely coupled, serial transmission is used.

Here a pair of wires is used, with data being transmitted or one wire while the second acts as a common signal ground. As the term implies, bits are transmitted serially, and so this form of transmission is more practical for long-distance communication because of the lower cost of the wiring required. In addition, it is simpler and less expensive to amplify signals rather than use multiple signals in order to overcome the problem of line noise, which increases as the distance between the transmitter and receiver grows. Data transmission frequently makes use of telephone lines designed for voice communication, and since the public voice

networks do not consists of parallel channels, serial transmission is the one practical solution.

Parallel data on multiple wires are converted to serial data by means of a device known as an interface. In its simplest form, an interface contains a register or buffer capable of storing the number of bits that comprise one character. In the case of data going from serial to parallel format, the first bit enters the first position in the register and is shifted along, thereby making room for the second bit. The process continues until the sampling block that is strobing the state of the line indicates that the correct number of bits has been received and that a character has been assembled. The clock then generates a signal to the computer which transfers the character in parallel format. The reverse process is carried out to convert parallel to serial data. This single-buffered interface does have limitations, however. The computer effectively has to read the character immediately, since the bit of a second character will start arriving to begin its occupation of the register. This makes no allowance for the fact that the computer may not be available instantly. Nor does it allow any time to check for any errors in the character received.

To overcome this problem, a second register is added, creating a double-buffered interface. Once a signal is received indicating that the requisite number of bits have been assembled, the character is parallel transferred to the second (or holding) register, and the process can continue. The computer now has as much time as it takes to fill the shift register in order to check and transfer (again in parallel format) the character.

Multiline Interface. With the development of technology, the transmitter and receiver functions are now carried out by an inexpensive chip. Therefore, the major costs in the interface are those of the mechanism used to interrupt the CPU when a character has been assembled and the connection to the computer's bus used to transmit the received data to the CPU, or in some cases direct to memory. The interrupt mechanism and the bus interface are not heavily used. Indeed, they function only when a character is received or transmitted. These facilities are shared in a multiline interface, sometimes (though not strictly correctly) known as a multiplexor. To achieve this, the device has several receivers and transmitters and a first-in, first-out (FIFO) buffer for received characters. The receivers are scanned, and when a flag is found indicating that a character has been received, the character is transmitted into the FIFO buffer, along with its line number. An interrupt tells the CPU that there are characters in the buffer, and these are communicated over the bus to the computer. Similarly, the scanner checks the transmitters and when it discovers a flag indicating that a transmitter buffer is empty, it interrupts the CPU. Typically, the number of lines supported by a multiline interface increases by powers of two for convenient binary representation, 4, 8, 16, 32, 128, 256 being common. The economies of scale in such an interface mean that further sophistications can be included, such as program-selectable formats and line speed, and modem control for some or all of the lines.

However, the term *multiplexing* actually refers to the function of sharing a single communications channel across many users. There are two commonly used methods of achieving this. One is a technique called time-division multiplexing (TDM), which consists of breaking down the data from each user into separate messages which could be as small as one or two bytes and meaningless when taken individually. The messages, together with identifying characters, are interleaved and transmitted along a single line. They are separated at the other end and the messages reassembled. This is achieved by use of devices known as concentrators or multi-

plexors. The second technique used to achieve this objective of making maximum use of a communication line is frequency division multiplexing. The concept is similar to that of time division multiplexing. It is achieved by transmitting complete messages simultaneously but at different frequencies.

Modem. A significant complication of using public voice networks to transmit data is that voice transmission is analog but data generated by computers are digital in format. Thus, an additional piece of equipment is required between the digital sender/receiver and the analog circuit. This device modulates and demodulates the signal as it enters and leaves the analog circuit, and is known by the abbreviated description of its functions, as a modem (Figure 5.3). Modems are widely available from computer shops.

Fiberoptic Cable. Cabling for transmissions has traditionally been constructed of a copper-based core, this being a viable compromise between cost and conductivity for anything other than the very shortest communication paths. It is difficult to imagine a gold cable being laid from New York to Boston and remaining in place for very long! However, copper has its own limitations, such as weight, resistance, noise, etc. The development of fiberoptic cable to the stage where a set light pattern can be sustained over long distances without distortion and then be sensed and interpreted at the other end has signaled the beginning of the end of copper as a standard communication medium. The main advantages of fiberoptic are:

1. Very lightweight
2. Greater communication capacity (number and speed of channels) for the same size
3. Digital transmission
4. Immunity from most causes of interference and noise associated with copper
5. Cost decreasing as volumes increase

However, developments in copper technology have seen the development of cables that can support over 100 megabits per second, and since copper is currently much cheaper to terminate, most of the cable to the desk is still copper.

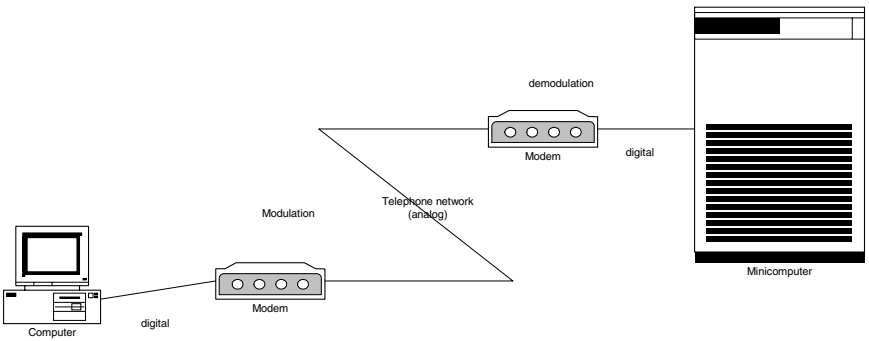


FIGURE 5.3 The use of modems in a communication link.

Laser. This works on exactly the same principle as fiberoptic, except that the light signal is passed between two laser/receivers on a point-to-point line-of-sight basis. It is ideal, therefore, in situations where communications are required between two different buildings but neither party owns or controls the land between them. The only other method would be to use a common carrier, resulting in a higher cost and probably lower speed and quality of communication.

Microwave. Where an organization requires extremely large volumes of data to be transmitted or a very high speed to be achieved, then it is sometimes viable for it to set up its own microwave network.

Transmission Techniques

There are two techniques commonly used to transmit data on serial lines. One varies the current and the other varies the voltage in order to indicate the presence or absence of bits on the line.

Current Variable. The current-based technique communicates binary data by turning on and off a 20 mA current flowing through both the transmitter and receiver. Current on indicates a mark or 1 bit and current off signifies a space or 0 bit. This technique of turning a current on and off is less susceptible to noise than the technique of varying the voltage. However, it does have some drawbacks. Optical isolators are needed to protect logic circuits from the high voltages which may be required to drive the loop. Since there is one current source, an active interface and a passive interface are required, and finally, since a 20 mA system cannot carry the necessary control information, it cannot be used with modems,

Voltage Variable. The HA (Electronic Industries Association) and CCITT (Comité consultatif internationale de télégraphique et téléphonique) systems contain specifications and recommendations for the design of equipment to interface data terminal equipment (computers and terminals) to data communication equipment (modems). The specific EIA standard to which most modem equipment is designed is RS232C. The CCITT equivalent of RS232C is known as V.24—List of Definitions of Interchange Circuits Terminating Equipment. The EIA/CCITT systems communicate data by reversing the polarity of the voltage; a 0 is represented by a positive voltage and a 1 by a negative voltage.

The signals in the EIA/CCITT specifications are not recommended for use over distances greater than 15.5 m.

Consequently, the modem and interface should not be more than 15.5 m, though in practice distances in excess of 300 m have been operated without problems.

Transmission Types

Different communications applications use one of two types of transmission: asynchronous or synchronous. Slower devices such as VDUs and low-speed printers typically use asynchronous (or start–stop) transmission, in which each character is transmitted separately. In order to tell the receiver that a character is about to arrive, the bits representing the character are preceded by a start bit, usually a zero. After

the last data bit and error-checking bit the line will return to the 1-bit state for at least one bit time—this is known as the stop bit.

Asynchronous transmission has the advantage that it requires relatively simple and therefore low-cost devices. It is, however, inefficient, since at least two extra bits are required to send eight data bits, and so it would not be used for high-speed communication.

In synchronous transmission, the transmitter assembles characters into blocks and so the stream of data bits travels along the line uninterrupted by start and stop bits. This means that the receiver must know the number of bits that make up a character so that it can reassemble the original characters from the stream of bits. Preceding the block of data bits, synchronization characters are sent to provide a timing signal for the receiver and enable it to count in the data characters. If the blocks of data are of uniform length, then this is all that is required to send a message. However, most systems would include some header information that may be used to indicate the program or task for which the data are destined and the amount of data in the block. In addition, if the messages are of variable length, some end-of-message characters will be required.

Because it does not contain start and stop bits for every character, synchronous transmission is more efficient than asynchronous. However, it can be inappropriate for some character-oriented applications since there is a minimum overhead in characters that can be high relative to small transmitted block sizes, and the equipment required to implement it is more expensive.

Direction of Transmission

There are three types of circuit available for the communication of data and, correspondingly, three direction combinations: simplex, half-duplex, and full duplex.

Simplex communication is the transmission of data in one direction only, with no capability of reversing that direction. This has limitations and is not used in the majority of data-communications applications. It can be employed, however, for applications that involve the broadcasting of data for information purposes in, for example, a factory. In this instance there is neither a need nor a mechanism for sending data back to the host. The simplex mode of operation could not be used for communication between computers. The television network is simplex in that it has been only necessary to send pictures to a person's television receiver—it has not been necessary to receive data back from the appliance (however, this requirement is changing with the advent of Internet TV applications).

Half-duplex permits the user to transmit in both directions, but not simultaneously. Two-wire half-duplex has a built-in delay factor called *turnaround time*. This is the time taken to reverse the direction of transmission from sender to receiver and vice versa. The time is required by line-propagation effects, modem timing, and computer-response time.

Full-duplex operation allows communication in both directions simultaneously. The data may or may not be related, depending on the applications being run in the computer or computers.

Error Detection and Correction

Noise on most communications lines will inevitably introduce errors into messages being transmitted. The error rates will vary according to the kind of transmission

lines being used. In-house lines are potentially the most noise-free since routing and shielding are within user control. Public switched networks, on the other hand, are likely to be the worst, though this problem has been addressed by most common carriers by the introduction of digital switching exchanges and fiber technology.

Whatever the environment, however, there will be a need for error detection and correction. Three systems are commonly used: VRC, LRC, and CRC.

VRC (vertical redundancy check) consists of adding a parity bit to each character. The system will be designed to use either even or odd parity. If the parity is even, the parity bit is set so that the total number of ones in the character plus parity is even. Obviously, for odd parity the total number will be odd. This system will detect single-bit errors in a character. However, if two bits are incorrect the parity will appear correct. VRC is therefore a simple system designed to detect single-bit errors within a character. It will detect approximately nine out of ten errors.

A more sophisticated error-detection system is LRC (longitudinal redundancy check), in which an extra byte is carried at the end of a block of characters to form a parity character. Unlike VRC, the bits in this character are not sampling an entire character but individual bits from each character in the block. Thus, the first bit in the parity character samples the first bit of each data character in the block. As a result, LRC is better than VRC at detecting burst errors, which affect several neighboring characters.

It is possible to combine VRC and LRC and increase the combined error detection rate to 99%. A bit error can be detected and corrected because the exact location of the error will be pinpointed in one direction by LRC and the other by VRC.

Even though the combination of LRC and VRC significantly increases the error-detection rate, the burst nature of line noise means that there are still possible error configurations that could go undetected. In addition, the transmission overhead is relatively high. For VRC alone, in the ASCII code, it is 1 bit in 8, or 12.5%. If VRC and LRC are used in conjunction it will be 12.5% plus 1 character per block.

A third method that has the advantage of a higher detection rate and, in most circumstances, a lower transmission overhead is CRC (cyclic redundancy check). In this technique the bitstream representing a block of characters is divided by a binary number. In the versions most commonly used for 8-bit character format, CRC-16 and CRC-CCITT, a 16-bit remainder is generated. When this calculation has been completed, the transmitter sends these 16 bits—two characters—at the end of the block. The receiver repeats the calculation and compares the two remainders. With this system, the error detection rises to better than 99.9%. The transmission overhead is less than that required for VRC/LRC when there are more than 8 characters per block, as is usually the case.

The disadvantage with CRC is that the calculation overhead required is clearly greater than for the other two systems. The check can be performed by hardware or software, but, as is usually the case, the higher performance and lower cost of hardware have made CRC more readily available and commonly used.

Once bad data have been detected, most computer applications require that they be corrected and that this occurs automatically. While it is possible to send sufficient redundant data with a message to enable the receiver to correct errors without reference to the transmitter, the effort of the calculation required to achieve this in the worst possible error conditions means that this technique is rarely used. More commonly, computer systems use error-correction methods that involve retransmission. The two most popular of these are stop-and-wait retransmission and contin-

uous retransmission. "Stop and wait" is reasonably self-explanatory. The transmitter sends a block and waits for a satisfactory or positive acknowledgment before sending the next block. If the acknowledgment is negative, the block is retransmitted. This technique is simple and effective. However, as the use of satellite links increases, it suffers from the disadvantage that these links have significantly longer propagation times than land-based circuits and so the long acknowledgment times are reducing the efficiency of the network. In these circumstances, 'continuous retransmission' offers greater throughput efficiency. The difference is that the transmitter does not wait for an acknowledgment before sending the next block, it sends continuously. If it receives a negative acknowledgment it searches back through the blocks transmitted and sends it again. This clearly requires a buffer to store the blocks after they have been sent. On receipt of a positive acknowledgment the transmitter deletes the blocks in the buffer up to that point.

Communications Protocols

The communications protocol is the syntax of data communications. Without such a set of rules a stream of bits on a line would be impossible to interpret. Consequently, many organizations, notably computer manufacturers, have created protocols of their own. Unfortunately, however, they are all different, and consequently yet another layer of communications software is required to connect computer networks using different protocols. Well-known protocols include Bisync and SDLC from IBM, DDCMP from the former Digital Equipment Corporation, ADCCP from the American National Standards Institute (ANSI), and HDLC from the International Standards Organization (ISO). The differences between them, however, are not in the functions they set out to perform but in the way they achieve them. Broadly, these functions are as follows.

Framing and Formatting. These define where characters begin and end within a series of bits, which characters constitute a message, and what the various parts of a message signify. Basically, a transmission block will need control data, usually contained in a header field text—the information to be transmitted—held in the body, and error-checking characters, to be found in the trailer. The actual format of the characters is defined by the information code used, such as ASCII or EBCDIC.

Synchronization. This involves preceding a message or block with a unique group of characters that the receiver recognizes as a synchronization sequence. This enables the receiver to frame subsequent characters and field.

Sequencing. This numbers messages so that it is possible to identify lost messages, avoid duplicates, and request and identify retransmitted messages.

Transparency. Ideally, all the special control sequences should be unique and therefore never occur in the text. However, the widely varied nature of the information to be transmitted, from computer programs to data from instruments and industrial processes, means that occasionally a bit pattern will occur in the text which could be read by the receiver as a control sequence. Each protocol has its own mechanism for preventing this, or achieving "transparency" of the text. Bisync employs a technique known as character stuffing. In Bisync the only control char-

acter which could be confusing to the receiver if it appeared in the text is DLE (data link escape). When the bit pattern equivalent to DLE appears within the data a second DLE is inserted. When the two DLE sequences are read, the DLE proper is discarded and the original DLE-like bit pattern is treated as data. This is character stuffing. SDLC, ADCCP, and HDLC use a technique known as bit stuffing, and DDCMP employs a bit count to tell the receiver where data begin and end.

Start-up and Time-out. These are the procedures required to start transmission when no data have been flowing and recovering when transmission ceases.

Line Control. This is the determination, in the case of half-duplex systems, of which terminal device is going to transmit and which is going to receive.

Error Checking and Correction. As described under Error Detection and Correction above, each block of data is verified as it is received. In addition, the sequence in which the blocks are received is checked. For data accuracy all the protocols discussed in this section are capable of supporting CRC (cyclic redundancy check). The check characters are carried on the trailer or block check character (BCC) section,

RS232C. This is a standard issued by the United States Electronic Industries Association (EIA) to define the interface between Data Circuit-terminating Equipment (DCE) and Data Terminal Equipment (DTE). In plain language these are usually referred to as the modem and terminal, respectively. The "C" at the end of the standard designation indicates the latest revision of this standard that is applicable. This standard is in widespread use in the United States and formed the basis for the European CCITT standard V.24, which defines the interchange circuits and their functionality. Thus V.24 can be considered a subset of the full RS232C standard. In Europe the other components of RS232C are covered by other standards, CCITT V.28 for the electrical characteristics and ISO 2110 for the pin connector allocations. The terms RS232C and V.24 are often interchanged, and for practical purposes an interface that is said to be "V.24-compliant" means that it also complies with RS232C.

The full interface specification deals with more than 40 interchange circuits, though in practice this number is almost never used. The most common form of connection is the D type connector, so called because of the shape of the male and female plugs used to terminate the cable. These interchange circuits are collated into two distinct groups. The 100 series is used for data, timing and control circuits, whereas the 200 circuits are used for automatic telephone calling.

The principle of operation is simple in that both the modem and the terminal are able to indicate their readiness or not to accept/transmit data by adjusting the voltage on a predetermined circuit. A positive voltage represents a binary 0 or logical OFF condition and a negative voltage a binary 1 or logical ON condition. The other end of the interface can then detect this change in voltage level. Some circuits are kept constantly in a defined state (usually ± 12 V) at all times during transmission to indicate that a piece of equipment continues to be available. Once readiness to transmit data has been achieved, then other circuits are used to pass data to/from each end of the interface. This is carried out by raising or lowering voltage levels on the send or receive circuits phased according to a clock source, which may be external to the modem or internal to it. Both instances use different circuits for the timing signals, and they may not be used together.

The physical arrangement of the connectors can vary, but the female connector (socket) is usually found on the modem and the male connector (plug) on the terminal. The connector design itself does not form part of the standards but the D type is in such widespread use throughout the world that it has, in practice, become a standard in its own right. The pin connections are defined in ISO 2110. Note that some pin allocations are left to the discretion of national bodies and thus complete compatibility is never certain, though this is not generally a problem in practice.

There are many instances in computing where it is desirable to connect terminals directly to computer or other equipment without physically routing through a modem device. This can be achieved through the use of a special switch-over device or, more simply, by crossing over some of the connections at either end. Earlier, such devices were often referred to as "null modems" and cables wired in this way are still called "null modem cables."

ISDN. Integrated Services Digital Network (ISDN) is an all-digital telephone service. It provides reliable voice and data networking using the existing telephone network cabling. ISDN provides each subscriber with two 64 kbps connections (B channels) that can be combined into a single 128 kbps connection. Each of the two 64 kbps connections operate independently of each other and can be used for voice or data communication. The two 64 kbps connections are termed basic rate interface (BRI). BRI includes a separate 16 kbps signaling channel called the D channel. The D channel is used to generate calls, reset calls, and receive information about incoming calls such as the identity of the caller. Large companies normally subscribe for 32 B channels, called primary rate access. ISDN will rapidly be overtaken by new technologies that will offer greater data speeds across standard telephone wiring (such as ADSL).

5.15 COMPUTER NETWORKS

In the early days of data communications information traveled along a single, well-defined route from the remote computer to the host. The reason for this was that the remote computer was fairly restricted in its computing and data-storage capabilities and so the serious computing was carried out at the data center. Most large organizations have retained their large data-processing centers but have changed emphasis on the use to which they are put. They are now used for batch processing of data where either the volume is too large to be processed by the remote systems or the processing itself is not time-critical or as large repositories for files so that the files can be accessed and changed by a number of microcomputers. The advent of very powerful microcomputers (some much more powerful than earlier mainframes), coupled with the marked increase in the reliability and speed of networks, has moved much of the data processing out to the world of the user onto the shop floor, into the laboratory, within an office department, and even to individuals on the desks in their own homes.

In the automobile industry, for example, the European headquarters of a U.S. corporation would have its own designs and engineering department with a computer capable of processing, displaying, and printing design calculations. However, it may still require access to the larger U.S. machines for more complex applications requiring greater computer power, particularly for the modeling of designs that can

be very CPU intensive. In addition, there may be a number of test units, testing engines, and transmissions, each controlled by its own micro and supervised by a host machine. If there is a similar engineering department in, for example, Germany, it may be useful to collect and compare statistical data from test results. Also, since people must be paid, it may be useful to have a link with the mainframe computer in the data center for the processing of payroll records. The demand for the linking of computers and the sharing of information and resources is increasing constantly. This has led to the advent of a data communications network (usually just called a network).

Network Types

There are a number of network types.

Point to Point. This is the simplest form of network and involves the connection of two devices—two computers or a computer and terminal. If the communication line goes down for any reason, then the link is broken, and so it is usual to back up lease lines with dial-up facilities.

Multipoint. As the name implies, multipoint describes the connection of several tributary stations to one host. It is usual for the host to poll the tributary stations in sequence, requesting messages. Not really used today for the interconnection of desktop computers, it is still used for systems such as building management, and fire alarms.

Centralized. Also known as a star network. In this type of network the host exercises control over the tributary stations, all of which are connected to it. The host may also act as a message switching device between remote sites.

Hierarchical. A hierarchical structure implies multiple levels of supervisory control. For example, in an industrial environment special-purpose microcomputers may be linked to the actual process equipment itself. Their function is to monitor and control temperature and pressure. These microcomputers will then be connected to supervisory computers that can store the programs and set points for the process computers and keep statistical and performance records. The next link in the chain will be the resource management computers, keeping track of the materials used, times taken, comparing these with standards, calculating replenishment orders, adjusting forecasts, and so on, Finally, at the top of the network, the financial control system records costs and calculates the financial performance of the process.

Fully Distributed. Here each station is inter-connected to many others in the network and can interact directly with each other. The possibility then exists to share resources such as specialized peripheral devices or large memory capacity and to distribute the database to the systems that access the data most frequently. It also provides alternative routes for messages when communication lines are broken or traffic on one link becomes excessive. However, the design of such systems requires sophisticated analysis of traffic and data usage, and even when set up is more difficult to control than less sophisticated networks. Most computer networks in use today are of this type.

Local and Wide Area Networks

A communications network that exists within a small geographic area (commonly a building or a campus) site is normally called a local area network (LAN). Its characteristics are that:

- It is normally owned by a single company for its own purposes.
- It operates at very high speeds (in excess of 1 gigabyte per second in many cases).
- It allows any device connected to interact with any other device similarly connected.

The interconnection of LANs over a city-wide geographic area is called a metropolitan area network, and the interconnection of LANs over a large geographic area is called a wide area network (WAN).

Previously, to interconnect LANs together usually required using the services of a public network operator (common carrier) with annual rentals due to the network operator for providing such a service. Interconnection is usually using technologies such as ISDN, X25, leased lines, ATM, and SDH. However, with the advent of laser line-of-sight devices and lower costs for microwave, it is now possible to interconnect LANs without using the public network operators provided that line-of-sight can be obtained between the points to be connected. Thus, an organization may connect its systems together to form its own WAN. Within Europe the deregulation of the telecommunications market has also made it much easy to obtain a licence to operate a microwave link.

For the factory environment many network manufacturers offer proprietary networks for connecting terminal equipment to circuits based on tree structures or loops. Connections to the circuit may be from video terminals for collection of, for example, stores data, special-purpose card and badge readers used to track the movement of production batches, or transducers for the control of industrial processes.

Ethernet

Ethernet was developed in its experimental form at the Xerox Palo Alto Research Center in 1972 and is now the most common technology in use for building LANs. The prime objective of the network is to enable high-speed communication between computer equipment and other hardware, irrespective of the make or design of that equipment. Until the arrival of Ethernet most inter-machine communication, except that between equipment from the same manufacturer, was limited in practice to around 9600 bps on twisted pairs.

Ethernet is a multiaccess communications system for transporting data between distributed computer systems that reside in close proximity to each other. The technique used to transfer data under controlled conditions is packet switching, whereby data are composed into discrete packets for onward transmission without regard to their logical use within an application. There is no central point of management in an Ethernet system. Each station may attempt to transmit when it needs to, and control of packet reception is ensured by the use of unique addresses for every Ethernet device ever manufactured. Only if the packet address matches its own address will a station pick up and use a packet on the network.

Communication occurs on a shared channel that is managed through a concept known as carrier sense multiple access with collision detect (CSMA/CD). There

are no predefined or preallocated time slots or bandwidth. Stations wishing to initiate a transmission attempt to acquire control of the communications channel (which is often referred to as the ''Ether'') by sensing the presence of a carrier on the network. If so, then the station delays its transmission until the channel is free, at which point transmission begins. A station that has detected collision will also jam the channel for a very brief period to ensure that all stations have detected and reacted to the collision it has itself detected.

During transmission the station will listen in to ensure that no other station has started to transmit at the same time. Should this be the case (i.e., a collision has been detected), then both stations will stop transmitting for a randomly generated delay period (called the collision interval). Since all stations will wait a different period of time before attempting to retransmit, the chances of further collision are considerably reduced. It is important that the collision interval is based upon the round-trip propagation time between the two stations on the network that are farthest apart. Software is available that will monitor the collision level on the network and advise on capacity planning and physical network structure to ensure maximum throughput. A CRC check is applied to all packets on transmission and is checked by the receiver before handing the packet over to the station for further processing. Damaged packets are generally retransmitted.

The first generation of Ethernet transmitted at a maximum of 10 mbps. However, this bandwidth quickly became saturated and faster networks of 100 mbps (fast Ethernet) and then 1 gigabit per second were developed.

Ethernet is defined by the Institute of Electrical and Electronics Engineers (IEEE) standard 802.3. The original 802.3 standard was published in 1985. Originally two types of coaxial cables were used, called Thick Ethernet and Thin Ethernet. Later unshielded and shielded copper twisted pair (TP), used for telephones, was added. The gigabit Ethernet was designed to operate over fiber as well. Most installations today use a mix of fiber and twisted pair for cabling. A new 10 gigabit per second standard is being drafted, but this standard will only operate over fiber.

ATM

Asynchronous Transfer Mode (ATM) is a means of digital communications that is capable of very high speeds. It is used for the transport of voice, video, data, and images. ATM is an International Telecommunications Union-Telecommunication Standardization Sector (ITU-T) standard.

Unlike Ethernet, ATM is a connection-orientated protocol—a call set-up phase establishes a virtual/logical channel between the device and the network. Once this is established, all data will be transferred over this logical channel until the data transfer is complete. This call-set phase also establishes the throughout and quality of service required.

Data are conveyed in small, fixed-size cells (packets) of 53 bytes—5 bytes of header information that define the logical channel and 48 bytes of actual data. The information to be sent is divided among these cells, transmitted, and then reassembled at their final destination.

ATM corresponds to the lower layers of the OSI model:

1. Physical layer of ATM corresponds to layer 1 of the OSI model and supports various transmission media operating at kilobits per second to gigabits per second.

2. ATM layer corresponding to the lower part of the OSI layer 2 defines short fixed-length cells with multiplexed logical channels within a physical channel. The fixed length cells enable very-high-speed switching hardware to be manufactured.
3. The ATM adaptation layer defines the type of service to be provided (constant bit rate services, variable bit rate services, connectionless services/data protocols, high-speed data protocol).

Open Systems Interconnect (OSI)

In the past few years much emphasis has been placed on a concept of a standard that would permit equipment from a manufacturer or supplier to communicate with any other equipment, irrespective of the supplier. This enabled the interconnection between systems in a completely open manner, which led to the name Open Systems Interconnect (OSI).

The concept breaks down the whole business of communicating between systems into seven different layers. Thus, the problem of physical connection is separated from the method of controlling the movement of data along that connection. Each layer is subject to an individual standard compiled by ISO. Some of these standards also incorporate earlier standards issued by other bodies such as the IEEE. The seven layers are as follows:

Layer 1. Application: The traditional computer program (application) that determines what need is to be met, what data are to be processed or passed by whom to whom for what purpose.

Layer 2. Presentation: Interfaces between the application and other layers to initiate data transfer and establish data syntax.

Layer 3. Session: Manages communication session connection/severance, synchronization and reports on exception conditions.

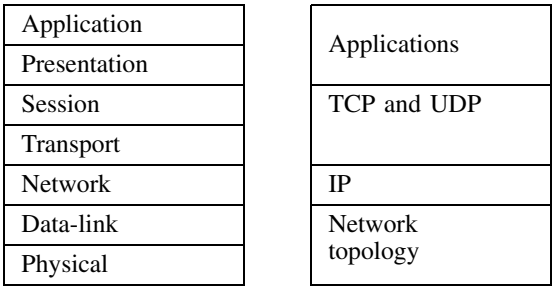
Layer 4. Transport: Manages end-to-end sequencing, data flow control, error recovery, multiplexing, and packeting.

Layer 5. Network: Maintains the availability and quality of the overall network and manages network flow and logical division.

Layer 6. Data link: Detects and attempts to correct physical errors, and manages data linkages, station identification, and parameter passing.

Layer 7. Physical: Provides the actual physical mechanical and electrical services required to establish and maintain the physical network. Examples include serial and parallel cables, telephone cabling.

OSI is a conceptual model and is very useful for discussing the various network services. However, not all networking technologies have seven layers, nor do they match the seven layers in the model. However, the model presents a good framework for comparing the various technologies. The most used protocol for network technologies is TCP/IP. TCP/IP does not strictly conform to the OSI conceptual model. The diagram below shows how TCP/IP fits into the framework.



TCP/IP

In the early 1960s most companies bought large mainframe computers that they linked together using proprietary networking technologies provided by the company manufacturing the computer. It was not easy to interconnect computers and peripherals from different manufacturers together. In an effort to increase the sharing of resources across different computers, the Advanced Research Projects Agency (ARPA) of the Department of Defense of the U.S. Government initiated a project to link their research computers together across many sites using packet-switching technology. The project was called ARPAnet.

In 1973, work began on a new protocol suite, TCP/IP, a set of network protocols to allow any system to connect to any other system over any network topology. In 1982, the ARPAnet adopted this protocol for their network. During the same period the University of California at Berkeley began incorporating TCP/IP in their freely distributed version of UNIX, already widely used in the academic community. This then allowed research computers across the United States to interact and share resources with themselves and the ARPAnet. This collection of computers all interconnected using some form of network topology using the TCP/IP protocol suite became known as the Internet. The TCP/IP protocol relates to the OSI model as shown above.

The Internet protocol is the network layer in the OSI reference model. Its job is sending to blocks of data (packets) from one point in the network to another using the lower layers to do so (commonly in a LAN Ethernet). Every device on a network is allocated a unique IP address, and an address resolution protocol resolves this IP address to the actual physical address of the device to receive the data. Every message/data of TCP/IP eventually ends up as an IP packet of data. Each packet contains the necessary headers and trailers, checksums, the IP source and destination addresses, a protocol version number, and the block of data to be sent. TCP/IP allows a maximum packet size of 64 KB of data but for practical purposes normally packets are of less than 512 bytes of data.

Every IP address is unique, and sites have to apply to a formal body to obtain their unique address range. Each address is 32 bits long, written as a 4-byte sequence separated by full stops. For example:

255.255.255.255

These addresses are used by routers in the network to move the packets of data to their eventual destination.

IP is a connection-less protocol—it has no concept of a job or session. Each packet of data is an entity in itself. This is like a postal worker sorting letters—the worker is not concerned with whether a letter (packet) is part of a batch of not. The worker simply routes each individual letter (packet), one at a time, to the next location on the delivery route.

On a LAN it is quite common for each computer on the network to be allocated its IP address at system start-up time. Under the UNIX operating system this was achieved by a program called BOOTP and a BOOTP name server. More commonly this service is now performed by a protocol called dynamic host configuration protocol (DHCP) and a corresponding DHCP name server. DHCP is a superset of BOOTP.

Transport Layer. Application protocols do not communicate with the IP directly but rather talk to one of two transport layer protocols—TCP and UDP. These transport layers then pass data to the IP, which then encapsulates the data into the IP packets to send over the network. The transport protocols hide the network from the applications protocols so they do not have to deal with the packeting of data and other issues while also shielding the network with having to multiplex many different application protocols (e.g., file transfer, mail, etc.).

Application Layers. The common application protocols used with TCP/IP include FTP (file transfer protocol), and SMTP (simple mail transfer protocol), and SNMP (simple network management protocol). However, since TCP/IP is used in some many networks, many other protocols have also been developed to handle, for example, telephony, video transmission, etc.

Network Concepts

Whatever the type of network, there are a number of concepts that are common.

File Transfer. A network should have the ability to transfer a file (or part of a file) from one node to another without the intervention of programmers each time the transfer takes place. The file may contain programs or data, and since different types (and possibly generations of computers) and different applications are involved, some reformatting may be required. This requires a set of programs to be written to cover all foreseen transfer requests and a knowledge of all local file access methods and formats. One good example of the need for this is the application known as archiving. This involves the transmission of copies of files held on computer to another system in another location.

Resource Sharing. It may be more cost-effective to set up communication links to share expensive peripheral devices than to duplicate them on every computer in the network. For example, one computer may have a large sophisticated flatbed printer/plotter for producing large engineering drawings. To use this, the other computers would store the information necessary to load and run the appropriate program remotely. This would be followed by the data describing the drawing to be produced.

Remote File Access/Enquiry. It is not always necessary or desirable to transfer an entire file, especially if only a small amount of data is required. In these circumstances what is needed is the ability to send an enquiry from a program (or

task) running in one computer and remotely load, to the other system. This enquiry program will retrieve the requisite data from the file and send them back to the original task for display or processing. This comes under the broad heading of task-to-task communications.

Logical Channels. Users of a computer network will know where the programs and data that they want to access exist. They do not want to concern themselves with the mechanics of how to gain access to them. They expect there to be a set of predefined rules in the system that will provide a logical channel to the programs and data they wish to reach. This logical channel will use one or more logical links to route the user's request and carry back the response efficiently and without errors. It may be that there is no direct physical link between the user's computer and the machine he or she is trying to access. In these circumstances the logical channel will consist of a number of logical links. The physical links, in some cases, may be impossible to define in advance, since in the case of dial-up communication using the public switched network the route will be defined at connection time.

Virtual Terminal. This is a very simple concept. It describes a terminal physically connected to computer A but with access (via A) to computer B. The fact that one is communicating via A should be invisible to the user. Indeed, to reach the ultimate destination, the user may unknowingly have to be routed through several nodes. The use of common systems such as Ethernet and the promotion of common standards such as Open Systems Interconnect has bred a new concept in connecting terminals to computers, with the emphasis placed more on the service that a user requires. Whereas previously the user had only to know to where the connection was required and not how to get there, with Ethernet-based servers he or she need only know the name of the service that is required and no longer need specify where it resides. The terminal will be connected to Ethernet through a computer acting as a router. The server will know on which machine or machines the service required is currently available, and needs to know if the service has been moved, whereas the user does not. Furthermore, if the service is available on more than one machine, then the server will be capable of balancing the terminal workload given to each machine, all without the user even having to know from where the service is being provided.

Many terminal servers are even capable of running more than one terminal to computer sessions simultaneously on the same terminal, enabling the user to switch between them as desired without the host computer thinking that the session has been terminated. Workstations are able to carry out this sessions service for themselves. In all these examples the terminal is considered to be virtual by any of the host machines to which it is connected via the terminal server, This concept and the facilities that it offers is quickly eroding many of the problems associated with previous methods of connecting terminals to computers, and the switching and physical patching that was required to connect a terminal to a new machine.

Emulator. As the name implies, this consists of one device performing in such a way that it appears as something different, For example, a microprocessor acting as a terminal to a remote host uses terminal emulation software to do so. This software emulates all the actions of a terminal onto the microprocessor.

Routing. As soon as we add a third node, C, to a previously point-to-point link from A to B, we have introduced the possibility of taking an alternative route from A to B, namely via C. This has advantages. If the physical link between A and B

is broken, we can still transmit the message. If the traffic on the AB link is too high we can ease the load by using the alternate route.

However, this does bring added complications. The designer has to balance such factors as lowest transmission cost versus load sharing. Each computer system has to be capable of recognizing which messages are its own and which it is required merely to transmit to the next node in the logical link. In addition, when a node recognizes that the physical link it was using has, for some reason, been broken, it must know what alternative route is available.

Network Design

Network design is a complicated and specialized science. Computer users do not typically want to reinvent the wheel by writing from scratch all the network facilities they require. They expect their supplier to have such software available for rent or purchase, and, indeed, most large computer suppliers have responded with their own offerings.

Standard Network Architecture. The most commonly used network architecture is the Internet, which uses the TCP/IP protocol suite to interconnect a wide variety of different manufacturers' computers together using many different types of physical media.

Packet Switching. Packet switching was designed to overcome the limitations of the telephone network. When one computer wished to interact to another, many different types of interactions were required at the same time, for example a file being transferred from one to computer to another at the same time as a terminal interaction. If the file was very large it was not desirable that the terminal interaction wait for the complete file transfer to complete, resulting, perhaps, in the human user of the terminal waiting perhaps minutes or hours for their interaction to continue. This led to the idea of breaking down messages into small parts (known as packets) that are addressed and then interleaved and transmitted over a network. The user has no influence over the route the packets take. Indeed, the complete contents of a message may arrive by several different routes.

This also overcame a reliability issue of early networks. If an error resulted in the transmission, only the packets lost in transmission needed to be resent rather than the whole message.

5.16 INTERNET

As stated above, the Internet is the worldwide connection of interconnected computer networks using the collection of protocols known as TCP/IP. Although it was initially used to interconnect research networks and later business networks, there are now many network operators providing services to customers interconnecting their home computers to the Internet. For a long time the Internet was primarily used for e-mail, file transfer, and terminal access to other computers. However, the last 10 years have seen the development of the World Wide Web, which has changed the way that people access documents and information.

World Wide Web

The World Wide Web (also known as W3) was originally developed to allow information sharing within internationally dispersed teams and the dissemination of information by support groups. Originally aimed at the High Energy Physics community, it was developed at the Centre for European Nuclear Research in Switzerland by Tim Berners-Lee in the early 1990s.

The Web uses hypertext and hypermedia to link together documents on different computers across the Internet. Hypertext is specially formatted text that allows various forms of linking and retrieval between multimedia files and text documents. Linked words or phrases in a hypertext document can be chosen by a reader, which causes another document or file to be retrieved and displayed. Hypertext is also described as nonlinear text. Many people consider the terms *hypertext* and *hypermedia* synonymous. Nominally *hypertext* refers to relating textual elements, while *hypermedia* encompasses relationships among elements of any media type. The concepts are identical, though hypertext is more difficult to implement in nontextual media.

You use a mouse to point and click on different parts of the document that may take you to another document located on a computer located elsewhere transparently. With the Web you are delivered a page of information where the page can contain text, graphics, photographs, video, etc. The software you use to view documents located on the Web is called a browser, the most commonly used being those provided by Netscape and Microsoft.

HTML. Hypertext markup language (HTML) is the computer language used for publishing hypertext on the World Wide Web. It is a nonproprietary format based upon SGML (Standardized General Markup Language) and can be created and processed by a wide range of tools, from simple plain text editors (typed in from scratch) to sophisticated WYSIWYG authoring tools such as Microsoft FrontPage Editor. HTML uses tags such as <h1> and </h1> to structure text into headings, paragraphs, lists, hypertext links, etc. For example, a title of a document would be represented as:

```
<title>my title</title>
```

and a heading for a chapter as

```
<h1>an important heading</h1>
```

A paragraph would be

```
<p>content of the paragraph</p>
```

and a link to another document on the WEB.

XML. XML is a language to define document structures and elements. Like HTML, XML makes use of tags and attributes, but while HTML specifies what each tag and attribute means (and often how the text between them will look in a browser), XML uses the tags only to delimit pieces of data, and leaves the interpretation of the data completely to the application that reads it. For example a "<p>" in an XML file is not necessarily a paragraph. Depending on the context, it may be a price, a parameter, a person, or another defined value.

XML is a subset of SGML and enables generic SGML to be served, received, and processed on the Web.

5.17 SOFTWARE

Introduction

Software is the collective name for programs. Computer hardware is capable of carrying out a range of functions represented by the instruction set. A program simply represents the sequence in which these instructions are to be used to carry out a specific application. However, this is achieved in a number of ways. In most cases the most efficient method of using the hardware is to write in a code that directly represents the hardware instruction set. This is known as machine code and is very machine-dependent. It requires a high level of knowledge of the particular type of computer in use, is time consuming, and is specific for each type of computer the program is being written for. In practice, therefore, programmers write in languages in which each program instruction represent a number of machine instructions. The programs produced in this high-level language clearly require to be translated into code that can operate upon the computer's instruction set.

It would be possible, of course, to buy computer hardware and then set out to write every program one needed. However, this would take a very long time indeed. Most users require their system to perform the same set of basic functions, such as reading, printing, storing, and displaying data, controlling simultaneous processes, translating programs, and many others. Consequently, most computers are supplied with prewritten programs to carry out these functions. These fall into four basic categories.

Operating Systems

The operating system is a core set of programs that sit between the application programs designed to solve a particular problem and the general-purpose hardware (monitors, disks, etc.). It allocates and controls the system's resources, such as the CPU, memory, storage, and input/output, and allocates them to the application program or programs. Part of the operating system will be permanently resident in main memory and will communicate with the user and the programs that are running. The functions it will carry out will typically be:

- The transfer into memory of non-resident operating system routines.
- The transfer into memory of application programs or parts of them. In some cases there is insufficient memory to hold an entire program and so little-used portions of the program are held on disk and overlaid into memory as they are required.
- The scheduling of processor time when several programs are resident in memory at the same time.
- The communication between tasks. For ease of programming, a large program can be broken down into sections known as tasks. In order to complete the application it may be necessary to transfer data from task to task.
- Memory protection, ensuring that co-resident programs are kept apart and are not corrupted.

- The transfer of data to and from input and output devices.
- The queuing of input/output data until the appropriate device or program is ready to accept them.

Capabilities of operating systems vary according to the tasks they have to perform. For example, a large operating system supporting many thousands of users has to provide a range of security features and administrative functions to keep track of the usage of the computer, whereas an operating system supporting just one user of a desktop computer can be much smaller and less complex.

Operating systems are generally classified in terms of the number of users they support, the number of simultaneous tasks they perform, and the type of processing they allow.

Single-User Operating System. Such an operating system allows only one user at a time to access the computer. Examples include the DOS operating system and the operating systems that enable PDAs to function. Most of the small computers now use a graphical user interface (GUI) to interact with the user. A GUI typically includes the following:

- Icons (graphical images) used to represent items of the computer such as files, printers, etc.
- A graphical pointer, controlled by a pointing device (normally a mouse) to select the above icon and move around on the screen
- Pull-down menus that appear/disappear on the screen and are controlled by the pointing device
- Windows that contain applications and objects on the screen

Batch Processing. This was the original processing method and is still heavily used where large amounts of data have to be processed efficiently without a major emphasis on timing. Data are transcribed onto some input medium and then run through the system to produce, typically, a report. Classical batch jobs include such applications as payroll and month-end statement runs.

The advantage of batch processing is its efficiency in processing large amounts of data. The major disadvantage is that once a user has committed a job he or she must wait until the cycle is completed before any results are received. If they are not correct the job must be re-submitted with the necessary amendments.

Time-Sharing Operating Systems. Time-sharing enables more than one user to use the computer at the same time. A single person using a keyboard does not usually use the power of a computer to any more than a fraction of its capacity. Consequently, the resources of the system maybe shared between many users in a process known as time sharing. This should not be apparent to the individual user, who should receive a response to a request in one or two seconds under normal loading of the CPU and other resources. Time sharing, as the name suggests, involves the system allotting time slices, in rotation, to its users, together with an area of memory. Some users may have a higher priority than others, and so their requests will be serviced first. However, all requests will be serviced eventually.

Requirements of interactive time-sharing operating systems are efficient system management routines to allocate, modify, and control the resources allocated to individual users (CPU time and memory space) and a comprehensive command language (normally provided by a graphical user interface).

Such systems are still used where a large number of users require access to a controlled environment where they are all sharing the same data, such as ATMs in banks, flight planning and booking, order input, and enquiry etc.

Examples of time-sharing systems include the VAX and UNIX operating systems.

Multitasking Operating Systems. A single-task operating system allows only one program to execute at a time, and the particular program executing must finish completely before a new program may be started. Multitasking operating systems appear to allow a single CPU to execute more than one program at once—they do this by using interrupts. Microsoft's Windows operating system is an example of a single-user, multitasking operating system.

Transaction Processing. This is a form of interactive processing which is used when the operations to be carried out can be predefined into a series of structured transactions. The communication will usually take the form of the operators filling out a form displayed on the terminal screen, a typical example being a sales order form. The entered data are then transmitted as a block to the computer, which checks them and sends back any incorrect fields for correction. This block method of form transmission back to the computer is very efficient from a communications perspective but can be inefficient from the point of view of the terminal operator if there are many fields in error or if the validation of any of the fields is dependent on the contents of other fields on the same form. Some systems, therefore, send back the input character by character and are able to validate any field immediately and not let the operator proceed past a field until it is correct. The options available to the operator will always be limited and he or she may select the job to be performed from a menu displayed on the screen.

Typical requirements of a transaction-processing operating system are as follows:

1. Simple and efficient forms design utilities.
2. The ability to handle a large volume of simultaneous interactive users.
3. Efficient file-management routines, since many users will be accessing the same files at the same time.
4. Comprehensive journaling and error recovery. Journaling is a recording of transactions as they occur, so that in the event of a system failure the data files can be updated to the point reached at the moment of failure from a previously known state of the system (usually a regular back-up).

Real-Time Operating Systems. Real-time operating systems are used to record and control processes. In such applications, the operating system must respond to external stimuli in the form of signals from sensing devices. The system may simply record that the event has taken place, together with the time at which it occurred, or it may call up a program that will initiate corrective action, or it may pass data to an analysis program.

Such a system can be described as event- or interrupt-driven. As the event signal is received it will interrupt whatever processing is currently taking place, provided that it has a higher priority. Interrupt and priority handling are key requirements of a real-time operating system. Some operating systems may offer the user many possible interrupt levels, and the situation can arise in which a number of interrupts of increasing priority occur before the system can return to the program that was originally being executed. The operating system must be capable of recording the

point reached by each interrupted process so that it can return to each task according to its priority level.

Common Concepts. There are some concepts that are common to most operating systems.

Foreground/Background. The simplest form of processing is single user, either batch or interactive. However, a more effective use of a computer's resources is to partition the memory into two areas. One (background) is used for low-priority interruptable programs, while the other (foreground) is occupied by a program requiring a faster response to its demand. The latter will therefore have higher priority. The recent increases in both the memories and power of many mini and micro systems has relegated the use of this technique to the smaller end of the micro range.

Multitasking. This is an extension of foreground/background in which many programs compete for the systems resources rather than just two. Only one task can have control of the CPU at a time. However, when it requires an input or output operation, it relinquishes control to another task. This is possible because CPU and input/output operations can take place simultaneously. For example, a disk controller, having received a request from the operating system, will control the retrieval data, thus releasing the CPU until it is ready to pass on the data it has retrieved. The operating system is normally stored on a systems disk or on a Read Only Memory (ROM) chip. When the computer is started up, the monitor (the memory resident portion of the operating system) must be read from storage into memory. The routine that does this is known as the "bootstrap."

System Generation (Sysgen). When a computer is installed or modified, the general-purpose operating system has to be tailored to the particular hardware configuration on which it will run. A sysgen defines such items as the devices attached to the CPU, the optional utility programs that are to be included, and the quantity of memory available and the amount to be allocated to various processes.

Data-Management Software

Data to be retained are usually held in auxiliary storage rather than in memory, since if they were held in memory without long-term power back-up they would be lost when the system was turned off. To write and retrieve the data quickly and accurately requires some kind of organization, and this is achieved by data-management software.

The most commonly used organizational arrangement for storing data is the file structure. A file is a collection of related pieces of information. An inventory file, for example, would contain information on each part stored in a warehouse. For each part would be held such data as the part number, description, quantity in stock, quantity on order, and so on. Each of these pieces of data is called a *field.* All the fields for each part form a record, and, all the inventory records together constitute the file. The file is designed by the computer user, though there will usually be some guidelines as to its size and structure to aid swift processing or efficient usage of the storage medium. With file-management systems the programs using the files must understand the type of file being used and the structure of the records with it. There are six types of file organization:

1. Sequential
2. Relative

3. Physical
4. Chain
5. Direct
6. Indexed

Sequential File Organization. Before the widespread use of magnetic storage devices, data were stored on punched cards. The program would cause a record (punched card) to be read into memory, the information would be updated, and a new card would be punched. The files thus created were sequential, the records being stored in numeric sequence. A payroll file, for example, would contain records in employee number sequence.

This type of file organization still exists on magnetic tapes and disks. However, the main drawback is that to reach any single record all the preceding records must be read. Consequently, it is efficient only when the whole file needs to be processed from beginning to end and random enquiries to individual records are rarely made.

Relative File Organization. Relative files permit random access to individual records. Each record is numbered according to its position relative to the first record in the file, and a request to access a record must specify its relative number. Unfortunately, most user data, such as part number, order number, customer number, and so on, does not lend itself to such a simplistic numbering system.

Physical File Organization. Another version of the relative technique is used to retrieve a specific block of data relative to the first block in a file from disk. This is done irrespective of where the actual data records reside in the block, and it would be the responsibility of the application program, not the operating system, to separate out individual records (unpacking). Consequently, situations where this method is advantageous are rare, but if the record size equals that of a physical block on disk then this technique offers considerable advantages in speed of retrieval of the data, particularly if the file is in a physically continuous stream on the disk. This type of file is often referred to as a *physically direct* file.

Chain File Organization. This is, in effect, a file that is required to be read sequentially but where not all the data are available at one time. Earlier file systems did not permit the extension of a sequential file once it was written, and adding data to a file meant reading the whole file, writing it out to a new file as it was read, and then adding the new data onto the end of the new file.

To overcome this limitation, the chain file technique was introduced. Each record was written to the file using relative file techniques, with the application specifying to where each record was to be written. However, each record contained a pointer to the location of the next record in logical (not physical) sequence in the file, or some method of indicating that there were no more records in the chain (usually a zero value pointer). This then enabled the application program to read the file in sequence, irrespective of where the data resided on disk or when the data were put there. The widespread use of sequential files that can be extended coupled with a considerable improvement in database and indexed file techniques has largely made this technique redundant.

Direct (Hashed) File Organization. This is a development of the relative file organization and is aimed at overcoming its record-numbering disadvantage. The actual organization of the file is similar. However, a hashing algorithm is introduced between the user number, identifying a particular record and the actual relative record number that would be meaningless to the user. The algorithm is created once and for all when the system is designed and will contain some arithmetic to carry out the conversion.

This file organization permits very fast access, but it does suffer from the disadvantage in that most algorithms will occasionally arrive at the same relative record number from different user record identification numbers, thus creating the problem of synonyms. To overcome this problem, the file management software must look to see if the record position indicated by the algorithm is free. If it is, then a new record can be stored there. If it is not, then a synonym has occurred and the software must look for another available record position. It is, of course, necessary to create a note that this has occurred so that the synonym can subsequently be retrieved. This is usually achieved by means of points left in the original position indicating the relative record number of the synonym.

The user-numbering possibilities permitted with direct files may be more acceptable to the user since they are not directly tied to the relative record number. However, the need for an algorithm means that these possibilities are limited. In addition, the design of the algorithm will affect the efficiency of recording and retrieval since the more synonyms that occur, the slower and more cumbersome will be these operations.

Indexed File Organization. The indexed method of file organization is used to achieve the same objectives as direct files, namely, the access of individual records by means of an identifier known to the user, without the need to read all the preceding records. It uses a separate index that lists the unique identifying fields (known as keys) for each record together with a pointer to the location of the record. Within the file the user program makes a request to retrieve part number 97834, for example. The indexed file management software looks in the index until it finds the key 97834, and the pointer it discovers there indicates the location of the record. The disadvantage of the system is fairly apparent; it usually requires a minimum of three accesses to retrieve a single record and is therefore slower than the direct method (assuming a low incidence of synonyms in the latter). However, there are a number of advantages:

1. It is possible to access the data sequentially as well as randomly, since most data-management systems chain the records together in the same sequence as the index by maintaining pointers from each record to the next in sequence. Thus, we have indexed sequential or ISAM (indexed sequential access method) files.
2. Depending on the sophistication of the system multiple keys may be used, thus allowing files to be shared across different applications requiring access from different key data (Figure 5.4).
3. Additional types of keys can be used. Generic keys can be used to identify a group of like records. For example, in a payroll application, employee number 7439 may identify K. Jones. However, the first two digits (74) may be used for all employees in the press shop. It is therefore possible to list all employees who work in this department by asking the software to access the file by generic key.

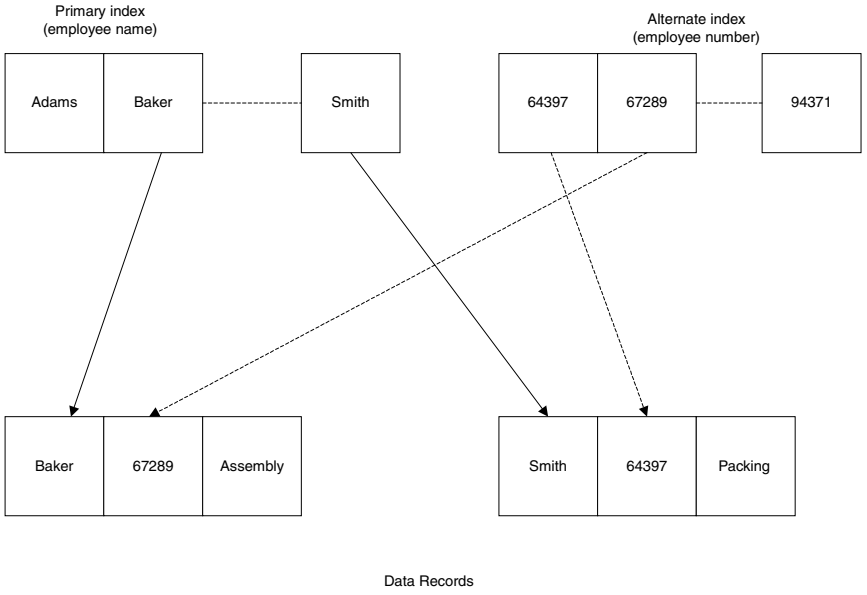


FIGURE 5.4 Multikey ISAM file organization.

4. Another possibility is that of asking the system to locate a particular record that contains the key value requested, or the next highest, if the original cannot be found. This is known as using approximate keys.
5. It is also possible to retrieve records within a given range of keys.
6. Most computer manufacturers provide multikey ISAM systems and so the users do not need to concern themselves about the mechanics of data retrieval.

5.18 DATABASE MANAGEMENT

Files tend to be designed for specific applications. As a result, the same pieces of information may be held several times within the same system, and are often held many times within the same business or organization. This has many disadvantages:

1. It is wasteful of space and effort.
2. It is very difficult to ensure that the information is held in its most recent form in every location.
3. Security maintenance is much more difficult with multiple dispersed copies than it is with a single copy.

It is, of course, possible to share files across applications. However, a program usually contains a definition of the formats of the data files, records, and fields it

is using. Changes in these formats necessitated by the use of the data within new programs will result in modifications having to be made in the original programs.

The database concept is designed to solve these problems by separating the data from the programs that use them. The characteristics of a database are:

1. A piece of data is held only once.
2. Data are defined so that all parts of the organization can use them.
3. It separates data and their description from application programs.
4. It provides definitions of the logical relationships between records in the data so that they need no longer be embedded in the application programs.
5. It should provide protection of the data from unauthorized changes and from hardware and software.

The data definitions and the logical relationships between pieces of data (the data structures) are held in the schema (Figure 5.5).

The database is divided into realms—the equivalent of files and the realms into logical records. Each logical record contains data items that may not be physically contiguous.

Records may be grouped into sets that consist of owner and member records. For example, a customer name and address records may be the owner of a number of individual sales order records.

When an application is developed, a subschema is created defining the realms to be used for that application. The same realm can appear in other subschemas for other applications (Figure 5.6).

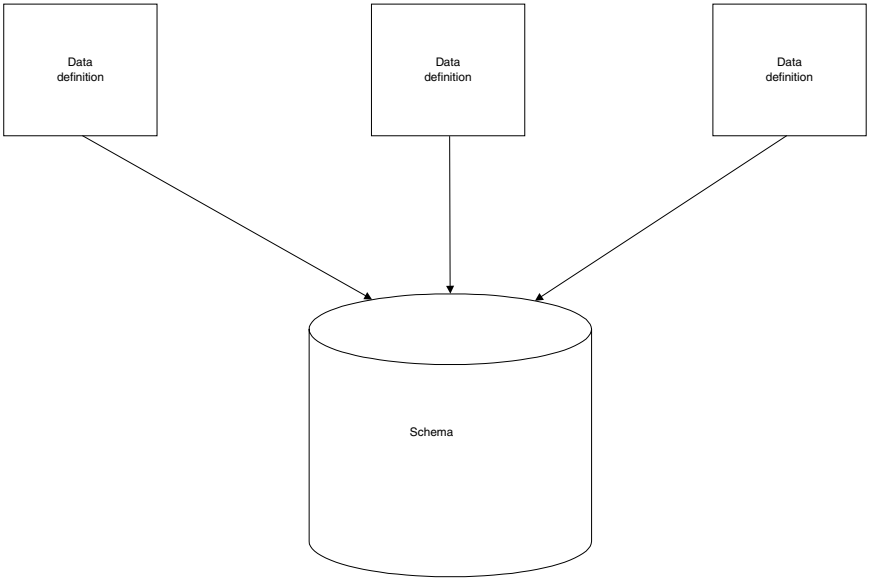


FIGURE 5.5 The schema.

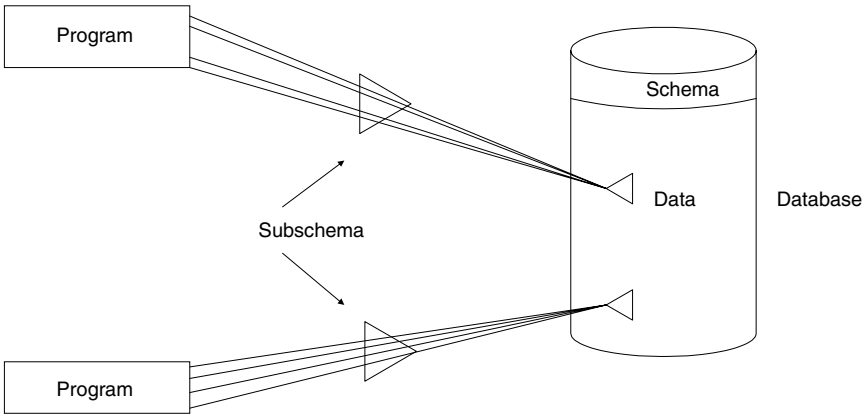


FIGURE 5.6 The subschema.

There are four major definitions of the logical relationships between the data (Figure 5.7):

1. *Sequential:* Here each record is related only to the immediately preceding and following records.
2. *Tree or hierarchical:* In this structure each record can be related to more than one record following it. However, records along separate branches are not directly linked with each other and the relationship can be traced only by traveling along the branches.
3. *Networks:* These are the most complex structures. They are effectively groups of trees where records can be related across branches. Any record can, in fact, be related to any other.
4. *Relational:* A relational database stores data in two-dimensional, tabular arrays. Each table (file) of the database is referred to as a relation, and each row of the table (record) is referred to as a tuple. Through the use of normalization (the successive breaking down of data into groups of two-dimensional arrays where each group's data are functionally dependent upon the group's key) the data are defined in a logical format suitable for use with a relational database. The result is a totally flat file that, using a relational database management system, has the flexibility to dynamically create new relations from extracts of one or more existing relations. This is the most common type of database structure.

Because, within a database management system, data are separated from the programs that use them, the data are regarded as a corporate asset. Management of this asset is in the hands of a database administrator. He or she maintains the schema and works with application programmers to define the parts of the database to which they may have access and to help them create subschemas for their particular applications.

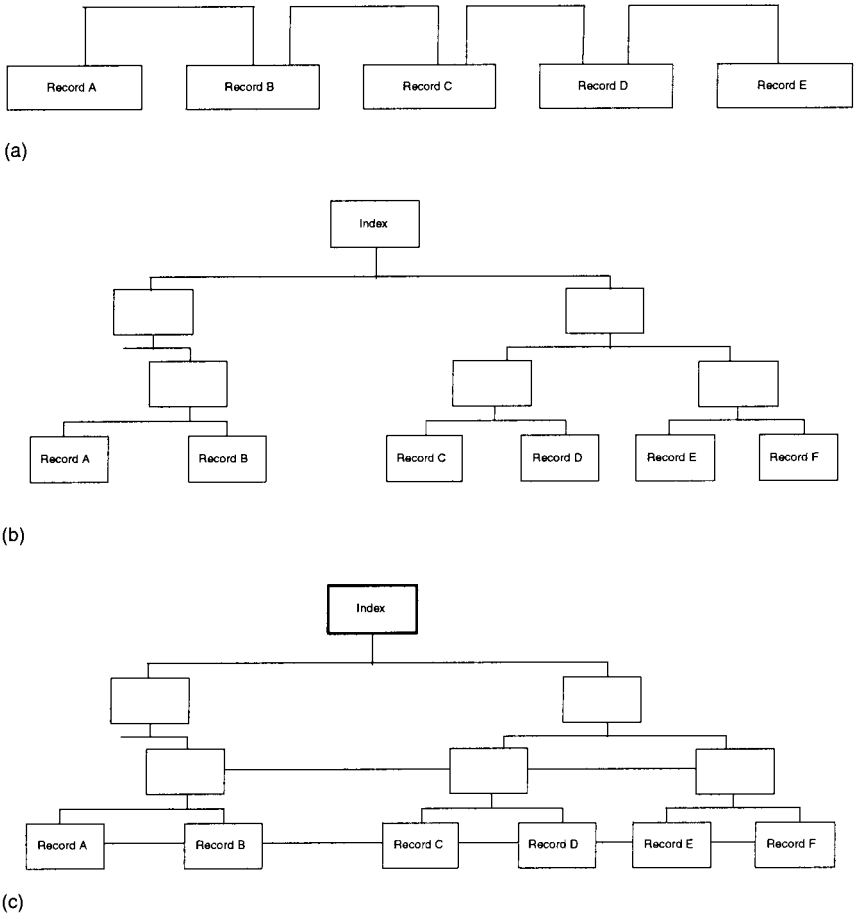


FIGURE 5.7 Types of logical data relationships: (a) sequential, (b) hierarchical, and (c) network.

5.19 LANGUAGE TRANSLATORS

A programming language is a convention containing words, letters, and symbols that have a special meaning within the context of the language. However, programs have to be translated into the binary language understood by computers. The programmer writes a source program that is converted by the language translator into an object program. Usually, during this process checks are made on the syntax of the source program to ensure that the programmer has obeyed the rules. Any errors discovered will be noted, usually in two categories—*terminal* and *warning*. Terminal errors indicate that the translator has found a definite error that is so serious as to either prevent translation from completing or indicate that it is not worth

doing so. A warning is generated when an item is encountered and the translator would have expected something different, but it may not actually be an error. Program errors are known as *bugs* and the process of removing them, as *debugging.* Bugs are sometimes humorously referred to as ''planned features''!

Programs are normally stored in both their original code format (known as *source*) and in their final format that is understandable, and executable directly by the computer (known variously as *object, binary,* or *executable* formats). The working program is the object program, but when changes have to be carried out, these will be made to the source program, which will then be translated to produce a new-object program.

There are two kinds of language translators: *assemblers* and *compilers.*

Assemblers

An assembler is a language processor designed for use on a particular type of computer. In assembly language there is generally a one-to-one relationship between most of the language mnemonics (abbreviated instruction codes which can be read and interpreted by humans) and the computer binary instructions, although predefined sets of instructions can be ''called'' from the assembly program.

There are four parts to a typical assembly language instruction:

- *Label:* This is a name defined by the programmer. When he or she wants to refer to the instruction, this can he achieved by means of the label. It is this facility that enables a programmer to alter the sequence in which instructions within a program are obeyed by jumping from the current position so another identified by its label.
- *Operation code:* This will contain a call or an instruction mnemonic. If a call is used the assembler will insert a predefined code during the assembly process. If the programmer used a mnemonic this will define the operation to be carried out.
- *Operand:* This represents the address of the item to be operated on. An instruction may require one or two operands.
- *Comments:* This is an optional field used for ease of interpretation and correction by the programmer.

Assembly languages are generally efficient and have been used for writing operating systems and routines that require particularly rapid execution. However, they are machine-dependent, slow to write, and demanding in terms of programmer skills.

Compilers

These are used to translate high-level languages into binary code. These languages are relatively machine-independent, though some modifications are usually required when transferring them from one type of computer to another. The instructions in the high-level language do not have a one-for-one relationship with the machine instructions. Indeed, in some languages a single statement could generate many hundreds of machine instructions. Most compilers read the entire source program before translating. This permits a high degree of error checking and optimization. An incremental compiler, however, translates each statement immediately into ma-

chine format. Each statement can be executed before the next is translated. Although it does not allow code optimization, it does check syntax immediately and the system prompts the programmer to correct errors as they occur. These incremental compilers are often referred to as interpreters, since they interpret and then act upon each instruction.

5.20 LANGUAGES

Introduction

All computers work with and understand instructions in the same format—binary. The content of an instruction to achieve the same objective may well differ from one machine to another, but the instruction will be coded in a binary format. This first format of computer language is called the first generation. It is almost unheard of today to find anyone other than compiler writers who work at this level.

Programming in this first generation of languages was complex, lengthy, skillful, and extremely prone to error. It became obvious that improvements were essential and that assistance was required. What better tool to assist than the very computer that was being programmed? If a code system could be developed that was easier to write and read, and that could then be translated by the computer into machine instructions, then programming would become easier. These languages formed the second generation of computer languages and are referred to as *assembly languages* (from the action of the translator of assembling everything together to validate it). Many of the original languages have long since gone or have been replaced by easier and more powerful alternatives, but assembly languages are still the best language for writing programs where flexibility and speed of execution are paramount (such as the operating systems themselves).

Third-Generation Languages (3GLs)

These languages were developed to bring the nature and structure of the instructions much nearer to the programmer's native language, principally to speed up the programming process and reduce the level of skill required. Since many of these languages were developed in the United Kingdom or the United States, the native language was English, and this is still the predominant, almost universal language of computing today. Some examples of third-generation languages are given below.

BASIC (Beginners' All-Purpose Symbolic Instructor Code). This is an easy-to-learn conversational programming language that enables beginners to write reasonably complex programs in a short space of time. The growth in the popularity of time-sharing systems has increased its use to the point where it is used for a whole range of applications, from small mathematical problems through scientific and engineering calculations and even to commercial systems.

A BASIC program consists of numbered statements which contain English words, symbols and numbers such as LET, IF, PRINT, INPUT. * (multiply), +, and so on. BASIC was developed at Dartmouth College, and while there is a standard there are many variations developed by different manufacturers. Both BASIC interpreters and compilers are available.

Visual Basic. Visual Basic is a programming environment that allows a programmer to build Windows applications by using a graphical user interface to choose and modify prewritten code modules, routines, and user-interface components. Programmers drag and drop code pieces to quickly build a new application. The language is popular because it is easy to learn and permits rapid application development.

Fortran (Formula TRANslation). This originated in the 1950s and was the first commercially available high-level language. It was designed for technical applications and its strengths lie in its mathematical capabilities and its ability to express algebraic expressions. It is not particularly appropriate when the application requires a large amount of data editing and manipulation. It is only now used in scientific programming.

A Fortran program consists of four types of statement:

1. Control statements (such as GOTO, IF, PAUSE, and STOP) control the sequence in which operations are performed.
2. Input/output statements (such as READ, WRITE, PRINT, FIND) cause data to be read from or written to an input/output device.
3. Arithmetic statements, such as * (multiplication), / (division), perform computation.
4. Specification statements define the format of data input or output.

The language became more structured with the introduction of FORTRAN 77.

COBOL (Common Business-Oriented Language). The first Codasyl (Conference of Data Systems Languages) specifications for COBOL were drawn up by 1960, the aims of which were to create a language that was English-like and machine independent.

COBOL is a structured language with well-defined formats for individual statements. A COBOL program consists of four divisions:

1. *Identification,* which names and documents the program
2. *Environment,* which defines the type of computer to be used
3. *Data,* which names and describes data items and files used
4. *Procedures,* which describes the processing to be carried out

The sentences within COBOL can contain ''verbs'' such as ADD, SUBTRACT, MULTIPLY, and DIVIDE and are readable in their own right. For example, in an invoicing program you may find the line:

IF INVOICE TOTAL IS GREATER THAN 500 THEN GO TO DISCOUNT ROUTINE

As a result, by intelligent use of the language the programmer can produce a program that is largely self-documenting. This is a significant advantage when modifications have to be made subsequently possibly by a different programmer.

A general-purpose, structured language like COBOL is not as efficient in terms of machine utilization as assembly language or machine code. However, in a commercial environment, programmer productivity and good documentation are gen-

erally the most important factors. The calculations are usually not complex and therefore do not require great flexibility in terms of number manipulation.

PASCAL. This is a strongly typed, block-structured, procedural third-generation programming language which became very popular in the 1980s due to its extensive use in educational establishments. PASCAL lends itself readily to modern design techniques (including top-down, stepwise refinement, etc.) and has built into it the necessary building blocks to make the most of structured programming. Additionally, because of its strong enforcement of data type, scope, and syntax rules, PASCAL is a language less prone to programmer errors than many others.

All of the above points assist programmers in writing well-designed programs and produce very supportable code. PASCAL was the forerunner of modern programming languages.

The C Language. The C language is a general-purpose language. It has a close association with the UNIX operating system, which is written in C. However, not being tied to any operating system or machine has made it a popular choice for the development of portable programs that can be run on a variety of hardware. This has been enhanced with the publication of an ANSI standard for the C language, based on the original C reference manual, which includes a definition of the language itself and also of a set of library routines for accessing the operating system such as file operations, memory allocation and string manipulation.

The basic data types available in C are characters, integers, and floating-point numbers. More complex data types can be created using structures, unions, and arrays. A fundamental data type is the pointer. Use of pointers allows programs to be written to take full advantage of dynamic memory allocation and for those programs to be independent of the underlying machine architecture.

The basic language constructs are ''if–else'' for decision making, ''for,'' ''while'' and ''do'' loops, and ''switch'' statements selecting from of a number of possible cases. Much of the C language is based upon a language called BPCL developed by Martin Richards.

ADA. This language was developed primarily to support the development and implementation of complex, real-time software for embedded (i.e., usually contained in ROM) software, military hardware, and applications. Its history can be traced back to a U.S. Department of Defense decision in January 1975 to set up the H Order Language Working Group to pursue the development of such a language. ADA uses PASCAL techniques for control and data typing. It also permits the segregation of regions within a program that enables separate development of each other. Routines can be written in a generic sense (say, for sorting) and rapidly customized to perform the same operation on different data. It is able to do this since explicit data typing is required initially.

ADA supports multitasking in the same program, but the sequence of operation is not guaranteed to the programmer. Exception handling is possible under programmer control, even to the extent of permitting re-entry into routines that under normal circumstances would have crashed the system! Each compiler created for ADA under licence has to be periodically revalidated against a defined ADA standard, in an attempt to ensure that this standard is not compromised. However, the standard ADA functionality is very large and has been a disincentive for many organizations to support development on its machines under ADA.

It has been adopted as the standard preferred language by many federal and international organizations, including the U.S. Department of Defense, the Federal Aviation Administration, NASA, NATO, and the ministries of defense in many western European countries. Despite these advances, the use of ADA is not widespread and is mainly limited to defense contracts, with very limited uses in commercial environments.

JAVA. During the early 1990s a team of programmers at Sun Microsystems developed a computer language called Oak. The language was developed to aid the programming of interactive devices such as cable TV devices. In 1994, Sun made a decision to adapt the language to make it suitable to meet the demands of the Internet, including meeting the demand of making the language computer neutral—the same program should be capable of running on computers from different manufacturers without having to be changed at all. In 1995, the language was renamed Java and had been developed into a powerful programming language for the development of Web-based applications. Sun made the language and development environment freely available across the Internet, and the programming community quickly adopted it. Now many commercial developers and software companies support the language and provide development tools.

Object-oriented programming is about the identification and manipulation of objects. An object allows a concept to be expressed and then used repeatedly without the need to worry about the actual detail of the object. For example, consider a car. There are many different types of car, but we all recognize the properties of what makes up a car—we can just refer to a car without having to continually describe all the details that enabled us to identify it as a car in the first place. Java supports this level of abstraction via what is known as the object-oriented mechanism of class—where a class allows an object to expressed once and then reused many times. The Java system comes with an integral library of objects in the form of a large class library covering many of the tasks that a programmer wishes to use.

Java programs come in two forms—applications and applets. JAVA applications are no different from the normal type of programs that are developed in other computer programming languages. Applets are mini-applications that can be embedded into Web pages. Applets are downloaded to an Internet browser and contain one or more small Java programs than start to execute when the Web page is displayed.

Java is a compiled language, but instead of producing instructions for a particular computer the compiler generates Java byte codes that are then interpreted by a Java virtual machine. The Java virtual machine is effectively an ideal computer containing all the normal features of a typical computer. The Java virtual machine is then implemented onto an actual computer. Programmers no longer need to worry about the features of a number of different computers but can target their programs towards the one architecture.

C++. C++ is a development of the C programming language that uses object-oriented technology. It is used widely to write operating systems and applications systems such as spreadsheets, word-processors, etc. It is one of the most widely used programming languages in use.

Fourth-Generation Languages

These are the closest yet to native language, enabling a PC with very little knowledge or skill in programming languages to write instructions performing complex

and lengthy tasks on the computer. They do not generally produce code that is as efficient as previous generations, and the amount of machine resources that is required to support their use is sometimes either misunderstood by those using them or misrepresented by those selling them and very often both.

Unlike previous generations of programming languages, 4GLs demand that the instructions given to them need to follow the exact procedure and sequence in which they are require to execute. They are capable of breaking complex English statements down into many component instruction parts, then deducing the optimum order in which to carry them out. For this reason they are often referred to generically as nonprocedural languages.

These languages are most powerful when used in conjunction with some form of database management system, since they can simplify the interface between the programmer/program by undertaking all necessary communication with the database system. Many of these language/data interfaces have developed so far that the user need not know, and will often be unaware of, how the data are organized or formatted within the system. Some examples of these languages are PowerHouse, Focus, SQL, and Query.

In broad terms, the increased speed of processing and data handling currently available and the low cost of memory have reduced the pressures on programmers to code for maximum speed and efficiency. It is frequently more economical to spend more money on hardware than to allow programmers to spend time optimizing the performance of their programs.

This, coupled with the shortage of trained programmers, has resulted in increased emphasis on simple languages, good program development tools, and a general emphasis on programmer productivity.

SECTION 6

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

Section Editor: Mark Davies

PART 1

SUMMARY OF NUMBER SYSTEMS

Charles J. Fraser

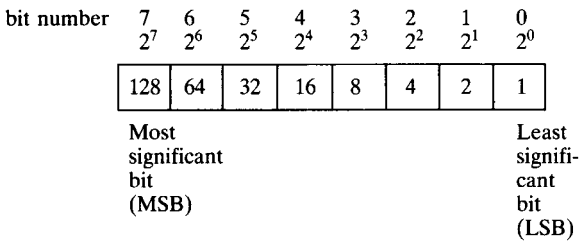
John S. Milne

In the manipulation of data within a computer a two-state numbering system is used. Termed the binary system, and it is based on a simple ON/OFF principle. For the semiconductor integrated circuits which make up the computer system, 5 V denotes ON (or logic level 1) while 0 V denotes OFF (or logic level 0). In practice, a tolerance band is adopted, with 2.4–5 V representing logic 1 and 0–0.8 V logic 0.

The microelectronic devices in the system handle the transfer of information in 1's and 0's, which are referred to as BITs, being a short form for BInary digiT. A group of eight bits is termed a byte and a number of computer systems are based on 8-bit technology with the handling of data codes as 8-bit "words." Sixteen- and 32-bit machines are also available.

The computer operates with three numbering systems, decimal, binary, and hexadecimal (often simply called "hex"). Numerical data would normally be entered by a human operator in decimal form, since this is the most familiar number system. The computer, however, must ultimately convert the decimal number into a binary code, since this is the eventual form in which the number will be processed and stored. The hexadecimal system is an in-between state and represents a particularly compact method of handling binary numbers as groups of four bits.

In binary representation the only possible logic levels are 0 and 1. The base is chosen as 2 and integer numbers can be represented using 8-bit codes as shown below:



The conversion from binary to decimal is illustrated as follows:

binary number	1	0	1	1	1	0	0	1
giving	128	+ 0 +	32	+ 16	+ 8	+ 0	+ 0	+ 1
	= 185 decimal							

Conversion from decimal to binary is the reverse process to the above.

It is apparent that the highest number that can be accommodated in 8-bit binary notation is 1111 1111, which is equivalent to 255 decimal. Generally, therefore,

computer systems handle integer numbers in four consecutive bytes, i.e., as 32-bits. The most significant bit is used to denote the sign of the number and the resulting range of integer numbers is

$$-2^{31} - 1 \text{ to } 2^{31}, \text{ or } -2,147,483,648 \text{ to } 2,147,483,647$$

Real numbers are handled in five bytes, with the most significant byte representing an exponent and a sign bit and the other four bytes the mantissa and a sign bit. The resultant range of real numbers is 2^{-128} to 2^{127}, with either a positive or a negative sign for the mantissa.

The handling of numbers in binary notation is extremely cumbersome for a human and a shorthand notation is adopted for convenience. This is the hex system, in which the binary number is arranged into groups of four bits. Four bits, which is half of a byte, is called a *nibble*. A byte therefore consists of an upper and a lower nibble.

Since there are only ten unique symbols in the decimal numbering system, the first six letters of the alphabet are used to denote the additional six symbols in the hexadecimal system, i.e.,

Decimal	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Hexadecimal	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F

Using hex notation then, 8-bit binary numbers may be replaced by two hex symbols, e.g.,

$$167 \text{ decimal} = 1010\ 0111 \text{ binary} = A7 \text{ hex}$$

Higher numbers are similarly handled:

$$6836 \text{ decimal} = 0001\ 1010\ 1011\ 0100 \text{ binary} = 1AB4 \text{ hex}$$

6.1 ASCII CODE

In the interchange of information between the constituent parts of a computer or a peripheral device, a binary code is used to represent the alphanumeric characters. The most commonly used code for digital communication links is the American Standard Code for Information Interchange (ASCII, pronounced Askey). ASCII is a 7-bit code that can accommodate 128 definable characters.

When communication takes place in a serial fashion, the ASCII code is extended to 8-bits, usually by inserting a zero in the most significant bit. Additionally, one or two start bits, a parity bit, and a stop bit are also included. The start bit(s) inform the receiving device that a character code follows. The parity bit provides a check that no bits have been corrupted during transmission, by ensuring that the sum of all the 1's in the ASCII group give either an even number for "even parity" or an odd number for "odd parity." The stop bit, set to logic "1," terminates the transmission of the character.

The transmission rate in bits/second is termed *baud*. Since there are 11 bits associated with the transmission of one character, a speed of 2400 baud corresponds to 2400/11 = 218 characters per second.

6.2 GRAY CODE

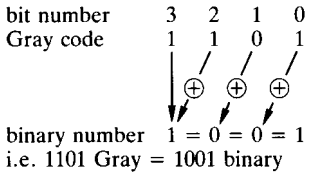
The Gray code is but one of many binary codes in which only one of the digit changes between successive consecutive numbers. The main application is in the sensing of rotational and translational position in mechanical systems.

In converting from Gray to binary code the most significant bit of the binary number, B, is equal to the most significant bit of the Gray code, G. For all other bits, the relationship between binary and Gray is given by

$$B(n) = G(n) \oplus B(n + 1) \tag{6.1}$$

where n denotes the bit reference number and $\oplus$ is an exclusive-OR logic comparison.

The conversion of 1101 Gray to binary is shown below:



In practical position-sensing applications the conversion process can be programmed in the software or implemented in a hardwired logic circuit using logic gates.

PART 2

MICROPROCESSORS

Charles J. Fraser

John S. Milne

6.3 SYSTEM ARCHITECTURE

Since the mid-1970s a number of different microprocessor designs have become available in several versions. The popular designs are produced by a few manufacturers, including such companies as Intel, Motorola, Rockwell, Texas Instruments, and Zilog.

Microprocessor-based systems require additional family support chips, and, in true digital form, all microelectronic components which constitute a microcomputer are designated numerically rather than by name.

Although it is unnecessary for the user of the technology to understand in detail how each individual chip actually functions, it becomes essential to have at least a working knowledge of the logical organization of the system hardware and how each component relates to each other. The composition of this hardware structure is known as the system architecture.

A digital computer system has three main constituent parts: the *microprocessor,* the *memory,* and the *input/output*. Digital signals which have a common function are transmitted between the main components by a group of wires or conduction tracks, termed a bus. In a microcomputer there are three buses: the data bus, the address bus, and the control bus. The interconnection between the basic hardware components in a microcomputer is illustrated in Figure 6.1.

The microprocessor is a very large-scale integrated circuit (VLSI), which is the brain of the microcomputer system and acts as the central processing unit (CPU). Integrated circuits are generally classified according to the number of components on the silicon chip.

The main feature of the microprocessor is the arithmetic and logic unit (ALU). The ALU allows the arithmetical manipulation of data, addition, and subtraction

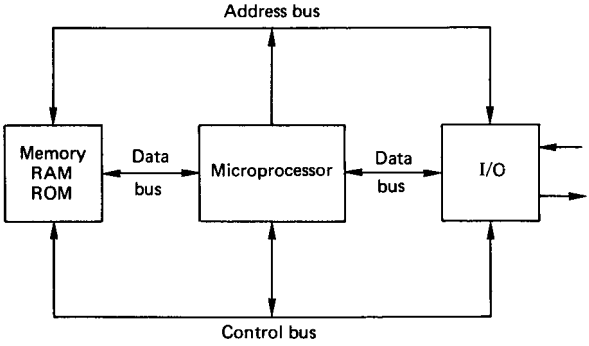


FIGURE 6.1 Basic components in a microcomputer.

with 8-bit systems and multiplication with 16-bit systems. Logical operations (AND, OR, etc.) can also be performed. In addition to the ALU, the CPU contains a number of temporary data storage registers to fetch, decode, and execute instructions, and to time and control circuits to manage the internal and external operation of the complete microprocessor system.

The processing power of the CPU is influenced by such factors as word length, instruction set, addressing modes, number of available registers, and information transfer rates. For word processors or the manipulation of large quantities of data as in CAD packages, 16- or 32-bit microprocessors are essential. In the field of measurement and control, 8-bit systems can be adequate.

The system clock, accurately controlled by a quartz crystal maintaining a constant frequency, acts as the heartbeat for the system and provides all the timing reference points.

All the basic components (CPU, memory, and I/O) and their interconnections may reside in a complete microcomputer system encompassing the traditional keyboard, monitor, etc. Alternatively, they may reside on a single card, or even on a single chip to give a single-chip microcomputer.

6.4 BUS STRUCTURE

The connection between the system components is made by an arrangement of three buses:

1. *The data bus* transmits the words in binary form representing the program instructions for the CPU to execute. It can also carry the information transmitted between the CPU and memory or I/O devices.
2. *The address bus* transmits the memory address related to the data bus information. In 8-bit systems this bus commonly has 16 lines to give 64K of addresses.
3. *The control bus* transmits the collection of timing and control signals which supervise the overall control of the system operation.

The physical format of a bussing system is basically a circuit board with a number of connectors. Different types of microprocessors require different hardware interfaces, and to alleviate the problems, standard bus structures have been developed in order to facilitate the connection of hardware components. In industrial-type systems, cards for various microcomputer functions such as processor, memory, digital and analog I/O, power switching, etc. slot into a standard backplane or motherboard rack. This offers the advantage of being able to plug any specific card, designed to the bus standard, into a free slot in the rack to build up the system as required.

The physical form of the bus is represented by its mechanical and electrical characteristics. Such information as card dimensions, input and output pin-out connections, signal levels, loading capability, and type of output gates must be known.

6.5 MEMORY DEVICES

Memory devices consist of those used to store binary data, which represents the user program instructions, and those which are necessary for the user to operate

the system. Memory takes the form of one or more integrated circuits. These basically hold locations capable of storing a binary word. Each location is assigned a unique address within the system, and data can be selected through the address bus. As a binary code is deposited by the CPU on the address bus, defining a specific location in memory, the contents of that location are selected and placed on the data bus. The appropriate piece of memory hardware and specific location is selected by means of an address-decoding circuit built up from logic gates within the microcomputer system. The end result is a highly flexible data manipulation arrangement.

In an 8-bit microcomputer (i.e., 8-bit data bus) the address bus is 16-bits wide. This enables $2^{16} = 65\,536$ locations to be addressed, and the total memory capacity of the machine is said to be 64K. The memory is further subdivided into pages, with the high-order byte of the address denoting the page number and the low-order byte indicating one of the 256 locations available on each page.

The types of memory chips built into the system basically divide into two categories:

1. *Random Access Memory* (*RAM*), where data can be read from or written to any specified location. RAM is more correctly defined as read/write memory and data retention is dependent upon power being applied to the device. This type of memory is normally employed for the temporary storage of the computer programs, at the editing or execution stage, or the storage of data from measuring transducers prior to permanent storage as a disk file. In a number of systems available, the RAM is made nonvolatile by providing battery back-up.
2. *Read Only Memory* (*ROM*), where data are held in a secure manner and can be read in any specified sequence. Once the chip is configured it cannot be overwritten, and the programs which specify the system operation (termed the monitor program) are ''burnt'' into ROM when they are known to operate in a satisfactory manner. Basic ROM is inflexible since the software contained therein is developed by the system manufacturer. It is often useful, however, to have all programs which are to be permanently stored in the microcomputer in a nonvolatile form, held in an Erasable and Programmable Read Only Memory (EPROM).

EPROMs are supplied in an uncommitted form with each location holding FF hex. They are configured using an EPROM programmer which ''burns'' or ''blows'' the required data, in machine code form, onto the chip. If an error in the data exists or an alteration is to be made, then the complete EPROM can be returned to its uncommitted state by exposing the small ''window'' in the device to intense ultraviolet light for about 20–30 min. EPROM erasers are available for this purpose. Once programmed as required, it is usual to cover the window with opaque material. If uncovered, it would normally take some months before program corruption was experienced through the effects of natural sunlight.

A similar type of memory device is an Electrically Erasable Programmable Read Only Memory (EEPROM or E^2PROM). This is essentially similar to the EPROM but enables the user to alter any particular byte of data rather than wiping the entire chip.

6.6 INPUT/OUTPUT (I/O) STRUCTURE

With the microprocessor acting as the brain of the microcomputer system and the memory chips storing the system-operating software and application programs, the

other essential hardware required is that associated with the input and output of data in essentially binary form. Interface support chips associated with the various microprocessor families are available to enable communication with such hardware essentials as keyboards, display monitors, disk drives, and printers.

The same I/O interface circuits are used in measurement and control applications. The main functions required of the devices are:

1. Digital I/O logic lines, which can be read or set by the microprocessor
2. Data direction register to configure lines as either input or output
3. Handshake lines to supervise data transfer via the I/O lines
4. Timing and counting facilities

The software used for controlling the communication between the microcomputer and other external devices is dependent upon the I/O interfacing technique employed. The two most common methods are *memory mapped* and *dedicated port addressed*.

Memory Mapped I/O

In this method the I/O chip is connected into the system in the same way as the memory illustrated in Figure 6.1. The I/O lines are contained in groups of 8-bits termed a *port,* and this byte is addressed in the same manner as any other location in memory. The port is accessed using memory transfer instructions like PEEK and POKE in high-level BASIC, or LDA and STA in low-level 6502 assembly language.

Since the interface is connected into the bus structure in exactly the same way as the RAM and ROM, no additional decoding hardware is required. Memory addresses are, however, used up for I/O, and as a result communication is slower than the port addressed alternative.

Dedicated Port Addressed I/O

This method involves a second dedicated I/O data bus as shown in Figure 6.2. When data are to be input or output the necessary control signals are sent from the CPU to the I/O interface chip and the port data are transmitted via the dedicated I/O data bus. This does not affect the addressing of memory within the system and results in faster data transfer than with the memory mapped technique. The ports (or channels) are assigned unique addresses (numbers) on the dedicated bus and

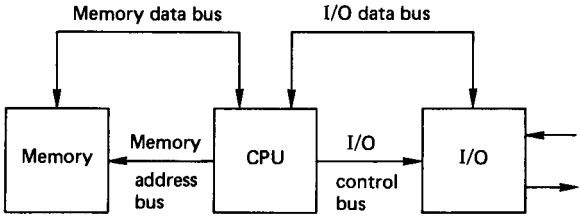


FIGURE 6.2 Port addressed I/O.

are accessed using the additional software instructions of IN (or INP) and OUT in both low- and high-level programming languages.

Direct Memory Access (DMA)

In data-acquisition systems involving analog and digital signals suitably conditioned for inputting to a microcomputer there is a frequency limitation on the sampling rate when using direct program control to transfer data to memory. If it was necessary to acquire the maximum amount of data at the highest speed, using the maximum amount of the computer's resources, then the DMA technique might be employed.

This is a hardware technique which causes the microprocessor to momentarily abandon control of the system buses so that the DMA device can directly access the memory. The DMA controller, connected to the I/O interface, needs to know how many bytes are to be transferred and where in memory the input data are to be stored. The data transfer rate is much faster than in an interrupt servicing method.

6.7 MEMORY MAP

The memory locations in RAM and ROM, which the processor can address, must accommodate space for such requirements as system monitor and utilities, user software, and input/output. The manufacturer of the microcomputer assigns an area of memory for each functional requirement and provides the necessary information in a system memory map.

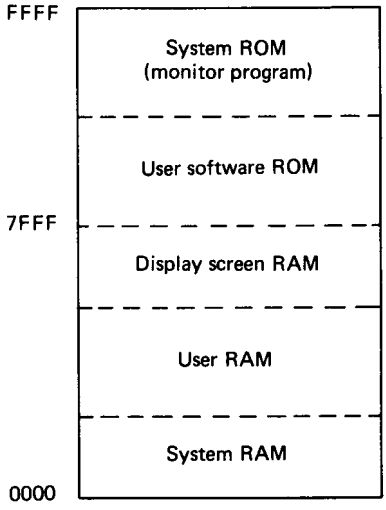


FIGURE 6.3 System memory map.

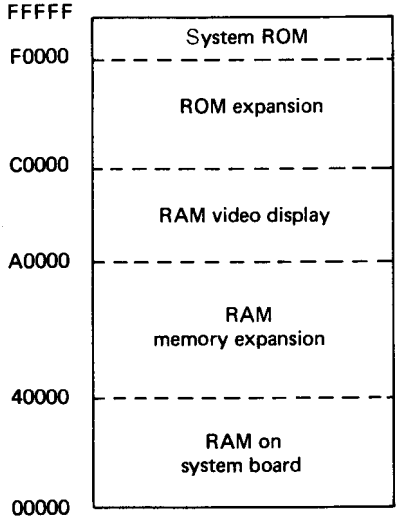


FIGURE 6.4 General memory map for a PC.

The ROM holds the operating system software, and normally some space is available in EPROM form for user firmware. In addition to providing space for user programs, the RAM area contains the system stack and the visual monitor data storage. The I/O facilities are also assigned an area of memory in a memory mapped system.

In a 6502 or 6800 based system the RAM is usually low down in memory and the ROM is high up. A typical memory map is shown in Figure 6.3. The I/O is accommodated anywhere within the above structure and varies from one manufacturer to another.

A general memory map showing the distribution of RAM/ROM for a PC is given in Figure 6.4.

A familiarity with the memory map of the system to be used in any data-acquisition application is essential, since it indicates the areas reserved for the operating system. The programmer can then knowledgeably determine the locations available for data storage and machine-code programs.

PART 3

COMMUNICATION STANDARDS

Charles J. Fraser

John S. Milne

Various standards have been drawn up to define the protocol for the transmission of binary data from within the microcomputer bus structure to external devices such as display monitors, printers, and other peripheral equipment. Most microcomputers are equipped with this facility and manufacturers of data measurement and control instrumentation usually offer an external communication port as an extra.

The most commonly accepted standards are those defined by the American Electronic Industries Association (EIA) and the Institute of Electrical and Electronics Engineers (IEEE). The standards fall into the two categories of serial and parallel data communication. The difference between the two relates to the number of bits of information transmitted simultaneously between the devices. The serial method is the slower of the two, with the bits denoting the characters of information travelling sequentially along a single path. In the parallel method the data word is sent as a parallel code, generally 8-bits wide, resulting in a "bit parallel, byte serial" transmission of information.

6.8 SERIAL COMMUNICATION

Serial communication is the most common method used for the interconnection of a microcomputer to the relatively slow peripheral hardware, or between two computers, when transferring a low volume of information. The (EIA) RS232C, or its successors the RS422 and RS423, have been widely adopted. This allows communication with one peripheral device only. Twenty-one of the signal lines are defined in the standard, although only five (or even three) are all that are usually required.

The three main connections *transmitted data* (pin 2), *received data* (pin 3), and *signal ground* or *common return* (pin 7). These would normally be connected as shown in Figure 6.5. For communication in both directions, i.e., full duplex, the two handshaking control lines—"request to send" (pin 4) and "clear to send" (pin 5)—are also required.

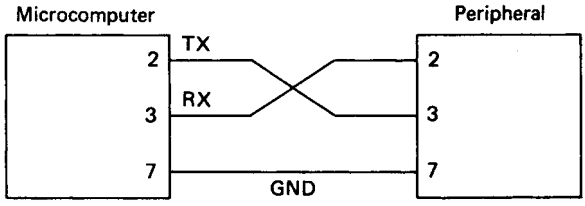


FIGURE 6.5

The standard applies to data transmission interchange, usually at rates between 110 and 9600 baud. A logic "1" is represented by a voltage in the range of −3 to −15 V and a logic "0" by a range of +3 to +15 V. This large differential between "1" and "0" ensures good immunity against electrical noise. However, the voltages used are not compatible with the TTL logic semiconductor family, and interconversion chips are required within the interface.

The RS232C is limited to short communication links of about 30 m, although the RS422 and RS423 standards, succeeding the RS232, have extended communication distances and increased transmission speeds. The RS423, which is compatible with the RS232, has superior driving and receiving interfaces, allowing communication over distances of up to 1500 m at 9600 baud, or 15 m at 100K baud.

It should be noted that while the voltages and signal connections for the plug are defined in the standard, the data protocol is not identified. This must be known for the devices which are to be connected and can be set accordingly by software. The requirements are:

1. Baud rate
2. Number of bits in the ASCII group defining the character being transmitted
3. Odd, even, or no parity
4. Number of stop bits

6.9 PARALLEL COMMUNICATION

The RS232 serial standard for communication was developed essentially for the connection of microcomputers via a telephone link. The parallel standard emerged from the need to establish a means of interfacing a variety of instruments for data-logging applications. A common standard for the integration of automated test systems, developed by Hewlett-Packard, is referred to as the IEEE-488 interface bus.

The bus consists of 24 lines, accommodated within standard stacked type connectors. The eight bidirectional data lines carry information as 7-bit ASCII codes between the microcomputer (controller), and an instrument (listener) on the bus. The roles may be reversed when data are being logged. To process the information on the data bus, up to eight control and status signals are available.

The bus is designed to interface with up to 15 instruments within a localized area, involving a total cable length of not more than 20 m. Each instrument is uniquely numbered within the range of 0–30 and the overall activity is controlled by one of the devices, termed the *Controller*. This is usually the microcomputer with an appropriate interface. Each device number is switch selectable within the instrument. Other functional aspects of the devices on the bus are that they must be capable of acting as a *Listener* or a *Talker*. A Listener is a device that can receive data over the bus, and a Talker is capable of transmitting data. There may be several Listeners active on the bus at any one time, but there can only be one Talker. Most devices, including the microcomputer Controller, can act as either Listeners or Talkers.

When setting up an instrument to measure some physical variable, codes devised by the instrument manufacturer are sent on the bus, in ASCII format, as a data

string to the numbered device. In the case of a multichannel DVM, this could take the form of the channel number to be monitored, voltage range to be selected and a terminating character. An example of the corresponding string to be put on the bus is

C9R2T

which denotes channel 9, range number 2 (say, 0–10 V) and "T" is the string terminating character recognized by the instrument.

Manufacturers of add-on cards, to give IEEE-488 facilities with microcomputers, usually supply software for initializing the bus, setting it up for transmitting data from controller to instrument and returning data from instrument to controller. The measured quantity is also sent to the computer in the form of an ASCII string from which the actual numerical value can be extracted.

One of the most important management control lines is the service request (SRQ). This is a type of interrupt line that is activated low by a device residing on the bus and needing service from the controller. It is used as a means of indicating that the instrument is ready to transmit the current reading onto the bus.

Thus, a typical software sequence for implementing the control of an instrument on the IEEE-488 bus for data acquisition is:

1. Initialize bus and set instrument as a Listener.

2. Put control string on the bus to set up the instrument as required.

3. Check for SRQ line to go low indicating that data can be read.

4. Set instrument as a Talker.

5. Read returned string and convert into a numerical value.

When operating in high-level BASIC, high data-collection rates are not possible. However, since most instrument manufacturers offer the standard as an option it provides an intelligently controlled flexible arrangement for test and measuring instruments.

PART 4

INTERFACING OF COMPUTERS TO SYSTEMS

Charles J. Fraser

John S. Milne

The serial and parallel communication standards are the basic interfacing links between computers and their associated peripheral devices, or between computers and a comprehensive range of measurement instrumentation. In general computer control applications, however, two other common interfaces are fundamental. These are the digital interface, which implements the controller output and the analogue interface associated with the measured variable input.

6.10 DIGITAL INTERFACING

The computer output port may be used to transmit control signals on any one of the available lines by writing the appropriate number to the port address. When a line (or bit) is set "high" (i.e., a logic level of 1) the voltage on the line is approximately 5 V. The current available, however, is fairly minimal (on the order of 1 mA) and no load can be connected directly to the port. There is in fact a danger of causing extensive damage to the computer by connecting a load directly to the port. An interface must therefore be provided to enable the computer to switch in power loads using the logic level control signals from the output port. The most common interface device used for this purpose is the power transistor. In typical applications the power transistor, operating on logic level control signals, switches in a mechanical relay, which in turn switches in the load.

The Darlington Driver is a popular power transistor available as an integrated circuit. It normally includes a number of separate stages. Figure 6.6 shows the wiring diagram for a single stage in a Darlington Driver. The Darlington Driver can switch up to 500 mA at 50 V, and each stage is diode protected for the switching of inductive loads.

To provide a total isolation from high voltages, port output signals can be coupled through an opto-isolator, which, interfaced between the computer output port and the power control device, is not an essential element in the digital interface. Opto-isolation, however, ensures that no hardwired connections are made between the computer and the power device. An added advantage is that the opto-isolator acts as a buffer to spurious noise signals which can corrupt the digital logic values being transmitted on the buses.

The opto-isolator (Figure 6.7) transmits signals by means of infrared radiation, emitting from a source and sensed at a photo-transistor. When a computer based on transistor/transistor logic (TTL) is powered up, the state of the lines of the output port float high. That is, each output line becomes set to a logic value of 1. Since a logic 1 is normally associated with the function of switching a device ON, a port which floats high could inadvertently activate some power device. Obviously this is a dangerous precedent, which requires an additional element in the digital

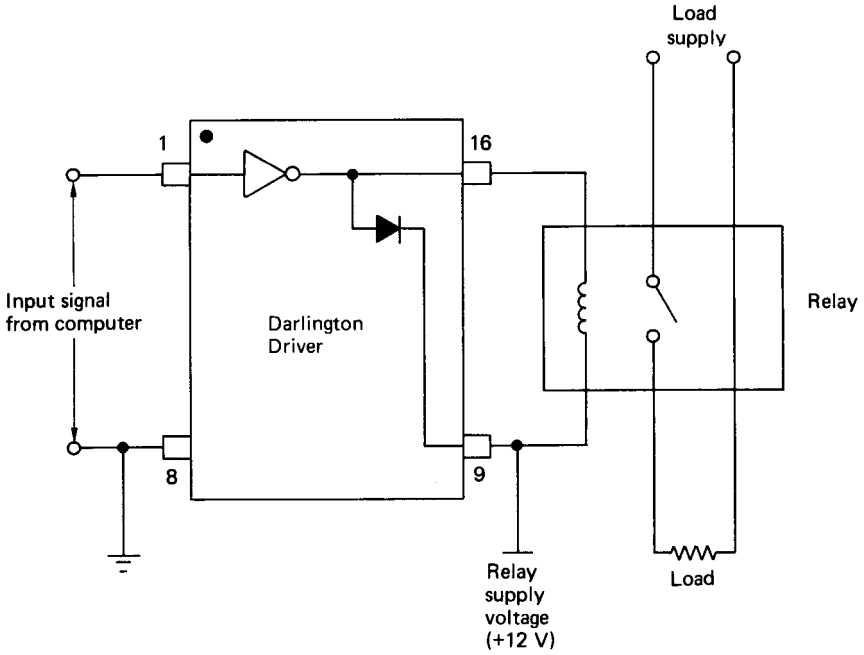


FIGURE 6.6 Power switching with a Darlington Drive.

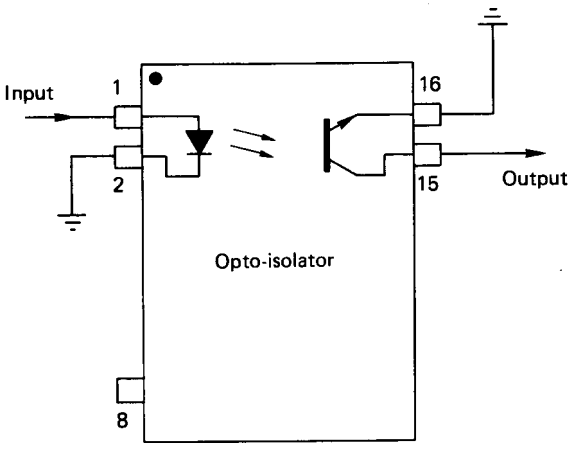


FIGURE 6.7 Opto-isolator.

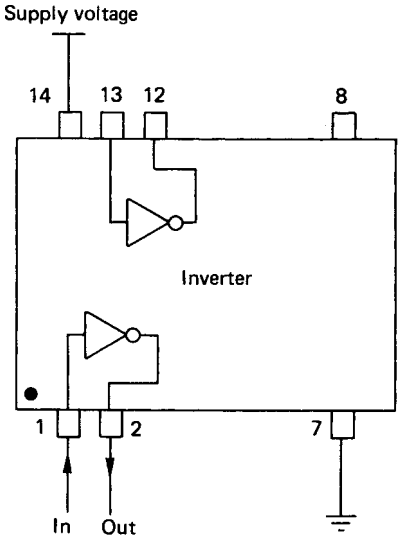


FIGURE 6.8 Inverter integrated circuit.

interface to counteract the effect. The device commonly used is the *inverter* (or NOT gate), which has the simple function of inverting all logic signals from 1 to 0 and vice versa (Figure 6.8). Following power up, a logic 0 must then be sent to the relevant line of the output port, to become a logic 1 after inversion and to operate the control function.

The composite digital interface for a computer output port suitable for power switching is depicted in Figure 6.9. In high-frequency switching applications electromechanical relays are not suitable. Semiconductor devices such as silicon controller rectifiers (SRCs, alternatively called thyristors or triacs) may be more appropriate. Also particularly suitable are the various solid-state relays which can operate directly from logic level signals.

6.11 CONTROLLER OUTPUT INTERFACE HARDWARE

The digital interfaces discussed above are suitable for switching in power loads in an ON/OFF control system. For a digital control algorithm based on a PID strategy, some means is required of discretely varying the output power supplied to the controlled device (Figure 6.10). A number of different methods are used to supply variable power to the system and these include the following.

The Digital-to-Analog Converter (DAC)

The required control effort value, U, is calculated in the program according to the control strategy employed. This value is converted to an equivalent binary number

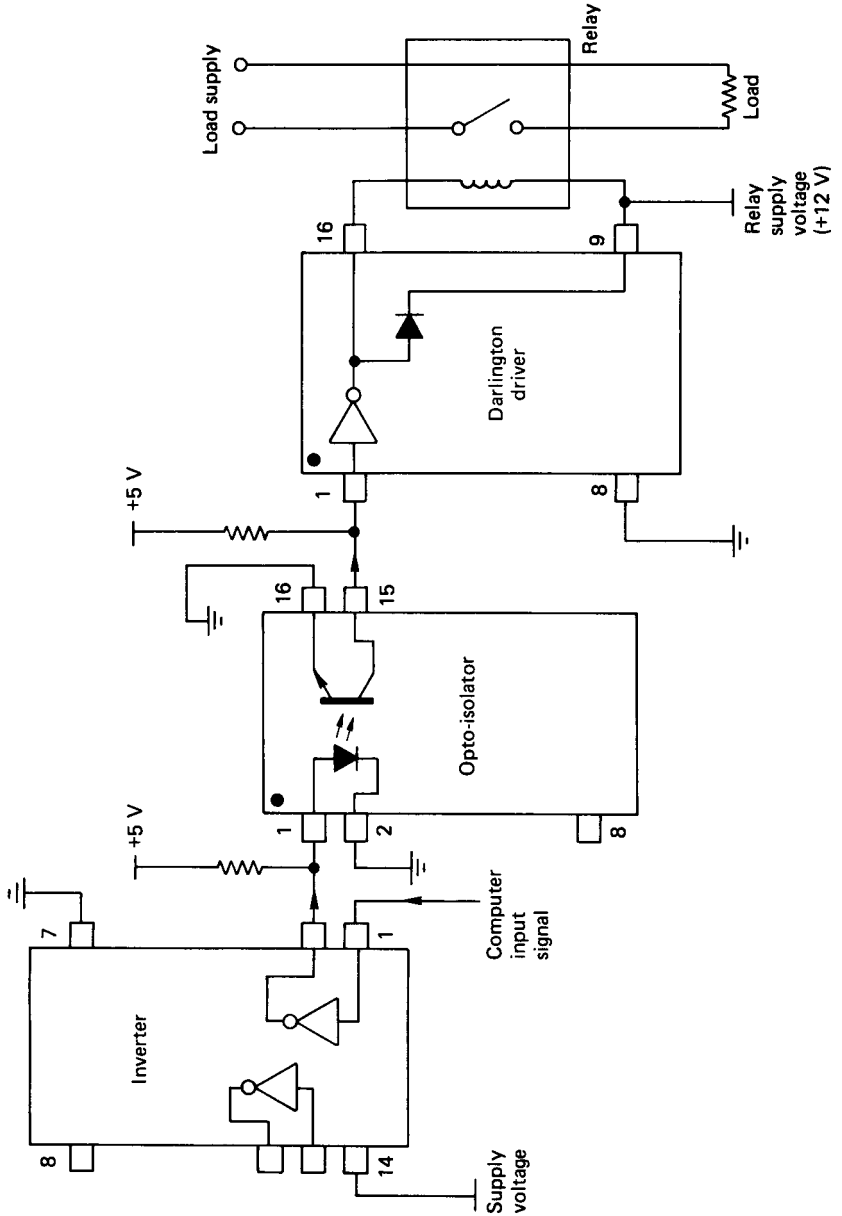


FIGURE 6.9 Output port digital interface for power switching.

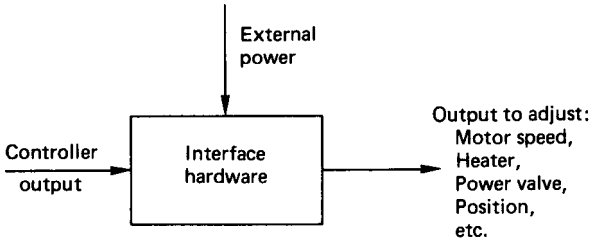


FIGURE 6.10 Variable power output device.

and output to the computer port where it is then transmitted to a DAC. The DAC converts the binary input into a proportional output voltage which may then be suitably amplified to drive the controlled device. The controlled device could, for example, be a d.c. motor whose speed is directly related to the supply voltage. The interface is illustrated in Figure 6.11.

Two basic types of DAC are available: the adder converter and the ladder converter. The adder converter can be illustrated as a simple example of Ohm's law. A 4-bit adder type DAC is shown in Figure 6.12. The resistance value of the line resistors are halved for each consecutive increasing bit and the supply voltage is common. The current drawn through each line, if it is connected, is 1, 2, 4, and 8 mA, corresponding with bits 0, 1, 2, and 3. The summation of the currents at output therefore equates on a decimal to binary basis with the input. Figure 6.12 shows a digital or switched input of 1011 binary. This gives a current summation of 11 mA, equivalent to the corresponding decimal number. The output, a proportional current, can be converted to a proportional voltage through an operational amplifier. Because of the range of resistor values required, the adder converter is less popular than the ladder converter, which uses only two resistor values, R and $2R$.

The ladder converter DAC (Figure 6.13) must be analyzed using the network theorems of Thevenin and Norton. The end result is similar, however, with a pro-

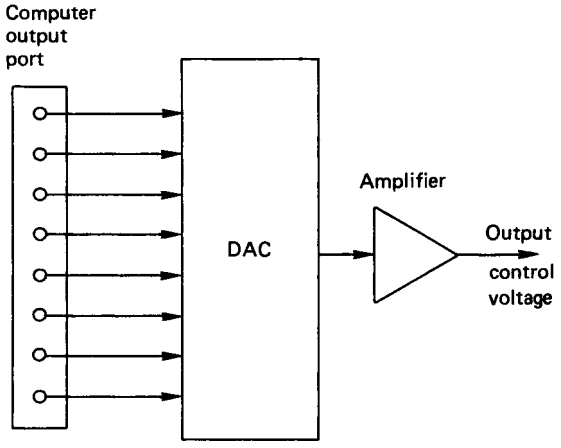


FIGURE 6.11 Variable power output using a DAC.

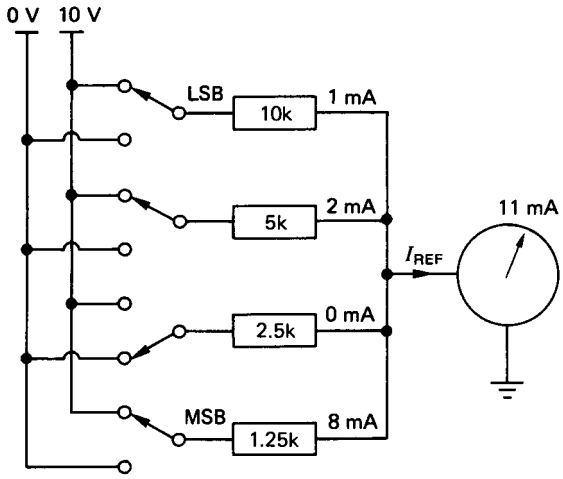


FIGURE 6.12 4-bit, adder-type DAC.

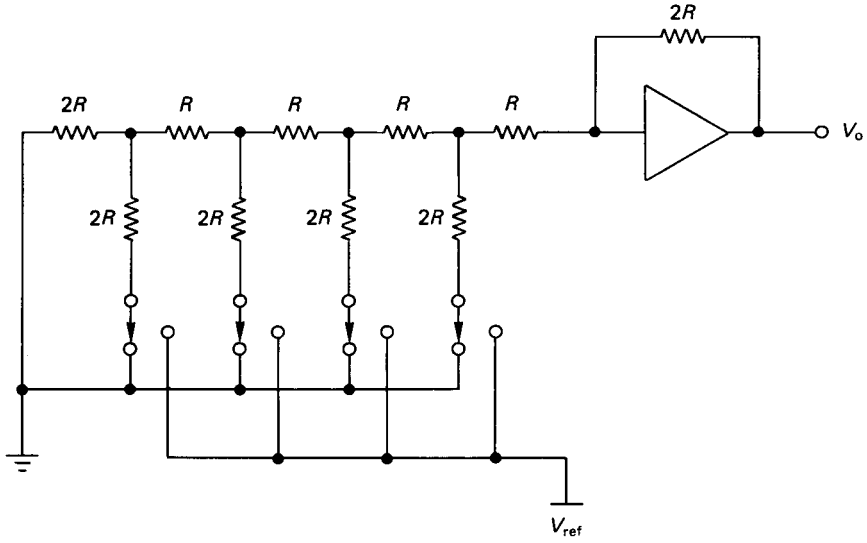


FIGURE 6.13 Ladder-type DAC.

portional voltage output corresponding to the digital switched input on a decimal-to-binary basis.

Pulse Width Modulation (PWM)

Output power from the interface hardware can be varied by sending ON/OFF pulses to the power device. The frequency range is normally between 2 and 10 kHz. If

the time ON and the time OFF periods are equal, then rapid switching of the power supply will transmit 50% of the total power available. Due to the relatively high switching frequencies and the levels of power transmitted, the solid-state relays are more suitable for this application. The time ON is called the MARK and the time OFF is the SPACE. The MARK/SPACE ratio can be evaluated from the controller output value calculated in the software. PWM can be achieved either by keeping the MARK fixed and varying the SPACE (i.e., varying the signal frequency) or by varying both the MARK and the SPACE within a constant period.

If the PWM output is supplied by the controller, then it is not possible to allow the MARK+SPACE period to extend over the complete control loop cycle. There must be some time allowed to sample the process transducer and to calculate the required controller output value. The problem is overcome by using a synchronization (sync) pulse signal with a longer period than the MARK+SPACE time. The excess time during the sync pulse is used to read the transducer output and to process the information. The MARK/SPACE ratio is then implemented over the remainder of the sync pulse. Two separate timing loops are used to control the MARK and SPACE times, respectively. Due to the relatively high frequency of the output signal, the control software must be written in assembly language for speed. An alternative to this is to use hardware support chips to output the PWM signal under the control of the computer's CPU.

Controlling a.c. Power by Control of Thyristor Phase Angle

Various applications, such as temperature control, require a.c. power adjustment and solid-state relays can be used effectively to vary output power between 0% and 100%. The power is controlled by varying the phase angle between the supply voltage and that which appears across the load when current conduction begins. A phase control device operating off a 0–5 V signal from a DAC can be used to alter the phase angle between voltage and current in the range 0–180°. The control of the power output to the load is nonlinear, but linearization between output and input can be accomplished in the software.

Controlling Flow Control Valves

In level-control systems, fluid flow rates are controlled by varying the degree of opening of a gate-type valve. Because of the forces involved, a pneumatic actuator is normally employed, working off a controlled pressure in the range between 3–15 pounds per square inch (psi). (*Note:* The U.S. aircraft industry still favors the Imperial unitary system. The approximately equivalent pressures in the metric system are 20–100 kN/m^2.) An applied pressure of 3 psi is equivalent to the valve being fully open, 15 psi corresponds to the valve being fully closed.

Most flow control valves are fitted with a positioner, which operates off a current signal in the range 4–20 mA. The current range generally corresponds on a linearly proportional basis to the pressure range. Since the computer interface usually involves a DAC, then an additional element (a voltage-to-current converter) is required to interface between the DAC and the valve positioner. The complete interface is shown in Figure 6.14.

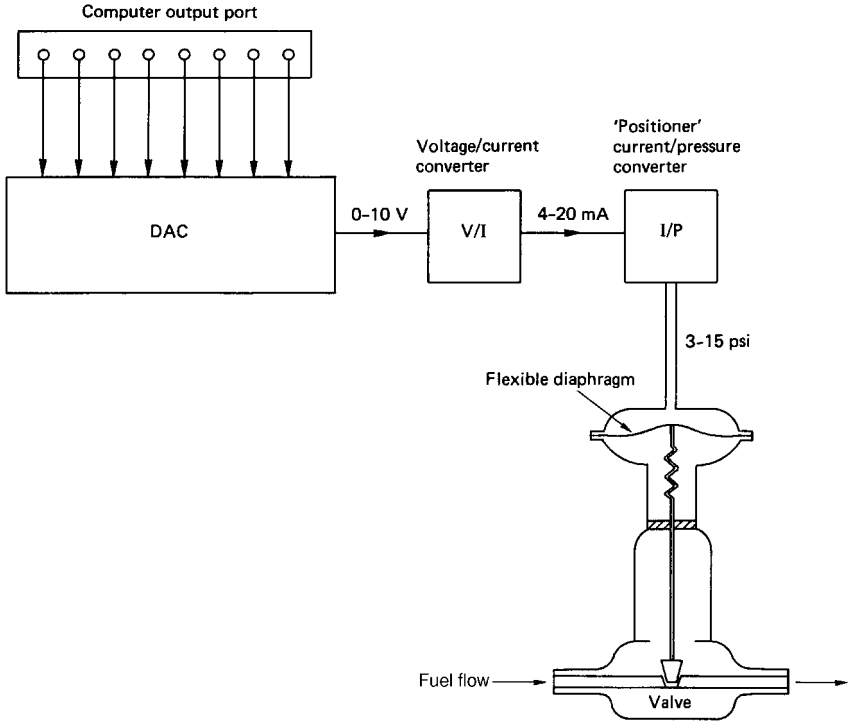


FIGURE 6.14 Digital interface for a flow control valve.

6.12 ANALOG INTERFACING

The basic role of the analog interface is one of conversion of the continuous analog signals, from process measuring transducers to the digital representation that the computer requires to operate on. In all practical applications, the monitoring and acquisition of the data is the necessary precursor to the subsequent control functions that might be actioned.

The process variables are ultimately represented as voltages. Using the appropriate signal-conditioning circuits, these voltages would ideally be processed to range between zero and some reference value. The final task is the digitization of the analog signal, which is accomplished through an analog-to-digital converter (ADC). The ADC samples the analog signal, performs the conversion, and outputs a digitally encoded binary number which is directly proportional to the magnitude of the input voltage. The essential elements in the signal train are shown in Figure 6.15.

This figure indicates a sample and hold (S/H) element between the signal conditioner and the ADC. Since the analog input may be varying while the conversion is taking place, there is a degree of uncertainty in deciding the instant in time which the output code represents. The sample and hold element removes this uncertainty

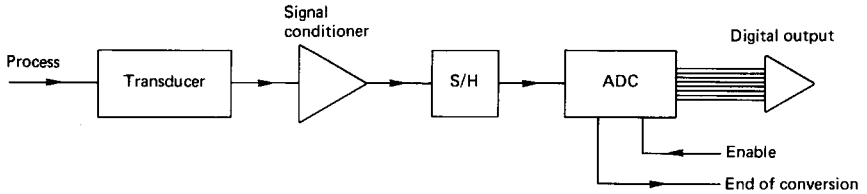


FIGURE 6.15 Analog-to-digital conversion.

by capturing the instantaneous snapshot of the input for the ADC to convert before moving on to the next sample. The S/H element is only essential if the input signal is varying very rapidly. The ADC and S/H functions are often packaged within a composite integrated circuit.

ADCs are available in a number of different forms and these include the following.

Staircase and Comparator

The staircase and comparator is the simplest form of ADC (Figure 6.16). The device incorporates a DAC which generates a voltage increasing in small steps as shown. At each step the staircase input is compared with the analog input. When the generated staircase is approximately equal to the input, the process is halted and a binary count is made of the number of steps taken during the process. The binary count from zero represents the coded digital output.

The staircase and comparator ADCs have relatively slow conversion times (typically 20 ms). They are, however, cheap and are essentially immune to electronic noise.

Integrating Type ADC (or Dual Slope)

The major elements comprising a dual-slope ADC are illustrated in Figure 6.17. At the start of conversion a voltage-to-current converter is switched to the integrator, causing it to ramp up a slope which is proportional to V_{in}. This occurs over a fixed period of time, at the end of which the input is switched over to the reference current source. At the instant of switching the integrator output voltage is propor-

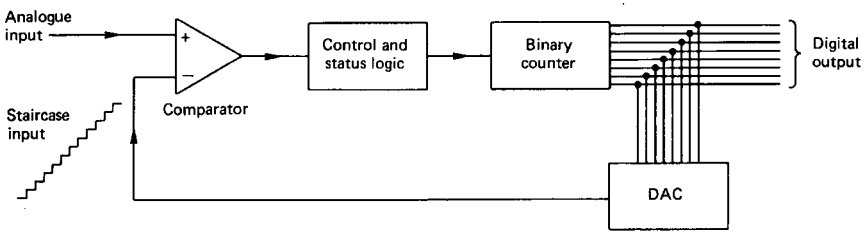


FIGURE 6.16 Staircase and comparator-type ADC.

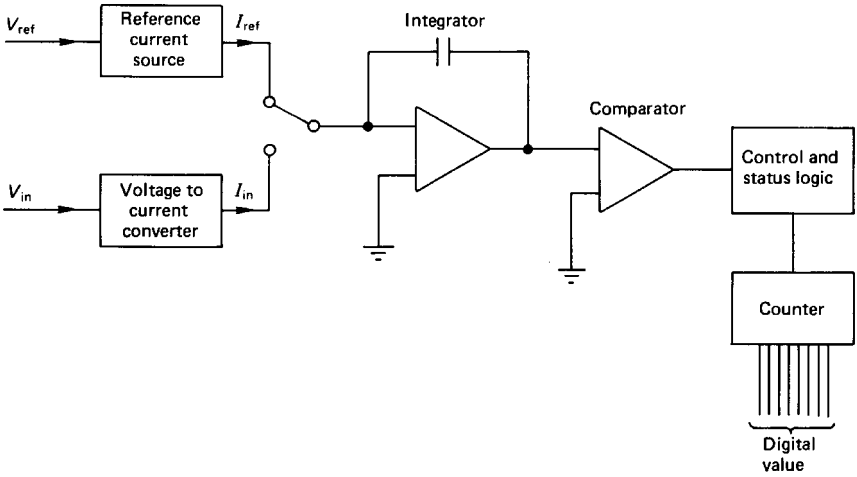


FIGURE 6.17 Dual-slope ADC.

tional to V_{in}, a counter is enabled and counting begins at a rate set by the internal clock. Meanwhile, the reference current causes the integrator to ramp down at a slope which is proportional to V_{ref} (i.e., a constant slope). When the integrator output again reaches ground the comparator switches the counter off and the counter then contains a digitally encoded value proportional to V_{in}. Figure 6.18 shows the voltage variation at the integrator output.

From this figure it can be seen that there are two similar triangles such that

$$V_{in} = V_{ref}(T_v - T_f)/(T_{max} - T_f) \tag{6.2}$$

T_v is directly proportional to the counter output, and with T_{max}, T_v, and V_{ref} all known, the input voltage, V_{in}, is determined by proportion.

The integrating types of ADC have similar operating characteristics and conversion times to that of the staircase and comparator types. For faster analog-to-digital conversion, the successive or counter type are generally employed.

Successive Approximation Type ADC

In this ADC the input signal is compared with a number of standard reference voltages, generated from a DAC, until the combination of standard voltages required to make up the input value has been determined. The main components of the converter are a clock, a counter, a comparator, and a DAC.

When an analog signal is input to the converter, the counter starts a count and passes a digital value to the DAC. The DAC generates a voltage to represent the most significant bit and the comparator assesses this against the analog input. If the analog signal is greater than the voltage from the DAC then the logic 1 in the MSB is retained. If the analog signal is smaller, then a logic 0 is assigned to the MSB. This process is then repeated on the next most significant bit and so on for all the other bits down to the LSB. The conversion time for these types of converters

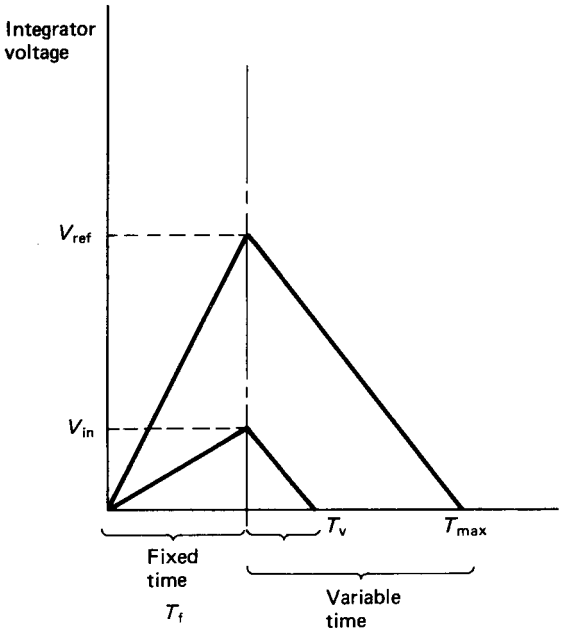


FIGURE 6.18 Integrator voltage variation.

may be of the order of 10–25 μs, but this will depend upon the hardware design. Figure 6.19 outlines the essential features of a successive approximation ADC.

Parallel Conversion Type ADC

The parallel type ADC has the fastest conversion time but it is also the most expensive. With parallel conversion, the analog input is fed simultaneously to a

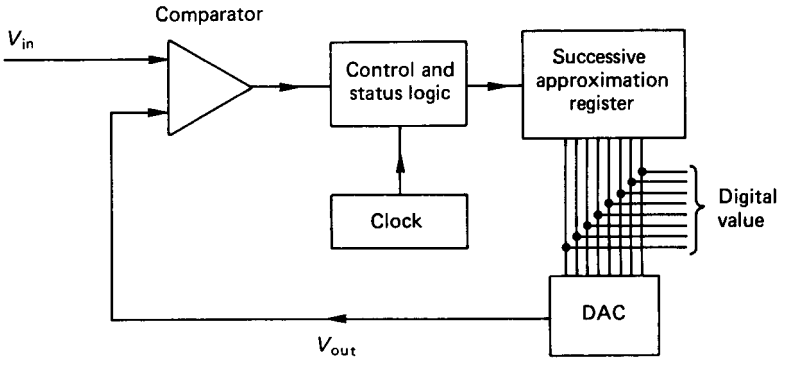


FIGURE 6.19 Successive approximation ADC.

number of comparator circuits, each having a different reference voltage. The resulting comparator outputs are fed to a logical coding network which generates the appropriate digital values to represent the state of the comparator outputs.

Regardless of the type of ADC used, the pin functions on the integrated circuit are basically similar and generally comprise the power supply, the data bits, the start conversion pin (SC or $\overline{\text{CONVERT}}$) and the end of conversion pin (EOC or STATUS). The overbar signifies that the pin is active low.

The conversion is software initiated by sending a "pulse" (logic 0, followed by logic 1) to the $\overline{\text{CONVERT}}$ pin. On the negative edge of this pulse the counter in the successive approximation ADC is set to zero and on the positive edge the counter starts incrementing. At the start of conversion the STATUS pin goes from low to high, and when it again goes low, the conversion is complete (Figure 6.20).

The end of conversion may be readily detected using suitable software. As an alternative, it is possible to include a software-generated time delay following the start conversion pulse to allow conversion to complete before reading the value at the input port. The length of the delay can generally be found by trial and error.

In choosing the appropriate ADC for a particular application the four main features to be considered are:

1. *Conversion time:* The conversion time is a measure of the operating speed of the converter and is the time taken for the complete translation of an analog signal to digital form. In many of the staircase and comparator and the integrating types of ADC this may be dependent on the level of the analog input signal. Faster conversion times are obtained with low-level inputs due to the manner in which the conversion is completed. Successive approximation and parallel conversion types of ADC have a fixed conversion time. This is because the same conversion process is performed, regardless of the analog input level.

   The conversion time of the ADC does not, however, indicate the fastest rate at which data can be captured. If the data are to be stored in the computer's RAM, then this must be done in a sequential and ordered manner. This involves setting a base address and incrementing various registers to step the storage addresses of each byte of data placed in memory. Further time delays could be accrued in a sample and hold device. The minimum data capture period is often therefore many times greater than the specified conversion time of the ADC.
2. *Resolution:* The resolution of an ADC is the number employed to represent the digital output in binary form. For example, the resolution of an 8-bit ADC is

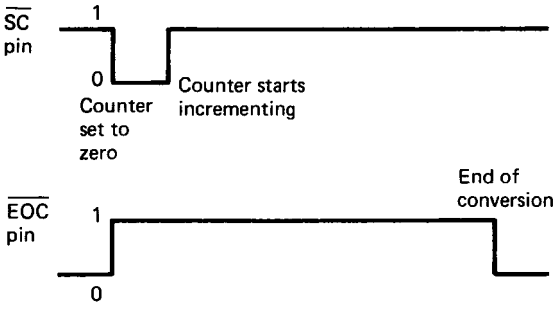


FIGURE 6.20 Start conversion and end of conversion pin signals.

TABLE 6.1

n-bits	2^n	Resolution (%)
8	256	0.4
10	1 024	0.1
12	4 096	0.025
16	65 536	0.0015

limited to one part in 256 of the maximum voltage corresponding to the full-scale setting. An improvement in resolution can be obtained with a 12-bit converter, with one part in 4096. Table 6.1 summarizes the relation between the number of bits and the resolution.

3. *Accuracy:* The accuracy is related to linearity defects, zero error, and calibration deficiencies in the electronics of the converter and should not be confused with the resolution.
4. *Cost:* Cost will depend on the quality required in the three areas previously described and on the means of conversion employed. It is closely associated with the speed of the conversion and with the resolution and accuracy. Cost generally rises with increases in all or either of the three other variables.

6.13 MULTIPLEXING

In applications where a number of transducers are to be sampled, a multiplexer (MUX) can be used to switch in various channels as and when required to a single

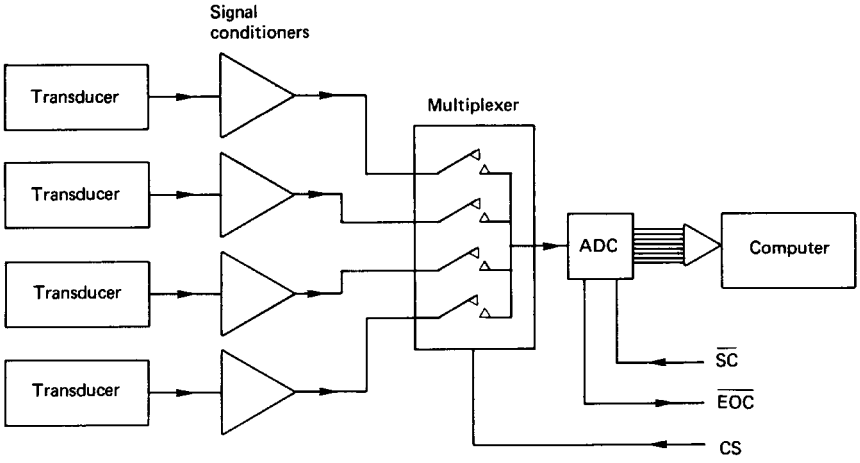


FIGURE 6.21 Multiplexer for multiple inputs.

ADC. The switching is software controlled from the computer. Figure 6.21 illustrates the basic principle.

The multiplexer and ADC often form an integral part of a complete system. In some cases the signal conditioning can be software controlled, with all the necessary hardware mounted on a single card or chip and plugged directly into the computer's bus system. Multiplexers (or analog switches) are available with various numbers of input channels.

Minimum cost conditions usually dictate whether multiplexing will be implemented or not, but the reduced cost must be balanced against an inevitable reduction in sampling rate. Figure 6.22 shows three possible arrangements of signal conditioning, multiplexing, and conversion for analog interfaces.

System A is the most common, while B and C can provide for virtually simultaneous sampling. System C gives the most representative snapshot at a particular period in time, but it is also the most costly.

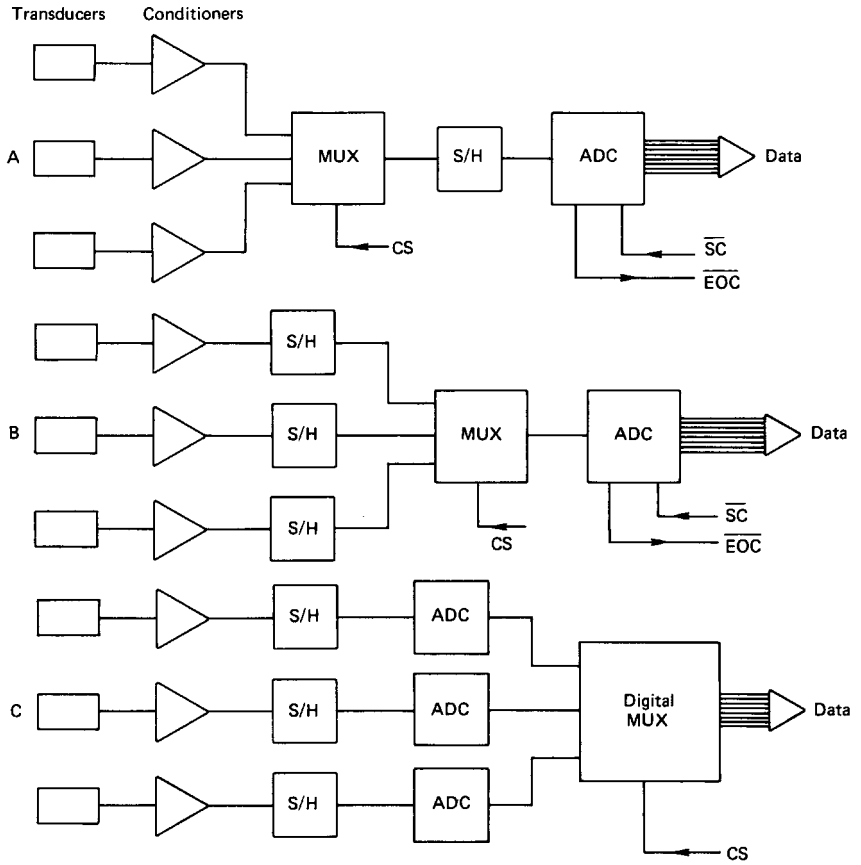


FIGURE 6.22 Multi-input systems.

6.14 MACHINE TOOL INTERFACES

The control system for a machine tool slide is shown in Figure 6.23. Typically, there are two negative feedback loops, one for position and one for velocity in a cascade arrangement as shown in the figure. The position sensor is usually an optical grating device (or an inductosyn) and the speed sensor a tachometer.

The CNC interface initially has to decode the manual or control tape input data. This consists of a sequence of commands, including feed and speed data, essential dimensional reference points, and other constraints to be observed by the machine during its operation. In operation, the interface is required to monitor the slide position and speed and check various limit switch settings for compliance with the sequential program instructions. The transducer input signals would normally be switched in through a multiplexer prior to digitization with a fast conversion type ADC. Limit switches would also be checked or set through additional digital I/O lines. If any errors are detected, the interface must be able to indicate these and take the appropriate action. The interface includes a real-time clock which generates an interrupt every few milliseconds. The clock acts as a monitor of operator actions, enables the output of error signals to the machine servos, and checks all current signals from each of the feedback sensors.

For a typical CNC milling machine there are three independent axes, and each would have the same monitoring and control functions applied to them. In addition, the spindle speed would be monitored and controlled and the machine might also incorporate a tool-changing facility based on a simplified robot arm.

Further refinements could include a load transducer in an additional feedback loop to measure the cutting forces during machining. Force sensing may be used as the basis for an adaptive control loop. In the context of machine tools, adaptive

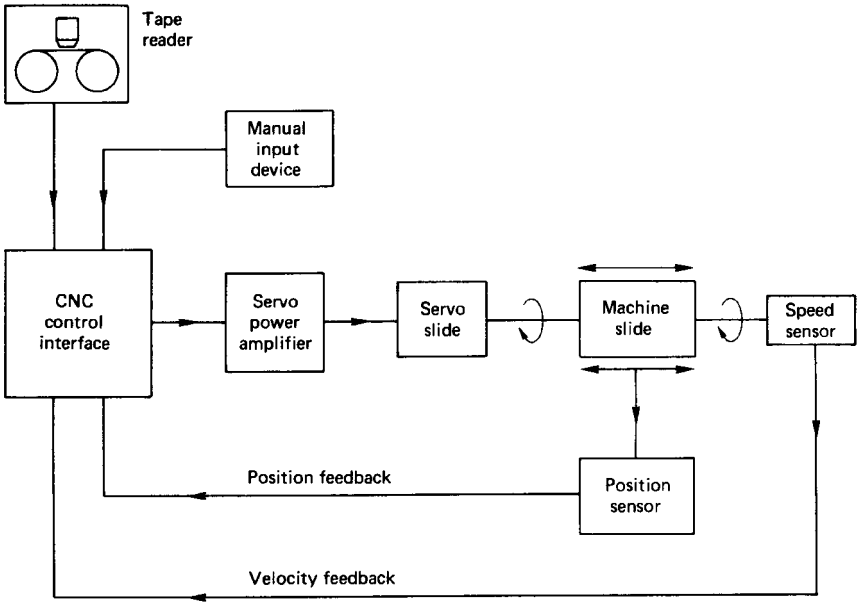


FIGURE 6.23 Machine tool slide control interface.

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

control is usually associated with the alteration of feed rates and cutting speeds to maximize the cutting power. Figure 6.24 shows an adaptive control option on an NC turning machine.

The adaptive loop can optimize the cutting operations, prevent spindle overload, maximize tool life, reduce time loss in "air cuts," and simplify the programming. The additional sensors and their protection in the harsh machining environment means, however, that the adaptive loop is much more costly to implement. The adaptive control interface, which has no manual input data facility, is also necessarily complex and requires considerable memory capacity.

6.15 ROBOT CONTROL INTERFACES

The machine tool interface described in the previous section can be programmed to perform a series of operations which might be described as *sequenced automa-*

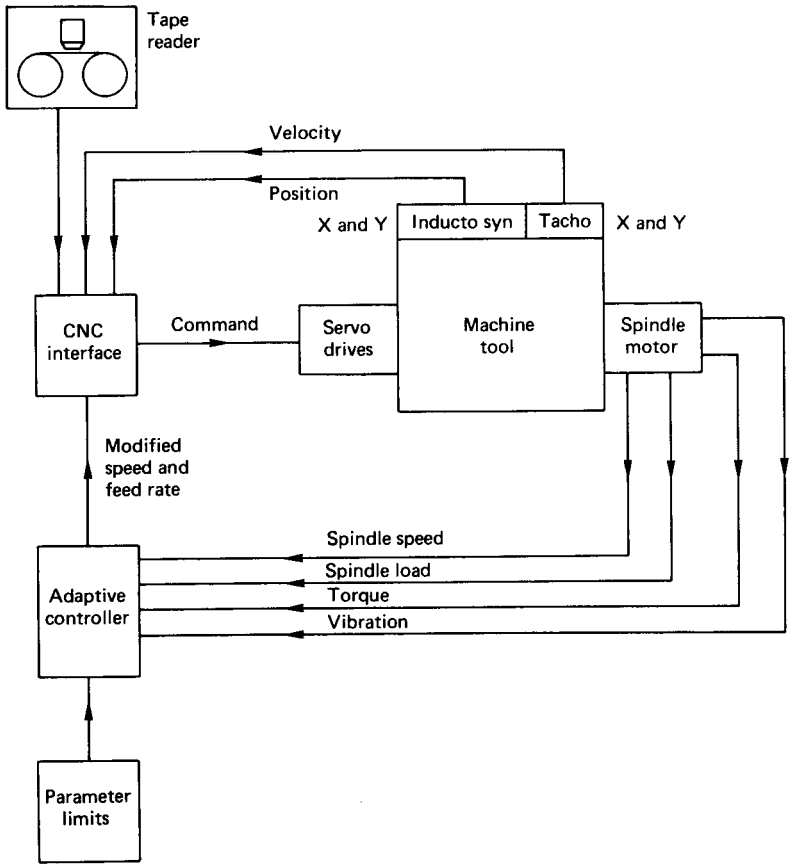


FIGURE 6.24 Adaptive control on an NC turning machine.

tion. Many of the simpler robots (e.g., pick-and-place machines) use the same technology and perform essentially similar tasks. These machines are not, however, robots in the strictest sense. The essential feature of a true robot is its capability of exercising independent control in each of its axes, or rotating joints, such that its ''hand'' can reach any position and any orientation within the working volume.

Each joint on the robot has an actuator, an associated position sensor, and a velocity sensor. Six actuators are required for full flexibility in position and orientation, although in most cases only five or less are used. The computer must at all times be able to ascertain the current and desired locations of the hand. The position sensor data processing therefore involves the manipulation of various coordinate transformation matrices in the definition and control of the hand location.

Force sensing within the gripper is commonly featured in robot arms for space applications. The obvious application is to prevent the proverbial vice-like grip on some fragile object. Force sensing is also incorporated by inserting a sensor between the gripper and the wrist. Many wrist force sensors are strain gauge based and can respond to applied forces and moments in each of the three Cartesian directions. The wrist sensor provides feedback signals for static and inertial loading on the robot arm. Robots which operate a control strategy based on wrist force sensor data are termed *active compliant*. The active compliant robot system has the capability of modifying the joint positions and motion to minimize the effect of external forces. Many wrist force sensors have been successfully developed for robot manipulators engaged in pin insertion type operations. Whitney (1982) reports on some detailed investigations on wrist force sensor applications.

The sensors producing signals related to position, velocity, and force form the main feedback loops in many industrial robot assemblies. These signals are converted through ADCs and the digital control algorithm, often based on a PID strategy, is then solved to determine the required controller output values at that given time. Power is subsequently output to the actuators either through DACs or using pulse width modulation techniques. In most respects the above process is exactly the same as that which would be used to control a CNC machine tool with an adaptive loop. The robot arm motion, however, is much more complex and the computation involved is far more extensive than that required for the CNC machine tool. The computer used in robotic manipulators therefore requires a large memory capacity and a fast processing speed to accommodate the multiplicity of tasks associated with the control functions. The control algorithm is itself complicated by the multiple feedback loops for position, velocity and force, and a hybrid control strategy is normally adopted. This strategy allows force to be controlled along certain Cartesian axes while position is controlled along the remaining degrees of freedom. Velocity is controlled along the same degrees of freedom as position. The computational effort is significant due to the many coordinate transformations required, and hybrid control can only be implemented when the robot is moving slowly.

Recent advances in robot technology are centered on the development of tactile sensors, range finders, and machine vision. Tactile sensors ideally simulate the human sense of touch. They should therefore be able to detect presence, shape, location, and orientation. In addition, the ideal tactile sensor should be able to respond to applied forces and moments in terms of magnitude, direction, and location. Varying degrees of success have been achieved with tactile sensors (see Bejczy 1977; Harmon 1982; and Rebman and Trull 1983).

Many range-finding devices are based on optical proximity sensors or on ultrasonic echo-sounding transducers in applications to short- and long-range finding,

respectively. The function of the rangefinder is twofold: either to locate the gripper in the correct position relative to the object to be picked up or to avoid potential collisions with other objects. Like tactile sensors, however, many range-finding devices have not yet been developed to the stage where they are consistently reliable for routine use.

Perhaps the most useful additional sensory attribute that a robot can have is vision. Machine vision is currently the subject of considerable worldwide research and development and a number of effective systems are already available.

The first problem with a machine vision system is the formation of a digital image. There are a number of different digital imaging devices in use, but they all have the common function of converting light energy into voltage (see Ballard and Brown 1983). The voltage output is inevitably processed through a fast ADC to produce the digital representation of the light intensity. The resolution of the digital image depends on the number of discrete lines and points per line which are used in the construction of the picture. The use of 512 lines, with 512 discrete points in each line, results in 262,144 individual picture cells (pixels) to define the image. Each pixel is represented as an 8-bit number and these can be stored in memory as a two-dimensional array. Note that this would use up 262,144 bytes of computer memory for the storage. The preprocessor interprets the visual information and then transmits the result to the robot controller.

Since the image can be represented numerically and stored in memory, it can also be retrieved and processed according to the particular requirements. In many robot vision applications the vision task is that of part recognition and orientation of the part. Often a two-dimensional projection of the part shape is all that is required, and a single solid-state camera can be used to generate the two-dimensional image. Part recognition is based on the part silhouette, and the image-processing technique requires only a brightness threshold level to distinguish between object and background. The level of vision is crude but can be sufficient for many automated assembly processes. Control of the lighting arrangements can greatly enhance the vision system, but, in general, machine vision still leaves much to be desired at its current stage of development.

The three major development areas of tactile sensing, range finding and machine vision are all being actively researched at present. It is apparent that these sensory functions all involve considerable amounts of additional input data and this imposes further limitations on the system in general. The limitations are related to the data-processing speed and the communication interfaces.

6.16 SIGNAL CONDITIONING

Although the output signal from many sensors and transducers is already in the necessary voltage form, only very rarely is the voltage high enough to be directly interfaced to a measurement system. In most instances the transducer output will require some form of amplification.

Amplifier is a shortened form of the full description *voltage amplifier*. Most amplifiers amplify voltage, but other types are encountered and these are normally given their full description (e.g., current amplifier, charge amplifier, and power amplifier). Amplifiers are further classified as a.c. or d.c. The d.c. amplifier will accept a.c. inputs, but the a.c. amplifier will block any d.c. input. The input signal may be either single-ended with one, or differential with two active signal lines.

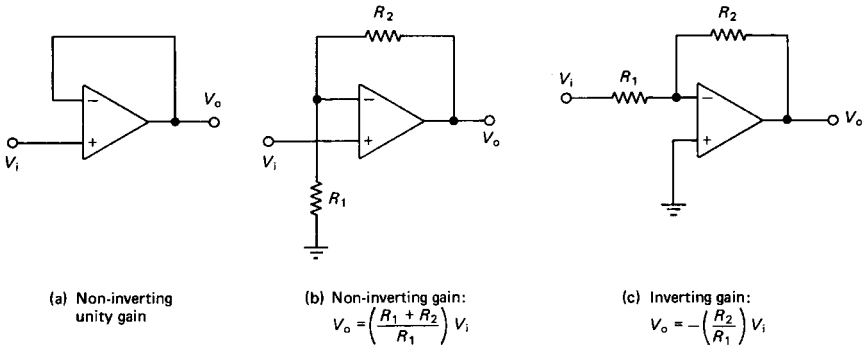


FIGURE 6.25 Operational amplifier circuits.

Finally, the amplifier may be *inverting,* with a reversal of sign at the output, or *non-inverting*.

The most common amplifier configuration embodies a differential input, single-ended output voltage amplifier with either inverted or noninverted output. This type is referred to as *operational amplifiers.* Figure 6.25 illustrates some typical idealized operational amplifier circuits.

The important characteristics of an operational amplifier are its gain and its bandwidth. The gain (which can be as high as 10^6) is related to the combination of passive resistors which make up the external circuit. The gain is constant only over a restricted range of input signal frequencies. Outside this range the gain is attenuated as shown in Figure 6.26.

The 3 dB cut-off frequencies are used to define the bandwidth of the amplifier. The product of gain and bandwidth (GBW) is quoted by some manufacturers as a quality parameter. Other descriptive parameters associated with amplifiers are:

1. *Common-mode rejection ratio:* Ability of an amplifier to reject differential input gain variation
2. *Offset voltage:* Voltage output attributed to input and output voltages generated by component variations
3. *Offset drift:* Temperature-dependent voltage output
4. *Nonlinearity:* Departures from linear input/output characteristics
5. *Distortion:* Frequency-dependent nonlinearities

For high-accuracy measurements, where low drift and low noise are also essential, an instrumentation amplifier with a high input impedance and high common-mode rejection ratio would normally be required (Figure 6.27).

The problem of impedance mismatch is frequently encountered in instrumentation systems. This occurs when a transducer having an internal resistance and functioning as a small voltage source is coupled to a recording instrument which also has an internal resistance. If the internal resistance of the transducer is many times smaller than that of the recording instrument, then the problem is insignificant. If, on the other hand, the internal resistances are of similar orders of magnitude, then the error in the measurement and the loss of the signal can be quite large. The standard solution to the impedance mismatch problem is to interface an amplifier

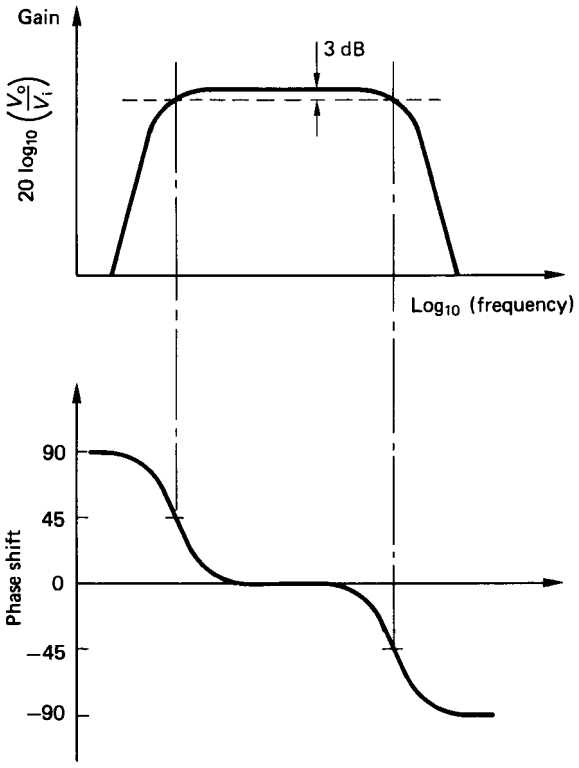


FIGURE 6.26 Gain and phase shift characteristics of an operational amplifier.

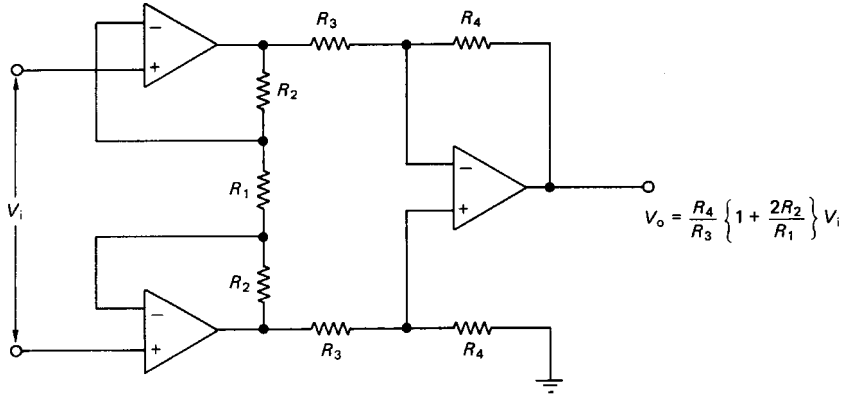


FIGURE 6.27 Instrumentation amplifier.

with a high-input impedance and a low-output impedance between the transducer and the recording equipment. Such amplifiers are usually configured to have unity gain since their primary function is the prevention of serious signal loss.

For many measurement applications, custom-built special-purpose amplifiers are available in integrated circuit form. These greatly ease and facilitate the signal-conditioning circuitry. Fraser and Milne (1990) give a practical guide to many of the signal-conditioning interfaces required for the amplification of signals from strain bridges, thermocouples, and other low-level voltage output sensors.

6.17 ANALOG AND DIGITAL FILTERING

Noise is inherently present in all physical systems where measurements are made. In sampled data systems (i.e., when an ADC is employed) the effect of noise, illustrated in Figure 6.28, can give rise to further misinterpretation in the form of aliases.

Discrete sampling (shown in Figure 6.28) results in an output signal which suggests that the measured variable is increasing linearly but with a superimposed sinusoidal fluctuation. The apparent sinusoidal variation is entirely the effect of background noise, and it is obviously good practice to try to eliminate noise in the measurement system. It is perhaps fortuitous in mechanical systems that background noise is generally manifested at much higher frequencies than that associated with the primary variable of interest.

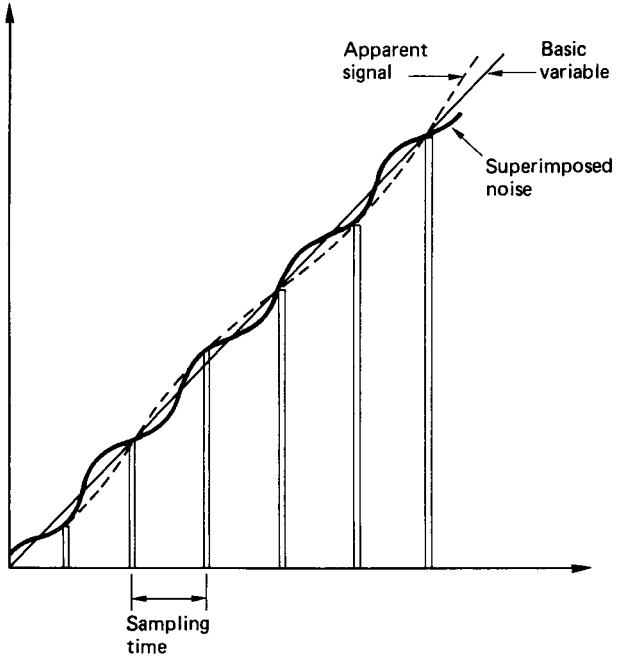


FIGURE 6.28 Noise generated aliases in a sampled-data signal.

The sources of noise are varied and may originate from thermoelectric effects, electrochemical action, electrostatic and electromagnetic pick-up, self-generated component noise, offset voltages, and common earth loops. If the frequency content of the signal to be measured is known beforehand, positive steps can be taken to eliminate most of the unwanted effects of noise by the inclusion of suitable filters. Filters exist in three broad categories: lowpass, high-pass, and bandpass. The gain characteristics for each type are shown in Figure 6.29.

A lowpass filter is one which allows the transmission of signals below a particular cut-off frequency. Signals whose frequencies are above the selected cut-off are progressively attenuated. The highpass filter, in contrast, transmits only that part of the signal whose frequencies are above the cut-off value. The bandpass filter transmits, without attenuation, the signal contained within an upper and a lower cut-off value. The cut-off frequency, is defined as that at which the signal attenuation is −3 dB.

The simplest forms of analog filter are those which incorporate only passive resistive, capacitive, or inductive elements:

1. *Lowpass filter* (Figure 6.30): The transfer function for a lowpass filter is given as equation (6.18) with the gain, $k = 1$. The time constant, τ, is equal to the product of the values of resistance and capacitance, RC. The −3 dB cut-off frequency is given by

$$f = 1/(2\pi RC) \tag{6.3}$$

   A suitable choice of resistor and capacitor can therefore allow any desired cut-off frequency to be imposed in the signal-conditioning train.

2. *Highpass filter* (Figure 6.31): The transfer function, in terms of the Laplacian variable, for the highpass filter is

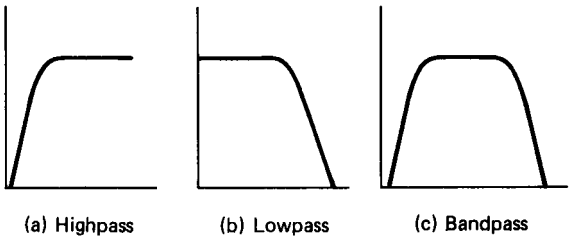


FIGURE 6.29 Filter performance curves.

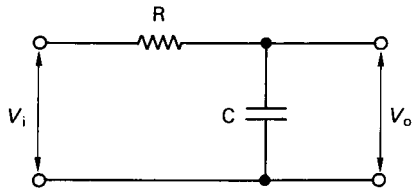


FIGURE 6.30 Lowpass filter.

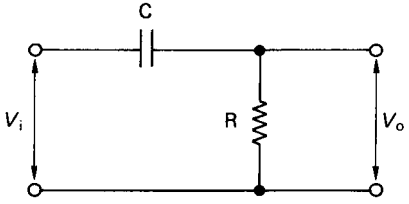


FIGURE 6.31 Highpass filter.

$$(V_o/V_i) = (sRC)/(1 + sRC) \tag{6.4}$$

The cut-off frequency is similarly selected through a judicious choice of resistor and capacitor.

Bandpass filters might be thought of as a series arrangement of a lowpass and a highpass filter. With the appropriate combinations of resistors and capacitors, the low- and high-frequency noise components in the signal can be suitably attenuated as required.

Filters which include an amplifier in the circuit are referred to as active, and the relationship between input and output is a much more complex function of time. Williams (1981) gives a comprehensive discourse on filter design.

Figures 6.30 and 6.31 illustrate the simplest forms of passive analog filter which are used to suppress background noise. The governing equations (6.4) and (6.18) may equally well be expressed in terms of finite differences. In finite difference form, the equations can be used to action the filtering process on a discretized version of the input signal. This is the basis of a digital filter which can be implemented in software and requires no external hardwired components.

The setting of the cut-off frequencies in the digital filter are achieved through adjustment of the constants appearing in the finite difference approximating function. These numerical constants are simply related to the physical time constant in the equivalent analog filter, and also the digital sampling rate.

The advantages that the digital filter has over its analog counterpart include the ease with which the cut-off frequencies can be adjusted. The -3 dB cut-off frequency can also be set exactly, since no hardwired components, with physical tolerance bands, are used. The digital signal may also be filtered any number of times simply by processing the data repetitively through the filtering algorithm. The disadvantage incurred is that digital filtering takes longer in real time to perform. The results of digital filtering on the signal are exactly the same as would be obtained using an analog filter. That is, any time-varying signal whose frequency is outside the cut-off value is subject to attenuation with a corresponding phase shift.

REFERENCES

Ballard, D. H. and Brown, C. M. 1983. *Computer Vision,* Prentice-Hall, Englewood Cliffs, NJ.

Bejczy, A. K. 1977. "Effect of Hand-Based Sensors on Manipulator Control Performance," *Mechanism and Machine Theory,* vol. 12, pp. 547–567.

Fraser, C. J. and Milne, J. S. 1990. *Microcomputer Applications in Measurement Systems,* Macmillan Education, Basingstoke, U.K.

Harmon, L. D. 1982. "Automated Tactile Sensing," *International Journal of Robotics Research,* vol. 1, no. 2, pp. 3–32.

Rebman, J. and Trull, N. 1983. "A Robot Tactile Sensor for Robot Applications," in *Proceedings of ASME Conference on Computers in Engineering,* Chicago.

Whitney, D. E. 1982. "Quasi-static Assembly of Compliantly Supported Rigid Parts," *Journal of Dynamic Parts,* March, pp. 65–77.

PART 5
INSTRUMENTATION

Charles J. Fraser
John S. Milne
Ben Noltingle

6.18 INTRODUCTION

Many aspects of aerospace engineering depend essentially on the ability to make measurements of relevant quantities. This is as true of measurements taken in flight as it is of research and development and manufacturing quality control. In some areas, such as nondestructive testing or experimental stress analysis, the techniques for making the measurements are linked so closely to the rest of the subject that they are better dealt with in their particular context. Here we consider the more general measurements encountered throughout the whole field and the instrumentation that makes them possible.

In later sections techniques are described under the headings of the quantities to be measured—sometimes referred to as measurands. First, it is useful to discuss topics that are of importance in many instrumentation systems.

Sometimes the process of measurement is a simple one, with the human operator playing a direct part: putting a rule alongside a component to determine a length, or measuring time using a stopwatch. More often a more complex system is used, which can be understood by thinking of the elements that make it up. Such a system is shown in Figure 6.32. Many variations are possible.

The measurand acts on a transducer and produces a signal corresponding to its value. This output signal is commonly electrical, but may be pneumatic or optical. The signal can then be transmitted as needed. We show the transmission or telemetry as a distinct part of the system in Figure 6.32 to indicate its importance. Some processing of the signal is often needed before it is displayed or recorded. Often a computer is used to control the whole operation, probably covering many transducers, partly in the light of the information given by the measurements.

For instance, temperatures might need to be determined at many points on a plant. Thermocouples could be used as transducers, other equipment being some

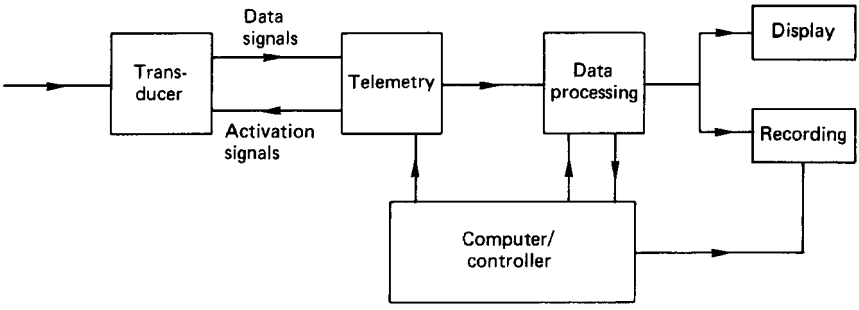


FIGURE 6.32 Typical instrumentation system.

distance away, more accessible, and in a less severe environment. Temperatures should be recorded at regular intervals, maxima, minima, and averages being extracted and an alarm given if any temperature exceeds a prechosen limit. Such a system might be part of the control of operational plant, signals being available as inputs for actuators. Alternatively, it might be logging the results of an experiment.

Most data processing and computing use digital signals, as does complex telemetry; simple, short-range telemetry can be analog. Most transducers have analog outputs, so an analog-to-digital converter (ADC) is needed at the input to data processors. If the transducer gives an output in the form of frequency, the ADC can operate simply by counting the number of cycles in a given time.

Communicating to operators the results of measurement is a distinct phase of the whole. It may amount to recording permanently for study later, possibly simply on paper, more probably in some form suitable for computer input. Alternatively, an indication (called Display) may be needed to show immediately the output of the instrumentation system. This is all part of the human/machine interface.

Accuracy

The accuracy of any instrumentation system is, of course, of great importance. When making a measurement it is possible to distinguish systematic and random errors. When a measurement of what should be the same thing is repeated it will be found that the readings are not identical; the spread of readings corresponds to random errors between individuals. The final error can be reduced by repetition and averaging (provided time allows). However, even the average is not necesssarily correct, because it contains the systematic error. Systematic errors can be greatly reduced by careful calibration under conditions resembling the operating ones as closely as possible. Calibration is in fact an important feature of all systems where it is required to relate the readings taken back to absolute values of the quantities concerned. The word ''traceability'' is used to describe building up links in this chain. Absolute values of measurands, are not always important. Sometimes it is ''repeatability''—the capability of giving the same reading (even if it is the wrong one!) under the same conditions—that matters. The precision of a reading relates to the smallest difference in value that can be detected.

When considering how faithfully an instrument represents what it is supposed to be measuring, the idea of ''influence quantities'' should be taken into account. The resistance of a strain gauge, for instance, varies with strain and so is used to measure it; but it also varies with temperature, so a one-to-one correspondence with strain will only hold provided temperature is constant. Temperature is an influence quantity. Reducing the errors introduced by influence quantities is facilitated by making bridge measurements. This idea is most familiar in electrical circuits, when two components are connected so that changes in them affect the output in opposite directions. It can then be arranged that influence quantities affect them equally and so cancel, while the measurand either only affects one component or affects them oppositely.

Another concept that is helpful in understanding transducer performance is that of gauge factor. This is the ratio of the fractional change in output to the fractional change in measurand. Again it is easily understood in the frequently quoted example of the resistance strain gauge, where it is given by the formula

$$\delta R/R = \delta l/l$$

However, the concept is more generally applicable, with the expectation that a transducer with a higher gauge factor will be more immune to influence factors.

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

Hysteresis is observed in many instruments. This is the name given to the defect of indicating a different output according to whether the quantity measured is increasing or decreasing (see Figure 6.33). Nonlinearity is also shown in Figure 6.33. This is the situation when the relation between input and output cannot be exactly represented by a straight line.

Reliability

The reliability of any equipment is often of the utmost importance, and instrumentation systems are no exception. Attempts have been made to develop the concept of reliability, putting it on a semiquantitative—though necessarily probabilistic—basis, so allowing the chances of failure to be calculated. While carrying out such an exercise, lessons can be learned of features that have been identified as significant. With large systems, the configuration in which different items are connected can make an unexpected difference. The schedule of servicing plays a large part in the down time that is to be expected. The consequences of any particular failure, which may be very diverse, should be analyzed. The idea of independence may be developed; this implies that it is safer for two measurements, if they are intended to corroborate one another, to be made with essentially different components—or even techniques—because they are then less likely to suffer from a common fault. In all considerations of instruments and their installations due weight must be placed on reliability. High accuracy has little value if it cannot be counted on, or perhaps it may be put that the possibility of a very large error can be more damaging than the probability of a small one.

Environmental Conditions

In a broad way, it is always recognized that there are limits to the conditions under which instruments can be used. The working temperature range may well be specified. Users should be aware of other significant conditions. The atmosphere can be harmful, from moisture or other corrosive effects; dust in the atmosphere is an

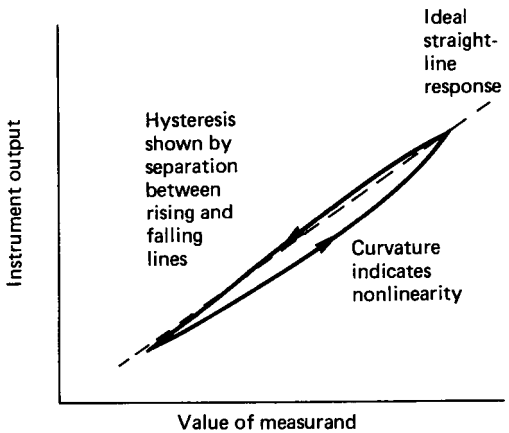


FIGURE 6.33 Hysteresis and nonlinearity errors.

enemy to many instruments. Subjection to mechanical abuse (notably excessive vibration) must also be guarded against. More sophisticated hazards include electrical interference and nuclear radiation. The capabilities of particular items to withstand all these are often not spelled out. Common sense is often called for to judge what is reasonable to expect. Sometimes precise limits are quoted by suppliers, but it should be noted that approaching these limits closely may reduce accuracy and especially reliability.

Frequency Coverage

A simple description of what an instrument does may imply that there is an indefinite amount of time in which to make a measurement—the steady-state or d.c. behavior. Sometimes this is not the whole relevant story. There are certain devices that cannot be used for steady-state measurements, but only for varying ones. There is always an upper limit to the frequency at which an instrument will operate faithfully. This may restrict the number of readings that can be taken in a limited time and is very often of concern when the measurement process is part of a closed control loop (see Part 7). It is therefore of great importance to take account of this feature of an instrument's performance.

6.19 DIMENSIONAL/GEOMETRICAL MEASUREMENTS

The lengths that people have been interested in and hence wanted to measure range from perhaps 10^{-15} to 10^{18} meters. At the extremes the interest is mainly from nuclear physicists and astronomers, and we shall here concentrate more on the middle distances.

Of course, there are simple, manual instruments that will continue to be used: rules, micrometers, calipers. A dial gauge incorporates gears to provide magnification so that a movement of even a fraction of a millimeter gives an observable rotation of a pointer. Gauge blocks should be mentioned—elements with accurately parallel faces a precise distance apart; they can be thought of as coming into a calibration exercise, or as forming references allowing the dimensions of other parts to be compared with them.

However, a large part of modern dimensional instrumentation involves converting lengths into electrical (or, occasionally, optical or pneumatic) signals. The three electrical quantities—resistance, inductance, and capacitance—are all used in transducers for this purpose.

Resistance Transducers

In one form of transducer a slider moves over an extended resistance element, which may be either wire-wound or made of some nonmetal. The resistance between the contact of the slider and either end of the element varies with the position of the former. Such devices are simple and only need simple electrical equipment. They can be made to obey some nonlinear law if that should be wanted. However, they are not of the highest accuracy, and, because rubbing and friction occur, cannot be expected to have indefinitely long lives.

A different form of resistance transducer is the strain gauge. In essence, a small element is bonded to a structural unit and changes its resistance with the dimensional changes brought about by the latter being strained. The length changes are small (typically a fraction of a percent) and the resistance changes are generally only a little larger. However, because it is possible to measure electrical resistance very precisely, strain gauges can still be used in accurate work. When directly measuring strain, they are, of course, measuring very small movements and need a large force to act on them. The range can be increased and the force reduced by having the movement to be measured applied to a flexible strip on which the gauge is mounted, as shown in Figure 6.34.

Inductance Transducers

The inductance of a coil is given approximately by

$$L = N^2 \frac{A}{l} \mu$$

where N is the number of turns, A the cross-sectional area, μ the effective permeability, and l the length of the magnetic path. For an iron-cored coil with an air gap as shown in Figure 3.35 two very different values of μ arise, and the formula becomes

$$L = \frac{N^2 A}{(l_{\text{iron}}/\mu_{\text{iron}}) + (l_{\text{air}}/\mu_{\text{air}})}$$

Even for small values of the air gap, $(l_{\text{iron}}/\mu_{\text{iron}} << (l_{\text{air}}/\mu_{\text{air}})$ and L is very sensitive to changes in l_{air}. Many inductance transducers make use of this principle.

FIGURE 6.34 Strain gauge on flexible strip.

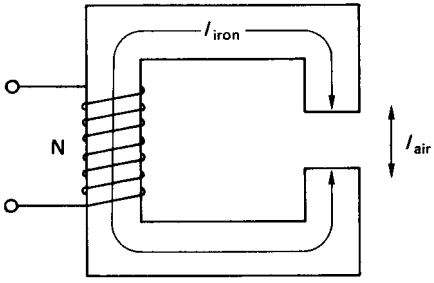


FIGURE 6.35 Coil with air gap.

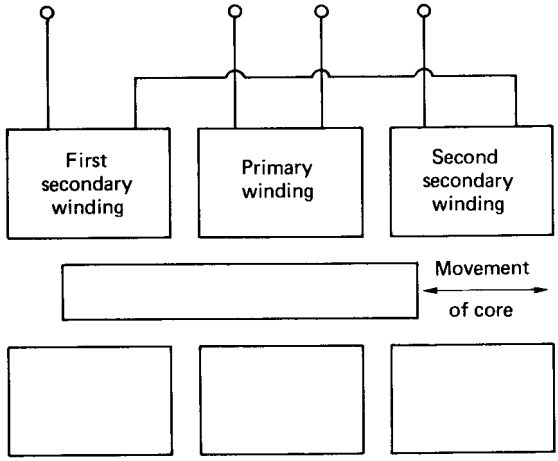


FIGURE 6.36 Linear variable differential transformer (LVDT).

If extreme sensitivity is not required, the iron circuit need not be so nearly closed. An example of an instrument of this type is the linear variable differential transformer (LVDT), where there are two further refinements shown in Figure 3.36:

1. A differential system is used so that the inductance of one winding increases at the same time as that of another decreases; and
2. By having a further winding, a mutual inductance or transformer replaces the self-inductance.

These additions increase the magnitude and linearity of the output and give an inbuilt bridge system, so increasing the effective gauge factor.

While most practical devices employ ferromagnetic (iron) cores, it is, in principle, possible to have an air-cored inductor serving as a transducer.

Capacitance Transducers

The electrical capacitance in farads between a pair of parallel plates (as shown in Figure 6.37) is

$$C = \varepsilon_0 \varepsilon \frac{A}{d}$$

where ε_0 = the permittivity of free space ($= 8.9 \times 10^{-12}$ F/m)
ε = the relative permittivity of the material between the plates
A = the area of either plate, or of their overlap if they are not exactly equal and opposite
d = the separation between the plates

C can therefore be changed by changing either A or d. Since d can be a millimeter

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

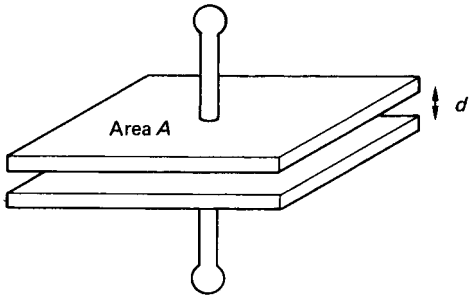


FIGURE 6.37 Parallel plate capacitor.

or less, while the lengths involved in A are likely to be a centimeter or more, and the percentage change in C equals the percentage change in A or d, it can be seen that variable-d transducers are more sensitive than variable-A ones, i.e., those in which the overlap is changed. On the other hand, the proportionality between C and A makes the second type inherently linear, unlike the inverse C/d relationship when the gap is varied. A differential arrangement, in which the same movement increases one capacitance and decreases another, is often used and can improve linearity.

Electrical Circuits and Comparison of Techniques

Resistance measurements are simple and straightforward for the large changes involved in sliding contact devices; the basic accuracy of the transducer may not be high enough to justify elaborate circuitry. The much smaller changes in devices using strain gauges call for the use of bridge circuits, and some form of bridge configuration is usual for inductance and capacitance transducers. Audiofrequency (AF) power supplies are used for resistance and inductance measurements, though d.c. is, of course, also effective for resistance.

Circuits used for capacitance measurement must take account of the stray capacitance that occurs between nearby conductors unless they are specifically screened from each other. Thus in Figure 6.38 capacitance variations between its lead and earth (either within or outside the screened cable) are indistinguishable (to the measuring circuit) from transducer capacitance changes. Various arrangements can be adopted to overcome this problem. The relatively small capacitance in most transducers corresponds to a very high impedance at lower frequencies,

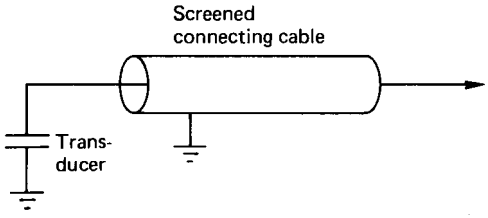


FIGURE 6.38 Effects of stray capacitance.

and this is an argument for working at higher frequencies, but in fact AF bridge systems with very high sensitivities are available and give good performance when spurious effects are eliminated.

Many considerations come into the choice of transducer technique. As a very simple summary, it may be suggested that while resistance devices are simple and inexpensive, inductance devices, while tending to be larger and more complicated, have a long history of development and mass production. Capacitance devices, simple and sensitive in principle, need more elaborate circuitry but may well give the best approach for particularly onerous requirements. Sometimes the force needed to move a transducer element is important. In general, the force is less for capacitors than for inductors, while with variable resistors it may be less repeatable.

Optical Methods of Position Measurement

Some classical experiments in physics depend on optical interference. If two coherent light beams are superposed they reinforce or cancel each other, according to whether they are in or out of phase, and this phase difference depends on the different lengths of the paths they have traveled. If they have traveled the same distance or their paths differ in length by an integral number of wavelengths, then they reinforce, while if their paths differ by an odd number of half wavelengths they cancel.

Figure 6.39 shows how this can be used for an accurate measurement of movement. As the mirror M moves, the light intensity changes from maximum to minimum and back for successive distances of half a wavelength—a fraction of a micrometer, making the system highly sensitive. Refinements are needed to determine the direction of motion and to give general stability; a corner cube reflector instead of a simple mirror eliminates the otherwise high sensitivity to the angle of the mirror. A laser is a convenient source of coherent radiation. The output signal, going through a succession of peaks, is essentially digital.

Moiré fringes are sometimes used to measure movement. Figure 3.40 shows two adjacent gratings as seen from above. If they are positioned as in (a) light can pass through, but if one is moved by half a wavelength as in (b) the path is blocked and the combination appears dark. The wavelength or pitch can be very short, as small

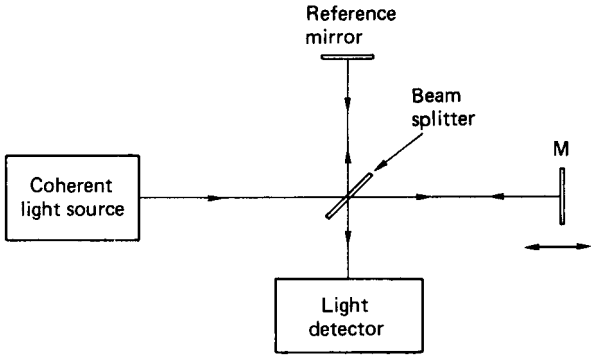


FIGURE 6.39 Movement measured by optical interference.

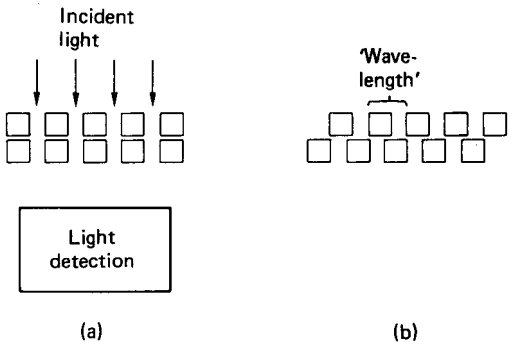


FIGURE 6.40 Principle of moiré fringes.

as a few micrometers (the name *moiré* comes from silk weave in which the effect can be observed), giving a high potential accuracy. Again there is a basically digital output and the need to determine the direction of movement.

As shown in Figure 6.40, the interrogating light is transmitted through the gratings; it is, of course, possible to have a mirror system when the light source and detector are both on the same side of the gratings. The gratings may be at an angle to each other (Figure 6.41), when the alternate bright and dark areas form fringes perpendicular to the gratings; the fringes move bodily with linear displacement of either grating, while the separation between them depends on the angle between the two.

If the gratings do not have quite the same pitch, there are fringes parallel to the grating elements (Figure 6.42). This principle is sometimes used in strain measurement when the strain to be measured is arranged to alter the pitch. In all these

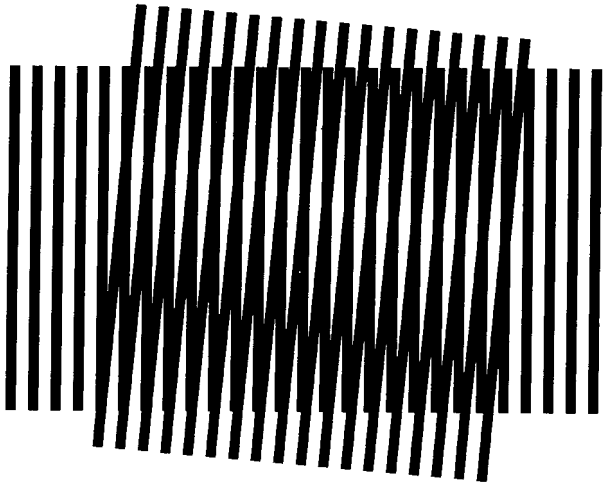


FIGURE 6.41 Moiré gratings at an angle.

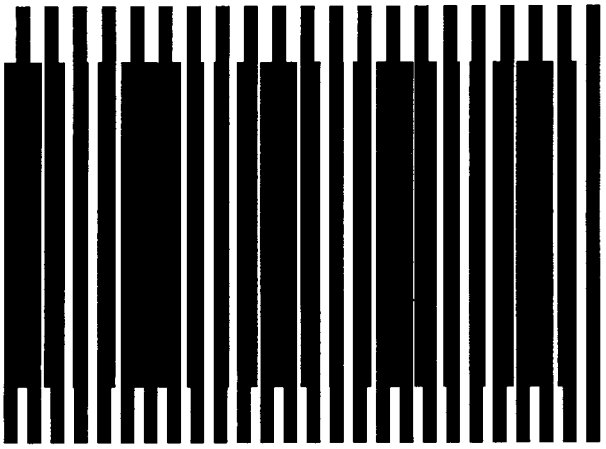


FIGURE 6.42 Unequally spaced moiré gratings.

arrangements there is an effective magnification, so that small movements, on the scale of the small pitch of the gratings, give rise to much larger movements of the fringes.

Pneumatics

Pneumatic instrumentation systems have the drawbacks of needing somewhat delicate mechanical devices and of introducing significant delays when signals are transmitted over long distances. However, they are by no means extinct and have the great safety advantage that there need be no question of their introducing electric sparks.

The heart of a pneumatic instrument is a flapper adjacent to a nozzle. As the separation between these is changed, the air flow through the nozzle changes markedly and hence also the pressure drop across the series restrictor. The effect can be amplified by the introduction of further elements in the shape of valves and a pressure-sensitive diaphragm. The primary behavior is inherently nonlinear, but the use of a pressure-feedback device with levers and a spring-controlled bellows allows a movement of the order of a millimeter to give a proportional pressure change of some tens of kilopascals.

Angular Displacement

The synchro—sometimes called a magslip or selsyn—is widely used in the measurement of angles. If a.c. is applied to the central element (rotor) of such a device (left-hand side of Figure 6.43) then the voltages induced in the three circumferential windings depend on the angular position of the rotor relative to them. This system has the particular advantage that if a second, identical unit is connected appropriately (right-hand side of Figure 6.43) forces will act within it until the two rotors

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

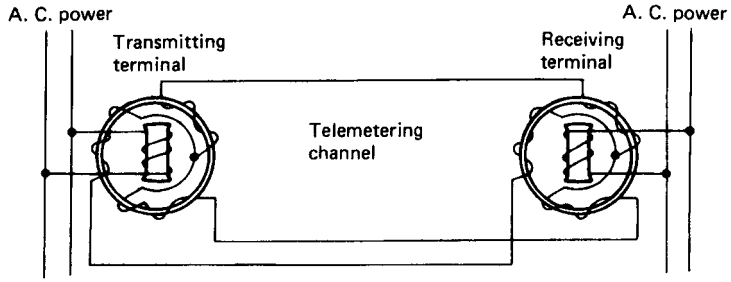


FIGURE 6.43 Principle of synchro.

take up identical angles. This is a robust and widely used technique for telemetering an angular position.

Capacitive transducers with variable overlap readily give a measurement of angle. The arrangement is in fact just that of the orthodox variable capacitor.

Encoders are used to give a digital signal corresponding to angular position. Moving clockwise round the disk shown in Figure 6.44, it can be seen that successive positions 1, 2, 3. . . correspond to successive binary numbers if black and white areas give digits 1 and 0, respectively, for powers of 2 starting at the largest radius. Black can be distinguished from white using six optical beams in the example shown, or a single beam can be traversed radially across the encoder. Alternatively, the distinction can be between conducting and insulating material, detected electrically.

A difficulty with this form of coding follows from imperfections of manufacture. Considering, for instance, the move from position 7 to position 8, if the outermost black should turn white slightly before the others, the configuration will momentarily correspond to position 6, while premature changes of the other blacks would indicate 5 or 3, respectively. The problem arises from the need for simultaneous changes at more than one radius, and to overcome this, codes have been devised

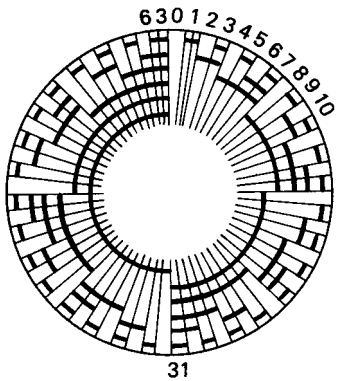


FIGURE 6.44 Encoding for angular position.

in which only one change occurs at a time. As indicated previously, small changes of angle can also be detected with moiré fringes.

Velocity Measurement

Angular velocity is commonly measured employing electrical induction. Using the fundamental law that induced voltage is proportional to rate of change of flux, generators, either d.c. or a.c., can be made for which output voltage gives a direct measure of the speed of rotation. Under a completely different principle, a technique is to mount markers on the circumference of a rotor and count the number passing a stationary point in a given time, or alternatively, the time lapse between successive passages, which can be detected optically, magnetically, or electrostatically. This system, of course, provides a digital output; it requires a finite time to give an indication.

Linear velocity is sometimes deduced from angular velocity as in a car's speedometer. It can also be calculated as the rate of change of position or as the integral of acceleration, and this is particularly relevant to vibration studies.

6.20 VOLUME AND LEVEL

Volume, as such, is a quantity that is seldom measured. Instrumentation for rate of change of volume (or flow) is widely applied and can be integrated to give total volume; this is dealt with in a later section. Volume and mass are simply related through density, and mass can be measured as weight. Again, the volume of material in a container can be inferred from the level it reaches, and this is a common measurement.

Measuring level, we can distinguish between continuous, normally analog methods and digital techniques, in which the action is really detection rather than measurement. The presence or absence of the material in question at a particular level is indicated. The second category can be used, as shown in Figure 6.45, to move a follower outside the container under study so that it remains opposite the internal interface, allowing the height to be measured in a more accessible place.

The level of a liquid conductor can be found from resistance measurement. Figure 6.46 shows two resistive wires that are effectively short-circuited where they enter the liquid, so that the resistance seen at their terminals decreases as the level rises.

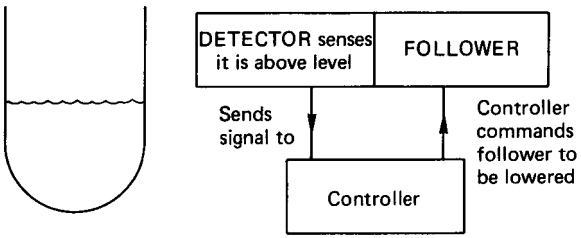


FIGURE 6.45 Level measurement with a follower.

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

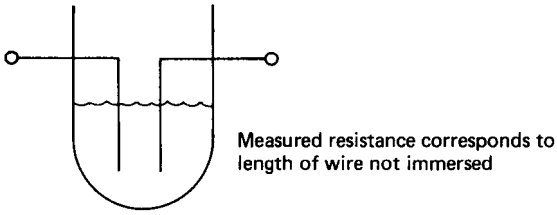


FIGURE 6.46 Level measurement using resistance.

For an insulating liquid, capacitance measurement is appropriate. With the arrangement of Figure 6.47, capacitance increases as the level rises and a larger area of the overlapping plates is separated by a dielectric of higher permittivity. A sonar-ranging system can also be used in which the time taken for an echo to return from the surface being studied gives an indication of its position (Figure 6.48).

A sophisticated single-point technique involves passing gamma rays through the container. These will be more attenuated if there is a denser material in their path, so the intensity of radiation received at the detector shows whether liquid (or solid) rather than just gas is present. In Figure 6.49 it can be recognized that the detector output will be larger if the level of liquid in the container falls below the line from source to detector.

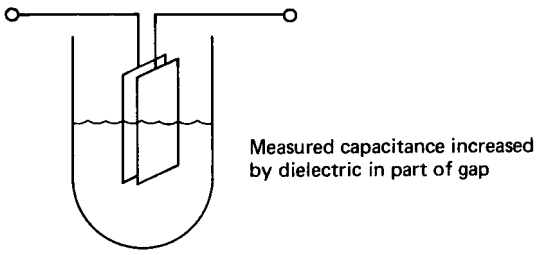


FIGURE 6.47 Level measurement using capacitance.

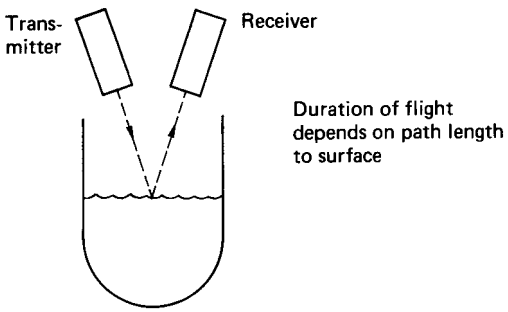


FIGURE 6.48 Level measurement by sound ranging.

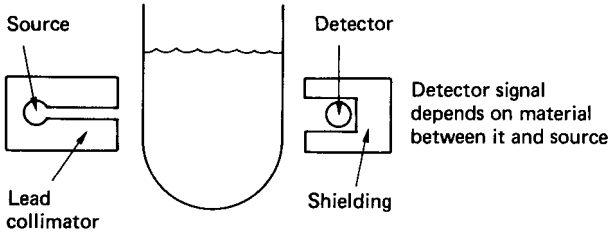


FIGURE 6.49 Level measurement by gamma rays.

Different types of probe have been devised for detecting the presence of a liquid that depend on refractive index or resistivity or permittivity, all of which may have different values above and below the interface whose level is to be detected. Yet another technique is to measure level as a differential pressure. If one pressure transducer is mounted internally at the bottom of a vessel and another at the top, then the difference between the pressures they show will depend on how much of the height between them is occupied by liquid and how much by gas (i.e., on the level of the former).

6.21 MEASUREMENT OF VIBRATION

Here we touch briefly on techniques of vibration measurement.

In sinusoidal motion at frequency ω, linear amplitude s, velocity v, and acceleration a are simply related as

$$a = \omega v = v^2 s$$

$$\nu = vs$$

Displacement can be measured using techniques described earlier, or velocity with a generator, commonly a coil moving in the field of an electromagnet. If a nearby point is known to be stationary, either measurement can be relative to this. Alternatively, part of the transducer can be an element with sufficient inertia not to move, when the measurement can be made relative to that. The criterion whether the inertia is large enough is that the resonant frequency of the element—decided by its mass and its flexible mounting—should be much lower than any frequencies in the vibration.

Vibrational accelerations are very often measured, using the associated force F, where $F = ma$ and m is an inertial mass. In this case, the element's resonant frequency must be much greater than the highest vibrational frequency to ensure that a is the same as the acceleration of the part on which the transducer is mounted. A piezoelectric device often forms the link to the inertial mass, the charges excited in it providing the output signal, while its high degree of stiffness gives a high resonant frequency. The wide range of mass and the variety of piezoelectric elements that may be used give scope for a wide range of applicability for vibration pick-ups.

We may want to know the frequencies contributing to the vibration studied. For this, some form of spectrum analyzer will be desirable.

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

6.22 FORCE/WEIGHT MEASUREMENT

This field of instrumentation gives good examples of some general principles mentioned earlier in this subsection. Concerning speed of response, in many instances the measurement called for is a static one, but sometimes a quickly varying force is to be studied, and this calls for a different approach. The potential accuracy varies more than a thousandfold—with corresponding price ranges for equipment. Sensitivity to extraneous influences is also a factor in accurate force measurement, where errors from temperature, wrong location of the force and other things must be guarded against. Weight, of course, is a force, and, in general it is measured in similar ways to other forces, though the measurement is always a static one.

Lever-type instruments, such as the classical analytical balance, are basically devices for comparing forces—often the weights of different masses. Unequal lever arms allow widely different forces to be compared; an arm of variable length allows a precise ratio to be established without the need for an adjustable force. The spring balance is the most familiar member of a large family of instruments in which the force to be measured is balanced by the reaction from an elastically strained member whose distortion can be measured.

The proving ring illustrated in Figure 6.50 is a refined form of spring balance. The applied force, which may be compressive or tensile, distorts the ring from its intially circular shape. The change in diameter (measured mechanically with a dial gauge or electronically) indicates the force. Much smaller movements are measured than with a coil-spring spring balance, making the total system more compact.

In a strain-gauge load cell the process is taken rather further, the elastic strains in a member being directly measured with strain gauges, allowing very compact devices to be constructed. The principle of a simple, columnar load cell is shown in Figure 6.51; the four strain gauges are connected into the four arms of a bridge. Structures in which shear strains are measured are also widely used; their readings are less dependent on the position where the load is applied. In hydraulic load cells the unknown force alters the pressure in a liquid system, allowing it to be measured as a pressure.

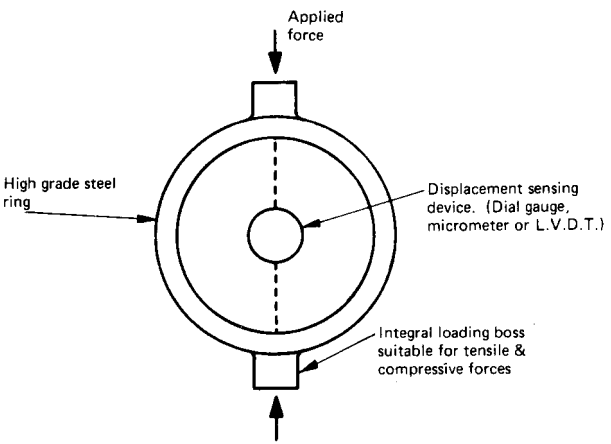


FIGURE 6.50 Proving ring.

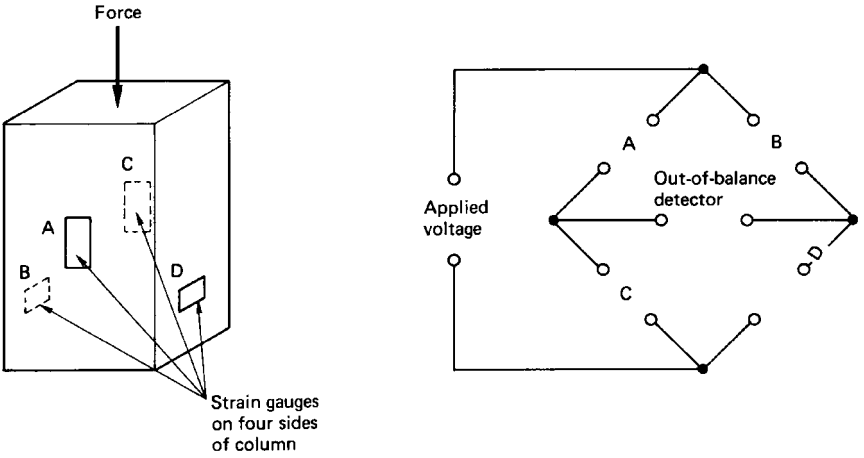


FIGURE 6.51 Principle of load cell.

Systems of particular value for dynamic measurement of quickly changing forces include piezoelectric elements, mentioned in connection with vibration instrumentation. Force balance systems are also used. Figure 6.52 shows how the displacement produced by a force to be measured can control the restoring force in a coil, the current in which gives a direct indication of the first force provided the gain is large and the displacement small—with due attention paid to stability.

Table 6.2 summarizes the features of different ways of measuring force.

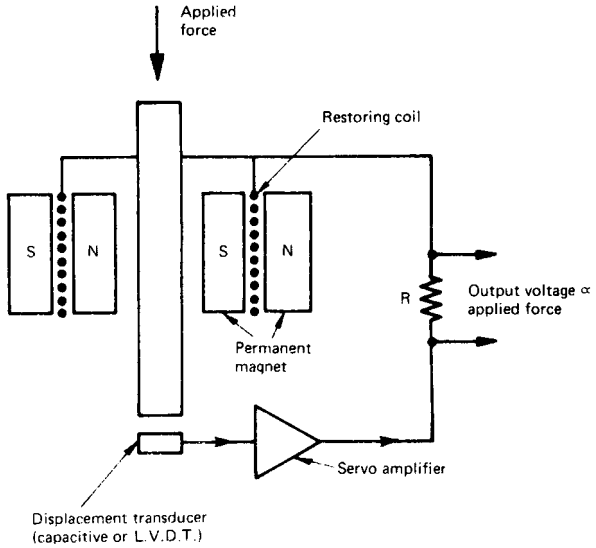


FIGURE 6.52 Force balance system.

TABLE 6.2

Method	Type of loading	Force range, N (approx.)	Accuracy % (approx.)	Size
Lever balance	Static	0.001 to 150 k	Very high	Bulky and heavy
Force balance	Static/dynamic	0.1 to 1 k	Very high	Bulky and heavy
Hydraulic load cell	Static/dynamic	5 k to 5 M	0.25 to 1.0	Compact and stiff
Spring balance	Static	0.1 to 10 k	Low	Large and heavy
Proving ring	Static	2 k to 2 M	0.2 to 0.5	Compact
Piezoelectric transducer	Dynamic	5 k to 1 M	0.5 to 1.5	Small
Strain-gauge load cell	Static/dynamic	5 to 40 M	0.01 to 1.0	Compact and stiff

6.23 PRESSURE

Pressure is easily measured from the difference in level of the liquid in two arms of a U-tube (Figure 6.53):

$$P_1 - P_2 = h\rho$$

where ρ is the density of the liquid. Mercury is commonly used as the working liquid. Its high density means that a large pressure difference can be measured without the equipment becoming too big in order to accommodate a large h. There is greater sensitivity if mercury is replaced by water. Further movement for a given pressure difference can be achieved if such a manometer is at a small angle to the horizontal instead of being held vertical. Care must be taken that the liquid can move freely over the inside surface of the tube (otherwise there will be hysteresis) and that any distortion of the surface from surface tension is the same in both limbs.

Note that it is pressure *differential,* $P_1 - P_2$ that is of concern. Three situations should be distinguished:

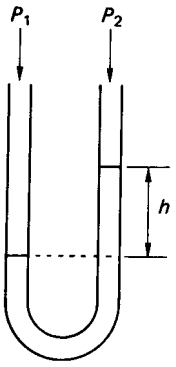


FIGURE 6.53 U-tube manometer.

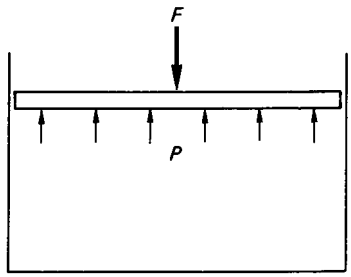


FIGURE 6.54 Absolute pressure measurement.

1. Absolute pressure, where P_2 is zero, corresponding to a vacuum
2. Gauge pressure, where P_2 is the atmospheric pressure in the neighborhood of the equipment
3. Differential measurements where both P_1 and P_2 may vary but it is their difference that is significant

Pressure can be measured in terms of its fundamental definition of force per unit area. In Figure 6.54, if the cross-sectional area of the cylinder with its piston is known, then the pressure, P, is directly given by the force, F (usually a weight), needed to balance it. This method can give high accuracy, but there are practical complications, notably to ensure that the piston can move freely without the liquid leaking, so it is used mainly to calibrate other pressure gauges.

Neither of these approaches uses compact equipment or leads directly to an output signal, so they are often replaced by the use of transducer elements. Various configurations change shape with pressure, and the consequent displacement can be used for measurement.

A Bourdon tube has an elliptical (or otherwise unsymmetrical) cross section and is bent into a circular arc. If the pressure inside the tube increases, it tends to make the cross section more nearly circular, and this in turn straightens the arc. With some further mechanical amplification, the movement is large enough to be read against a scale.

A metal diaphragm distorts according to the difference in pressure of the fluid on either side of it. The sensitivity varies widely with the dimensions. A wide, thin diaphragm moves appreciably under small pressures, and the danger of its rupturing under overload can be greatly reduced by the provision of stops. The movement can be detected pneumatically or by capacitive or inductive devices.

An alternative approach is to measure strain in the diaphragm. Since different parts are strained in different senses, strain gauges connected in different arms of an electrical bridge can be mounted on a diaphragm so that their pressure-induced outputs sum while the spurious changes from, for example, temperature variation cancel each other out. A development from this is to have the strain gauges integral with the diaphragm. Using appropriate fabrication techniques, a silicon member can serve as diaphragm and can have certain parts modified and electrically insulated so that their strain-sensitive properties can be used to give an electrical output.

Stiff diaphragms—with a high natural frequency—allow rapidly changing pressures to be measured. Other devices have much slower responses, but often the measurement required is only of quasistatic pressure. Pressure transducers including piezoelectric force measurement have a quick response but cannot be used statically.

Vacuum

Some widely different methods are available for measuring vacuum (i.e., a pressure less than atmospheric). As conditions approach the zero of absolute vacuum, measurements become increasingly difficult.

In a low vacuum (i.e., an appreciable fraction of an atmosphere) instruments described in the previous section can be used. The McLeod gauge is a development from the U-tube manometer, in which a sample of gas is compressed by a known amount before its pressure is measured; this allows much lower initial pressures to be measured. Two other broad techniques are used for high vacua.

With thermal conductivity instruments (notably the Pirani) use is made of the fact that the larger number of molecules in a gas at higher pressure increase its

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

heat transfer, so that measurement of the temperature of a heated member can indicate the degree of vacuum surrounding it. In ionization instruments (Buckley, Penning, Bayard-Alpert) the current resulting from ions in the vacuum is measured. This gives the population density of ions and hence the pressure.

The ranges over which different techniques can be used are shown in Table 6.3.

6.24 FLOW

An important and widely applied field of instrumentation is the measurement of fluid flow. Sometimes the concern is to measure velocity at a point. More often the requirement is for a single measurement representing the total volume of fluid passing along a pipe or other container—though this can be achieved by integrating from point values. Instantaneous readings for flow are of primary interest; often their time integral (i.e., the total volume that has passed) needs to be known. Gases, as well as liquids, come under study. Occasionally the main interest is with the mass rather than the volume that is passing.

Conceptually, the most direct form of instrumentation for flow measurement is the positive displacement meter, in which it is arranged that a known volume is repeatedly filled and emptied and the number of times that this takes place is counted. The capacities of chambers A and B are altered by a known volume as the diaphragm between them moves between the limits of its travel. A suitably phased slide valve ensures that the two chambers are connected alternately to the inlet manifold and to the outlet. For smoother operation, A and B are duplicated by C and D, running out of phase with them.

In other types, the volumes that are alternately filled and emptied are defined by rotating parts, either on a single axis or by two meshing rotors as shown in Figure 6.55. Care must be taken that sealing—usually by a liquid—is effective

TABLE 6.3 Comparison of Vacuum Gauge Techniques

Technique	Pressure range (Pa)	Accuracy (±%)	Cost[a]	Advantages	Limitations
Bourdon tube	10^5–10^2	10	A	Simple robust	Poor accuracy below 100 Pa
Diaphragm	10^5–10	5	B	Good general-purpose gauge	Zero setting varies
Liquid manometer	10^5–10^2	5–10	A	Simple, direct reading	Vapor may contaminate vacuum
McLeod	10^5–10^{-3}	5–10	C	Wide range. Used for calibration	Intermittent. Measures *gas* pressures only
Thermal conductivity	10^3–10^{-2}	10–20	C	Can be robust, with fast response	Risk of zero variation
Ionization	10^2–10^{-8}	20	D	Sensitive, fast response	Care needed in use

[a]Scale of costs: A, low; B, medium; C, high; D, very high.

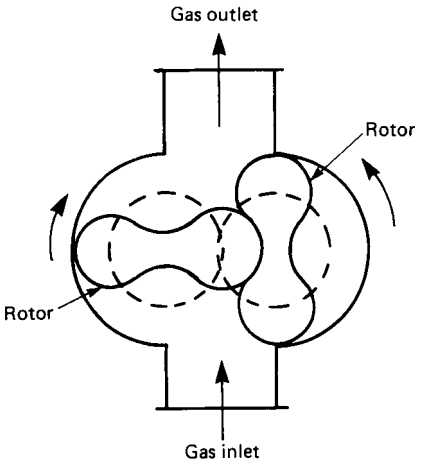


FIGURE 6.55 Positive displacement flow meter.

while still allowing free movement under the small forces associated with gas at low pressure.

Positive Displacement for Liquid Flow

Similar devices are available as flow meters for liquids. A rotating piston, mounted eccentrically in a larger cylinder, is a common arrangement. Reciprocating pistons are also used as well as the sort of rotary systems described for gases. A turbine meter (Figure 6.56) may be thought of as a positive displacement instrument, having been designed so that the angle its bladed rotor turns through is proportional to the volume of liquid that has passed (axially) through the meter.

With all these meters, the number of rotations or excursions must be counted, the flow rate, of course, being given by the number occurring in a particular time. Information about internal movements must be conveyed through a container wall to the outside, and this is often done by the passage of permanent magnetic poles past external pick-ups.

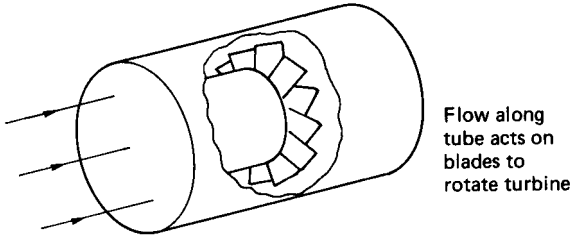


FIGURE 6.56 Turbine meter.

Differential Pressure

Where the cross-sectional area of a pipe changes, so does the pressure of a liquid flowing in the pipe, and the magnitude of the pressure change depends on the flow rate. This is often used as the principle of a flow meter. Two configurations of changing cross section may be distinguished: the Venturi throat and the orifice plate.

In the former, a smooth profile serves to reduce the area (Figure 6.57); in the latter, changes are more abrupt. The Venturi has the advantage that less energy is absorbed, but at the cost of greater size and expense. Profiles different from either of these are sometimes used. In the Venturi, the difference in pressure between the throat and a point upstream is measured. With an orifice plate, the two relevant pressures are simply those upstream and immediately downstream of the plate, since for some distance downstream the effective area is still that of the orifice.

There is a square-law relation between the flow-rate Q (m^3/s) and the differential pressure, Δp:

$$Q^2 \propto \Delta p \quad \text{or} \quad Q = k\sqrt{\Delta p}$$

In some circumstances this is an inconvenience.

One of the factors in the proportionality is the area A at the point of restriction:

$$Q = cA\sqrt{\Delta p}$$

and in one type of flow meter, the rotameter, Δp is kept constant and A made variable. This is achieved by having the liquid flow up through a tapered tube in which is placed a plummet whose weight causes the differential pressure. Increasing flow carries the plummet to a point where the annular area around it is such as to satisfy the equation.

Superficially similar to differential pressure types is the target flowmeter, in which the force exerted on a body obstructing the flow is used as an index of that flow. This again follows a square-root law.

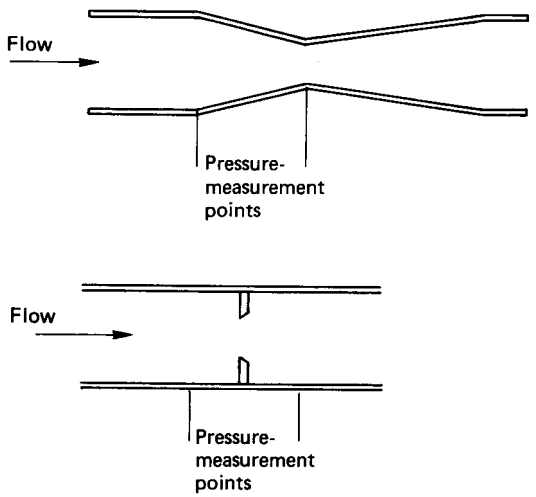


FIGURE 6.57 Venturi and orifice plate.

Three techniques may be mentioned that have more recently been developed to measure flow. When an obstruction is mounted in a pipe, the flow can be disturbed so that vortices are shed alternately from its opposite sides, and the frequency of this shedding is accurately proportional to flow rate. Sensitive detectors are needed to detect the vortices, commonly using their pressure or cooling effects or their modulation of an ultrasonic beam. The method has the advantage of not being dependent on the exact sensitivity of the detector; in fact, it uses a digital signal, namely frequency.

Electromagnetic flow meters use the principle of Faraday's law of electromagnetic induction. This states that a conductor moving in a magnetic field will give rise to an electromotive force (i.e., potential or voltage). The field, the movement, and the potential are all mutually perpendicular. In a conventional electrical generator the conductor is a wire, but it can equally be a conducting liquid such as water. All that is needed is to provide a magnetic field—commonly nonsinusoidal at a low frequency—and suitably insulated electrodes in contact with the liquid.

A time-of-flight ultrasonic flow meter depends on the fact that sound pulses are transmitted more quickly downstream than upstream. The transmission time across flowing liquid which is the medium thus depends on which is transmitter and which receiver. The transmission path must not be straight across a pipe, but it does not have to be strictly axially along it, and an arrangement as in Figure 6.58 is adopted. Various forms of electronic processing are usefully applied to the primary time-of-flight data.

Measurement of Velocity at a Point

The flow meters described so far are concerned with measuring the total flow rate in a stream. It is also sometimes of interest to measure the local velocity in an extended volume of fluid. This approach can be used to measure total flow by having a representative number of points to cover the complete cross section of a stream.

The Pitot tube comes into this category. This has a small orifice facing into the fluid flow, and closed so that fluid cannot escape. The pressure build-up will then be velocity dependent, in fact with a square-root relationship

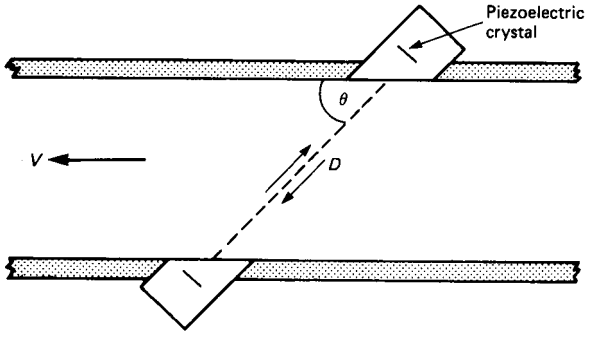


FIGURE 6.58 Time-of-flight ultrasonic flow meter.

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

$$v \propto \sqrt{p}$$

The Pitot can thus be thought of as a differential pressure device, with p the difference between what is measured in the tube and the static pressure in the fluid nearby, only the primary measurement is of point velocity rather than total flow rate, as is the case with an orifice plate.

The hot-wire anemometer uses the fact that the cooling effect of a flowing fluid increases with the fluid's velocity. A wire is heated and the temperature excess above its surroundings is measured. With a fine wire, the instrument can respond very quickly, particularly if a feedback system is introduced so that the heating power supplied is altered to maintain a nearly constant excess temperature.

The turbine meter is used to measure point velocities, particularly in surveys of large stretches of water.

Doppler techniques, where the frequency of some radiation is changed at reflection from a moving object, should also be mentioned under this heading. The ultrasonic Doppler flow meter is shown in Figure 6.59. A frequency, f_t, is sent out from the transmitter and the receiver registers a frequency f_τ. The flow velocity V is then given by

$$V = \frac{c(f_t - f_\tau)}{2f_t \cos\theta}$$

where c is the velocity of sound in the flowing fluid—which must contain some discontinuities to give reflections. As described, this refers to reflection from a single point; in practice the method is used for total flow measurement taking the mean f_τ as corresponding to the average of the different values of V across a pipe.

Laser Doppler techniques are a development from this, using electromagnetic instead of acoustic radiation. The closer positional control that is possible with laser beams allows more precise location of the point of reflection—at the cost of more complicated equipment. Laser techniques are now used extensively in wind tunnel testing.

6.25 TEMPERATURE MEASUREMENT

Although we daily experience its effects, the scientific concept of temperature is an involved one, linked with energy on an atomic scale and not directly accessible.

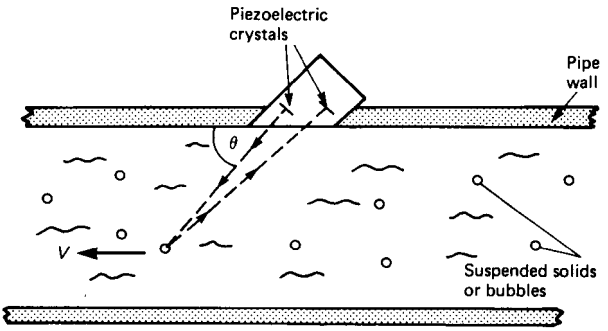


FIGURE 6.59 Ultrasonic Doppler flow meter.

To measure it we therefore need some property that varies consistently as temperature varies.

Thermal Expansion

A simple effect is thermal expansion. Mercury-in-glass thermometers depend on the expansion coefficient (defined as the fractional change in volume for one degree change in temperature) of mercury being larger than that of glass. The well-known shape of a bulb opening into a fine capillary tube allows the change in relative volumes to show as a change in position of the top of the thread of mercury in the capillary. Note that the space above the mercury must be evacuated so that pressure in it does not build up, and that the whole of the bulb should be at the temperature to be measured.

The linear expansion of a solid can also be used for a thermometer. Unless a reference frame at a known temperature should happen to be available, it will again really be a question of relative expansions of different materials. A bimetal strip, as shown in Figure 6.60, is used to magnify the movement. In such a strip, two metals of very different expansion coefficients, e.g., brass and Invar, are bonded together, when their combined curvature will alter as they warm up in order to equalize tensile and compressive forces in the strip. A direct pointer thermometer can be made by coiling the bimetal into a suitable spiral or helix.

Thermocouples

Other temperature-dependent properties are electrical. Thermocouples are widely used. They depend on an e.m.f. being set up in the circuit if two different metals are connected in series and their two junctions are not at the same temperature. Knowing one temperature and the controlling law, the other temperature may be deduced.

The provision of measuring circuitry is greatly simplifed because additional series junctions at intermediate temperatures (but with the same pair of metals) do not affect the aggregate e.m.f., and also further metals can be included provided all their junctions are at the same temperature. As shown in Figure 6.61, a parabolic relation generally holds, for a given cold junction, between e.m.f. and hot junction

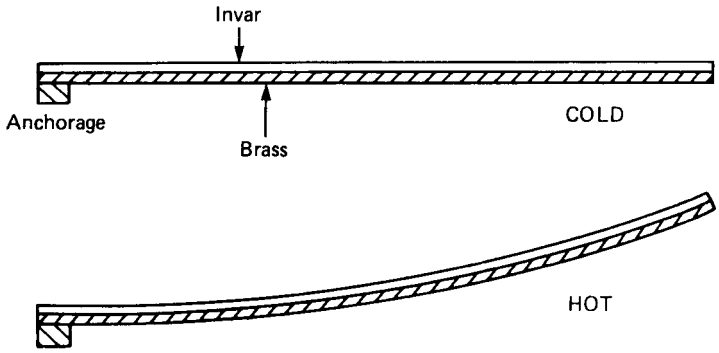


FIGURE 6.60 Bimetal strip.

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

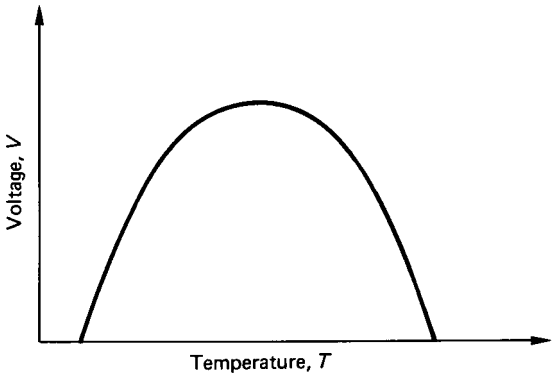


FIGURE 6.61 Thermoelectric potential.

temperature. If the peak of the parabola is far enough away from the temperatures to be measured (as is the case for some metal pairs) the characteristic will be approximately linear. However, a firm limit for the working range is set by the reduced sensitivity in the neighborhood of the peak and, worse, by the ambiguity that arises for temperatures beyond the peak, because two temperatures then correspond to the same voltage output.

Questions of corrosion also limit the type of thermocouple that can be used at high temperatures. Base metals are satisfactory below perhaps 1000°C; the more expensive noble metals and alloys are necessary to cover higher ranges. Some characteristics of common thermocouples are given in Table 6.4.

Potentials generated by thermocouples are never more than a small fraction of a volt. To measure this so as to give an acceptable accuracy calls for a precision of a few microvolts, but great sensitivity is easily achieved in electrical measure-

TABLE 6.4 Characteristics of Some Common Thermocouples

Type	Composition	Approx. temp. range (°C)	Typical output (cold junction at 0°C)
B	Platinum–30% rhodium/ platinum–6% rhodium	0–1500	1.2 mV at 500°C
E	Nickel–chromium/constantan	−200–850	6.3 mV at 100°C
J	Iron/constantan	−200–850	5.3 mV at 100°C
K	Nickel–chromium/nickel–aluminum (called chromel/alumel)	−200–1100	4.1 mV at 100°C
R	Platinum–13% rhodium/platinum	0–1500	4.5 mV at 500°C
S	Platinum–10% rhodium/rhodium	0–1500	4.2 mV at 500°C
T	Copper/constantan	−250–400	4.3 mV at 100°C
	Rhodium–iridium/rhodium	0–2000	6 mV at 1200°C
	Tungsten–5% rhenium/ tungsten–26% rhenium	0–2300	9 mV at 500°C
	Tungsten/molybdenum	1250–2600	5 mV at 2000°C

ments so that problems arise more from the errors introduced by spurious effects than from the absolute low signal level.

Calculations are often based on the cold junction being held at 0°C. Rather than do this physically with a thermostat, it may be more convenient to introduce compensation from a component whose resistance varies with temperature in a known way (see Figure 6.62). Since it is only necessary to cover a small temperature range, changes can be thought of as linear.

If the e.m.f. of a thermocouple were translated into a current, the resistance of the leads in the circuit would come into the equation and unknown variations along their length, caused by the temperature changes there, would cause errors. This is eliminated if the working current is reduced to zero, either by using the null technique of a potentiometer (when an equal and opposite potential opposes the output of the couple) or by having a detector with a high-input impedance.

With base-metal couples or short cable runs it is not impossible to use the same metals from hot junction to cold, and the additional circuitry can be at nearly constant room temperature. The expense of a long run of noble metal can be avoided by replacing it, for most of its length, by a cheaper alloy, chosen so that its thermoelectric behavior matches that of the noble metal over the limited temperature range to which most of the cable is subjected.

The two metals making up the thermocouple are commonly in the form of wires, forming leads as well as a junction. Bare wires welded or even twisted together can make an effective device, having, in fact, the advantage of a quick response. Note that if there are several points of contact it will be the coldest that is measured—the higher potential of the hotter junction being short-circuited through the colder. It is often convenient, however, for the couple to be supplied in a sheath having appropriate internal insulation and the whole forming a robust, replaceable unit. Metallic sheaths are the most common, but at the highest temperatures a ceramic construction gives greater protection against corrosion.

A mineral-insulated form of construction (MI) is widely used. In this, wires of the two different materials are located within a metal sheath and insulated from it and from each other by ceramic powder. It has been established that good insulation and stability are maintained even when the whole combination is drawn down to a

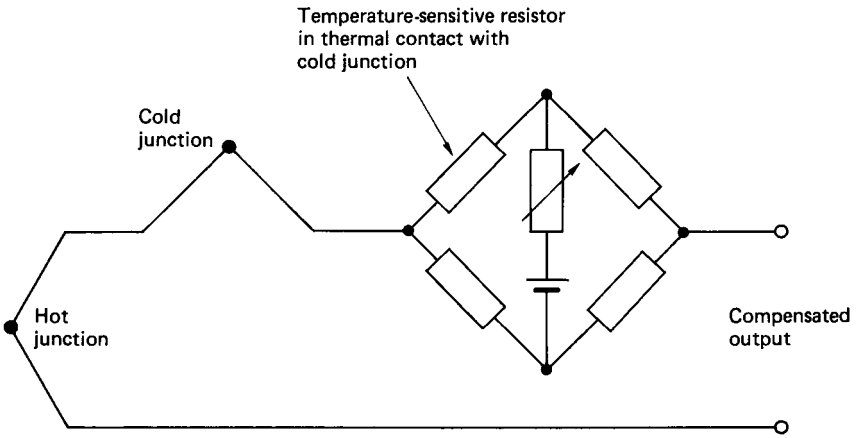


FIGURE 6.62 Cold-junction compensation.

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

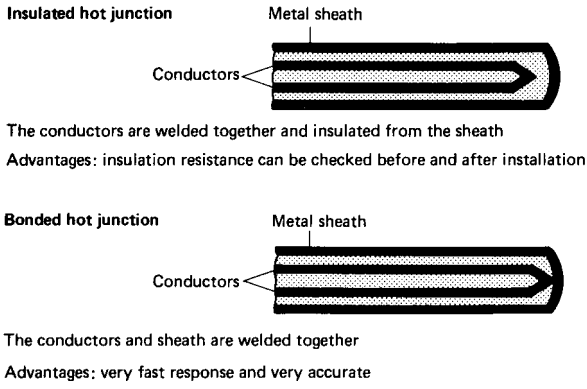


FIGURE 6.63 Thermocouple construction.

very small cross section, perhaps as little as 1 mm overall. Figure 6.63 shows how the junction may either be insulated or welded to the tip of the sheath; the latter gives a quicker thermal response but consequences for the electrical circuit may be undesirable.

Spot checks on metal temperatures can be made with special devices. In one of these a two-pronged fork has its sharpened prongs made from the two different couple materials. When the prongs are pressed against the metal surface whose temperature is to be measured, it follows from the law of intermedate materials that the thermal e.m.f. will be that of a couple at the required temperature.

To measure temperatures as high as those of molten steel (which is often called for) a permanently protected probe would be both expensive and slow-acting. It is therefore economic to have an expendable hot-junction, with the rest of the probe arranged for easy replacement after a reading has been taken.

Resistance Thermometers

One of the important properties that varies with temperature and so is used in thermometers is electrical resistance. This shares with thermocouples the advantage of leading into the convenience and precision of electrical measurements.

The electrical measurement of a resistance is easier when it is not too small—100 Ω is a convenient value for a resistance thermometer. That means that comparatively long, fine wire will be needed. A consequent danger is the susceptibility to corrosion, because even a very shallow, surface attack would make a proportionately large change in cross section, making the resistance rise:

$$R = \frac{\rho l}{A}$$

i.e., resistance is given by resistivity times length, divided by area, and we are looking for changes in ρ, not spurious changes in A.

Because of these considerations, platinum is the favorite material for accurate resistance thermometry. A further advantage is that platinum can be made very pure, allowing reproducible characteristics, even though resistivities are highly sus-

ceptible to small impurities. For less accurate, lower-temperature operation, nickel may be used.

It is important to avoid unpredictable stressing of the element—which would alter the resistance. Metalized film tracks on glass or ceramic are alternatives. The protective outer housing that is added means that, to outward appearance, there is not much difference between thermocouple and resistance transducers.

Conventional bridge and other circuits are used with resistance thermometers. Whereas with thermocouples it can be arranged that lead resistance does not play a large part, when the active element is itself a resistance, the lead resistance cannot be ignored, and unless the associated circuitry is very close, compensation must be introduced for variations in lead resistance due to unknown temperature changes along the length. It can be achieved by introducing one or two extra wires (according to the bridge configuration) in the cable that connects the transducer to the measuring circuit. A.C. circuits have the advantage that they avoid introducing errors from thermal e.m.f.s in leads or temperature elements but with them some care must be taken over stray inductance or capacitance in the transducer.

In the act of measurement a finite current flows through a resistance thermometer element. This generates heat and raises its temperature above that of its surroundings—which, of course, is what is really to be measured. This error, which depends on the thermal insulation of the element, can generally be neglected if the power dissipation is below 10 mW. Film-type resistors have the advantage of introducing particularly small errors of this type.

Semiconductors, with a different physical basis for their electrical conduction, can show a much greater change of resistance with temperature than metals do. This is exploited in thermistors, whose behavior is illustrated in Figure 6.64.

Thermistors can be supplied as beads or in rod, disk, washer or film form; they have the advantage that they can be very small. Their characteristic is given approximately by the law

$$\alpha = -B/T^2$$

where α is the temperature coefficient (ohms per Kelvin), B is a constant, and T is absolute temperature. Their high sensitivity makes thermistors attractive for many applications, though an individual's characteristics cannot be predicted to a tight tolerance and may show a drift equivalent to the order of 0.1 K over a year.

The characteristic of a different device, the switching thermistor, is also shown in Figure 6.64. It is used for protection purposes rather than continuous control. The switching device, unlike a conventional thermistor, has a positive temperature coefficient. Over a small span of temperature its resistance increases a thousandfold, so drastically reducing the current flowing through it. The critical temperature at which this happens can be chosen, for instance, to prevent electrical insulation being burned out.

Radiation Thermometers

Radiation thermometers (formerly known as pyrometers) are not based on any change of property with temperature but use the electromagnetic radiation from a body to be measured. As the body warms up, the total radiation it emits increases rapidly (with the fourth power of the absolute temperature) and the spectral distribution shifts to shorter wavelengths. The temperature can thus be determined by measuring the radiation, and there is the clear advantage that all the detecting

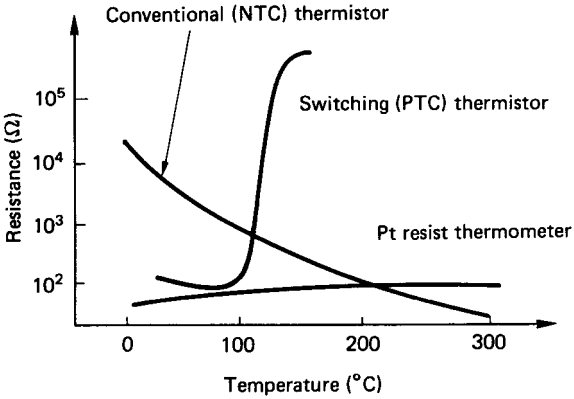


FIGURE 6.64 Thermistor characteristics.

equipment is remote from the hot body. Limitations to the technique are that it is more difficult to measure lower temperatures, where the energy emitted is much less, and that the emissivity of the radiating surface comes into the equation as well as its temperature.

In this type of thermometer the radiation is focused on a detector. A lens may be used for this purpose (it must be made of a material that transmits the appropriate radiation) or sometimes a mirror to give complete spectral coverage. A thermopile, consisting of a number of thermoelectric junctions connected in series to increase their output, may be used as detector. Alternatively, a pyroelectric device may be employed; in this, charges are liberated as the temperature changes. These latter devices do not respond to steady-state signals, so the radiation must be chopped, which is commonly effected by having a segmented disk rotating in its path. The semiconductor photodiode is another detector that is sometimes used at shorter wavelengths.

Surface emissivity is much less important when the radiation to be measured has emerged from a window in a hollow body. This makes the technique particularly applicable to furnaces. Dependence on emissivity is also reduced in the arrangement shown in Figure 6.65. If the reflectivity of the hemispherical mirror there ap-

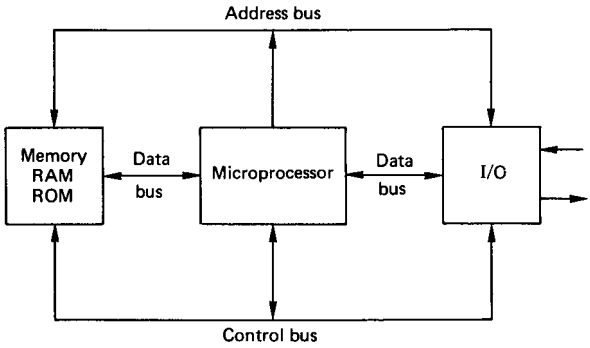


FIGURE 6.65 Arrangement to reduce emissivity dependence.

proaches unity, the effective emissivity of the surface also tends to unity, though with this set-up the advantage of having all equipment remote from the hot surface is, of course, sacrificed.

Sometimes the measurement is of total radiation, sometimes of that within a particular band of wavelengths, which are chosen from considerations of detector sensitivity and material transmission. Shorter wavelengths are appropriate for hotter bodies, longer for colder. By working in the far infrared (at 30 μm wavelength) temperatures as low as -50°C have been measured, but applications are much more common upwards from 50 to 100 K higher.

When, as is often the case, radiation thermometers are used in dusty atmospheres, an air purge will be desirable to keep the front optical surface clean. In some designs, the detector is kept further outside a hostile environment by using optical fibers as links.

Gas and Vapor Thermometers

In pneumatic instrumentation systems there is an advantage in having the temperature signal, not in an electrical form, but as something more immediately compatible with pneumatics. This is an attraction of gas and vapor thermometers.

With both of these a sealed bulb is situated where temperature is to be measured and connected by capillary tubing to a pressure-sensitive device. The bulb is either completely filled with a permanent gas or partially filled with a suitable liquid, which means that the pressure in the bulb changes with bulb temperature, according either to the gas laws or to the liquid's vapor pressure. Pressure measurements can thus indicate bulb temperature.

Practical Considerations

With such a multiplicity of methods for measuring temperature, the choice between them depends on many factors. The following line of thought may be helpful as a first, crude guide:

1. Use gas or vapor techniques if and only if they have to feed into a pneumatic system.
2. Failing them, use thermocouples (base metals for lower temperatures, noble metals for higher) unless
3. The highest accuracies are needed, when resistance elements have attractions or
4. Contact with the object studied is difficult or impossible, when radiation thermometers are the solution.
5. Remember that the large signals from thermistors may be an advantage if their limited range is acceptable.
6. For the highest temperatures, radiation techniques may prove much cheaper than thermocouples.

Speed of response is sometimes an important matter, but it depends more on details of construction than on the basic type of thermometer.

When deciding the details of an installation, several points should be borne in mind. The requirement is often to measure the temperature of a fluid, but when a

transducer is immersed in a fluid its equilibrium temperature is decided not only by the conductive and convective heat exchange with the fluid but also (if the fluid is transparent, as are most gases) by radiative heat exchange with the walls. Therefore, the transducer takes up a temperature that is intermediate between fluid temperature and wall temperature. The error so caused can be greatly reduced by introducing radiation shields. A further error arises when, as is normal, the transducer is mounted from the walls of the container, so providing a heat-conducting path through the mounting. The tendency to bring the transducer's temperature closer towards that of the walls is especially marked if it is housed inside a more permanent pocket (the American term is *Thermowell*) in order to facilitate replacement.

There can be sampling errors when measuring temperature. For instance, if the fluid of concern is flowing inside a pipe and has a temperature different from that of its surroundings, there will be a temperature profile which relates the local fluid temperature to its (radial) position in the pipe. As shown in Figure 6.66, at only one radius will the temperature correspond strictly to that of the mean. However, such errors can be positive or negative, and it is sometimes possible to make them offset the radiation and conduction errors referred to above.

6.26 BAR CODE READERS

A form of optical data input which is finding increasing application is the bar code reader. The code consists of a series of black and white vertical lines which are printed onto the object (Figure 6.67).

The code is read by an optical sensor which incorporates a lamp, a phototransistor, and a number of optical focusing lenses. The decoding software is, however, necessarily complex, since the speed at which the code is read can vary.

A typical bar code might consists of a start pattern, 101, five 7-bit characters, a check sequence, a second group of five characters and an end pattern. Two consecutive black bands represent a bit value of 1, while two consecutive white bands

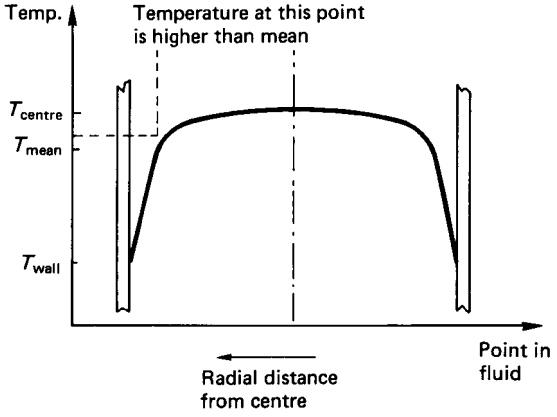


FIGURE 6.66 Errors caused by temperature profile.

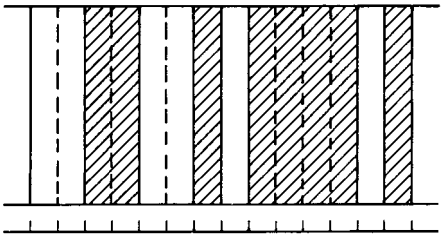


FIGURE 6.67 Section of a bar code.

represent a bit value of 0. The code is designed such that every character starts with a white band and ends with a black band. This ensures that every character starts and ends with a 1-0 transition. Additionally, every character code includes at least one 1-0 transition within the code.

The decoding program must include a timing loop to determine the speed of reading. The timing is usually based on a count of the 1-0 transitions. With the reading speed established, the code can then easily be translated. Bar code readers are very much evident in large supermarkets and libraries, but they have applications to manufacturing stock control and automatic assembly lines for component counting and identification purposes.

REFERENCES

Williams, A. B. 1981. Electronic Filter Design Handbook, McGraw-Hill, New York.

PART 6

CLASSICAL CONTROL THEORY AND PRACTICE

Charles J. Fraser

John S. Milne

6.27 INTRODUCTION

Control engineering is based on the linear systems analysis associated with the development of feedback theory. A control system is constituted as an interconnection between the components which make up the system. These individual components may be electrical, mechanical, hydraulic, pneumatic, thermal, or chemical in nature, and the well-designed control system will provide the best response of the complete system to external, time-dependent disturbances operating on the system. In the widest sense, the fundamentals of control engineering are also applicable to the dynamics of commercial enterprise, social, and political systems and other nonrigorously defined concepts. In the engineering context, however, control principles are more generally applied to much more tangible and recognizable systems and subsystems.

Invariably, the system to be controlled can be represented as a block diagram, as in Figure 6.68. The system is a group of physical components combined to perform a specific function. The variable controlled may be temperature, pressure, flow rate, liquid level, speed, voltage, position, or perhaps some combination of these. Analog (continuous) or digital (discrete) techniques may individually (or simultaneously) be employed to implement the desired control action. In more recent times the advances made in microelectronics have resulted in an emphasis towards digital techniques, and the majority of modern control systems are now microprocessor based.

Classification of Control Systems

Engineering control systems are classified according to their application, and these include the following:

1. *Servomechanisms:* Servomechanisms are control systems in which the controlled variable (or output) is a position or a speed. D.C. motors, stepper motor position control systems, and some linear actuators are the most commonly encountered examples of servomechanisms. These are especially prevalent in robotic arms and manipulators.

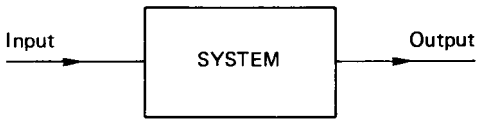


FIGURE 6.68 System to be controlled.

2. *Sequential control:* A system operating with sequential control is one where a set of prescribed operations is performed in sequence. The control may be implemented as *event based,* where the next action cannot be performed until the previous action is completed. An alternative mode of sequential control is termed *time based,* where the series of operations are sequenced with respect to time. Event-based sequential control is intrinsically a more reliable fail-safe mode than time based. Consider, for example, a process in which a tank is to be filled with a liquid and the liquid subsequently heated. The two control systems are depicted in Figure 6.69.

The time-based sequential control system is the simplest. The pump is switched on for an interval which would discharge enough liquid into the tank to fill it to approximately the correct level. Following this, the pump is switched off and the heater is switched on. Heating is similarly allowed to continue for a preset time, after which the liquid temperature would approximately have reached the desired value. Note that the control function is inexact and there are no fail-safe features. If the drive shaft between the motor and the pump becomes disengaged or broken, the heater will still come on at the prescribed time, irrespective of whether there is liquid in the tank or not. The event-based sequential control system has fail-safe features built in and is much more exact. In operation the pump is switched on until the liquid-level sensor indicates that the tank is filled. Then (and only then) is the pump switched off and the heater switched on. The temperature of the liquid is also monitored with a sensor and heating is applied until such time that the temperature reaches the desired value.

Obviously, with two additional sensors, the event-based system is the more expensive. The advantages it offers over the time-based system, however, far outweighs its disadvantages, and event-based sequentially controlled systems are by far the most common. Time-based systems do exist, nonetheless, and they are found in applications where the results of malfunction would be far less potentially catastrophic than those occurring in the example described. The essential difference between the two systems is that event-based sequential control incorporates a check that any operation has been completed before the next is allowed to proceed. The modern automatic washing machine and automatic dishwasher are good examples of sequentially controlled systems.

3. *Numerical control:* In a system using numerical control the numerical information, in the form of digital codes, is stored on a control medium, which may

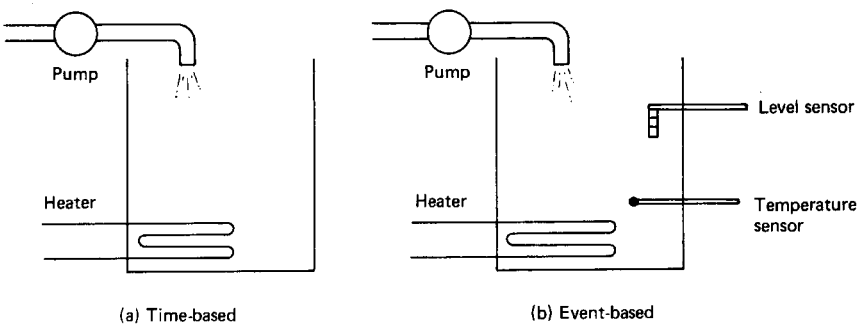


FIGURE 6.69 Simple sequential control systems.

be a magnetic sensitive tape or a magnetic sensitive disk. This information is used to operate the system in order to control such variables as position, direction, velocity, and speed. A large variety of manufacturing operations involving machine tools utilize this versatile method of control.

4. *Process control:* In this type of control the variables associated with any process are monitored and subsequent control actions are implemented to maintain the variables within the predetermined process constraints. The word "process" is all-encompassing and might include, for example, electrical power generation. The generation of electricity can be considered as a manufacturing process where the product is kilowatt hours. In the control of power generation, the variables which are measured include temperature, pressure, liquid-level, speed, flow-rate, voltage, current, and a range of various gas concentrations. This is further complicated by the need to satisfy the power demand, and it is apparent that the control of such a system is necessarily complex. Similarly complex examples exist in the oil and paper-making industries, in aerospace assembly plants and in any entity which aspires to the designation of a flexible manufacturing system.

Open- and Closed-Loop Control

The basic open-loop system is shown in Figure 6.68 and is extended in Figure 6.70 to illustrate a more complete picture. The input element supplies information regarding the desired value, X, of the controlled variable. This information is then acted on by the controller to alter the output, Y.

External disturbances are fed in as shown and will cause the output to vary from the desired value. The open-loop system may be likened to the driving of a vehicle where the driver constitutes the input element. Essentially, two variables are controlled by the driver—the speed and the direction of motion of the vehicle. The controller, in the case of speed, is the engine throttle valve and in the case of direction, is the steering system.

In order that the system become closed-loop, two further elements must be added:

1. A monitoring element, to measure the output, Y
2. A comparing element, to measure the difference between the actual output and the desired value, X

The monitoring and comparing elements are connected through the feedback link as shown in Figure 6.71.

It can be argued that the driver in the previous example also performs the functions of monitoring and comparing. The vehicle driver, therefore, if considered to

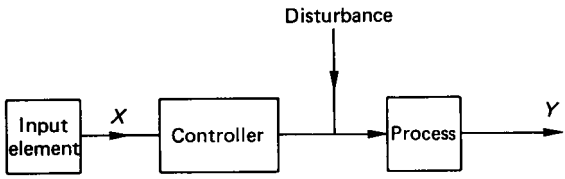


FIGURE 6.70 Open-loop control system.

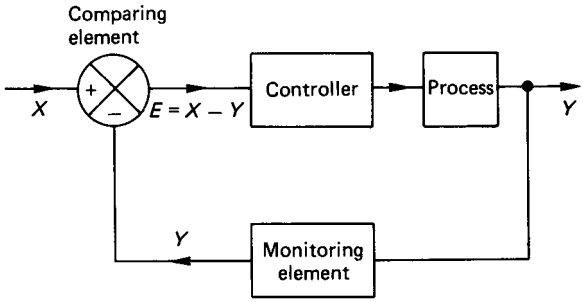


FIGURE 6.71 Closed-loop feedback control system.

be part of the complete system, constitutes a closed-loop feedback control. For the purpose of definition, however, any system which incorporates some form of feedback is termed closed-loop. With no feedback mechanism, the system is categorized as open-loop. For the most practical engineering purposes, control systems are of the closed-loop variety to take advantage of the benefits of feedback, which may be either positive or negative. A positive feedback signal aids the input signal. It is possible therefore to have output with no input when using a positive feedback signal and, since this is detrimental to the control function, positive feedback systems are very rare.

Linear and Nonlinear Control Systems

For a control system to be linear it must satisfy both the amplitude proportionality criteria and the principle of superposition. If a system output at a given time is $Y(t)$ for a given input $X(t)$, then an input of $kX(t)$ must produce an output of $kY(t)$ if amplitude proportionality is satisfied. Similarly, if an input of $X_1(t)$ produces an output of $Y_1(t)$, while an input of $X_2(t)$ produces an output of $Y_2(t)$, then if an input of $(X_1(t) + X_2(t))$ produces an output of $(Y_1(t) + Y_2(t))$ the superposition principle is satisfied. Nonlinear systems do not necessarily satisfy both these criteria, and generally these systems are compensated such that their behavior approaches that of an equivalent linear system.

Characteristics of Control Systems

The characteristics of a control system are related to the output behavior of the system in response to any given input. The parameters used to define the control system's characteristics are stability, accuracy, speed of response, and sensitivity.

The system is said to be stable if the output attains a certain value in a finite interval after the input has undergone a change. When the output reaches a constant value the system is said to be in steady state. The system is unstable if the output increases with time. In any practical control system, stability is absolutely essential. Systems involving a time delay or a dead time may tend to be unstable, and extra care must be taken in their design to ensure stability. The stability of control systems can be analyzed using various analytical and graphical techniques. These include the Routh–Hurwitz criteria and the Bode, Nichols, and Nyquist graphical methods.

The accuracy of a system is a measure of the deviation of the actual controlled value in relation to its desired value. Accuracy and stability are interactive, and one can in fact be counterproductive to the other. The accuracy of a system might be improved, but in refining the limits of the desired output the stability of the system might be adversely affected. The converse also applies.

The speed of response is a measure of how quickly the output attains a steady-state value after the input has been altered.

Sensitivity is an important factor and is a measure of how the system output responds to external environmental conditions. Ideally, the output should be a function only of the input and should not be influenced by undesirable extraneous signals.

Dynamic Performance of Systems

The dynamic performance of a control system is assessed by mathematically modeling (or experimentally measuring) the output of the system in response to a particular set of test input conditions:

1. *Step input:* This is perhaps the most important test input, since a system which is stable to a step input will also be stable under any of the other forms of input. The step input (Figure 6.72) is applied to gauge the transient response of the system and gives a measure of how the system can cope with a sudden change in the input.
2. *Ramp input:* A ramp input (Figure 6.73) is used to indicate the steady-state error in a system attempting to follow a linearly increasing input.
3. *Sinusoidal input:* The sinusoidal input (Figure 6.74) over a varying range of input frequencies is the standard test input used to determine the frequency response characteristics of the system.

Although the three standard test inputs may not be strict representations of the actual inputs to which the system will be subject, they do cover a comprehensive range. A system which performs satisfactorily under these inputs will, in general, perform well under a more natural range of inputs. The system response to a par-

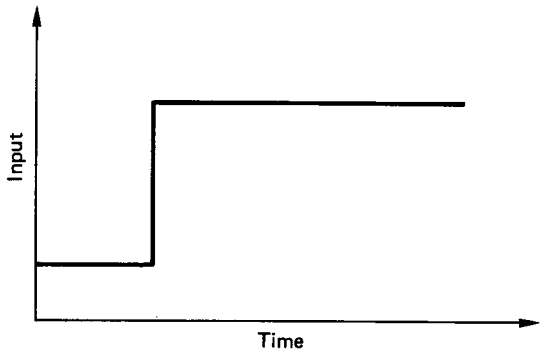


FIGURE 6.72 Step input.

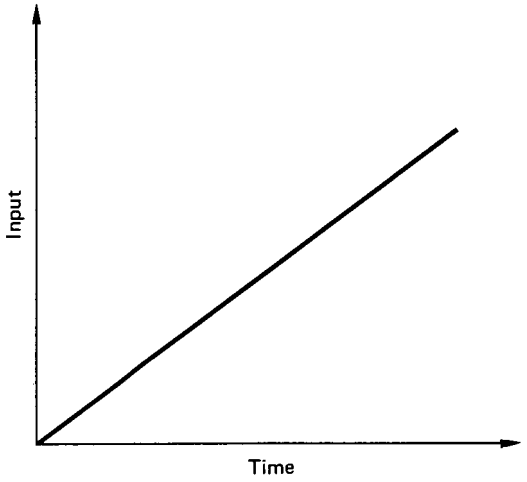


FIGURE 6.73 Ramp input.

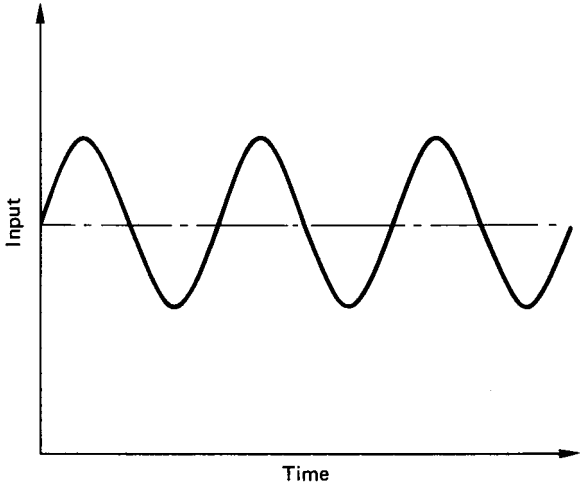


FIGURE 6.74 Sinusoidal input.

abolically varying test input can also be analyzed or measured, but this is a less commonly used test signal compared to the previous three.

Time Domain and Frequency Domain

The time domain model of a system results in an output $Y(t)$ with respect to time, for an input $X(t)$. The system model is expressed as a differential equation, the solution of which is displayed as a graph of output against time.

In contrast, a frequency domain model describes the system in terms of the effect that the system has on the amplitude and phase of sinusoidal inputs. Typically, the system performance is displayed in plots of amplitude ratio, $(Y(t)/X(t))$ or $20 \log_{10} (Y(t)/X(t))$, and phase angle, against input signal frequency.

Neither system model has an overriding advantage over the other, and both are used to good effect in describing system performance and behavior.

6.28 MATHEMATICAL MODELS OF SYSTEMS—TIME DOMAIN ANALYSIS

Differential equations are used to model the relationship between the input and output of a system. The most widely used models in control engineering are based on first- or second-order linear differential equations.

First-Order Systems

Some simple control systems (which includes the control of temperature, level, and speed) can be modeled as a first-order linear differential equation:

$$\tau \frac{dY}{dt} + Y = kX \tag{6.5}$$

where X and Y are the input and output, respectively. τ denotes the system time constant and k is the system gain. When the input X is a step of amplitude A then the solution to equation (6.5) gives the result shown in Figure 6.75. The solution curve shown in this figure has the analytical form

$$Y(t) = kA[1 - e^{-t/\tau}] \tag{6.6}$$

Equation (6.6), which is the time-domain solution, is an exponential function which approaches the value (kA) as t approaches infinity. Theoretically, the output

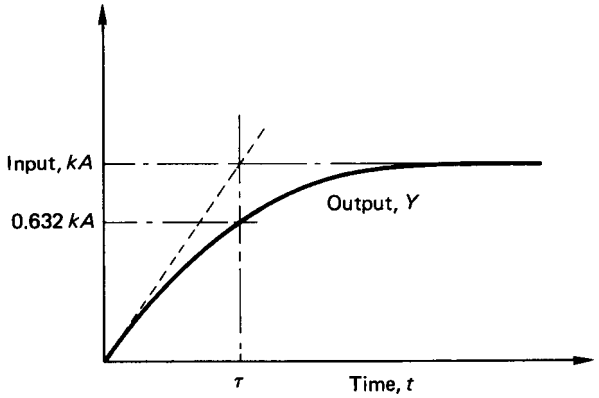


FIGURE 6.75 Response of a first-order system to a step input.

never reaches (kA) and the response is termed an exponential lag. The time constant τ represents the time which the output would take to reach the value (kA) if the initial rate of response were maintained. This is indicated by the broken line which is tangent to the solution curve at time, $t = 0$. For practical purposes the final steady-state output is taken to have been reached in a time of about (5τ).

If the input is a ramp function then the response of a first-order system is as shown in Figure 6.76. The ramp input is simulated by making the right-hand side of equation (6.5) a linear function of time, i.e., kAt. With this input, the time domain solution becomes

$$Y(t) = kA[t - \tau(1 - e^{-t/\tau})] \tag{6.7}$$

The solution equation shows that as t becomes large the output tends to $kA(t - \tau)$. The output response is asymptotic therefore to a steady-state lag ($kA\tau$).

The response of a first-order system to a sinusoidal input can be obtained by setting the right-hand side of equation (6.5) equal to $kA \sin(\omega t)$, where ω is a constant circular frequency in radians/second. The time-domain solution yields

$$Y(t) = \frac{kA}{\sqrt{(1 + \tau^2\omega^2)}} [\sin\alpha \cdot e^{-t/\tau} + \sin(\omega t - \alpha)] \tag{6.8}$$

where $\alpha = \tan^{-1}(\tau\omega)$. The response is shown in Figure 6.77. The output response exhibits a decaying transient amplitude in combination with a steady-state sinusoidal behavior of amplitude, $kA/[\sqrt{(1 + \tau^2\omega^2)}]$ and lagging the input by the angle α.

Second-Order Systems

While some control systems may be adequately modeled as a first-order linear differential equation, many more practical systems, including position control, are

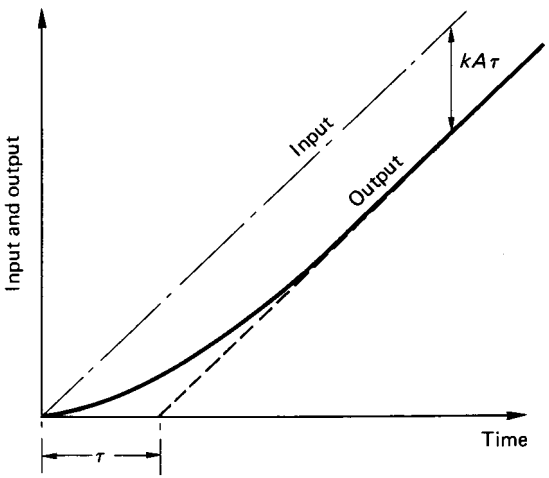


FIGURE 6.76 First-order system response to a ramp input.

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

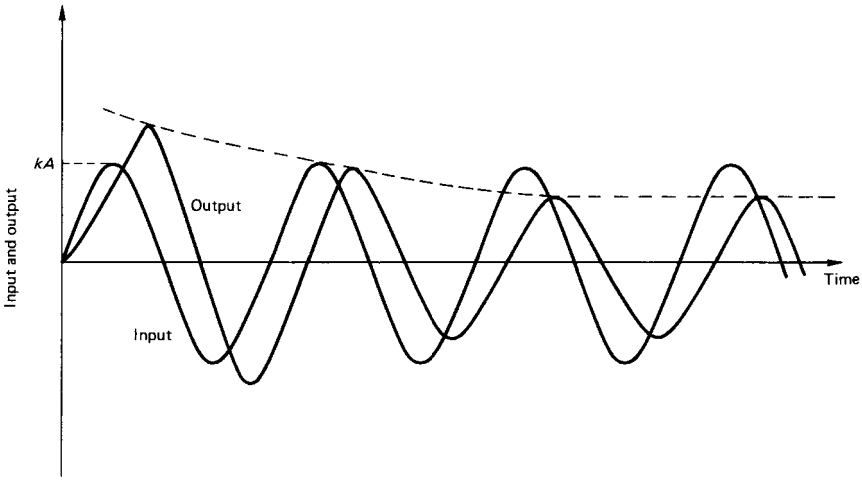


FIGURE 6.77 First-order system response to a sinusoidal input.

more conformably represented by a differential equation of the second order. The second-order differential equation has the general form:

$$\frac{d^2Y}{dt^2} + 2\zeta\omega_n \frac{dY}{dt} + \omega_n^2 Y = kX \tag{6.9}$$

where ζ is termed the damping ratio and is defined as the ratio of the actual damping in the system to that which would produce critical damping. ω_n is the undamped natural frequency of the system and k, again, is the system gain.

The time-domain solution depends on the magnitude of ζ and three solutions for a step input are possible:

1. *Light damping,* $\zeta < 1$.

$$Y(t) = \exp(-\zeta\omega_n t)[A \cos(\omega_n \sqrt{(1 - \zeta^2)}t) + B \sin(\omega_n \sqrt{(1 - \zeta^2)} \cdot t)] \tag{6.10}$$

2. *Critical damping,* $\zeta = 1$.

$$Y(t) = \exp(-\zeta\omega_n t[At + B] \tag{6.11}$$

3. *Heavy damping,* $\zeta > 1$.

$$Y(t) = A \exp\{-\zeta\omega_n t - [\omega_n \sqrt{(\zeta^2 - 1]t}\} + B \exp\{-\zeta\omega_n t + [\omega_n \sqrt{(\zeta^2 - 1)]t}\} \tag{6.12}$$

The three possible solutions for a step input are shown in Figure 6.78. The output, $Y(t)$, is plotted as a percentage of the step input, X, against the parameter, $(\omega_n t)$.

For ζ equal to unity, the system is critically damped and the steady-state value is attained in the shortest possible time without any oscillatory response. With ζ

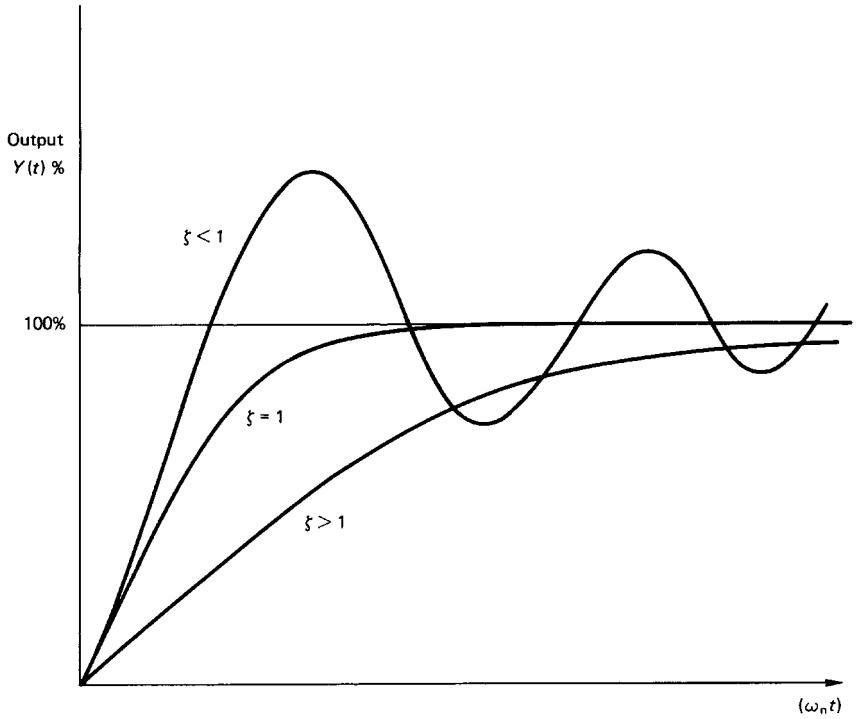


FIGURE 6.78 Response of a second-order system to a step input.

greater than unity, the system is overdamped and the response curve is again exponential in form. Overdamped systems may have an undesirably sluggish response. Indeed, since the effect of $\zeta > 1$ simply delays the response to the steady-state value there is no real advantage to be gained in using high ζ values.

For cases where ζ is less than unity, the system is said to be underdamped and the response curve is oscillatory with an exponential decay. A number of performance measures are used to describe the response of an underdamped system to a step input, and these are illustrated in Figure 6.79.

The speed of the response is reflected in the rise time, τ_R, and the peak time, τ_P. For underdamped systems, the rise time is the time taken for the output to reach 100% of the step input. The peak time is that taken to the first maximum in the output response. For critically damped and overdamped systems, the time taken for the output to change between 10% and 90% of the input is used alternatively as a measure of the speed of the response.

The degree in which the actual output response matches the input is measured by the percentage overshoot, *PO*, and the settling time. The percentage overshoot is defined as

$$PO = \frac{M_{PT} - 100}{100} \tag{6.13}$$

where M_{PT} is the peak value of the output.

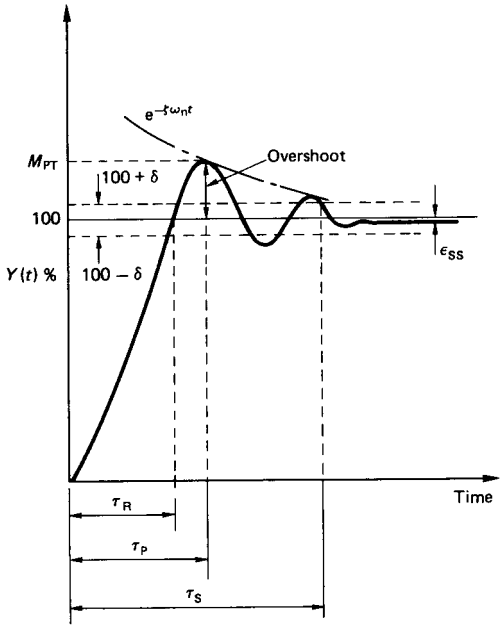


FIGURE 6.79 Response curve for an underdamped system to a step input.

It may further be shown that the percentage overshoot is given analytically as

$$PO = 100\, e^{-\zeta\pi/\sqrt{(1-\zeta^2)}} \tag{6.14}$$

Another useful relation is derived from the ratio of successive peaks, i.e.,

$$\ln\left(\frac{A_n}{A_{n+1}}\right) = \frac{2\pi\zeta}{\sqrt{(1-\zeta^2)}} \tag{6.15}$$

where n is an integer to denote the peak number (i.e., first, second, etc.). Equation (6.15) is referred to as the *logarithmic decrement*. The settling time, τ_S, is the time taken for the oscillatory response to decay below a percentage of the input amplitude, δ, often taken as $\pm 2\%$. Finally we have the steady-state error, ε_{SS}, which is self-explanatory.

The response of the second-order system to a ramp input is shown in Figure 6.80. The form of the response curves again depends on the value of the damping ratio, but in each case the output asymptotes to a steady-state lag. The lag is not the same in each case, however, since this is also dependent on the damping ratio.

The response of a second-order system to a sinusoidal input may also be considered. Generally, the output response will lag behind the input with a transient decaying amplitude depending on the nature of the damping ratio. It is more informative, however, to study the response of second-order systems to sinusoidal inputs using frequency-domain methods.

6.29 LAPLACE NOTATION FOR DIFFERENTIAL EQUATIONS—FREQUENCY-DOMAIN ANALYSIS

For analyses in the frequency domain it is customary to write the differential equation in terms of the Laplace operator, s. This gives rise to the system *transfer function,* which is formed by replacing the input and output (X and Y, respectively) with their corresponding Laplace transforms, $X(s)$ and $Y(s)$. The method applies only to linear differential equations. In practice, many systems would contain some degree of nonlinearity, and various assumptions would have to be made to simplify and approximately linearize the governing equation.

The advantage in using the Laplace transform method is that it allows the differential equation to be expressed as an equivalent algebraic relation in s. Differentiation is represented by multiplication with the Laplace variable, s. Thus dY/dt becomes $sY(s)$ and d^2Y/dt^2 is replaced with $s^2Y(s)$.

First-Order Systems

The governing equation is rewritten with the appropriate Laplace transforms replacing the differential operators. Thus, equation (6.5) becomes

$$\tau \cdot s \cdot Y(s) + Y(s) = kX(s) \tag{6.16}$$

Hence,

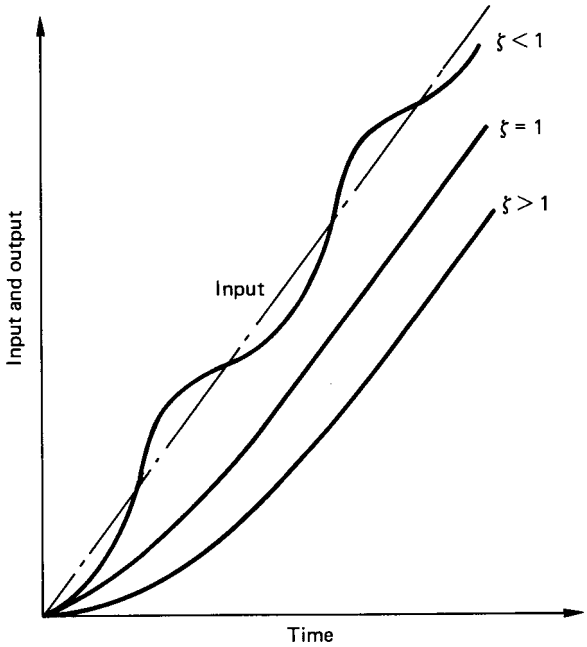


FIGURE 6.80 Response of a second-order system to a ramp input.

$$Y(s)[1 + \tau s] = kX(s) \tag{6.17}$$

The system transfer function is defined as the ratio of the output to the input and is

$$\frac{Y(s)}{X(s)} = \frac{k}{1 + \tau s} \tag{6.18}$$

Equation (6.18) enables the convenient facility of incorporating the transfer function within the usual block structure representation of a control system. Thus a first-order, open-loop control system can be systematically depicted as shown in Figure 6.81.

For analyses in the frequency domain we are predominantly concerned with the system response to sinusoidal inputs. Differentiation (or integration) of a sinusoidal function does not alter the shape or frequency. There is simply a change in amplitude and phase, e.g.,

$$\text{Input} = A \sin(\omega t) \tag{6.19}$$

$$\frac{d}{dt}(A \sin(\omega t)) = \omega A \cos(\omega t)$$

$$= \omega A \sin\left(\omega t + \frac{\pi}{2}\right) \tag{6.20}$$

Comparing equations (6.19) and (6.20) it can be seen that differentiation has changed the amplitude from A to ωA and that there is a phase shift of 90° associated with the process.

Equation (6.20), in fact, describes the steady-state output from a first-order, open-loop control system. The transient part of the output, which the time-domain solution illustrated in Figure 6.77, is not apparent in the frequency-domain solution.

The Laplace operator, s, may be replaced with $j\omega$, where j is the complex operator $\sqrt{-1}$. Equation (6.18) then becomes

$$\frac{Y}{X} = \frac{k}{1 + j\omega\tau}$$

Using the complex conjugate it may be shown that the modulus of the amplitude ratio is

$$\left|\frac{Y}{X}\right| = \frac{k}{\sqrt{(1 + \omega^2\tau^2)}} \tag{6.21}$$

Equation (6.21) shows how the output amplitude will be influenced by the input sinusoidal frequency. Note that this agrees with the time domain solution (equation (6.8)) for large values of t after which the steady state is achieved.

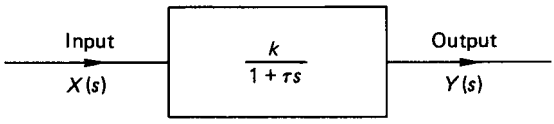


FIGURE 6.81 First-order open-loop control system.

The technique shown above is general and may be used to determine the amplitude ratio for any second- or higher-order system. *Note:* Common practice, especially in graphical representations, is to express the amplitude ratio in decibels, i.e.,

$$\text{dB} = 20 \log_{10}(Y/X) \tag{6.22}$$

Second-Order Systems

Using the Laplace transfer operator the governing equation (6.9) may be rewritten as

$$s^2 \cdot Y(s) + 2\zeta\omega_n \cdot s \cdot Y(s) + \omega^2{}_n \cdot Y(s) = kX(s) \tag{6.23}$$

$$\therefore Y(s)[s^2 + 2\zeta\omega_n \cdot s + \omega_n^2] = kW(s) \tag{6.24}$$

The system transfer function is

$$\frac{Y(s)}{X(s)} = \frac{k}{s^2 + 2\zeta\omega_n \cdot s + \omega_n^2} \tag{6.25}$$

Thus, a second-order, open-loop control system can be represented schematically as in Figure 6.82.

For a sinusoidal input of the form $X = X_0 \sin(\omega t)$, the frequency-domain analysis gives the following steady-state solutions for the amplitude ratio and the phase lag:

$$\left|\frac{Y}{X}\right| = \frac{1}{\sqrt{[(1 - r^2)^2 + (2\zeta r)^2]}} \tag{6.26}$$

and

$$\phi = \tan^{-1}[(2\zeta r)/(1 - r^2)] \tag{6.27}$$

where $r = (\omega/\omega_n)$.

The above frequency response characteristics are shown in Figure 6.83. When the input signal frequency is equal to the system's natural frequency the amplitude ratio has the value of $(1/2\zeta$ and the phase lag is $-90°$. Note that if the damping ratio is zero the amplitude ratio theoretically approaches infinity under this resonance condition. In practice, if the damping ratio is moderately low, very large output amplitudes can be expected if the input frequency is in the vicinity of the system natural frequency.

Thus far we have considered the open-loop system response for first- and second-order systems. Such systems are unconditionally stable. The addition of a feedback loop, however, increases the order of the system and there is always the

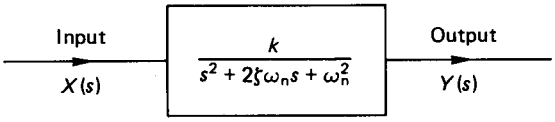


FIGURE 6.82 Second-order open-loop control system.

possibility that the second-order system with feedback may be unstable. Furthermore, if any system, first or second order, incorporates a time delay (also known as a dead time or a transportation lag), then unstable operation is more likely to occur.

6.30 STABILITY CRITERIA

Time delays are very difficult to handle mathematically when they occur in differential equations, and the inclusion of multiple feedback loops can greatly increase the order of the governing equation. For these two reasons solutions in the time domain become extremely difficult, and frequency domain methods are almost ex-

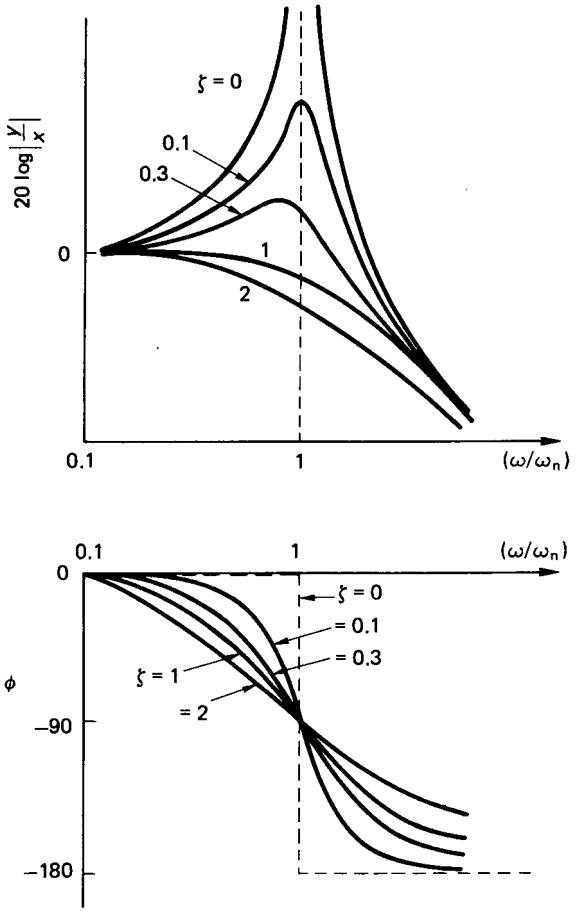


FIGURE 6.83 Second-order open-loop control system frequency response.

clusively used to assess the behavior of the more complex control systems. The main consideration in frequency-domain analyses is the stability of the system and how it can be adjusted if it happens to be unstable. Various graphical methods are used and these include the Bode and Nyquist plots.

The Bode plot is a graph of amplitude ratio and phase angle variation with input signal frequency. The resulting normalized plot for an open-loop first-order system is shown in Figure 6.84. Note that when the input frequency is equal to the inverse of the system time constant, the output amplitude has been decreased (or attenuated) by 3 dB. The phase lag at this point is −45°. This is characteristic of first-order systems.

The Nyquist plot represents the same information in an alternative form. The plot is in polar coordinates and combines the amplitude ratio and phase lag in a single diagram. Figure 6.85 shows the Nyquist plot for the open-loop, first-order system.

Bode and Nyquist Stability Criteria

The Bode (1940) criterion for stability is that the system is stable if the amplitude ratio is less than 0 dB when the phase angle is −180°. This is illustrated graphically

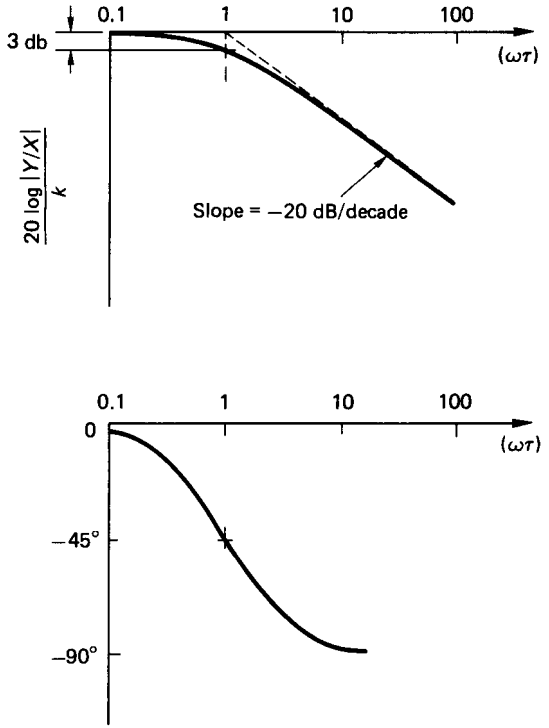


FIGURE 6.84 Bode plot for an open-loop first-order system.

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

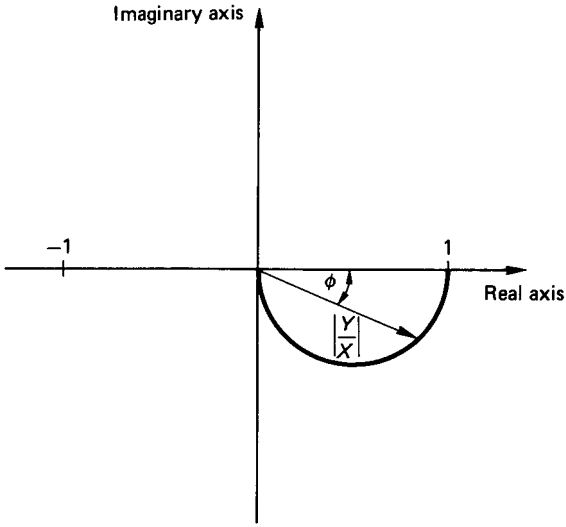


FIGURE 6.85 Nyquist plot for an open-loop first-order system.

in Figure 6.86. This figure represents a stable system since Bode's criterion is satisfied. The gain margin (GM) and phase margin (PM) are used as measures of how close the frequency response curves are to the 0 dB and −180° points and are indicative of the relative stability of the system.

The Nyquist (1932) criterion for stability is that the system is stable if the amplitude ratio is greater than −1 at a phase angle of −180°. In effect, this means that the locus of the plot of amplitude ratio and phase angle must not enclose the point −1 on the real axis. A stable response curve is shown plotted in Figure 6.87. Also indicated in this figure are the gain margin and phase margin in the context of the Nyquist plot.

System Stability with Feedback

In a closed-loop system the transfer function becomes modified by the feedback loop. The first task therefore is to determine the overall transfer function for the complete system.

For simple open-loop systems the transfer functions are combined according to the following rules:

1. For elements in series, the overall transfer function is given by the product of the individual transfer functions (see Figure 6.88).
2. For elements in parallel, the overall transfer function is given by the sum of the individual transfer functions (see Figure 6.89).

For a system with feedback, the overall transfer function can be evaluated using a consistent step-by-step procedure. Series and parallel control elements are com-

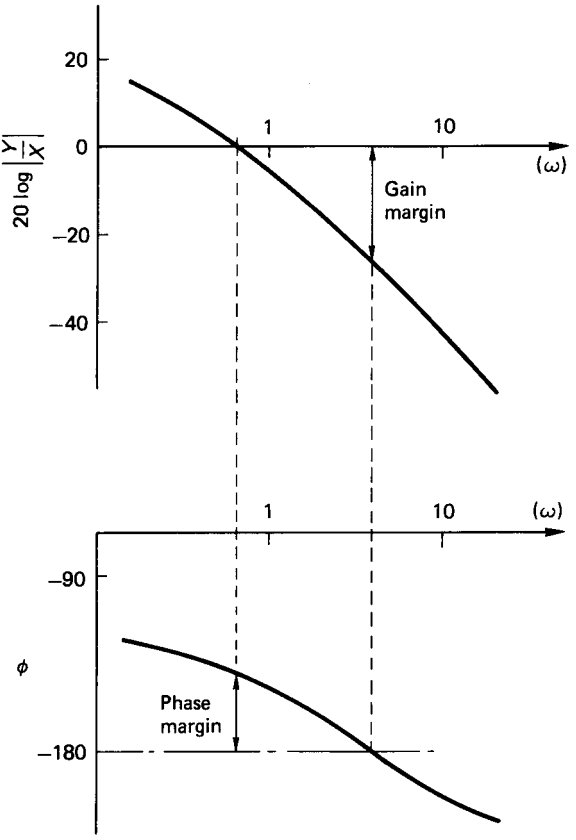


FIGURE 6.86 Bode's stability criterion.

bined in the manner as shown above to reduce the system to a single block, which then represents the overall transfer function.

Consider the simple control system depicted in Figure 6.90. Since the feedback line does not include any transfer function it is termed a *unity feedback* system, i.e., the output is compared directly with the input to produce the error signal. The closed-loop transfer function is obtained as follows:

$$Y = G(s) \cdot E = G(s)[X - Y]$$

Thus,

$$\frac{Y}{X} = \frac{G(s)}{1 + G(s)} \tag{6.28}$$

If the element whose open-loop transfer function, $G(s)$, is a first-order subsystem, then $G(s)$ may be replaced with the expression given in equation (6.18). The closed-loop transfer function may then be written as

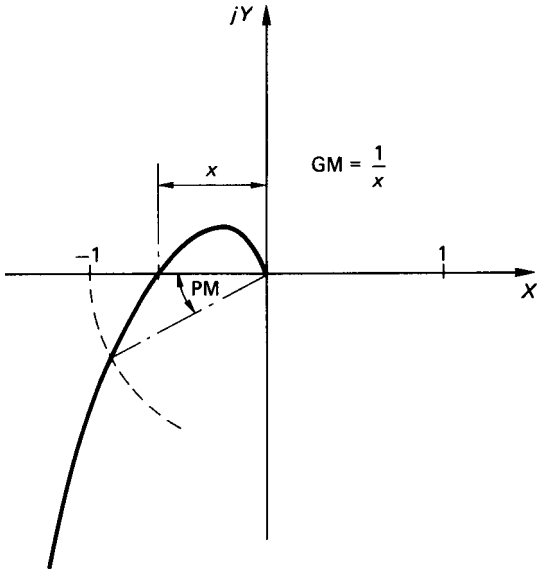


FIGURE 6.87 Nyquist's stability criterion.

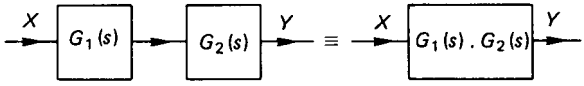


FIGURE 6.88 Transfer functions in series.

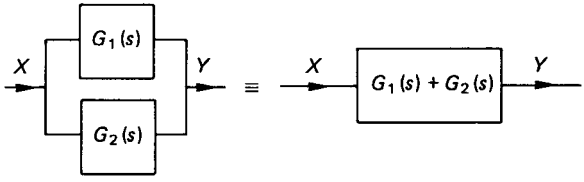


FIGURE 6.89 Transfer functions in parallel.

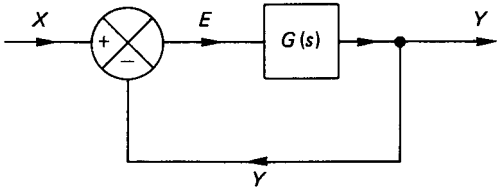


FIGURE 6.90 Control system with unity feedback.

$$\frac{Y}{X} = \left(\frac{k}{1 + \tau s}\right) \Big/ \left(1 + \frac{k}{1 + \tau s}\right)$$

$$= \frac{k}{1 + k + \tau s} \tag{6.29}$$

Dividing top and bottom by $(1 + k)$ results in

$$\frac{Y}{X} = \left(\frac{k}{1 + k}\right) \Big/ \left(1 + \frac{\tau s}{1 + k}\right) \tag{6.30}$$

We define the following terms as

$$k_c = \frac{k}{1 + k} \tag{6.31}$$

and

$$\tau_c s = \frac{\tau s}{1 + k} \tag{6.32}$$

where k_c is the closed-loop gain and τ_c is the closed-loop system time constant. The final closed-loop transfer function may be expressed as

$$\frac{Y}{X} = \frac{k_c}{1 + \tau_c \cdot s} \tag{6.33}$$

Equations (6.31) and (6.32) show, respectively, that both the closed-loop system gain and the time constant are less than those associated with the open-loop system. This means that the closed-loop response is faster than the open-loop one. At the same time, however, the closed-loop gain is reduced.

Using the procedures outlined in the example, any other complex control system may be similarly analyzed to determine the closed-loop transfer function. Thus, knowing the gain constants and other characteristics of the elements which make up the system, the frequency response may be obtained. The stability of the system may then be assessed and any corrective measures taken as necessary. In practice, it is often found that the gain of some of the system elements must be altered in order to ensure stable operation. Another commonly applied corrective measure is to add a phase advance circuit into the system. The procedure might also be operated in reverse, where, starting with a desired response, a suitable control system can be configured and adjusted to meet the response.

The practising control engineer will use many techniques to assess system stability. These might include the numerical Routh–Hurwitz criterion, which determines only whether a system is stable or not. Alternative graphical methods include the use of Hall charts, Nichols charts, inverse Nyquist plots, and root locus plots. The graphical methods additionally indicate the relative stability of a system.

Numerous commercial computer packages are available (Golten & Verwer Partners 1988; Cambridge Control Ltd. 1988; Arthur F. Saunders 1989) to assist the designer of control systems. These include the usual graphical representations and can be obtained from the suppliers whose addresses are given in the references. The reader is also referred to the Further Reading at the end of this Section for a more comprehensive coverage of these methods and techniques.

Effect of Transport Delay

The influence of a transport (or time) delay on the response of an underdamped second-order system to a step input is shown in Figure 6.91. Although it is virtually impossible to account for a time delay in a differential equation, it is simply accommodated in the frequency-domain model as an additional element in the system block diagram. In the frequency-domain model, the time delay effects a phase shift of $-\omega T$ and can be expressed as

$$\text{Time delay} = e^{-sT} = e^{-j\omega T} \tag{6.34}$$

Consider the open-loop response of a first-order system incorporating a time delay as illustrated in Figure 6.92. The open-loop transfer function becomes

$$\frac{Y}{X} = \frac{Y(s)}{X(s)} = G(s) = \frac{ke^{-sT}}{1 + \tau s} \tag{6.35}$$

$$\text{Amplitude ratio} = \left|\frac{Y}{X}\right| = \frac{k}{\sqrt{(1 + \omega^2\tau^2)}} \tag{6.36}$$

$$\text{Phase lag} = \phi = -[\omega T + \tan^{-1}(\omega\tau)] \tag{6.37}$$

The first-order, open-loop system response with a time delay is shown in Figure 6.93.

Similarly, a time delay in a system with feedback does not alter the amplitude ratio but adds a phase shift to the frequency response. The system therefore becomes less stable, and in some cases it may be necessary to reduce the closed-loop gain in order to obtain a stable response. The penalty to be paid for increasing the stability in this manner is an increase in the steady-state error.

6.31 CONTROL STRATEGIES

The basic closed-loop system with common symbol representation is given in Figure 6.94. The nomenclature used in this figure is defined as follows:

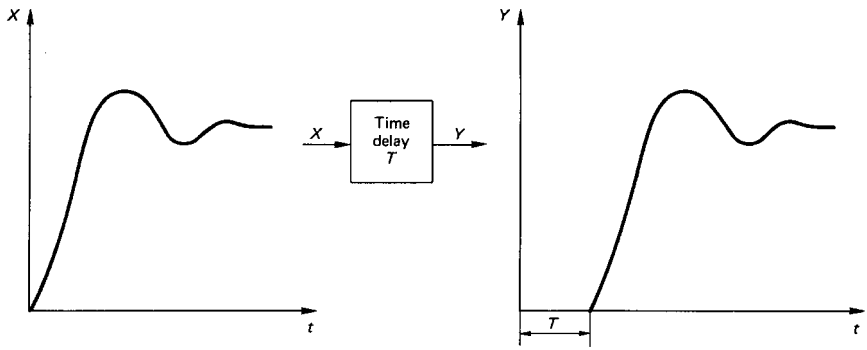


FIGURE 6.91 Effect of transport delay.

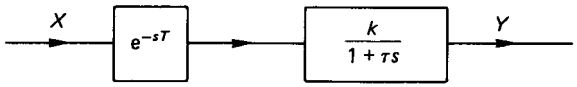


FIGURE 6.92 First-order system with a time delay.

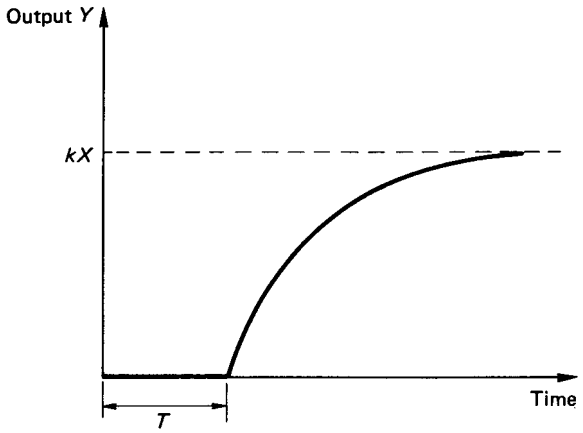


FIGURE 6.93 First-order open-loop system response with a time delay.

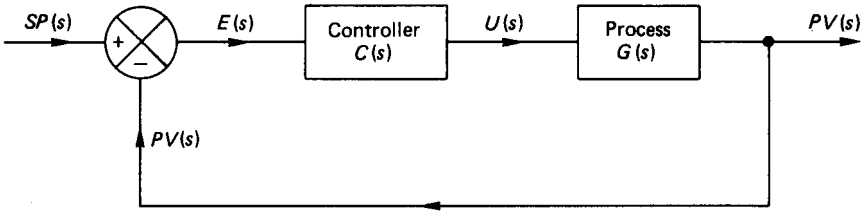


FIGURE 6.94 Basic closed-loop control system.

$SP(s)$ is the set point (required value, $r(t)$, is sometimes used).
$PV(s)$ is the process value (corrected value, $c(t)$, is sometimes used).
$E(s)$ is the error signal, which is the difference between SP and PV.
$U(s)$ is the control effort output from the controller to the process.
$C(s)$ is the controller transfer function.
$G(s)$ is the process transfer function.

The transfer function for the closed-loop system is obtained as before:

$$\begin{aligned} PV(s) &= C(s) \cdot G(s) \cdot E(s) \\ &= C(s) \cdot G(s) \cdot [SP(s) - PV(s)] \end{aligned} \tag{6.38}$$

Hence,

$$\frac{PV(s)}{SP(s)} = \frac{C(s) \cdot G(s)}{1 + C(s) \cdot G(s)} \tag{6.39}$$

ON/OFF Control

In many applications a simple ON/OFF strategy is perfectly adequate to control the output variable within preset limits. The ON/OFF control action results in either full or zero power being applied to the process under control. A mechanical type of thermostat provides a good example of an ON/OFF-based controller. The ON/OFF control strategy results in an output which fluctuates about the set point as illustrated in Figure 6.95.

ON/OFF controllers usually incorporate a dead band over which no control action is applied, which is necessary to limit the frequency of switching between the ON and OFF states. For example, in a temperature-control system, the ON/OFF control strategy would be:

If temperature $< T_{min}$, the heater is to be switched ON;

If temperature $> T_{max}$, then heater is to be switched OFF.

The dead band in the above case is $(T_{max} - T_{min})$ and while the temperature remains within the dead band no switching will occur. A large dead band will result in a correspondingly large fluctuation of the process value about the set point. Reducing the dead band will decrease the level of fluctuation but will increase the frequency of switching. The simple ON/OFF control strategy is mostly applicable to processes

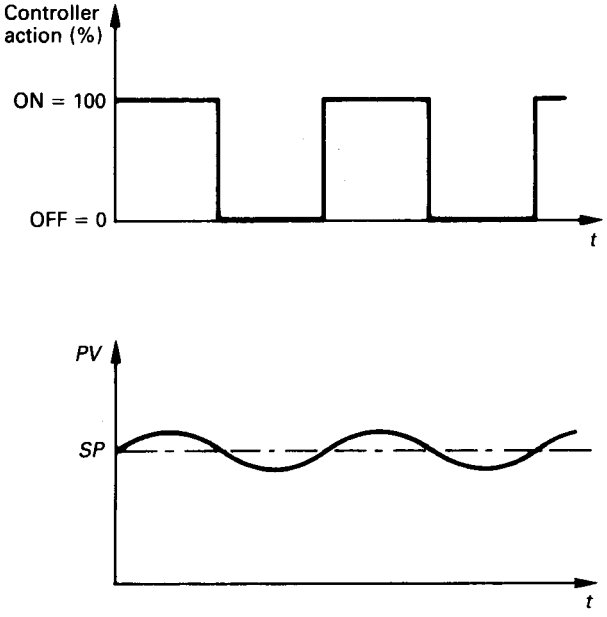


FIGURE 6.95 Output variation with ON/OFF control.

and systems which have long time constants and in consequence have relatively slow response times (e.g., temperature and level control).

While simple in concept, ON/OFF control systems are in fact, highly nonlinear and require some complex nonlinear techniques to investigate their stability characteristics.

Three-Term or PID Control

Since complicated transfer functions can be very difficult to model, the most common strategy used to define the controller transfer function is the so-called three-term or PID controller. PID is the popular short form for Proportional, Integral, and Derivative. The three elements of the controller action, U, based on the evaluated error, E, are as follows.

Proportional Action

$$\text{Controller output} = K \cdot E \tag{6.40}$$

where K is the controller gain. Manufacturers of three-term controllers tend to favor the parameter proportional band (PB) in preference to gain, K. The proportional band represents the range of the input over which the output is proportional to the input. The PB is usually expressed as a percentage of the input normalized between 0 and 100% (see Figure 6.96).

To illustrate the concept of proportional band a temperature-control application can be considered where the set point is, say, 80°C and the proportional band is set to, say, 5% over a measured temperature span of 0–100°C. The actual propor-

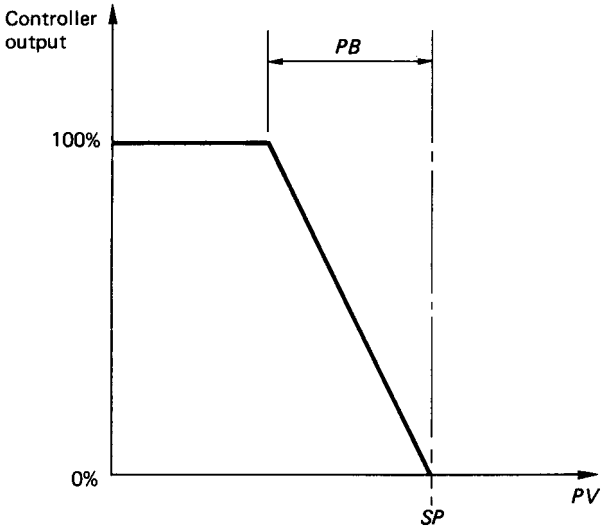


FIGURE 6.96 Illustration of the proportional band.

tional band is therefore 5°C and proportional action will apply over the temperature range between 75°C and 80°C. If the temperature is below 75°C then 100% of the available power will be supplied to the heating device. Between 75°C and 80°C, a proportion of the available power will be applied to the heating device as shown in Figure 6.96. For temperatures in excess of 80°C, 0% of the available power is supplied.

It should be apparent that proportional band is a more meaningful term than gain. The two parameters are, however, very simply related, i.e.,

$$\text{PB}\% = 100/K \tag{6.41}$$

It is also apparent from Figure 6.96 that as the proportional band is decreased, the control action is tending towards an ON/OFF strategy. A very large proportional band will result in a somewhat sluggish response.

It must also be noted that for proportional control only, there must always be an error in order to produce a control action. From equation (6.39) proportional control only gives a transfer function of the form

$$\frac{PV(s)}{SP(s)} = \frac{KG(s)}{1 + KG(s)} = \frac{1}{(1/KG(s)) + 1} \tag{6.42}$$

For steady-state conditions, s tends to 0 and $G(s)$ tends to a constant value. Equation (6.42) shows therefore that the gain must theoretically tend to infinity if $PV = SP$ and the steady-state error is to approach zero.

This is simply another manifestation of the classical control problem, i.e., stability at the expense of accuracy and vice versa. With a very high gain (i.e., low proportional band) the steady-state error can be very much reduced. A low proportional band, however, tends to ON/OFF control action and a violent oscillation may result in sensitive systems.

Integral Action. The limitations of proportional control can be partly alleviated by adding a controller action which gives an output contribution that is related to the integral of the error value with respect to time, i.e.,

$$\text{Controller output} = K_i \int E \cdot dt \tag{6.43}$$

where K_i is the controller integral gain ($= K/T_i$) and T_i is the controller integral time or reset.

The nature of integral action (equation (6.43)) suggests that the controller output will increase monotonically as long as an error exists. As the error tends to zero the controller output tends towards a steady value. The general behavior of the controller output with integral action is shown in Figure 6.97.

If T_i is very large, the integral action contribution will be low and the error may persist for a considerable time. If, on the other hand, T_i is too small the magnitude of the integral term may cause excessive overshoot in the output response. Unstable operation is also possible when T_i is too small and the controller output value then increases continuously with time.

Derivative Action. The stability of a system can be improved and any tendency to overshoot reduced by adding derivative action. Derivative action is based on the rate of change of the error, i.e.,

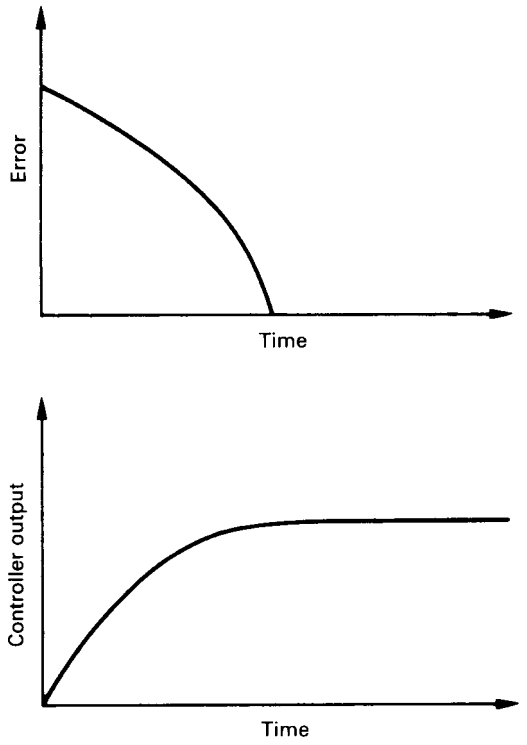


FIGURE 6.97 Controller output with integral action.

$$\text{Controller output} = K_d \cdot \frac{dE}{dt} \tag{6.44}$$

where K_d is the controller derivative gain ($= K \cdot T_d$) and T_d is the controller derivative time or rate.

Equation (6.44) indicates that the derivative action is dependent on how quickly or otherwise the error is changing. Derivative action tends therefore only to come into operation during the early transient part of a system's response.

The full three-term control strategy may be written as

$$K\left[E + \frac{1}{T_i}\int E \cdot dt + T_d \cdot \frac{dE}{dt}\right] \tag{6.45}$$

To summarize, the proportional action governs the speed of the response, the integral action improves the accuracy of the final steady state, and the derivative action improves the stability. Note that derivative action may result in poor performance of the system if the error signal is particularly noisy. In Laplace notation, the three-term controller transfer function is as shown in Figure 6.98.

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

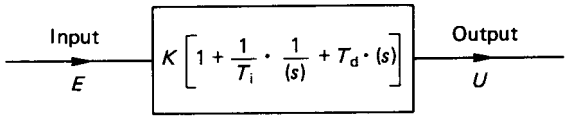


FIGURE 6.98 Three-term or PID control.

Empirical Rules for PID Controller Settings

A simple and still popular technique for obtaining the controller settings to produce a stable control condition is due to Ziegler and Nichols (1942). The method is purely empirical and is based on existing or measurable operating records of the system to be controlled.

Open-Loop Reaction Curve Method. The process to be controlled is subjected to a step-input excitation and the system open-loop response is measured. A typical open-loop response curve is shown in Figure 6.99. Any system which has a response similar to that given in the figure has a transfer function which approximates to a first-order system with a time delay, i.e.,

$$G(s) = \frac{ke^{-sT_1}}{1 + T_2 \cdot s} \tag{6.46}$$

In general industrial applications, oscillatory open-loop responses are extremely rare and Figure 6.99 is in fact representative of quite a large number of real practical processes. In the figure, N is the process steady-state value for a controller step output of P. The system steady-state gain is

$$k = N/P \tag{6.47}$$

From the process response curve the apparent dead time, T_1, and the apparent time constant, T_2, can be measured directly. The three parameters, k, T_1, and T_2, are then used in a set of empirical rules to estimate the optimum controller settings. The recommended controller settings are given in Table 6.5.

In fast-acting servomechanisms, where T_1 may be very small, the method is none too successful. For moderate response systems, however, the method will yield very reasonable first-approximation controller settings.

Closed-Loop Continuous Cycling Method. The process to be controlled is connected to the PID controller and the integral and derivative terms are eliminated by setting $T_d = 0$ and $T_i = \infty$. In some industrial controllers the integral term is eliminated with $T_i = 0$. A step change is introduced and the system run with a small controller gain value, K. The gain is gradually increased for a step input until constant-amplitude oscillations are obtained as illustrated in Figure 6.100.

The gain, K_u, which produces the constant-amplitude condition is noted and the period of the oscillation, T_u, is measured. These two values are then used to estimate the optimum controller settings according to the empirical rules listed in Table 6.6.

For a temperature-control system, typical values of T_u are about 10 s for a tungsten filament lamp, 2 min for a 25 W soldering iron, and from 10 to 30 min for a 3 kW heat treatment furnace.

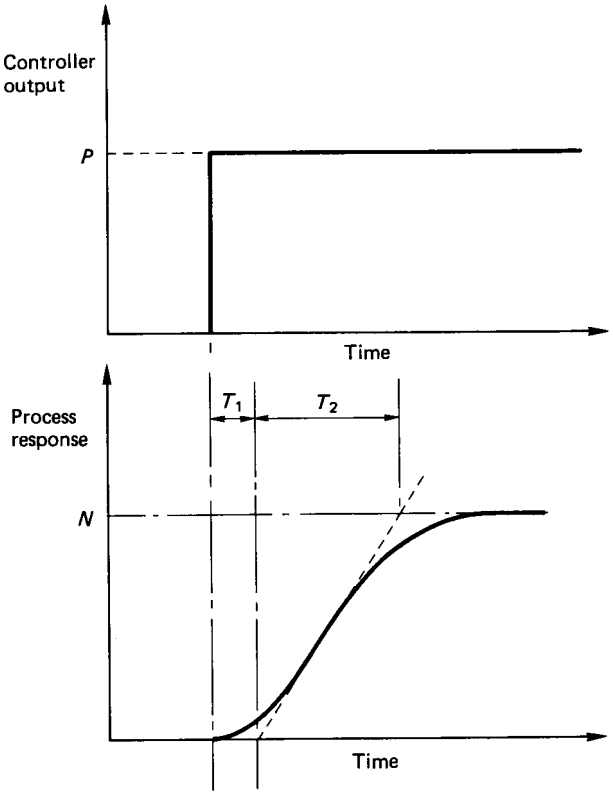


FIGURE 6.99 Open-loop system response to a step input.

TABLE 6.5 Optimum Controller Settings According to Ziegler and Nichols

Control action	K	T_i	T_d
P	$T_2/(T_1k)$	—	—
$P + I$	$(0.9T_2)/(T_1k)$	$T_1/0.3$	—
$P + I + D$	$(1.2T_2)/(T_1k)$	$2T_1$	$0.5T_1$

The PID settings obtained according to the methods of Ziegler and Nichols are approximate only, and some fine tuning would almost certainly be required in practice.

Three-Term Controller with a First-Order System. The block diagram of the system is depicted in Figure 6.94 and equation (6.39) defines the closed-loop transfer function. If a $P + I$ controller is to be used (i.e., no derivative action) the controller transfer function is

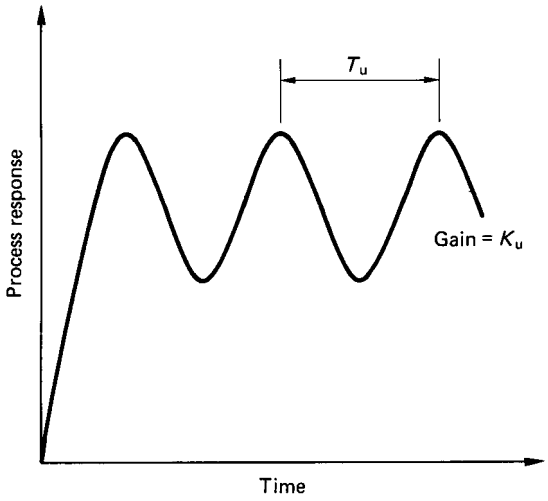


FIGURE 6.100 Continuous cycling method.

TABLE 6.6 Optimum Controller Settings According to Ziegler and Nichols

Control action	K	T_i	T_d
P	$0.5K_u$	—	—
$P + I$	$0.45K_u$	$T_u/1.2$	—
$P + I + D$	$0.6K_u$	$T_u/2$	$T_u/8$

$$C(s) = K\left[1 + \frac{1}{T_i s}\right] \tag{6.48}$$

The process is modeled as a first-order system and its open-loop transfer function is given by equation (6.18).

Substituting equations (6.18) and (6.48) into equation (6.39) results, after some manipulation, in

$$\frac{PV(s)}{SP(s)} = \frac{(kK/\tau T_i)(1 + sT_i)}{\{s^2 + [(1/\tau) + (kK/\tau)]s + (kK/\tau T_i)\}} \tag{6.49}$$

Comparing the denominator with that for the generalized second order system (i.e., equation (6.25)) it can be shown that

$$2\zeta\omega_n = \left(\frac{1}{\tau} + \frac{kK}{\tau}\right) \tag{6.50}$$

and

$$\omega_n^2 = (kK/T_i\tau) \tag{6.51}$$

For the system being controlled, both k and τ are known either via a mathematical model or an open-loop test. The controller settings, K and T_i, can then be calculated for a chosen damping ratio, ζ, and natural frequency, ω_n. Alternatively, a controller gain can be imposed and the corresponding natural frequency evaluated.

For full PID control, an initial value of $T_d = T_i/4$ can be used. Other systems can be similarly handled to obtain the approximate PID controller settings. In all cases some fine adjustment would probably be necessary to obtain the optimum output response.

Disturbance Sensitivity

The main problem with the classical single-loop control system is that it is not truly representative of the natural environment in which the system operates. In an ideal single-loop control system the controlled output is a function only of the input. In most practical systems, however, the control loop is but a part of a larger system and is therefore subject to the constraints and vagaries of that system. This larger system, which includes the local ambient, can be a major source of disturbing influences on the controlled variable. The disturbance may be regarded as an additional input signal to the control system. Any technique, therefore, which is designed to counter the effect of the disturbance must be based on a knowledge of the time-dependent nature of the disturbance and also its point of entry into the control system. Two methods commonly used to reduce the effect of external disturbances are *feedforward* and *cascade* control.

Feedforward Control. The principle of a feedback loop is that the output is compared with the desired input and a resultant error signal acted upon by the controller to alter the output as required. This is a control action which is implemented after the fact. In other words, the corrective measures are taken after the external disturbance has influenced the output. An alternative control strategy is to use a feedforward system where the disurbance is measured. If the effect of the disturbance on the output is known, then theoretically the corrective action can be taken before the disturbance can significantly influence the output. Feedforward can be a practical solution if the external disturbances are few and can be quantified and measured. The block diagram illustrating the feedforward concept is shown in Figure 6.101.

Feedforward control can be difficult to implement if there are too many or perhaps unexpected external disturbances. In Figure 6.101 the path which provides the corrective signal appears to go back. The strategy is still feedforward, however, since it is the disturbance which is measured and the corrective action which is taken is based on the disturbance, and not the output signal. Some control systems can be optimized by using a combination of feedforward and feedback control.

Cascade Control. Cascade control is implemented with the inclusion of a second feedback loop and a second controller embodied within a main feedback loop in a control system (see Figure 6.102). The second feedback loop is only possible in practice if there is an intermediate variable which is capable of being measured within the overall process. Cascade control generally gives an improvement over single-loop control in coping with disturbance inputs. The time constant for the

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

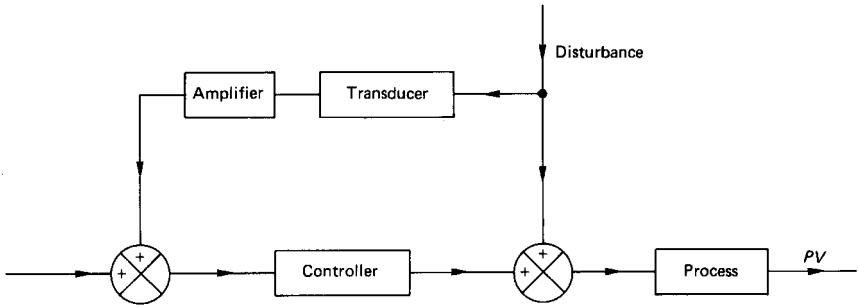


FIGURE 6.101 Feedforward control system.

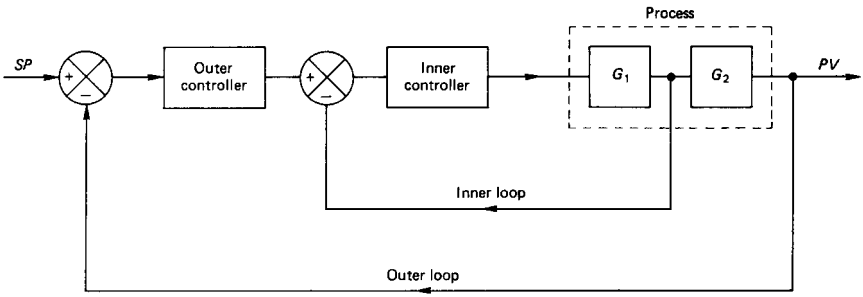


FIGURE 6.102 Cascade control system.

inner loop is less than that for the component it encloses, and the undamped natural frequency of the system is increased. The overall effects of cascade control are an increase in the system bandwidth and a reduction in the sensitivity to disturbances entering the inner loop. Disturbances entering the outer loop are unaffected. Cascade control works best when the inner loop has a smaller time constant than the outer one.

Direct Digital Control

Most of the standard texts on control engineering are centered on the mathematical modeling of systems and processes and, subsequently, the stability considerations of these entities. This approach requires detailed knowledge of the system constituent parts to enable the formulation of a suitable differential equation to describe the dynamic behavior. It is often only in the idealized world of servomechanisms that adequate models can be derived. For many real processes an adequate system model can be difficult (if not impossible) to obtain. The modern emphasis is therefore on an application of computer-based control strategies which can be made to work with real systems.

Developments in microelectronics, particularly microprocessors, has made microcomputer devices the natural choice as controllers for many systems. The microcomputer provides the ability to implement such functions as arithmetic and

logic manipulation, timing, and counting. With many analog input/output modules available to interface to the microcomputer, the overall intelligence of the system is greatly enhanced.

The basic elements of the computer-based control system include the microprocessor, memory, an input interface to measure the process variable, and an output interface to supply power to the controlled process. The control effort output to the process is determined by the control strategy, which takes the form of an algorithm incorporated within the computer software. The fundamental digital-based control system is depicted in Figure 6.103.

In the generalized layout given in this figure the microcomputer performs a number of tasks which would require separate elements in an equivalent analog system. The two inputs to the microcomputer are the desired set point and a signal from the process via a feedback loop. The term *process* is being used in a quite arbitrary sense in this context.

The microcomputer first performs the function of comparing the process value with the set point to establish the error. The control strategy is then applied to the error value to determine the corrective action necessary. The microcomputer subsequently outputs the appropriate signal to the process via other additional elements in the system. These include the input/output interfaces between the digital-based computer and the otherwise analog-based control system.

Adaptive and Self-Tuning Control

The concept of adaptive control is based on the ability to measure the system behavior at any time and to alter the controller settings automatically to provide an optimum system response. Adaptive control has been a very active aerospace research topic over the last years, but it is only recently that practical applications using adaptive controllers have appeared.

The simplest approach to adaptive control is the so-called gain scheduling method (Figure 6.104). The principle of gain scheduling is that some relevant external parameter is measured and an appropriate value of gain is then selected for the controller. Gain scheduling was first developed for aileron control in high-

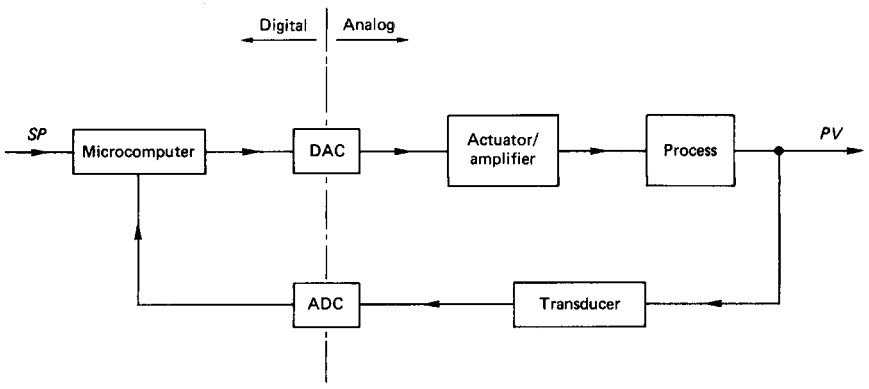


FIGURE 6.103 Fundamental digital-based control system.

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

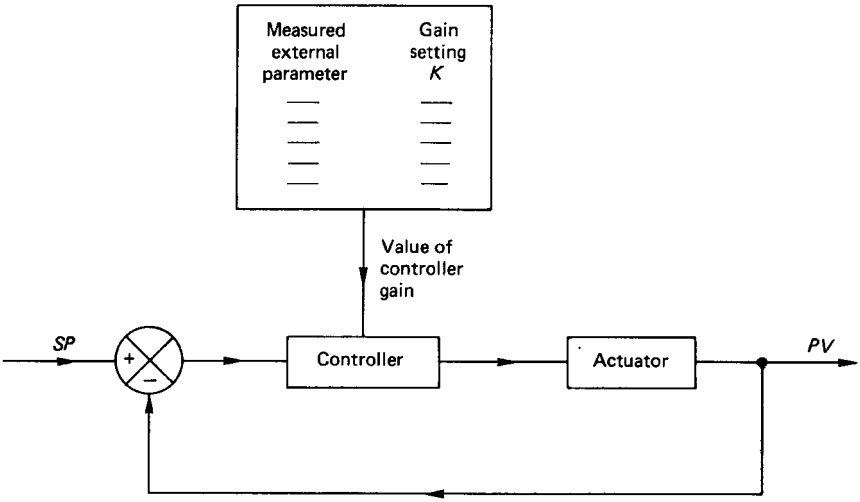


FIGURE 6.104 Adaptive control by gain scheduling.

altitude aircraft. The low air density at high altitude has a profound effect on the in-flight dynamics, and the purpose of gain scheduling was to provide the pilot with a more consistent feel for the aircraft's handling independently of altitude. Gain scheduling has the advantage that the system stability margins can be well established for any value of gain and the technique is generally fast acting. The method is limited, however, since the gain adjustment is a function of only one measured parameter. In most systems the process may be subject to any number of external parameters, and the more modern adaptive controllers use some mathematical model as a basis of comparison with the actual control system (Figure 6.105).

The mathematical model in Figure 6.105 receives the same input as the actual system, and an error is created relating the difference between the actual and the model system output. The error may then be used as a basis for altering the con-

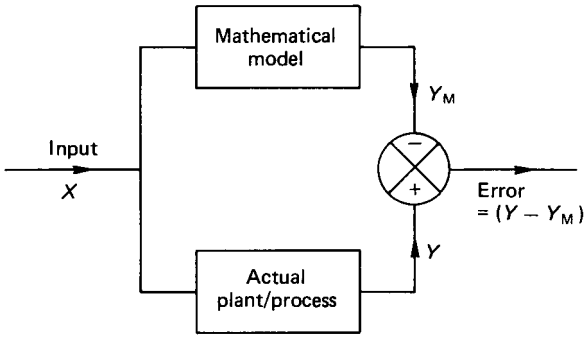


FIGURE 6.105 Model/actual error generation.

troller settings. Obviously, the quality of the control will depend on how well the model reflects the actual system. The usual implementation of model reference adaptive control is illustrated in Figure 6.106.

It is worth noting that the original feedback loop is left intact such that failure of the adaptive loop will not render the system inoperative. External disturbances operating on the actual plant will change the actual/model error signal and provide the basis for retuning the controller settings via the adaptive loop. The adjustment of the controller settings implies that there must be some well-defined strategy to determine the level and nature of the adjustments made.

Self-tuning control takes the adaptive concept one stage further in that the mathematical model of the system is also updated as more input and output data from the actual system are acquired. The schematic diagram of a self-tuning controller is shown in Figure 6.107.

The computer-based self-tuning controller estimates the system dynamics and then uses this estimate to implement the optimum controller settings. The continuous updating of the system parameters at each sampling interval is called recursive parameter estimation. Previously estimated parameters are also available, and these can be used in perhaps a least-squares method to provide some overall smoothing of the control function. With the latest system parameters available, the self-tuning controller then goes through a design procedure to optimize the controller settings. This design is usually based on the desired output response of the system. One particular design procedure is based on the root locus method for stability analysis. By adjustment of gains and time constants in the control algorithm, the method seeks to tune the transfer function and thereby govern the output response. Other procedures are often based on the rules of Ziegler and Nichols (1942). The final process in the self-tuning control cycle is the physical imposition of the optimized controller settings on the actual system.

Self-tuning control is generally applied to the more complex processes where transportation delays, nonlinearities, and multiple-control loops greatly add to the complexity. The stability of such systems is, in most cases, nondeterministic since

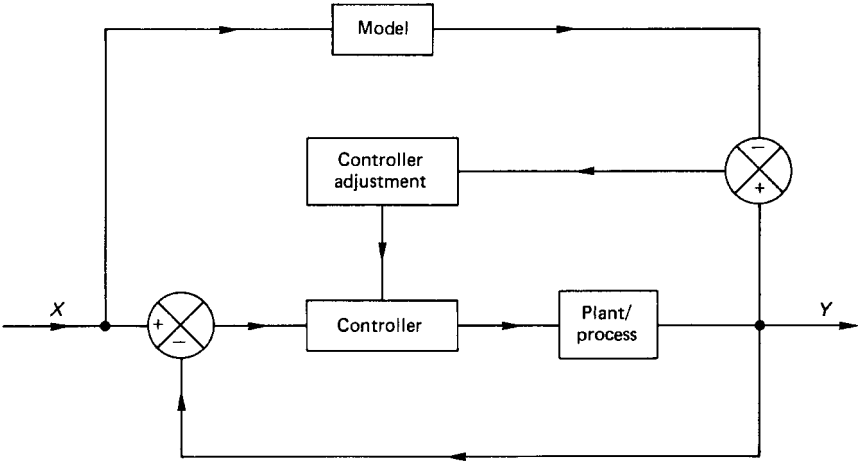


FIGURE 6.106 Model reference adaptive control.

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

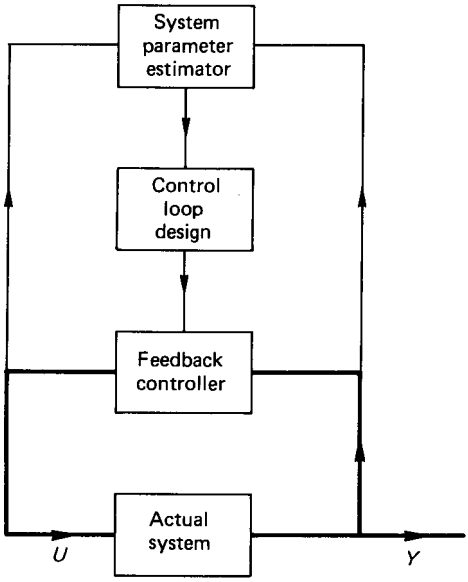


FIGURE 6.107 Self-tuning controller.

there is no generalized theory available. Traditionally, most self-tuning controllers are based on well-established three-term control principles, but with the added enhancement of adaptability. A number of proprietary self-tuning controllers are available commercially.

Sampled-Data Systems

The two previous subsections gave an overview of direct digital control and the natural progression to adaptive and self-tuning controllers. The common factor which relates these concepts is the use of a computer (or microcomputer) as a central feature of the control system. The computer acts as the compensator in the control loop and the analog-to-digital and digital-to-analog interfaces provide the link between the digital-based computer and the otherwise analog-based controlled system. Being digitally based, the computer operates in discrete time intervals, and indeed the control strategy, which exists in the software, must also take a finite time for its evaluation and implementation.

Time delays are also inevitable in the analog-to-digital and the digital-to-analog conversion processes and these cumulative time delays result in what is called a *sampled-data system.* The difference between a sampled-data or discrete signal and its continuous counterpart is shown in Figure 6.108. In the figure, the closure time, q, is the time taken to complete the digitization of the instantaneous signal. Generally, $q << T$.

It is apparent that much less information is available in the sampled-data signal as it exists only as a pulse train, interspaced with gaps in the information between the sampled points. If the sampling frequency is high enough then this need not be

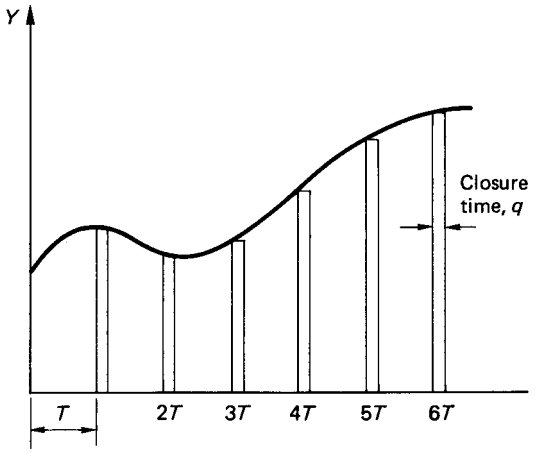


FIGURE 6.108 Digitization of a continuous signal.

troublesome. The inevitable additional time delays in a sampled-data system, however, have implications regarding the overall stability of the system.

The Laplace transform method cannot be used to analyze a sampled-data system, but there is a related transform which is applicable to discrete time systems known as the z-transform. The relation is

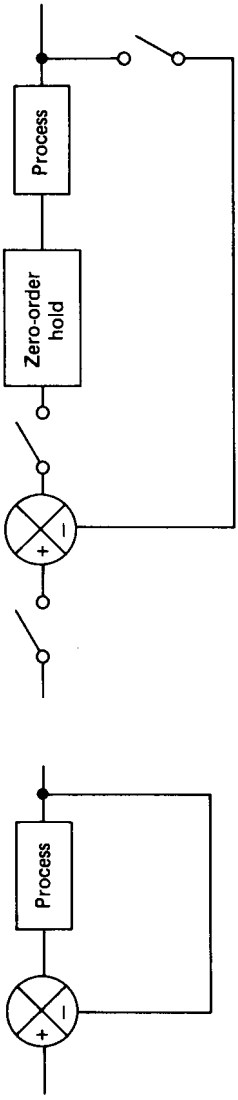
$$z = e^{sT} \tag{6.52}$$

The symbol z is associated with a time shift in a difference equation in the same way that s is associated with differentiation in a differential equation. Equation (6.52) then gives a conformal mapping from the s-plane to the z-plane and provides the means for the analysis of discrete time systems. The general method of solution involves the derivation of the closed-loop transfer function in terms of the Laplace variable. The equivalent discrete time system is then represented by introducing a zero-order hold to account for the additional time delays in the discrete system (Figure 6.109).

The transfer function for a zero-order hold is

$$\frac{1 - e^{-sT}}{s} \tag{6.53}$$

The zero-order hold is simply included in the evaluation of the closed-loop transfer function for the discrete time system. The next step is to replace the Laplace transforms with their equivalent z-transforms. The resulting transfer function in terms of z-transforms can then be analyzed for stability in much the same manner as the root locus method is used for continuous systems. Sampled-data systems and the application of the z-transform method are considered in Part 9. A comprehensive coverage of z-transform techniques and their application to the stability analysis of sampled-data systems is given by Leigh (1985).



(a) Continuous system (b) Discrete time system

FIGURE 6.109 Continuous and discrete closed-loop control systems.

Hierarchical Control Systems

The ultimate aim in aircraft flight optimization is the efficient control of complex interactive systems. Recent hardware developments and microprocessor-based controllers with extensive data-handling power and enhanced communications have opened up the possibilities for the control of interlinked systems. What is required (but not yet realized) is a theoretical framework on which to base the analysis of such systems. Nonetheless, and in the absence of theory, hierarchical control systems do exist and are currently being used effectively in the control of various large-scale plant and processes.

The usual approach adopted is to subdivide the complex system into a number of more manageable parts. This is the concept of hierarchical control, which might be thought of as a subdivision in decreasing order of importance. Hierarchical control exists in two basic forms: multilayer and multilevel. Multilayer control is that in which the control tasks are subdivided in order of complexity. Multilevel control, on the other hand, is that where local control tasks are coordinated by an upper echelon of supervisory controllers.

Multilayer control is illustrated concisely in an elaborate adaptive-type controller, and the hierarchy is depicted in Figure 6.110. The first level is that of regulation, which is characterized by the classical single closed-loop control system. Moving up the hierarchy, we have optimization of the controller parameters. Optimization is representative of the basic adaptive controller, using simple gain scheduling or a model reference criterion. The next-highest level is that of parameter adaptation. Parameter adaptation is embodied in the self-tuning controller, which represents the

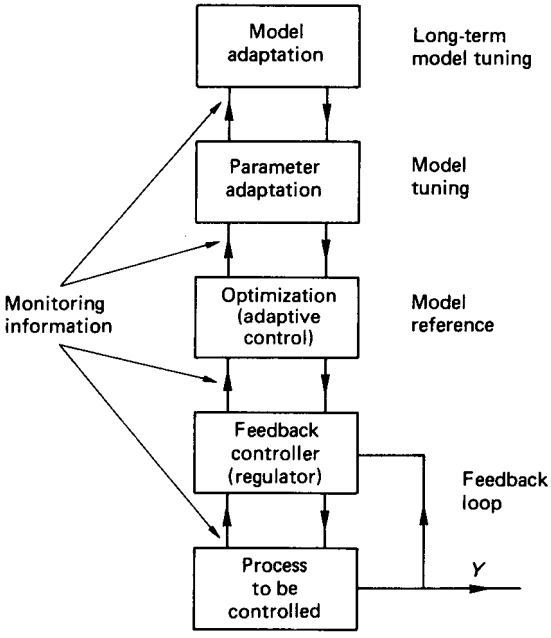


FIGURE 6.110 Multilayer control system.

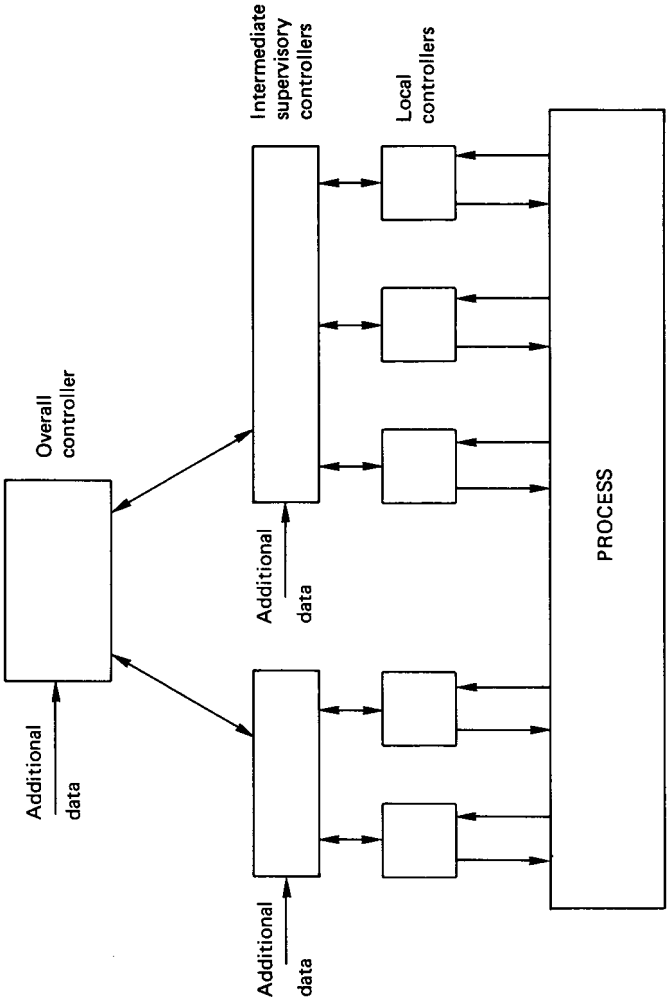


FIGURE 6.111 Multilevel control system.

beginnings of an expert system approach. The highest level is that of model adaptation, which is based on long-term comparisons between the model and the actual performance. If the system is modeled accurately to begin with, the model adaptation level might only rarely be entered.

Multilevel control is characterized as local controllers whose actions are governed by higher levels of supervisory controllers. The local controllers operate independently to achieve local targets. The function of the supervisory controller is to reconcile the interaction of the local controllers to achieve the best overall performance. The multilevel concept has some similarity with cascade control but is not so amenable to analysis.

Multilevel control gives rise to a pyramid-like structure, typified by that in Figure 6.111. At the base of the pyramid are the local controllers, monitoring and adjusting individual parameters in the overall process. At the next-highest level the supervisory controllers oversee a more complete picture of the process. The intermediate supervisory controllers have more input data to contend with, and they might perhaps relax the control of one of the process variables while tightening up on another. This, of course, would only be done to benefit the process overall.

The highest level of supervisory controller has the responsibility for the entire process. This controller may have access to additional input data which are not available to any of the lower-level controllers. The main supervisory controller is then in overall command and can influence any of the subordinate controllers. The similarity between multilevel control and the organizational structure of an industrial company is not just coincidental. The latter is the structural model upon which the former is based.

REFERENCES

Bode, H. W. 1940. "Relations between Attenuation and Phase in Feedback Amplifier Design," *Bell Systems Technical Journal,* July, pp. 321–454; also in *Automatic Control: Classical Linear Theory,* ed. G. J. Thayer, Dowden, Hutchinson & Ross, Stroudsburg, PA (1974), pp. 145–178.

Cambridge Control Ltd. 1988. *SIMBOL 2—Control System Design and Simulation on IBM-PC or PS/2,* Cambridge Control Ltd., Cambridge, U.K.

Golten & Verwer Partners. 1988. *Control System Design and Simulation for the PC* (*CODAS*), Golten & Verwer Partners, Cheadle Hume, Cheshire, U.K.

Kraus, T. W. and Myron, T. J. 1984, "Self-Tuning PID Controller Uses Pattern Recognition Approach," *Control Engineering,* June.

Leigh, J. R. 1985. *Applied Digital Control,* Prentice-Hall, Englewood Cliffs, NJ.

Nyquist, H. 1932. "Regeneration Theory," *Bell Systems Technical Journal,* January, pp. 126–147; also in *Automatic Control: Classical Linear Theory,* ed. G. J. Thayer, Dowden, Hutchinson & Ross, Stroudsburg, PA (1974), pp. 105–126.

Saunders, Arthur F. 1989. *Laplace Systems Analysis Program,* Laplace Systems–Arthur F. Saunders, Woolavington, Bridgewater, U.K.

Ziegler, J. G. and Nichols, N. B. 1942. "Optimum Settings for Automatic Controllers," *Transactions of the ASME,* vol. 64, p. 759.

PART 7

MICROPROCESSOR-BASED CONTROL

Charles J. Fraser

John S. Milne

Technological developments in microcomputers with their associated input/output hardware and software tools have enabled the designers of automatic control systems to incorporate a higher degree of intelligence than was possible in the past. Digital computers are now used extensively to control machines and processes. The physical appearance of these controllers vary considerably, according to the application, and may range from single-chip microcontrollers (SCMs), where all microcomputer components reside on one IC, to desktop personal computers (PCs).

SCMs provide very cheap computing power and are mainly associated with high-volume applications such as washing machines, automotive electronics, taxi meters, ticket machines, and time-attendance recorders. They can just as easily, however, be used in the control of manufacturing processes in the same way as PLCs, industrial rack-based controllers, and PCs.

Since its first appearance in 1981, the IBM PC and its associated compatibles has been adopted as an industry standard. In addition to an increase in processing power, there are a number of advantages in using a PC-based control system.

This integration of the disciplines of microelectronics, computer science, and mechanical engineering is the basis of the developing technology of mechatronics. It has been defined as the synergetic combination of mechanical engineering, microelectronics, and systems thinking in the design of products and processes.

There are many aerospace examples from spacecraft robotics to machine tools.

6.32 DIRECT DIGITAL CONTROL

Direct digital control (DDC), as outlined in Part 6, is employed in systems where such physical quantities as temperature, pressure, flow, position, and speed are to be constantly monitored and regulated. The design and operation of a DDC system incorporate formal control theory with computer-related hardware and software to achieve a chosen control strategy. The end product of a real-time control system is a computer-based system which runs the plant efficiently and safely under all operating conditions.

In addition to providing the necessary control functions related to the direct manipulation of valves, drives, and other actuators, there are a number of non-control functions available from the system:

1. Logging and storage of data
2. Processing and display of data
3. Communicating with operators through graphic displays of the system
4. Informing of abnormal operating conditions through alarm facilities

The benefits of DDC are directly identifiable with more precise control, which leads to increased production, improved quality of product, and the efficient use of energy and raw materials. There is also a reduction in maintenance requirements and a savings in capital investment through the minimization of recorders, instruments, and other control elements. The intangible benefits include an improvement in the handling of engineering and accounting data, increased knowledge of the process from the data collected, and displayed, and improved safety.

6.33 HARDWARE REQUIREMENTS

Figure 6.103 displays the elements of a real-time direct digital control system which constitutes a microprocessor-based controller with the associated input/output interfaces. Digital and analog interfacing techniques for connecting to external sensors and actuators are described in Part 4.

In industrial applications the transducers used for measuring the loop variable normally transmit a current in the range 4–20 mA in proportion to the measured value. This is in preference to using a voltage due to the attenuation over long distances. Also, current signaling can offer a better performance than voltage signaling methods in rejecting electrical noise. At the immediate controller input, the 4–20 mA current signal is converted into a voltage prior to the digital conversion. In control loops the actuators may also operate on a current standard. Since a variable voltage is usually generated through a D/A convertor, voltage to current, (V/I), conversion is also required.

The function of the controller is to monitor the operating conditions of the process and to evaluate, according to a specified strategy, the necessary output control action to ensure that the controlled system operates in a safe and efficient manner. A large variety of digital control devices are available to the system designer, ranging from single-chip microcontrollers, single-card computers, and programmable logic controllers (see 6.39) to complete personal computer systems.

Low-Level Control Devices

Following the inception of the microprocessor in the mid-1970s, the logical progression was to increase component density and incorporate memory and input/output interface facilities onto the one device. These so-called single-chip microcomputers (SCMs) are usually referred to as microcontrollers, and they provide the intelligence required for such applications as measurement and industrial control.

The architecture of a typical SCM is similar to that of a traditional single-card computer and includes:

1. A CPU
2. Memory—RAM and ROM
3. Parallel and serial ports
4. Timers
5. Hardware interrupt lines
6. A/D converters

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

There are a large number of SCMs now available, and some display all the above features. This greatly simplifies the designer's task, since many of the functions that previously necessitated separate chips in a microprocessor system are now included on a single integrated circuit.

Industrial Controllers

These are usually modular industrial microcomputers which are built to internationally recognized bus standards such as STD or EURO. Selected cards are held in a standard rack. The ability of these systems to acquire data, control equipment, and analyze the logged data make them particularly suited to process-control applications where performance monitoring is required. The system also lends itself well to laboratory data acquisition where the function is to collect data from a range of different instruments.

The development of applications software, usually in some form of real-time high-level language, is now usually carried out by communication between the controller and a PC or compatible. This makes life easier for the system builder by adding enhanced editing and data-storage facilities. These rack-mounted industrial controllers are generally more expensive than PLCs, although they offer more flexibility.

Since most practical industrial processes tend to be poorly defined and it is often difficult to derive an accurate mathematical model of the system, the control strategy usually adopted is the three-term one. This has led to the emergence of microprocessor-based PID controllers, whose control output is based on the error signal (evaluated from the declared set point and measured process variable) and the chosen settings of gain, integral and derivation time. Self-tuning PID controllers are now available (Kraus and Myron 1984). Some of these employ a process pattern-recognition technique with an expert system approach based upon the tuning rules usually employed by skilled control engineers.

6.34 SOFTWARE CONSIDERATIONS

The power in the digital control of a system is in the software, and any controller requires a real-time language. This is one which can synchronize its operations with events occurring in the so-called real world and thereby respond to and control interfaced mechanisms and processes. The facilities required of a real-time software language are:

1. Ease of switching external devices on and off using software
2. Ease of deciding whether external devices are on or off at any time
3. Ease of timing the duration of a process
4. Ease of making a process run for a predetermined time
5. Ease of making the program respond to things that happen in the outside world

Although high-level languages such as BASICA, QUICK-BASIC, TURBO-BASIC, FORTRAN, and PASCAL can, to some extent, meet the above require-

ments, there are a variety of real-time languages which have been specifically designed for control applications.

Real-Time High-Level Languages

BASIC. This is a programming language commonly used for engineering applications, and real-time interpreted dialects such as CONTROL BASIC and RTBASIC (Control Universal Ltd.) figure prominently. These can be used with both memory mapped and port addressed I/O organization and provide keywords which enable the individual bits on a digital port to be read or written to for switching devices on or off. A/D and D/A converters are also easily operated through keywords not usually available in ordinary dialects of BASICs. Timing operations can also be easily implemented.

In addition to the digital and analog I/O keywords available to assist in the development of control software, floating-point arithmetic is included for data-processing requirements.

FORTH (Brodie 1981). This is an interpreted threaded language developed specifically for control applications. The instructions which constitute an application are stored as a list of previously defined routines. This list is threaded together during the entry of source code from either the computer keyboard or the mass-storage buffers. The process of producing the list is often termed *compilation,* but this is not strictly correct since the result of the true compiling of source code produces pure machine code.

FORTH is a most unusual language since arithmetic calculations do not follow traditional methodologies. Before calculations can be made, it is first necessary to understand how the stack operates in FORTH. Most high-level languages use one or more stacks for their internal operations, but languages such as PASCAL and FORTRAN are designed so that the user does not have to understand how the stack functions. FORTH allows the user full control of the values stored in the stack and their manipulation.

One of the main features of FORTH (and one of the most powerful) is that once the program has been written, a single word which defines the complete program can be entered into the FORTH dictionary as a command. This dictionary contains words defining, say, routine control operations, and a single word could trigger a series of actions as its threads its way throughout the dictionary.

Programs can be typed in at the keyboard and executed directly, but to save programs for execution at a later date the mass-storage buffers and the FORTH editor must be used. Unlike other high-level languages, FORTH performs only a very limited number of error checks. The errors detected are those which are most likely to cause the system to crash if allowed to pass undetected. The main reason for the lack of error checking is that it would slow FORTH down, and since the main use of FORTH is in time-critical control situations, this would be prohibitive.

C. This is a general-purpose programming language which cannot be truly classed as either high-level or low-level. It has all the features expected of a high-level language but allows the programmer to access the computer's hardware and has the high performance usually expected of low-level languages.

Although the language was developed during the 1970s, it has now gained popularity in a range of industrial applications which include real-time control. There

are a number of versions of C available, but a standard for the programming language can be found in Kernighan and Ritchie (1988).

Accessing the input and output of data at external ports is easily accomplished, and another exceptionally important feature is the use of timing operations which have a resolution of the computer system clock period.

Using C, it is possible to construct concise, well-structured, and well-documented programs that can include a variety of useful library functions such as are needed for DOS input/output calls for hardware, screen handling, and real-time clock-timing facilities.

Low-Level Languages

At the machine level the program is stored in the memory in a binary format. It is conceivable to write the program immediately in machine code, but this would entail an unjustifiable amount of labor. It is easier to use mnemonic programming, where each instruction is accorded a symbolic name close to colloquial language. The name is easier to remember and the resulting program is much easier to read than pure machine code. This so-called assembly language program can be translated into the necessary machine code for a particular microprocessor by using an assembler program.

The main attraction of using machine code for a particular application is its speed of operation, and if a single-chip microcontroller is to be used the problem of software development must, nonetheless, be faced. This requires the writing of a program into the system memory when no keyboard, display monitor, or operating system software exists.

The inexpensive approach is to choose a SCM which supports an external EPROM, in piggyback fashion. The software code for the specific task may then be developed using an appropriate assembler program which operates on a PC or compatible. The resulting code is then used to blow the EPROM. This method can, however, be very time consuming since the debugging process will require the continual updating of the EPROM data as the program is edited to achieve the specified requirements.

An alternative is to use an EPROM emulator inserted into the target system to temporarily hold the program which is transmitted from the host computer memory. This enables the user to easily alter the program as required and then finally blow the EPROM for permanent use. EPROM emulation should not be confused with the technique of In Circuit Emulation, which involves the removal of the microprocessor from the target system. The In Circuit Emulator then takes complete control by emulating in real time all the functions of the removed microprocessor it replaces.

A software development system should therefore contain the following:

Hardware

A host microcomputer with keyboard and display monitor

A printer for a hardcopy of listings and disassemblies

Disk drives for the permanent storage of data and programs

An EPROM emulator

An EPROM programmer

Software

An editor

Translation tools such as an assembler or cross-assembler, disassembler, debugger, and linker

Traditional cross-assemblers will only assemble for one microprocessor, and it can be very expensive if a number of upgrades or different types are to be accommodated. These very sophisticated and dedicated development systems are generally too expensive and complicated for producing the operational programs to control a fairly basic and ordered sequence of events, as usually occurs in mechanical systems.

An alternative is to employ a PC with an assembler which can handle multiple instruction sets. The processor to be used is specified in the source code and the assembler adapts automatically to the correct format.

Tailoring a general-purpose and readily available machine such as a PC to a semidedicated role requires a relatively low investment and low software production costs. For one-offs this approach is the only real economic proposition.

6.35 SAMPLING FREQUENCY IN DIGITAL CONTROL LOOPS

A DDC loop contains both hardware and software contributions. Transducer measurements of the process variable must be regularly made and the invariably analog values converted into sequences of numbers that can be handled by the controller. The measured value is compared with some set condition and a control algorithm implemented in software evaluates the necessary control effort. This effort is calculated as a numerical value in the computer, and a conversion of the output is necessary to obtain a form which is suitable to drive the required control elements or actuators. A sampling rule for the measured variable is used to determine the rate at which the sampling is performed.

Digital sampling gives a sequence of snapshots of an analog variable. The controller only holds representations of the variables at the times when the samples are taken as illustrated in Figure 6.112. The sampling rate must obviously be matched to the rapidity of the variations in the process variable. High-performance control systems require the sampling interval, Δt, to be short, although very rapid sampling will increase the computational load. Δt is usually specified in terms of other system parameters.

The classic reference related to digital sampling is that due to Shannon and Weaver (1972), which states that a signal that is bandwidth limited can be recovered without any distortion provided samples are taken at a rate of twice that of the highest frequency. Twice may be regarded as an absolute minimum, and a value of five to ten times the highest frequency produces a more realistic digital representation of a sinusoidally varying continuous signal.

Real systems have limited bandwidth, and in applications with long time constants a sampling frequency of twice the highest plant frequency is adequate. In practice, for the great majority of control loops involving the control of such variables as temperature, pressure, flow, and level, sampling rates of 0.2–1 s usually prove to be fast enough. These are the rates normally fixed in the commercially available industrial PID controllers.

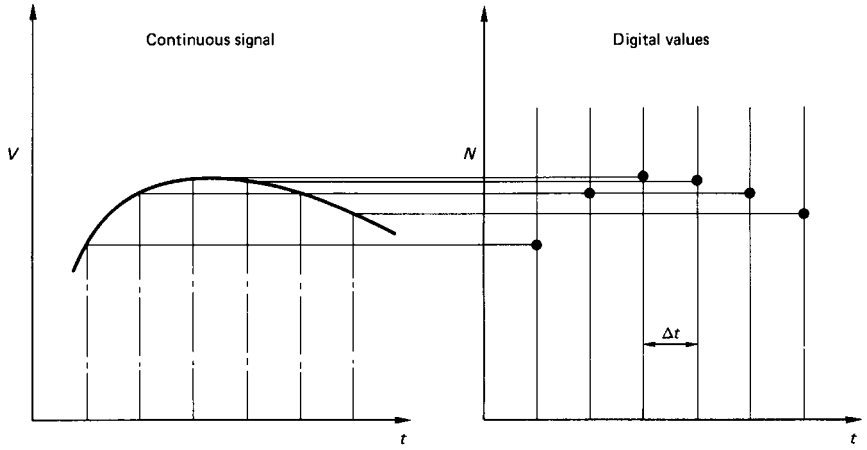


FIGURE 6.112 Digital sampling of a continuous signal.

Fast-acting electromechanical servosystems require much shorter sampling intervals (down to a few milliseconds). Shannon's sampling theorem cannot strictly be applied if the system maximum frequency is unknown. However, practical experience and simulation have produced useful empirical rules for the specification of the minimum sampling rate. One such rule is based on the dominant time constant, which can be obtained from an open-loop test on the process. A suitable sampling interval for use in a closed-loop control digital algorithm is

$$\Delta t < (\tau/10) \tag{6.54}$$

This may, however, prove to be unsatisfactory with systems which have a large time constant when a fast closed-loop response is forced by the controller settings.

6.36 PID DIGITAL CONTROL ALGORITHM

A mathematical model which produces a transfer function for the process to be controlled can usually be derived for servomechanisms, and specifications in terms of damping ratio, natural frequency, and bandwidth can realistically be defined. This is not the case, however, for the majority of industrial processes. For this reason, a three-term controller implementation is preferred in practice.

The five essential steps required to achieve good process control are:

1. The variable, *PV*, that best represents the desired condition of the final product must be measured. Measurements such as temperature, pressure, flow, position, and speed are commonly used.
2. This measurement must then be compared with the desired set-point, *SP*, value of the variable to yield an error signal, *E*, where $E = SP - PV$.
3. This error is then applied as an input to the controller.
4. The controller output, *U*, must then be applied to a final control device such as a powerstat, a valve, or a motor-drive system.

5. A source of energy or material flow is then adjusted in response to the controller output until the deviation between *PV* and *SP* will be as near zero as the sensitivity of the system permits.

This is the basic principle of negative feedback as used in most automatic control systems. The sequence of operations at a chosen sample rate is measure, compare, and correct.

The function of a digital controller is to apply a control algorithm based on the error value. An algorithm is a computer procedure that performs mathematical functions on a given set of variables. The computational procedure that converts ($SP - PV$) into a controller output is commonly called the PID algorithm. This algorithm is robust and performs well in practice. Although it may appear inferior as a scientific method, it is difficult to improve on significantly.

The controller output from a PID control strategy is as given earlier in equation (6.45):

$$U = K[E + (1/T_i) \int Edt + T_d(dE/dt)] \tag{6.45}$$

For implementation on a digital controller this must be transformed into the appropriate software for the system being controlled. This transformation from a continuous to a discrete form can be achieved by either the use of z-transforms or difference equations. The latter are easier to understand and implement, and the steps in deriving any digital algorithm by the difference equation method are:

1. Express the output requirement in a differential equation form.

2. Replace the equation in difference form using the discrete digital approximation:

$$dY/dt = (Y_i - Y_{i-1})/\Delta t \tag{6.55}$$

where Δt is the sampling interval.

3. Solve for the present value of the variable Y_i from the previous value Y_{i-1}.

Applying these steps to the three-term controller yields:

$$U_i = K[E_i + (\Delta_t/T_i) \cdot \Sigma E_i + (T_d/\Delta t) \cdot (E_i - E_i{-}1)] \tag{6.56}$$

The error at any particular time, i, is best evaluated as a percentage of the transducer span, and the above algorithm would give the controller output as a percentage value which must be contained within the range of 0–100%.

This PID digital algorithm can be easily programmed directly in a high-level language, and the framework for a program in BASIC is as follows:

1. The values for the chosen set point and the controller setting requirements K, T_i, and T_d must first be input to the program.

2. A sampling time Δt is often chosen and specified in the program.

3. Since a summation of the error value is required, a variable termed, say, *SUM* must be set to zero. This also applies to the previous error value E_{i-1} in order to start the control loop.

4. Numerical constants can be calculated, e.g.,

$$IC = \Delta t/T_i; \quad \text{and} \quad DC = T_d/\Delta t$$

5. The program listing continues:

```
 100 REM a routine is required to measure the process variable value
200  E=100*(SP−PV)/SPAN
300  SUM=SUM+E
400  OUTPUTI=IC*SUM
500  OUTPUTD=DC*(E−EP)
600  U=K*(E+OUTPUTI+OUTPUTD)
700  IF U<OTHEN U=0
800  IF U>100 THEN U=100
900  REM the output U must then be output from the controller in a form
1000 REM which is appropriate to the power requirement of the process
1010 REM being controlled.
1020 REM, e.g., an A/D converter or PWM (see above, Controller Output Interface
     Hardware)
1030 EP=E
1040 REM repeat control loop from line 100
```

Note: It is essential that the sampling interval Δt be greater than the time taken to complete the control algorithm. The above routine would generally be called from a main program which contained a time adjustment such that the complete sequence was contained exactly within the specified sampling time Δt.

The program reveals the general method of programming PID control algorithms for use with a computer and forms the basis of the software incorporated into various commercially available controllers.

6.37 SPEED CONTROL

Many systems require variable speed drives using electric motors. Electromechanical methods of motor speed control have largely been superseded by using digital controllers interfaced to power semiconductor devices. The common types of motor are those which are operated by either d.c. or a.c. Alternating machines are also referred to as synchronous motors.

D.C. Drives

D.C. motors can rotate in either a clockwise or an anti-clockwise direction, depending upon the direction of current flow to the coils via the brushes. The speed is load dependent and it can be varied by altering the supply voltage. Arcing at the brushes generates interference which may cause problems when using computer control through corruption of logic levels representing input/output data. It is essential that the spikes generated at the high interference frequencies are filtered out. This can be achieved by fitting a capacitor across the contacts close to the motor and adding an inductance in series with the power supply. A typical direct digital control loop is illustrated in Figure 6.113.

Details of typical methods of measuring rotational speed are given in Part 5. The digital controller invariably implements a three-term strategy and the numerical value calculated must be converted into a motor input through a final control ele-

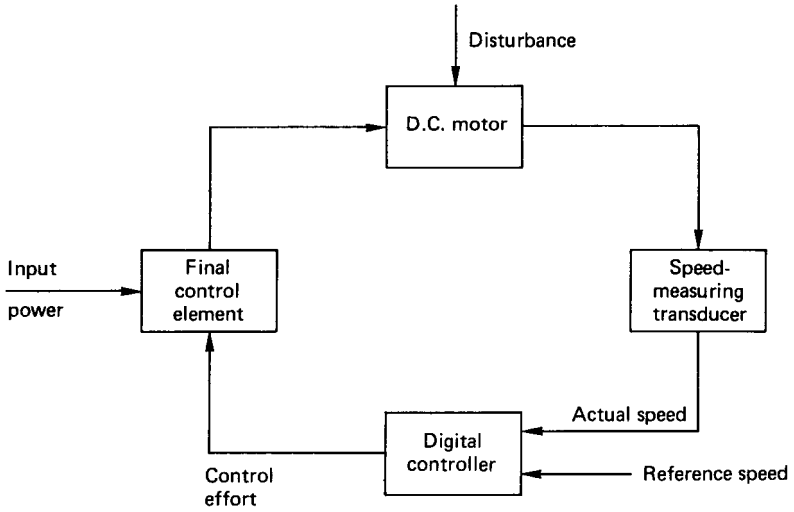


FIGURE 6.113 Closed-loop control of a d.c. motor.

ment. This may take the form of a D/A converter and a power amplifier. D.C. servomotor control modules are commercially available for a range of motors, and with a variable input from the D/A converter the motor speed can be varied over its complete operating range. The speed and nature of response to an external disturbance such as a change in load are dependent upon the controller settings.

A.C. Drives

The a.c. motor does not have brushes and hence the problem with interference when using digital control does not exist. They are also less expensive, more robust, and require less maintenance, since they do not have a commutator. However, since it is a synchronous device, the speed is locked to the frequency of the supply. To control the speed of an a.c. induction motor over a wide range the frequency as well as the amplitude of the applied voltage must be varied. The classic method of achieving this is to use a d.c./a.c. inverter in which a positive and negative d.c. supply is alternatively switched to the motor. This switching arrangement is shown in Figure 6.114.

In order to avoid short-circuiting the d.c. supply any one of the two switches must be off before the other is switched on. The switching may be achieved by using either transistors or thyristors which can operate with a turn-on/turn-off time of typically 5 μs. Thyristors have been developed for this purpose and have current- and voltage-handling capabilities of 1000 A and 4000 V, respectively. The corresponding figures for power transistors are 400 A and 800 V.

The control loop for the a.c. drive is identical in principle to that shown in Figure 6.113, where the final control element constitutes the high-frequency switching of an inverter circuit to produce a pulse width modulation technique for varying the input power to the machine. The transistor or thyristor based on a.c. drive at

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

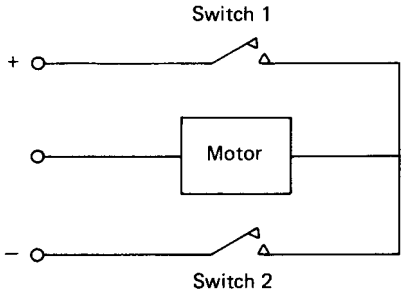


FIGURE 6.114 Principle of a d.c./a.c. inverter.

powers up to 500 kW is a viable proposition and has the advantages of close speed control in both directions, controlled acceleration or deceleration and high efficiency with a good power factor.

Stepper Motors

An alternative method of obtaining a variable speed is to use a stepper motor. This is a power device which converts a d.c. voltage pulse train into a proportional mechanical movement of its shaft. There are a range of commercially available motors which are designed for a variety of operating conditions.

The motor is designed with a number of phases which must be supplied with current in a specified preset continuous sequence according to the number of phases. Reversing the sequence causes the motor to rotate in the opposite direction. The motor shaft rotates by a finite amount for each pulse received. Thus if a motor is designed for 48 pulses/revolution then a pulse train received at the rate of 96 pulses/second will cause the motor to rotate at a speed of 120 rev/mm. Stepper motors having outputs in the order of kilowatts are available and they are used for applications involving accurate position or speed control such as *X*–*Y* plotters, numerically controlled machine tool slide drives, and in carburetors for internal combustion engines.

Interfacing a stepper motor to a digital computer is relatively straightforward, and a number of integrated circuits and power driver cards are available for a variety of popular digital controllers.

6.38 THE PC AS A CONTROLLER

The PC and its compatibles are firmly established as the standard for industrial computing worldwide. They are relatively cheap and plentiful and fairly easy to use. Since they are integrated into the industrial environment it is evident that they have a high potential as general control devices.

In addition to an increase in processing power, there are many other advantages in using a PC-based control system:

1. There is large choice of software which is not available for dedicated controllers.
2. There is a good selection of available tools to produce applications software efficiently.
3. The PC is available in a variety of forms, including from a single card, a portable, and a desktop.
4. Expansion plug-in slots to the PC bus structure are available and a large range of cards for digital or analog I/O have been produced by a number of manufacturers.
5. The PC-based controller is more flexible than the dedicated or minicomputer system and can be easily configured indefinitely to suit different applications.

Data acquisition and control add-ons for PCs are either external rack-mounted systems or plug-in boards. The external box approach usually involves attaching a separate rack-type enclosure with power supply to the host PC. The connection is via either the included serial or the parallel data communication link. Various modules based on a standard card format (such as the Eurocard) can be plugged into the enclosure housing as required.

There are two options for capturing data with PCs. The first is to use an analog-to-digital converter card that plugs directly into the host computer's backplane. The cards are generally port addressed and may be driven by any language having IN/OUT commands. The base address is usually switch selectable on the card. The second option uses instruments such as digital voltmeters and frequency meters which have an interface board that enables data transfer from or to the controlling PC. A standard is the IEEE-488 (GPIB), where data in bit-parallel, byte-serial manner are transmitted from the PC (as an ASCII string) to the instrument, informing it of the settings required for the measurement to be made. Once read, a control signal is sensed and a string representing the measured value is returned to the PC. It is, of course, necessary to have a card installed into an available slot in the PC. Up to 15 instruments can be accessed in this way from one card.

The quickest, easiest, and least expensive way to get measured data into the PC or control signals out is to use I/O cards. These are available for many applications and include:

1. Multichannel digital I/O with opto-isolation and Darlington Driver facilities
2. Pulse counting and timing
3. Multiplexed analog-to-digital conversion with programmable gain
4. Digital-to-analog conversion
5. Thermocouple input

Recently rapid advancements have been made in the software available for data capture and control. Packages are now available which provide the user with a development system having an environment of windows and pull-down menus. The National Instruments* Lab View, for use on PCs, is a library of function modules for programming specific instruments with the IEEE-488 interface. These modules are accessed in the development program via function panels to interactively set up and acquire data from instruments. In addition, a suite of programs is included for

*National Instruments, 21 Kingfisher Court, Hanbridge Road, Newbury RG14 5SJ, U.K.

data presentation, analysis and formatting in an interactive environment using either QUICKBASIC or C.

A number of manufacturers of rack-mounted microcomputer controllers and programmable logic controllers (PLCs) have united their product with a PC and provide user-friendly software to assist in the development of the required control programs. Such aids are invaluable for displaying, storing, and printing PLC ladder relay diagrams as an alternative to the hand-held programmer and EPROM blower. Logged data can also be transferred from the controller to the PC, where they can be displayed on a mimic diagram of the plant with animated symbols, updated values of controlled parameters, alarm messages, and bar graphs.

Software development for PCs has at last begun to catch up with the hardware, and the data acquisition and control market has been long overdue for reliable development systems.

REFERENCES

Brodie, L. 1981. *Starting FORTH,* Prentice-Hall, Englewood Cliffs, NJ.

Control Universal Ltd. *Real Time BASIC,* Cube Technical Manual, Cambridge, U.K.

Kerninghan, B. W. and Ritchie, D. M. 1988. *The C Programming Language,* Prentice-Hall, Englewood Cliffs, NJ.

Kraus, T. W. and Myron, T. J. 1984, ''Self-Tuning PID Controller Uses Pattern Recognition Approach,'' *Control Engineering,* June.

Shannon, C. E. and Weaver, W. 1972. *The Mathematical Theory of Communication,* University of Illinois Press, Urbana.

PART 8

PROGRAMMABLE LOGIC CONTROLLERS

Charles J. Fraser

John S. Milne

Automation systems used in the aerospace industry generally involve the application of such microprocessor-related equipment as CNC machine tools, pick-and-place machines, and industrial robots. As shown in Table 6.7, there has been a greater involvement in the application of programmable logic controllers (PLCs) in manufacturing industry than any other automation devices. (Northcott and Walling 1988). The figures in Table 6.2 have been obtained from a survey of all U.K. manufacturing industries employing 20 or more people.

PLCs are mostly employed in the relatively straightforward control of a single process or piece of equipment. They are particularly common in the food processing, chemical, and automotive industries. PLCs first appeared in the 1970s in place of relay circuits, and they have been continually developed as a result of the rapid progress in microelectronics. Some of the current more powerful PLCs overlap with microcomputers or process computers, and it is often impossible to distinguish between them.

6.39 THE PLC IN AUTOMATION SYSTEMS

The PLC is particularly useful for controlling manufacturing processes which constitute a sequence of operations that are either event or time driven. PLCs have had a significant impact on industrial control because of their ruggedness, versatility, and reliability. They have virtually replaced hard-wired relay, counter, and timer logic systems due to their cost, flexibility, and relative size. Additionally, the number of applications in which PLCs are being used has increased substantially in recent times.

The principal criteria for determining size is the input/output availability and they can broadly be divided into the following categories:

TABLE 6.7 Usage of Automation Equipment

Equipment used	1983 ('000)	1985 ('000)	1987 ('000)	1989 ('000: estimate)
CNC machines tools	17	19	27	42
PLCs	34	59	73	110
Pick-and-place machines	4	11	10	15
Robots	1	2	3	6

Source: Northcott and Welling 1988.

1. Small PLCs, up to 128 I/O lines, with typically 12 inputs and 8 outputs in the basic form. These are designed as stand-alone items and can usually be expanded with respect to digital I/O requirements. Such systems basically perform logic, counting, and timing operations but have no arithmetical manipulation capabilities.
2. Medium-size systems (128 to 1000 I/O lines), which are generally contained in a rack arrangement. These offer more extensive I/O, including analog-to-digital and digital-to-analog options with some enhanced programming features.
3. More powerful rack-mounted systems, with over 1000 I/O lines.

The latter PLCs offer data communication for operation within a complete computer-integrated manufacturing arrangement. They embody quite sophisticated man/machine interfaces which employ real-time operating systems and advanced computer graphics. Three-term PID control, digital filtering and vision capabilities can be implemented in many modern PLCs and a variety of operator interfaces are available to allow easy entry of desired settings and process variables.

Recent developments have seen a move from the automation of single machines towards that of the whole manufacturing process. The PLC is a typical representative. The modern concept is the automated factory, where computer-integrated manufacturing (CIM) combines the basic functions of production into a single system. The system incorporates the highest level of automation economically feasible for each function involved. The PLC constitutes one of the building blocks in such a distributed hierarchical control system, which is based on a pyramid structure. The peak of such an arrangement contains the mainframe computer, which is responsible for handling the databases for production scheduling, sales, and other management needs. The PLC or microcomputer immediately precedes the machine and process applications located at the base of the pyramid.

However, although the small single-purpose basic PLC is relatively secure in its role in the market, the medium and large PLCs face considerable competition from microcomputers. This is particularly the case in applications where much arithmetical manipulation must be performed. It is possible to get PCs and other microcomputers in a 19-inch rack format which can be mounted within standard control panels. This represents a serious competitor to many PLC-based systems.

6.40 THE PLC VERSUS THE MICROCOMPUTER

The PLC and the microcomputer contain the same basic constituent components such as central processing unit, memory devices, input/output interfaces, decoding logic, and a connecting bus structure. They are, however, entirely different in physical appearance and operation.

The Case for the Microcomputer

Since the appearance of the PC, several manufacturers have produced so-called compatibles, which are based on the associated 16- and 32-bit technology. Processors now allow the programmer to write programs that use more memory than is actually available in a given system. This is done by exchanging data between the

main memory and secondary storage devices. It is this processing power, memory availability, and peripheral hardware that has made a significant impact on the engineer's acceptance of PCs for control applications.

In addition to this increase in processing power, there are many other advantages in using a PC-based control system:

1. The large choice of available storage ranging from high-level language interpreters and compilers to sophisticated man/machine software interfaces which include a range of selectable data processing and graphical display routines
2. The selection of available tools to produce applications software efficiently
3. The variety of forms available, which range from the small portable type for use in the field to the rugged industrial version for the factory floor
4. The large range of plug-in cards available for use with the PC bus structure to provide both digital and analog I/O facilities
5. The ease with which the system can be reconfigured to enable, for example, the application of word processing, spreadsheets or databases to be carried out. There is no need to purchase a new system for every new application.

However, in order to truly harness the flexibility that accompanies the personal computer, some knowledge of microprocessor technology is required, and the ability to understand or develop applications software is essential.

The Case for the PLC

The alternative to the microcomputer for control applications is the PLC, which may be described as less intelligent but is ideally suited for carrying out logical sequential operations to produce outputs which are conditional on input states. The advantages of using a PLC for controlling machines and processes may be summarized as follows:

1. The PLC is relatively inexpensive compared with a microcomputer.
2. The construction is exceptionally rugged and it requires very little space in the control cabinet.
3. Reliability is high, immunity to electrical noise is good and maintenance is low.
4. Memory is used economically due to the method adopted for the processing of the data. The PLC processor is tailor-made to execute logical operations (AND, OR, etc.) as they occur in the control program.
5. The operating speed is fast.
6. A built-in interface provides for the easy connection to a variety of input/output devices connected directly to the machine or process.

The PLC is, however, less interactive than the microcomputer, partly due to the fact that relay ladder logic is generally used as the programming language because of its alleged wide acceptance by those who design, operate, and maintain control systems. During program execution the ladder is continuously scanned and outputs are set in accordance with input conditions. It is reputed that ladder logic is straightforward and easily understood, but software development can be a problem for the nonspecialist in systems which involve a number of timer and counter requirements.

In conclusion, there is probably equal support, at present, for both PC and PLC systems.

6.41 LADDER LOGIC PROGRAMMING

The most commonly used programming language for use with a PLC is the ladder diagram. For the popular small PLCs it is usually possible to enter the input, output, timer, and counter instructions directly into the controller using a hand-held programmer. The mnemonic codes are displayed on a small screen within the programmer prior to conversion into machine language and transfer into the PLC battery back-up memory. Although the program can be edited, the main disadvantage of using this method is that there is no way of saving and printing the developed program.

A preferable alternative is to use a PC with an appropriate software package available from the PLC manufacturer to develop and display the ladder diagram on the PC monitor. Use is made of the PC drives and printer as required. Once completed, the resulting file is transmitted from the PC to the PLC for execution. This is by far the best method of developing ladder logic software.

In a ladder diagram the power source is represented by the two vertical rails of the ladder and the various instructional requirements which represent the control circuits make up the rungs. The symbolic ladder circuit layout is constructed using standard graphic logic symbols to represent input contacts, output loads, timers and counters, as shown in Figure 6.115. When using a PC to develop the program, comments can be added to each instruction rung. This makes the program easier to understand and facilitates fault finding.

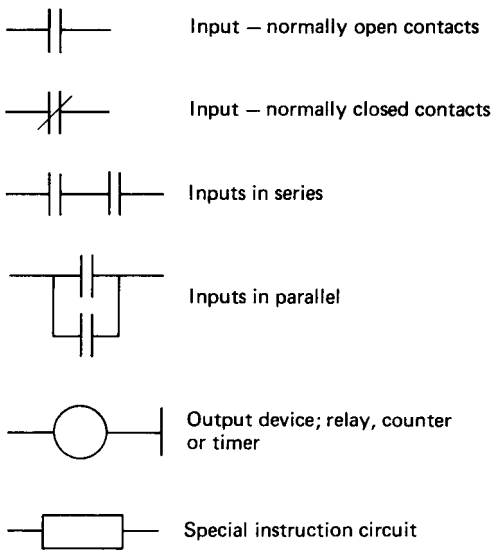


FIGURE 6.115 Instruction set ladder logic symbols.

Each rung on the ladder starts with an input condition and finishes with an output to a load, timer, or counter. When used to automate a machine, the ladder is continuously scanned in a sequential manner and sets the outputs on or off according to the requirements based on the sensed input conditions.

Inputs

The most common symbol used in programming a PLC is the input contact, which may be active in either the normally open (NO) or normally closed (NC) mode of operation. The NO contact symbol is a request to examine if the switch is ON in order to proceed along the run. For the NO condition to be satisfied the contact must be OFF.

A popular real-world input device is the switch. This is basically used to open and close an electrical circuit. The PLC supplies the voltage across the switch which is usually either 24 V d.c. or 100 V a.c. A large number of switch designs are used to control systems. The limit switch is one such example, and it is designed to open and close when a machine part reaches a specified position or a physical variable obtains a required value. Once the limit switch condition has been sensed, the PLC takes appropriate action as defined in the ladder program.

Outputs

The most common output devices are the electromagnetic relay and the solenoid. The relay provides isolation of the load voltage and current from the PLC and is ideal for the switching of high loads via the controller's built-in I/O interface. Alternatively, the solenoid, which consists of wire surrounding a movable plunger, allows electrical control of any mechanical device attached to the plunger. This arrangement is applied in electropneumatic systems to control the position of a valve spool in order to direct the air path through a pneumatic system.

Due to the nature of the sensed input and the switched output load, the PLC I/O module normally provides a safe electrical separation between internal and external circuits. This is effectively carried out by an opto-isolator, which ensures that there is no physical hard-wired connection between the PLC and the external I/O device.

A typical rung in a ladder diagram including I/O with logic operations is shown in Figure 6.116. Thus rung may be written for a particular PLC as:

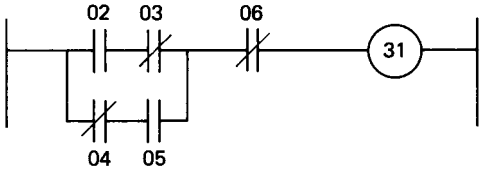


FIGURE 6.116 A typical ladder rung.

Step	Instruction	*I/O* number
000	LD	02
001	AND NOT	03
002	LD NOT	04
003	AND	05
004	OR BLK	
005	AND NOT	06
006	OUT	31
007	END	

Note: the "OR BLK" statement "ORs" the blocks grouped under the preceding two load (i.e., LD) statements.

Different PLCs use various ways of translating the ladder programs into mnemonics, but the principle is basically the same as that illustrated. This translation is necessary if the program is to be entered into the controller via a hand-held programmer. However, if a PC-based software development system is used the ladder diagram is entered directly by pressing the appropriately specified function keys.

Timers

Time delays are commonly used in control applications, and in a PLC program a timer is used to activate an output following an elapsed time as specified by a value stored in the memory. The timing base units are usually tenths or hundredths of a second.

There is no standard way of specifying timer routines and manufacturers use different methods, but the same basic principles of operation apply. This entails an enabling input condition to start the timer. At the end of the timed period the timer must be reset by disenabling the input.

It must be noted that during the program execution the ladder is *continuously* scanned rung by rung and outputs set according to input conditions. There can be no hold on any particular rung. However, in implementing timing operations the timer output can be used as an input on any rung, thus making outputs dependent upon the specified time period to have expired. This is illustrated in Figure 6.117. In this figure the timer number 50 is set for a timing operation of 1 s and is enabled by switching input number 05 to an ON state. In the next rung the output number 30 will not be energized until the specified time has elapsed. The timer must then be reset by setting the output 05 to OFF.

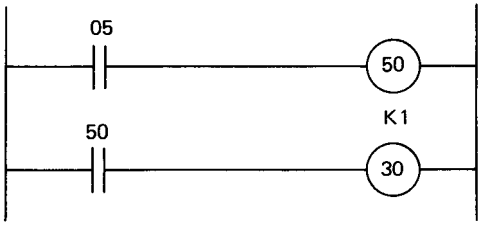


FIGURE 6.117 Example of a timing operation.

Counters

A counter is used to activate an output after a predetermined number of counts as stored in the PLC memory. It operates in much the same manner as a timer but, rather than counting time, the counter counts the number of times that an event occurs. The event is some specified input condition in the process. The counter is reset from another input switch and it can be used as an input to any other rung in the ladder. This is illustrated in Figure 6.118. Each time that switch 02 is activated a count is produced up to the specified value of 10. Once the count is completed the output load number 30 is switched ON. The counter is then reset by switching input 01 OFF. An incrementing or decrementing mode for the counter operation is usually possible with most PLCs.

Auxiliary Relays

Most PLCs are equipped with various special function input/output facilities that exist in the development of safe and efficient control programs. The auxiliary relay is one such device which operates internally and is assigned an address that is different from that of any other real output device in the system. It is basically a dummy relay which is used as an aid to developing the program. For example, consider an output condition which is dependent upon, say, six inputs and a rung is restricted to four inputs. The auxiliary control relay can be used to connect two rungs into a logical equivalence of the requirement as shown in Figure 6.119.

Input/Otput Numbering

Different PLC manufacturers use different numbering systems for the controller's available I/O. The addresses assigned for the small Mitsubishi type F-20M PLC are typical:

12 inputs:	00–07; 10–13
8 ouputs:	30–37
8 timers with a range of 0.1–99 s:	50–57

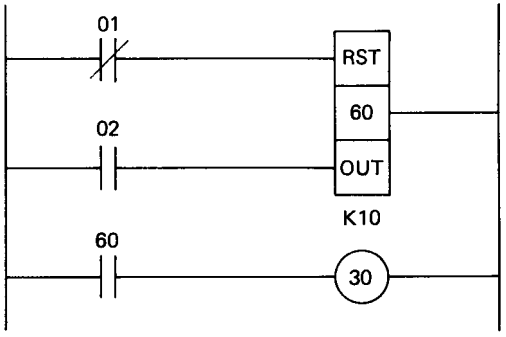


FIGURE 6.118 Example of a counter operation.

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

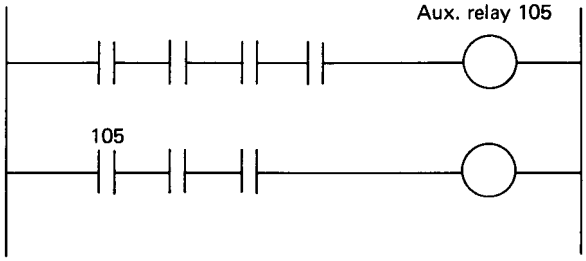


FIGURE 6.119 Use of an auxiliary relay.

8 counters with a range of 1–99 counts: 60–67
48 auxiliary relays: 100–157

(Note the octal numbering system adopted.)

6.42 CONTROLLING PNEUMATIC AND HYDRAULIC SYSTEMS

Consider the sequencing of a double-acting pneumatic cylinder with digital sensing devices of mechanical, optical, or magnetic construction to detect the end-of-stroke condition. The direction of motion is controlled by an electropneumatic five-port spool valve which is solenoid/pilot operated. The air supply to the system manifold is enabled by an ON/OFF switch connected to a solenoid-operated three-port valve.

The control program is required to switch on the air, extend the cylinder rod to the end of the stroke, wait 5 s then retract the rod and repeat to complete 10 cycles. The main ON/OFF switch is to be used to reset the counter and it must also stop the cycling at any time. The input/outputs are assigned as follows:

INPUTS
ON/OFF switch −00
Sensor to detect retracted position −01
Sensor to detect extended position −02
OUTPUTS
Air supply valve −30
Direction control valve solenoid to extend cylinder rod −31
Direction control valve solenoid to retract cylinder rod −32
TIMER − 50 and COUNTER −60

Figure 6.120 shows the ladder diagram to satisfy the specified sequence. Note that deactivating the ON/OFF switch will stop the system at any time and that the rod will not extend once the counter has completed the required number of cycles.

With solenoid/pilot valves the solenoid need only be pulsed since the pilot line in the valve supplies air to move the spool, and it is usually unnecessary to maintain

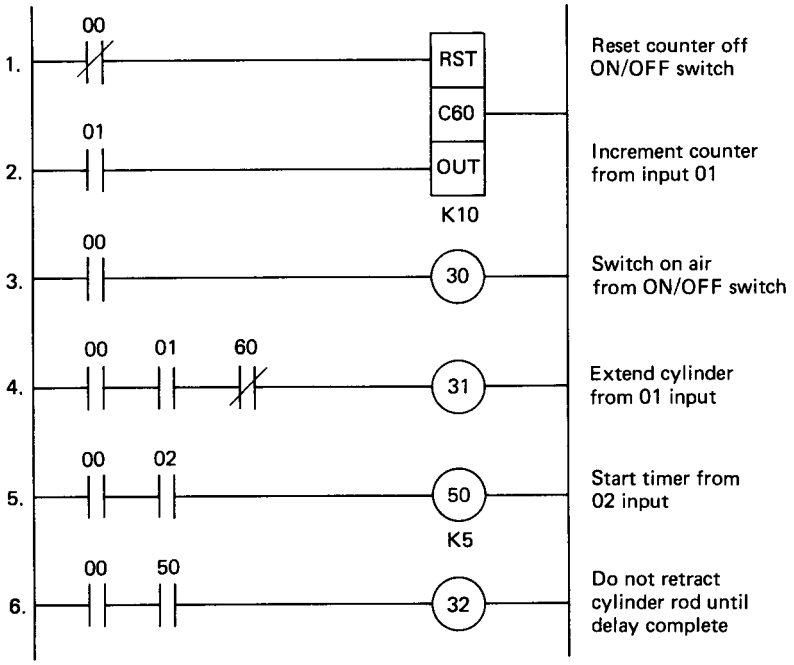


FIGURE 6.120 Ladder diagram for a specified example.

the solenoid in an energized condition. This is evident in the ladder diagram with inputs 0l and 02 from the end-of-stroke sensors returning to the OFF state once the cylinder rod moves from the end condition.

With solenoid only operated valves, however, it is necessary to latch the output load ON in order to maintain the rod motion. The ladder relay logic for rung 4 to latch on solenoid 31 in order to extend the cylinder rod is illustrated in Figure 6.121. Once the rod has left the sensor 01, power is transmitted to the output via the switched output acting as an input. This latches the control solenoid ON and it is switched OFF once the end of stroke is reached and sensor 02 is switched ON.

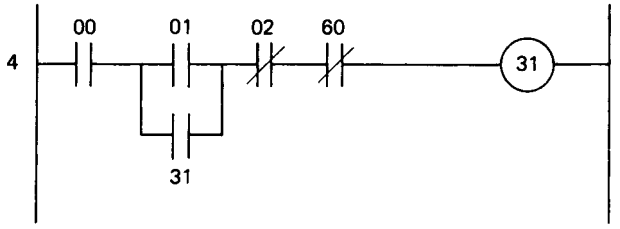


FIGURE 6.121 Latching of output.

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

6.43 SAFETY

Since the equipment generally associated with PLCs often involve high voltages and currents, electrical safety procedures must be followed at all times. In addition, when testing a developed program to perform a specific task it is is essential that a simulation first be carried out in order to avoid the incorrect movement of large pieces of machinery which would result in dangerous operating conditions. Most PLCs contain a test-mode facility where outputs can be simulated by LED indication for a switched input requirement. The process is simulated by the manual manipulation of switches at specific times. The effects of the control program on the process may then be observed by the status of the PLC LED indicators on the I/O modules. If necessary, the program can be edited in the program mode and then retested prior to connecting to the actual system which is to be controlled.

6.44 NETWORKING OF PLCs

Computer-integrated manufacturing incorporates the highest level of automation economically feasible from each computer-based activity within a complete system. The overall control task is divided into a number of individual routines assigned to specific control elements such as CNC machine tools, robots, and PLCs. The PLC thus constitutes one of the building blocks of the automated factory pyramid concept. This requires a suitable network to enable the various elements to exchange information and utilize a control database.

The manufacturers of PLCs have addressed the problem of networking, but the major disadvantage is that each manufacturer adopts their own standards. Communicating with a variety of other control devices is not a strength of PLC networks due to the number of data-transmission protocols and data formats adopted. However, most manufacturers provide a dedicated local area network (LAN) for communication between controllers within their own product range and examples are given in Table 6.8.

The basic concepts relating to the standards of communication in digital systems are covered in Part 3. In a decentralized network structure the sensing, actuating, and processing associated with each controlled element in a system is connected to a supervisory or master computer. This evaluates signals supplied by each component and returns the necessary control requirements. A PLC, PC, single-board, or single-chip computer could be used as the controller for the controlled element

TABLE 6.8 Typical PLC Networks

Manufacturer	Network
Allen Bradley	Allen Bradley Data Highway
Festo	PC-IFS
General Electric	GE Net Factor LAN
Mitsubishi	Melsec-NET
Texas Instruments	TI-WAY

and a process computer, PC, minicomputer, or mainframe can be employed in the supervisory role.

Network Structures

Although there are various methods adopted for the networking of computer-based products, they all possess certain common features. Each element must be equipped with a suitable interface associated with the LAN chosen and each element in the system must be connected into the network by means of wire or fiber-optic cabling to transmit the data from one station to another. Software is also necessary to handle all data transfer within the system correctly.

The physical arrangement of the elements or stations on the network are usually of a star or ring pattern. In the star-shaped network, as shown in Figure 6.122, the stations are connected on a line parallel to one another and connected to a central computer referred to as the *file server*. With this method, each user station must decide on whether data sent by the file server are for itself or another station. The cable length is limited in this system, but signal amplifiers can be used if required.

In the ring structure as shown in Figure 6.122 the file server master computer transmits data to the first station on the network. These data are checked, evaluated, and passed on to the second station if not required. The data are hence transmitted from one station to another until the user for whom they are intended is found. Data can therefore be passed around great distances but the failure of any one station causes the system to break down.

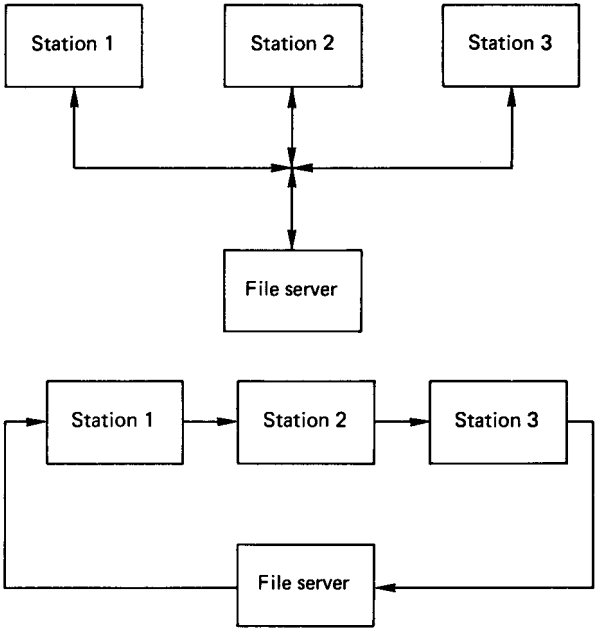


FIGURE 6.122 A star- and ring-shaped network.

The integration of all intelligent elements into an automation system requires that all devices communicate with each other. Due to the different suppliers having different communication specifications, international communication standards based on the Open System Interconnection Model (OSI) have evolved. Widespread adoption of such standards will make it easier and cheaper to link devices together.

One such standard is the Manufacturing Automation Protocol (MAP), which was initiated by General Motors to integrate all levels of control systems such as PLCs, robots, welding systems, vision systems, etc., irrespective of the manufacturer. Another very popular LAN standard adopted for industrial and commercial application is Ethernet, which was created by Digital Equipment Corporation, Intel, and Xerox.

Users of automation systems will make their purchasing decisions based on the amount of software packages and support services that the distributed control system vendors can provide (Babb 1989). Only those vendors who are capable of supplying such products and services to provide a complete solution to a control problem will be successful in the future.

REFERENCES

Babb, M. 1989. "Implementing Distributed Control in the 1990's," Control Engineering, August, pp. 2–4.

Northcott, J. and Walling, A. 1988. *The Impact of Microelectrics—Diffusion, Benefits and Problems in British Industry,* PSI, London.

PART 9

THE *z*-TRANSFORM

Charles J. Fraser

John S. Milne

As outlined in Part 7, the availability of inexpensive and continually improving microprocessors has led to their application to direct digital control techniques in the closed-loop feedback control of processes. With the introduction of these sampled-data systems mathematical techniques have been developed specifically to provide a design method for digital controllers. The technique generally employed is similar to the Laplace transform used in continuous systems.

The Laplace transform, through the s operator, is used to transform linear differential equations into a linear algebraic form which can then be appropriately manipulated to produce open- and closed-loop transfer functions for a system. This method forms a basis for the design of continuous controllers.

In a discretely sampled system the z-transform is used to provide an algebraic way of representing a sequence of digital data with respect to time measurements. The symbol z for the digital system is analogous to s in the continuous system. The z-transform is associated with the time shifting in a difference equation as s is associated with differentiation in a differential equation.

The operators may be mathematically defined as follows:

$$s = d/dt = j\omega$$

$$z = e^{sT}$$

where T is the discrete time sampling interval and z^{-1} is a delay operator which represents a delay of one sampling period.

6.45 REPRESENTATION OF DISCRETELY SAMPLED DATA

The characteristics of a digital control loop are related to the sampling process associated with the conversion of transducer analog data into a digital format and the converse between the controller output and the system being controlled. The sequence of events are as follows:

1. The process variable (PV) is measured and the corresponding analog signal is converted into a digital value using an A/D converter.
2. This value is compared with the declared set point (SP) to give on error value (E) on which the controller transfer function operates to produce a control effort (U) as a numerical value preferably expressed as a percentage within the range 0–100%.
3. The control effort derived is then converted by means of a final control element into the actual power supplied to the plant or process. This element invariably involves a D/A converter which holds the value of the impulse received until

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

the next value arrives in a staircase fashion. This is referred to as a *zero-order-hold* device.

The overall time for the sequencing cycle is the sampling frequency

$$f_s = 1/T$$

where T is the sampling interval.

It is most important to appreciate the physical nature of the variables at each stage in the loop and this is illustrated in Figure 6.123. The discretely sampled data associated with the process variable, *PV*, and its related derived values of error and control effort are displayed at time intervals of T. These are suffixed by an asterisk to indicate that they are snapshots of the plant state at a particular time. The z-transform provides an algebraic way of mathematically presenting such sequences of digital information.

Consider in detail the variation in the derived error signal, E, as shown in Figure 6.124. Numerical values, quoted as percentages of the range of the measuring transducer, are inserted to illustrate the mathematical form of the z-transform equation to represent the variation of E with time. The sequence is shown as:

$$E(n) = 80, 60, 40, 25, 12, 6, 0, 0 \ldots$$

and the corresponding z-transform of the sequence is defined by the sum

$$E(z) = 80 + 60z^{-1} + 40z^{-2} + 25z^{-3} + 12z^{-4} + 6z^{-5}$$

This method of representation is particularly useful for handling time delays. For example, if the sequence was delayed by one sampling interval then it becomes

$$0, 80, 60, 40, 25, 12, 6, 0, 0, \ldots$$

and the corresponding z-transform

$$80z^{-1} + 60z^{-2} + 40z^{-3} + \ldots$$

which is the first sequence multiplied by z^{-1}. Thus multiplication by z^{-1} denotes a one sampling period delay. Similarly, multiplication by z^{-2} would denote a delay of two sampling periods.

Generally, z-transform models of digitally sampled data or its derivatives can be obtained provided the relationship between the continuous variation of the data with time is known in a mathematical form. For example, say that the error value can be expressed in the form

$$E(t) = 100e^{-t}$$

which for the data sampled at intervals of T would yield:

$$E(z) = 100(1 + e^{-T} z^{-1} + e^{-2T} z^{-2} + e^{-3T} z^{-3} + \ldots)$$

This is the sum of a geometric series given by

$$\text{SUM} = 100/(1 - e^{-T} z^{-1})$$

$$= 100\, z/(z - e^{-T})$$

which is the z-transform $E(z)$.

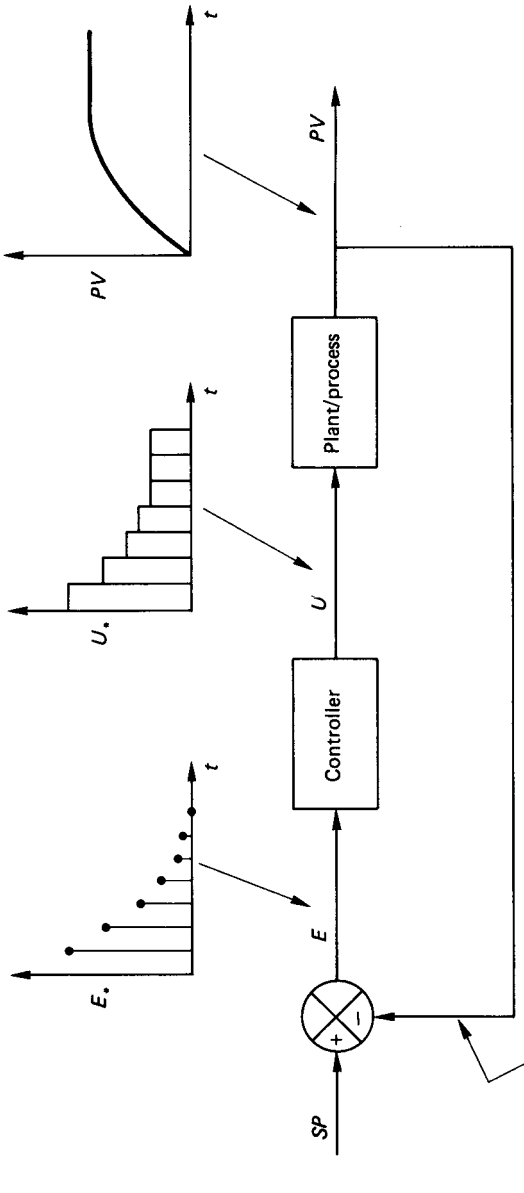


FIGURE 6.123 Nature of the variables in a digital control loop.

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

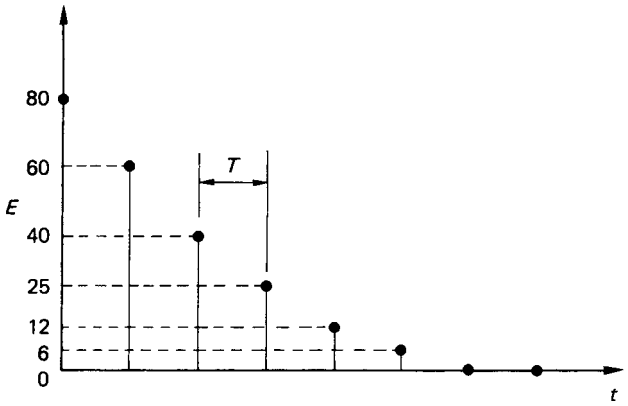


FIGURE 6.124 Variation of discrete error values with time.

If the data were sampled at a frequency of 10 Hz (i.e., $T = 0.1$ seconds) then

$$E(z) = \frac{100z}{(z - 0.905)}$$

In order to aid the design process associated with digital techniques, z-transforms of common sampled functions are tabulated in a way similar to Laplace transforms. Examples of some common time functions with their corresponding Laplace transform and z-transforms are given in Table 6.9.

When applying digital techniques the z-transform, $G(z)$, must usually be derived from the corresponding transfer function for the operation expressed in terms of s, $G(s)$. It should be noted that it is not simply a case of substituting z for s in the function. The basic approach is generally to obtain the partial fraction expansion of $G(s)$ to yield individual components which translate directly into z-transforms using a standard table such as Table 6.9.

For example, consider the transfer function $G(s)$ given by

TABLE 6.9 Table of z-Transforms

Time function	Laplace transform	z-transform
Unit step input	$1/s$	$z/(z - 1)$
t(ramp input)	$1/s^2$	$Tz/(z - 1)^2$
$t^2/2$	$1/s^3$	$\frac{T^2z(z + 1)}{2(z - 1)^3}$
e^{-at}	$1/(s + a)$	$z/(z - e^{-aT})$
te^{-at}	$1/(s + a)^2$	$\frac{Tze^{-aT}}{(z - e^{-aT})^2}$
$1 - e^{-at}$	$a/s(s + a)$	$\frac{z(1 - e^{-aT})}{(z - 1)(z - e^{-aT})}$

$$G(s) = 1/s(1 + 2s)$$

Although this form appears explicitly in Table 6.9 it can be used to illustrate the technique of partial fraction expansion. Using partial fractions it can be assumed that

$$G(s) = A/s + B/(1 + 2s)$$

from which $A = 1$ and $B = -2$. Hence, $G(s) = 1/s - 2/(1 + 2s)$. Rearranging the second term to conform to the appropriate Laplace transform in the table gives

$$G(s) = 1/s - 1/(s + 0.5)$$

The constitutent parts produce the z-transform from the table as

$$G(z) = z/(z - 1) - z/(z - e^{-0.5T})$$

and after algebraic manipulation gives

$$G(z) = z(1 - e^{-0.57T})/(z - 1)(z - e^{-0.57T})$$

This is as expected from the Laplace transform, $a/s(s + a)$, in Table 6.9.

6.46 THE z-TRANSFORM OF A CLOSED-LOOP SYSTEM

Before z-transforms for the digital control loop illustrated in Figure 6.123 can be derived, it is first necessary to consider the nature of the input and output to the plant/process. The input, associated with digital-to-analog conversion hardware, is basically a zero-order-hold element which effectively supplies a control effort to the process in a pulsed form. The nature of the output from the controller relates to the analog-to-digital conversion which effectively is indicative of the discrete sampling process. Thus, the input is a step and the output sets the sampling characteristics of the loop on the basis of a sample time T.

This arrangement, with the nature of the appropriate signals, is shown in Figure 6.125. What is required is the z-transform for this complete arrangement, $G(z)$, and not simply the equivalence of $G(s)$ alone. This is usually referred to as the *pulse transfer function*. The total sampled response consists of the difference between the positive stepped input to the process and the same value delayed by one sample period. This will give the required pulse transfer function, and it can be written as

$$G(z) = \frac{1}{s} \cdot G(s) - e^{-sT} \frac{1}{s} G(s)$$

$$= (1 - z^{-1}) * \text{the } z\text{-transform of } \left(\frac{1}{s} \cdot G(s)\right)$$

or

$$G(z) = (z - 1)/z * \text{the } z\text{-transform of } \left(\frac{1}{s} G(s)\right) \tag{6.57}$$

which is the z-transform equivalence of all the elements shown in Figure 6.125.

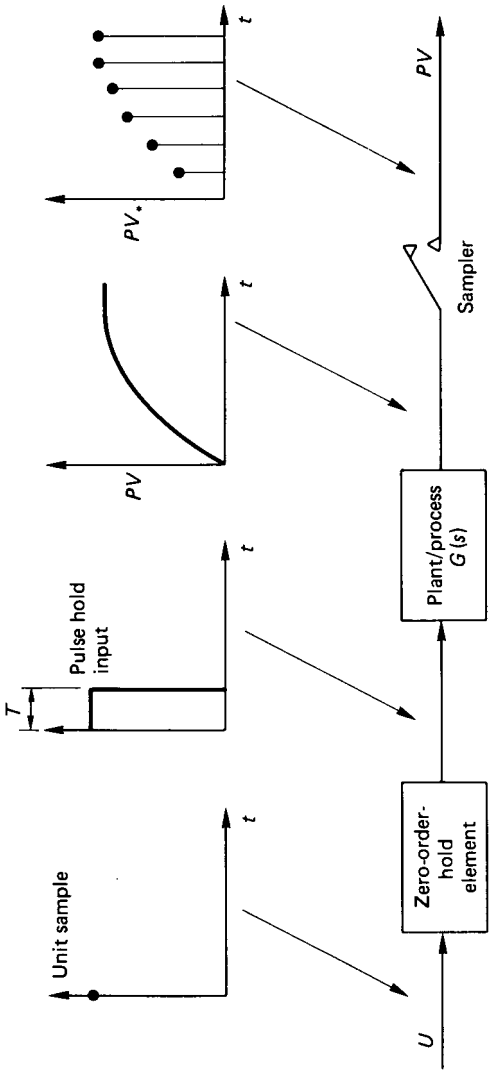


FIGURE 6.125 Discrete sampling in a control loop.

It should be noted that the discrete sampling process effectively introduces a phase lag into the system with its associated inherent stability problems. The closed-loop transfer function in terms of the z-transform is derived in a way similar to that for a continuous system, with $G(z)$ representing the pulse transfer function associated with the hold, plant/process and sampling routine.

Consider the basic loop as shown in Figure 6.126. The forward path transfer function is

$$\begin{aligned} PV(z) &= G(z)U(z) \\ &= G(z)C(z)E(z) \end{aligned}$$

and the closed-loop transfer function is

$$PV(z)/SP(z) = (C(z)G(z))/(1 + C(z)G(z)) \tag{6.58}$$

which has a form similar to that for the transfer function of a system with a continuous feedback control loop.

The procedure for evaluating the response of a plant/process to a step input with digital control using a sample time of T is as follows:

1. Obtain the pulse transfer function $G(z)$ from the known form for $G(s)$:

$$G(z) = (z - 1)/z * \text{the } z\text{-transform of } [G(s)/s]$$

2. Express the controller transfer function $C(s)$ in z-transform format.

3. Evaluate the closed-loop transfer function for $PV(z)/SP(z)$.

4. Algebraically manipulate this expression to yield a discrete time-ordered equation which formulates the current value of PV (say, PV_i) in terms of the past values of PV and SP. This is obtained using the relationships:

$$PV(z)z^{-1} = PV_{i-1} \text{ where } i - 1 \text{ denotes one period back in time}$$

$$PV(z)z^{-2} = PV_{i-2}$$

etc.

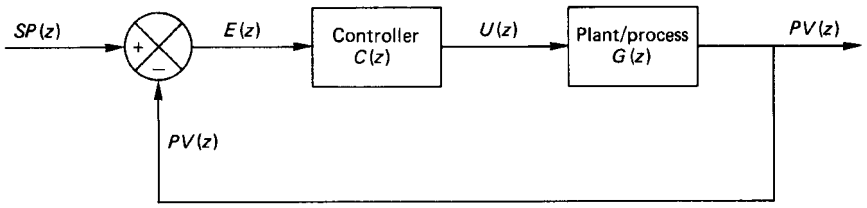


FIGURE 6.126 Closed loop in terms of a z-transform.

6.47 PROPORTIONAL CONTROL USING DIGITAL TECHNIQUES

For a proportional control strategy the controller transfer function $C(s)$ and the corresponding z-transform is simply the gain, K. Taking as an example the plant transfer function to be that for a first-order system of the form

$$G(s) = 1/(1 + \tau s)$$

the procedure gives the required pulse transfer function as

$$\begin{aligned} G(z) &= (z - 1)/z * \text{the } z\text{-transform of } [1/s(1 + \tau s)] \\ &= \frac{(z - 1)}{z}\left[\frac{z}{(z - 1)} \cdot \frac{(1 - e^{-T/\tau})}{(z - e^{-T/\tau})}\right] \\ &= C_1/(z - C_2) \end{aligned}$$

where $C_1 = 1 - e^{-T/\tau}$ and $C_2 = e^{-T/\tau}$. Hence the z-transform for the closed loop is

$$\begin{aligned} PV(z)/SP(z) &= (C(z)G(z))/(1 + C(z)G(z)) \\ &= ((KC_1)/z - C_2))/(1 + KC_1/(z - C_2)) \\ &= a/(z + b) \end{aligned}$$

where $a = KC_1$ and $b = (KC_1 - C_2)$.

It is now necessary to transform this equation into a discrete time-stepped form which expresses the current value of the process variable, PV_i, as a function of past values. This can be obtained by dividing the numerator and denominator by z to the power corresponding to the maximum value in the denominator. In this example the dividing factor is simply z, i.e.,

$$PV(z)/SP(z) = az^{-1}/(1 + bz^{-1})$$

or

$$PV(z) = aSP(z)z^{-1} - bPV(z)z^{-1}$$

Writing this as a time-difference equation yields

$$PV_i = aSP_{i-1} - bPV_{i-1} \tag{6.59}$$

For a step input, $SP = 1$ for $i > 1$ and starting values are, say,

$$i = 1; \quad t = 0; \quad SP = 1; \quad \text{and} \quad PV = 0$$

Equations of this form are easily programmed on a computer to produce a graphical time response of PV to a step input for declared values of system time constant τ, controller gain K, and sample time T.

Figures 6.127 and 6.128 illustrate the effect of the sample time, T, on the predicted response of the system using the z-transform technique outlined with $\tau = 1$ and $K = 10$. These are clearly illustrative of the introduction of a delay into the

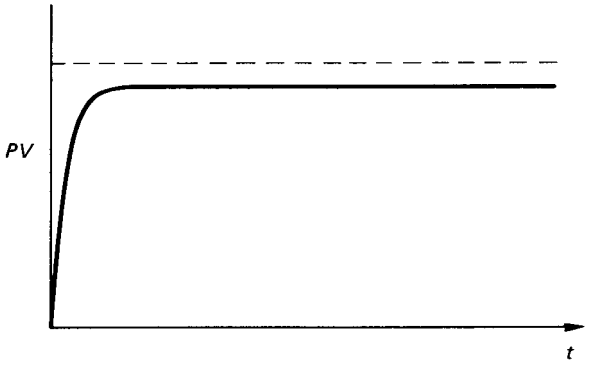


FIGURE 6.127 System response with $T = 0.01$ s.

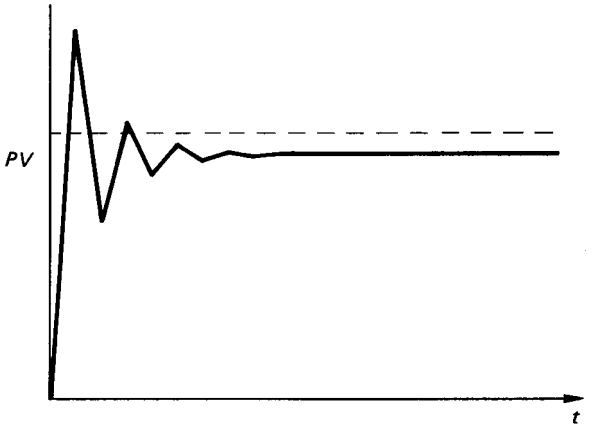


FIGURE 6.128 System response with $T = 0.15$ s.

system due to the incorrect sampling of the process, and leads to a transient behavior or, eventually, an unstable condition. The choice of a satisfactory sampling interval, T, was covered earlier, but it should be noted that it is dependent upon the system time constant and the controller settings. Hence one advantage of the z-transform technique in obtaining a simulation of the performance of a mathematically modeled system is the ability to investigate the effect of the sample rate for chosen controller settings.

6.48 THE z-TRANSFORM FOR A PID CONTROLLER

The PID control law for a continuous controller is

$$U = K[E + 1/T_i \int Edt + T_d dE/dt]$$

MICROPROCESSORS, INSTRUMENTATION, AND CONTROL

The integral term can be approximated numerically as a summation of rectangular elements which at any time t can be expressed as

$$I_i = I_{i-1} + E_i T$$

where I_{i-1} is the value of the integral term up to the instant $t - T$, or in z-transform notation:

$$I(z) = I(z)z^{-1} + E(z)T$$

i.e.,
$$I(z) = (Tz/(z - 1))E(z)$$

Similarly, the derivative term can be approximated as the slope of the line joining the current E value and the previous value, i.e.,

$$D_i = (E_i - E_{i-1})/T$$

or
$$D(z) = (E(z) - E(z)z^{-1})/T$$

i.e.,
$$D(z) = [(z - 1)/Tz]E(z)$$

The integral and derivative terms in z-transform notation can now be substituted into the full three-term expression to give $C(z)$:

$$C(z) = U(z)/E(z) = K\left[1 + \frac{1}{T_i} \cdot \frac{Tz}{(z - 1)} + T_d \frac{(z - 1)}{Tz}\right]$$

6.49 A P + I STRATEGY USING DIGITAL TECHNIQUES

For a $P + I$ strategy

$$C(z) = K\left[1 + \frac{1}{T_i} \cdot \frac{Tz}{(z - 1)}\right]$$

Applied to a first-order system, $G(s) = 1/(1 + \tau s)$ gives a corresponding pulse transfer function of

$$G(z) = C_1/(z - C_2)$$

with $C_1 = (1 - e^{-T/\tau})$ and $C_2 = e^{-T/\tau}$.

Substituting for $C(z)$ and $G(z)$ into the closed-loop transfer function and manipulating the result in a manner similar to that illustrated above yields the time-sequence equation as:

$$PV_i = aSP_{i-1} - bSP_{i-2} + cPV_{i-1} + dPV_{i-2} \tag{6.61}$$

where $a = K(1 - e^{-T/\tau})(1 + T/T_i)$
$b = K(1 - e^{-T/\tau})$
$c = (1 + e^{-T/\tau}) - K(1 - e^{-T/\tau})(1 + T/T_i)$
$d = K(1 - e^{-T/\tau}) - e^{-T/\tau}$

For $t < 0$ then $SP = 0$. At $t = 0$: $PV = 0$ and $SP = 1$ for a unit step input.

Figure 6.129 illustrates the response of a first-order system with $\tau = 1$ when a $P + I$ control strategy is applied using a gain K of 2 and an integral time T_i of 0.1. When the discrete sample time T is taken as 0.01 seconds, the result compares favorably with that for continuous control. If, however, the time step is increased to 0.35 seconds then the plant responds in a much more oscillatory manner for the same controller settings. This is illustrated in Figure 6.130. When T is increased beyond the value of 0.35 then the system behaves in an unstable manner. With digital control the sampling time, T, which would cause the system to go unstable depends upon the value of the system time constant and the controller settings.

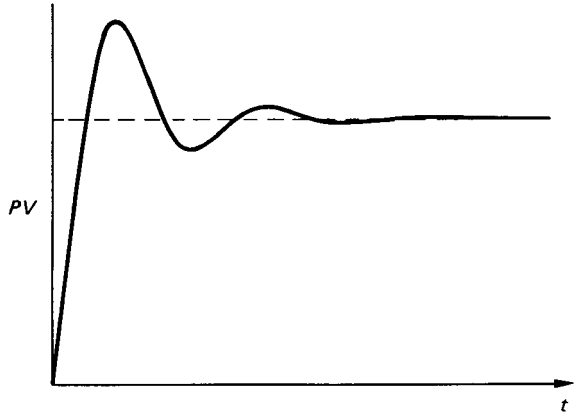


FIGURE 6.129 System response with $K = 2$, $T_i = 0.1$ s, and $T = 0.01$ s.

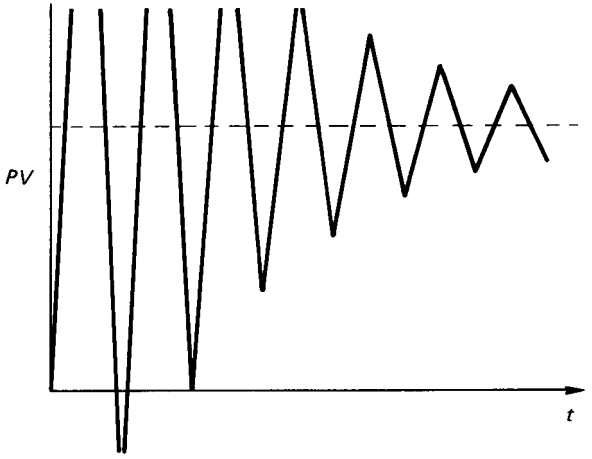


FIGURE 6.130 System response with $K = 2$, $T_i = 0.1$ s and $T = 0.35$ s.

6.50 STABILITY IN DISCRETE TIME SYSTEMS

As seen in the previous sections, an incorrectly chosen sample time can result in a system becoming unstable due to the imposed delay. This initial value of T can be obtained by considering the representation of the poles of the pulse transfer function $G(z)$ in the z-plane.

If the closed-loop transfer function in z-transform notation is written in terms of a numerator $N(z)$ and a denominator $D(z)$ then the poles are defined as the roots of the equation $D(z) = 0$. These roots may contain a real and imaginary component and take the form $a + jb$ with the resulting plot termed the z-plane. Systems whose poles are all inside the unit circle in the z-plane are stable and those with any poles outside it are unstable. The position of the poles within the unit circle are indicative of the degree of the resulting oscillatory motion which may result following a step input. A z-plane plot is shown in Figure 6.131.

When using the z-transform method of analysis for the digital control of a system the position of the poles in the z-plane are dependent upon the magnitude of the sample time taken. For example, consider a process having a transfer function $G(s) = 1/s$ and controlled using a proportional control strategy. Using the method previously the corresponding pulse transfer function and closed-loop z-transfer function are, respectively,

$$G(z) = T/(z - 1)$$

and

$$PV(z)/SP(z) = KT/(z - 1 + KT)$$

i.e.,

$$D(z) = z - 1 + KT$$

For the loop to be stable the poles of the real axis should lie in the range of $-1 < z < +1$, which means, by putting $z = -1$, that $KT = 2$ for the limit of stability. If $K = 5$ then the sample time T must be less than 0.4 for the system to remain stable.

It can be seen that the position of the poles in the z-plane is a useful concept in determining whether a system will be stable for a chosen sample time.

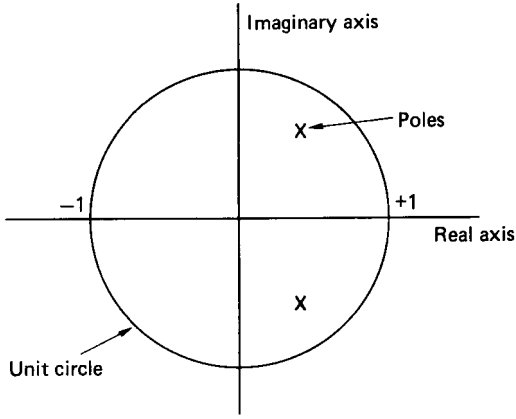


FIGURE 6.131 The z-plane.

PART 10
STATE VARIABLE TECHNIQUES

Charles J. Fraser

6.51 STATE VARIABLE REPRESENTATION OF SYSTEMS

The state variable approach to system modeling involves the use of matrix and vector methods which provide a consistent solution procedure for the analysis of complex control problems. While these methods can be applied to any single-input–single-output control system, their real power lie in their application to multiple-input–multiple-output systems. The state space representation also finds application to the analysis of nonlinear control systems where the non-linear elements of a system are accounted for in a self-consistent manner.

In order to generate a system of state variable equations the system must be adequately defined in terms of the governing differential equations in a continuous system or of the difference equations in a discrete one. A convenient method of representing the mathematical operations and equations is by means of a block diagram made up of the appropriate sequence of operators. In a continuous control system the most common operators are shown in Figure 6.132. The simple damped spring and mass system serves as a useful illustrative example (see Figure 6.133).

The relation which governs the motion of the mass as a function of time is given as a second-order linear differential equation, i.e.,

$$\frac{d^2y}{dt^2} + 2\zeta\omega_n \frac{dy}{dt} + \omega_n^2 y = u \tag{6.62}$$

or

$$\ddot{y} + 2\zeta\omega_n \dot{y} + \omega_n^2 y = u \tag{6.63}$$

Solving for the highest derivative gives

$$\ddot{y} = u - 2\zeta\omega_n \dot{y} - \omega_n^2 y \tag{6.64}$$

where u is the input forcing function.

The block diagram representation of equation (6.64) may be developed by first integrating y twice, as shown in Figure 6.134. The loop may then be closed by satisfying the requirement of equation (6.64) as shown in Figure 6.135.

The quantities y and $\dot{y}$ are called state variables. These are not unique and there is an endless range of possible combinations which may be chosen. It is normal, however, to select as state variables those quantities which are significant to the problem under consideration. In this respect the output, the first derivative of the output and any higher-order derivatives of the output would make the most suitable choice of state variables. In choosing the set of state variables it should be recognized that the state variables must be related through a differential or a difference equation. In a continuous-control system the normal practice is to select the output from any integrator as an appropriate state variable.

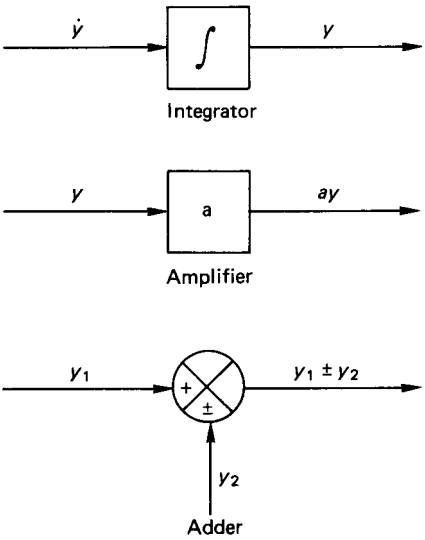


FIGURE 6.132 Mathematical operators in continuous-control systems.

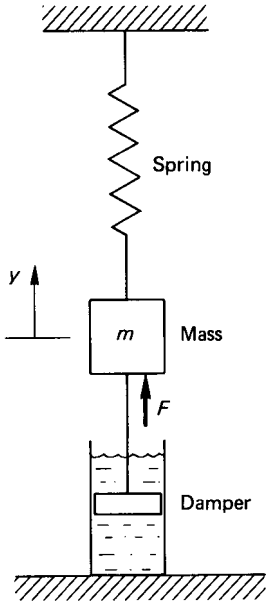


FIGURE 6.133 Simple damped spring and mass system.

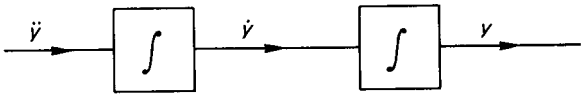


FIGURE 6.134 Double-integration process.

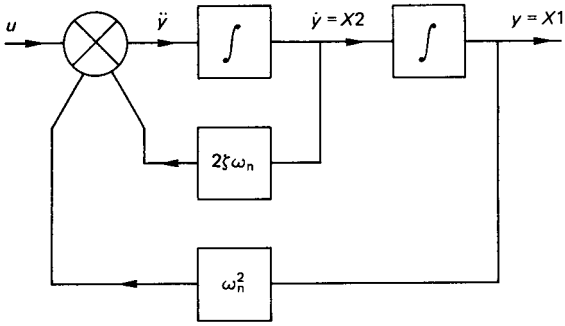


FIGURE 6.135 Block diagram representation of equation (6.64).

For the damped spring mass system the following state variables are selected:

$$X1 = y \quad \text{and} \quad X2 = \dot{y}$$

Consideration of the block diagram of Figure 6.135 shows that

$$\dot{X}1 = X2 \tag{6.65}$$

$$\dot{X}2 = -\omega_n^2 X1 - 2\zeta\omega_n X2 + u \tag{6.66}$$

It transpires therefore that the state variable representation reduces the original second-order differential equation to two simultaneous first-order differential equations. Similarly, a third- or higher-order differential equation could be represented as a set of three or a higher number of first-order differential equations.

Writing equations (6.65) and (6.66) in matrix notation gives

$$\begin{bmatrix} \dot{X}1 \\ \dot{X}2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega_n^2 & -2\zeta\omega_n \end{bmatrix} \begin{bmatrix} X1 \\ X2 \end{bmatrix} + u \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{6.67}$$

and

$$y = [1 \quad 0] \begin{bmatrix} X1 \\ X2 \end{bmatrix} + \mathrm{u}[0] \tag{6.68}$$

Any suitable matrix method can be used to solve for $X1$ which is equivalent to the system output (i.e., the position of the mass as a function of time) for any given input u. Alternatively, equations (6.65) and (6.66) may be written as simple finite difference approximations and solved simultaneously over an appropriate time increment. For free vibration following a step disturbance to the system, $u = 0$, the finite difference approximation may be written as

$$\Delta X2_i = \Delta t[-2\zeta\omega_n X2_{i-1} - \omega_n^2 X1_{i-1}] \tag{6.69}$$

$$X2_i = X2_{i-1} + \Delta X2_i \tag{6.70}$$

$$\Delta X1_i = \Delta t[X2_i] \tag{6.71}$$

$$X1_i = X1_{i-1} + \Delta X1_i \tag{6.72}$$

With suitable values assigned to ζ and ω_n and initial values designated to $X1$ and $X2$ at time = zero, the solution can march forward in time using a suitably small time increment Δt. A plot of $X1$ (which is the system output) against time gives the solution to the equation in the time domain. The more usual state variable representation, however, is shown as a plot of the state variables $X2$ against $X1$. For the damped spring mass system considered, this state space plot is a record of the instantaneous values of the velocity of the mass against its position with time as parameter. Figure 6.136 shows a set of results for selected values of $\sim$ ranging between 0 and 1, with initial conditions of $X1(0) = -0.25$ and $X2(0) = 0$.

The set of curves which result from the solutions of the state variable equations are referred to as trajectories in the state space. For a second-order system, the plot is more generally called the phase plane. The values of the state variables are said to represent the state of the system. With a knowledge of future inputs and the

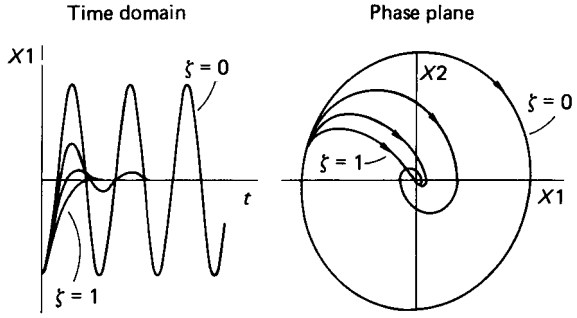


FIGURE 6.136 Time-domain and state-space representation of the damped spring mass system.

governing state space equations, the phase plane trajectory can be predicted and hence also the behavior of the system in general. Figure 6.137 shows another set of phase-plane trajectories for the damped spring mass system. In these plots the system is subject to heavy damping and the trajectories start from a range of different initial conditions. Plots such as these are referred to as phase-plane portraits.

One further example is included where a negative damping factor of -1 has been imposed and the computation started from $X1(0) = -0.10$ and $X2(0) = 0$. The resultant phase-plane plot is shown in Figure 6.138.

Negative damping renders the system unstable and the amplitude of the oscillation increases with time as shown in the time-domain plot. The phase-plane plot indicates a trajectory which is spiraling outwards, and this is the general state-space representation of an unstable system.

Other information which may be easily gleaned from the phase-plane trajectory are an indication of whether the system response is over- or underdamped and the extent of any system overshoots.

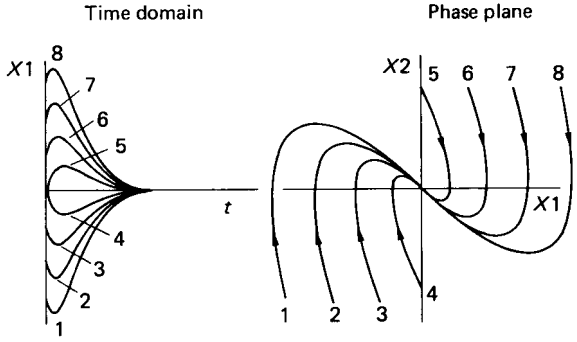


FIGURE 6.137 Phase-plane characteristic for the spring mass system with very heavy damping.

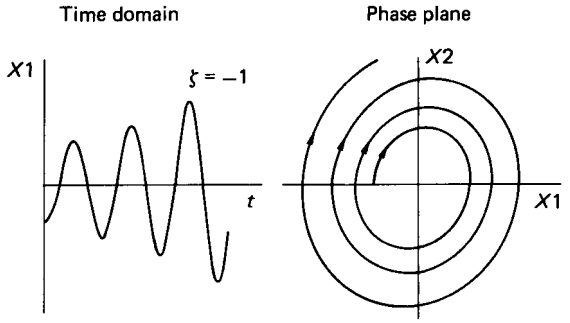


FIGURE 6.138 Phase-plane trajectory for the spring mass system with negative damping.

6.52 APPLICATION TO A FIRST-ORDER SYSTEM WITH A P + I CONTROLLER

Figure 6.139 shows the control block diagram for a process which has an open-loop transfer function which can be modeled as a first-order differential equation. The control effort, U, is determined through a controller which utilizes a combination of proportional and integral action as shown. The closed-loop transfer function may be written as

$$\frac{PV(s)}{SP(s)} = \frac{K[1 + 1/(T_i s)][k/(1 + \tau s)]}{1 + K[1 + (1/T_i s)][k/(1 + \tau s)]} \tag{6.73}$$

Some algebraic manipulation results in

$$\frac{PV(s)}{SP(s)} = \frac{(kK/\tau T_i)(1 + T_i s)}{s^2 + [(1/\tau) + (kK/\tau)]s + (kK/\tau T_i)} \tag{6.74}$$

Cross-multiplication gives

$$\frac{d^2(PV)}{dt^2} + \left[\frac{1}{\tau} + \frac{kK}{\tau}\right]\frac{d(PV)}{dt} + \frac{kK}{\tau T_i}(PV) = \frac{kK}{\tau}\left[\frac{d(SP)}{dt} + \frac{(SP)}{T_i}\right] \tag{6.75}$$

The differential equation (6.75) may be represented in block diagram form as shown in Figure 6.140.

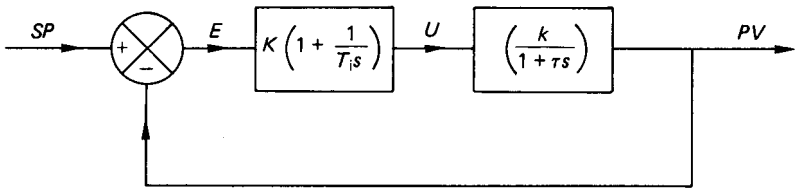


FIGURE 6.139 First-order system with $P + I$ controller.

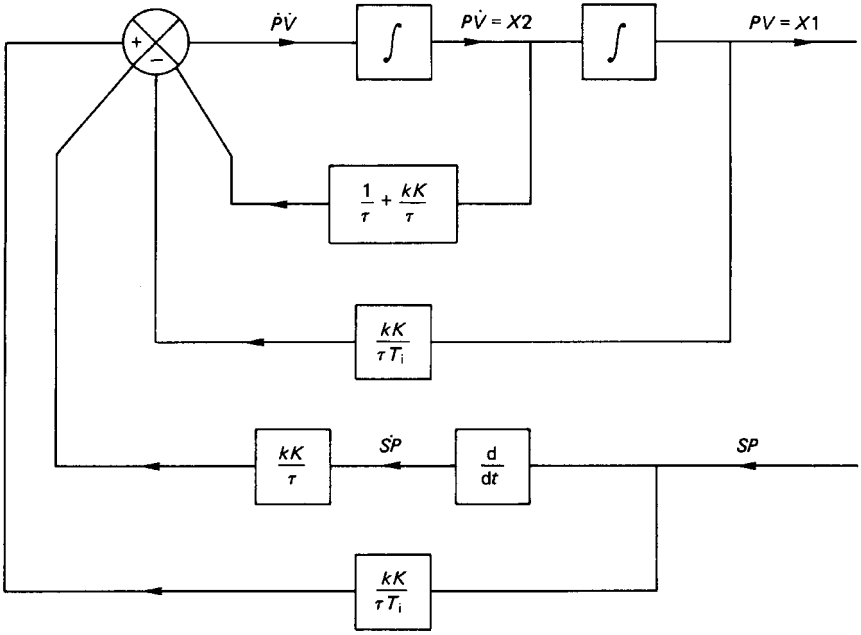


FIGURE 6.140 Block diagram representation of equation (6.75).

Choosing the state variables as the outputs from the two integrators gives

$$X1 = (PV) : X2 = (\dot{PV})$$

Thus,

$$\dot{X}1 = X2 \tag{6.76}$$

$$\dot{X}2 = \frac{-kK}{\tau T_i} X1 - \left[\frac{1 + kK}{\tau}\right] X2 + \frac{kK}{\tau}\left[\dot{SP} + \frac{SP}{T_i}\right] \tag{6.77}$$

The state variable equations can be written again as finite-difference approximations and a solution marched forward in time from specified initial conditions. The step response of the system for various settings of system gain, controller gain, integral time setting, and system time constant are illustrated in Figure 6.141.

6.53 APPLICATION TO A SECOND-ORDER SYSTEM WITH A P + I CONTROLLER

If the process to be controlled can be represented as a second-order differential equation, then the open-loop transfer function for the process is

$$G(s) = \frac{k}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{6.78}$$

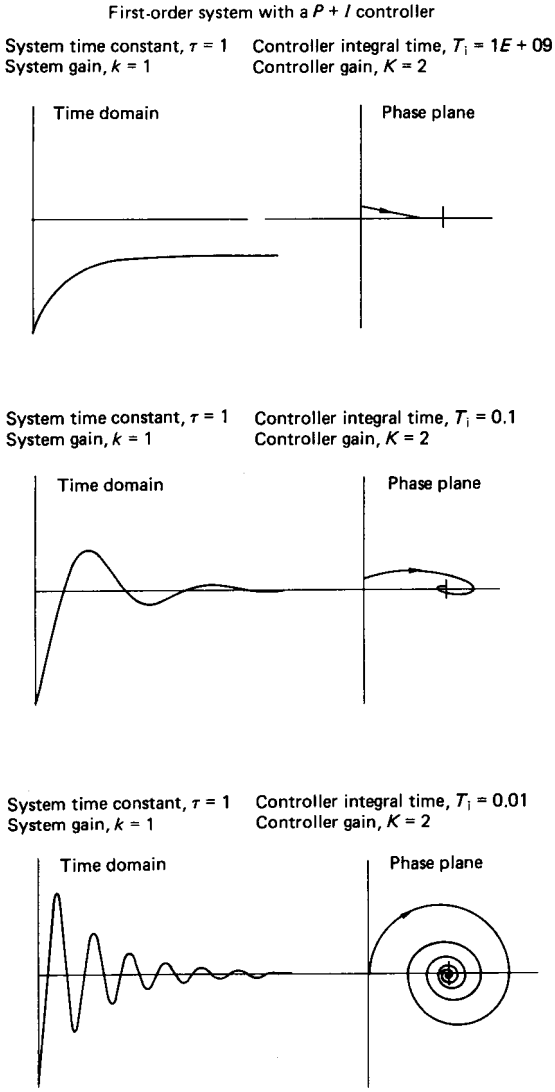


FIGURE 6.141 System response to a step input.

The closed-loop transfer function may be written as

$$\frac{PV(s)}{SP(s)} = \frac{kK(1 + T_i s)}{(T_i s^3 + 2\zeta\omega_n T_i s^2 + \omega_n^2 T_i s + kKT_i s + kK)} \tag{6.79}$$

or

$$\frac{d^3(PV)}{dt^3} + 2\zeta\omega_n \frac{d^2(PV)}{dt^2} + (\omega_n^2 + kK)\frac{d(PV)}{dt} + \frac{kK}{T_i}(PV) = kK\left[\frac{d(SP)}{dt} + \frac{SP}{T_i}\right] \tag{6.80}$$

The corresponding state variable representation is

$$\dot{X}1 = X2 \tag{6.81}$$

$$\dot{X}2 = X3 \tag{6.82}$$

$$\dot{X}3 = -2\zeta\omega_n x3 - (\omega_n^2 + kK)X2 - \frac{kK}{T_i}X1 + kK\left[(S\dot{P}) + \frac{(SP)}{T_i}\right] \tag{6.83}$$

where $X1 = PV$, $X2 = (\dot{P}V)$, and $X3 = (\dot{P}\dot{V})$.

The three ordinary differential equations involving the three state variables may be solved simultaneously and a phase-plane trajectory plotted as before. Strictly, the trajectory should be plotted, for this example, in three-dimensional state space. This would involve a plot of the trajectory in terms of the state variables $X1$, $X2$, and $X3$ with time as parameter. This would be difficult to plot and for still higher-order systems might be difficult even to visualize. Note, however, that the simple two-dimensional phase plane still has some significance and can be used as an indication of the system's general stability.

6.54 NONLINEAR SYSTEM ELEMENTS (METHOD OF ISOCLINES)

In many real control systems there is often some form of nonlinear element within the control loop which eliminates the use of linear theory to predict the closed-loop system response. In some cases, a linear approximation can be devised for the nonlinear element but the approximation will always imply that there is some degree of error involved in the system model.

The phase plane represents one viable method of accounting for those nonlinearities which cannot be either ignored or approximately linearized. It is not always necessary to plot the phase-plane trajectory in its entirety but rather to approximate the curve by estimating its slope at a number of locations in the phase plane. The approximate trajectory can then be fitted to match the known slopes in the phase plane. This is the so-called method of isoclines, which is best illustrated by an example.

Saturation Characteristics

Saturation can occur, for example, when an amplifier is operated outside its linear range. This nonlinear effect can be depicted as shown in Figure 6.142.

There are three distinct regions over which the amplifier has a significantly different characteristic:

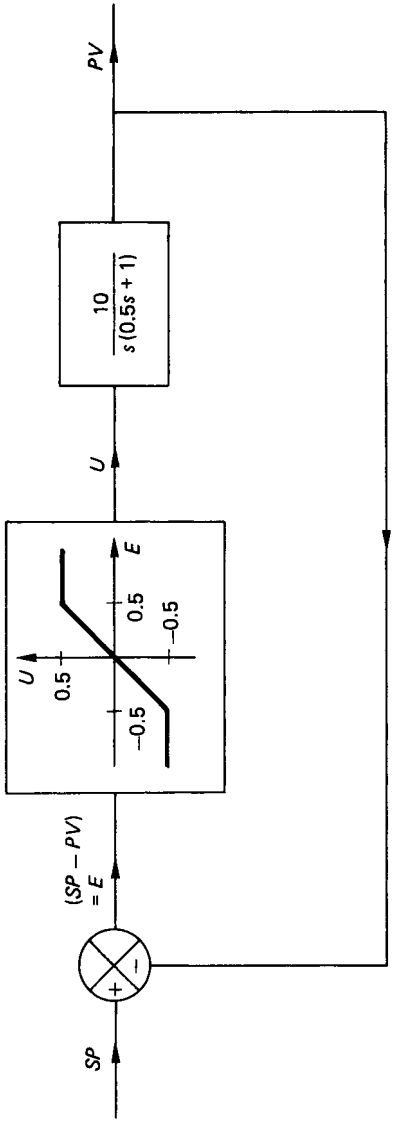


FIGURE 6.142 Position-control servo with saturation characteristic.

1. $E \geqslant 0.5$; $U = 0.5$
2. $-0.5 < E < 0.5$; $U = E$
3. $E \leqslant -0.5$; $U = -0.5$

In all regions

$$PV = 10U/[s(0.5s + 1)]$$

or

$$10U = (0.5s^2 + s)PV = (0.5s^2 + s)(SP - E)$$

Thus,

$$10U = \frac{1}{2}\frac{d^2(SP)}{dt^2} + \frac{d(SP)}{dt} - \frac{1}{2}\frac{d^2E}{dt^2} - \frac{dE}{dt}$$

For a step input at time $t > 0$,

$$\frac{d^2(SP)}{dt^2} = \frac{d(SP)}{dt} = 0$$

The equation reduces to

$$\frac{d^2E}{dt^2} + 2\frac{dE}{dt} + 20U = 0$$

In region (1), (U = constant = 0.5)
Thus,

$$\frac{d^2E}{dt^2} + 2\frac{dE}{dt} + 10 = 0$$

Choosing as a suitable state variable, $X1 = dE/dt$,

$$\dot{X}1 = -10 - 2X1$$

Defining the slope S as

$$S = \frac{\dot{X}1}{X1} = \frac{-10 - 2X1}{X1} = -2 - \frac{10}{X1}$$

Alternatively,

$$X_1 = \frac{-10}{S + 2}$$

Equation (6.84) is the isoclinal equation which applies in the region $U > 0.5$. The isoclinal equation defines a set of curves along which the slope of the trajectory in the phase plane of dE/dt against E will be constant. In selecting a suitable range of values for S, the isoclines may then be plotted in the phase plane. Equation (6.84) shows that these isoclines will be manifested as a set of lines which would be parallel to the horizontal axis.

In region 2 ($U = E$)
The governing equation becomes

$$\frac{d^2E}{dt^2} + \frac{2dE}{dt} + 20E = 0 \tag{6.84}$$

Again choosing $X1 = dE/dt$ gives

$$S = \frac{\dot{X}1}{X1} = \frac{-20E - 2X1}{X1} = \frac{-2 - 20E}{X1}$$

i.e.,

$$X1 = \frac{-20E}{S + 2} \tag{6.85}$$

The isoclinal equation in region 2 (equation (6.85)) defines a family of curves of constant slope, passing through the origin, with a gradient given by $-20/(S + 2)$.

In region 3 ($U = constant = -0.5$)
Region 3 is similar to region 1 with the isoclinal equation given by

$$X1 = \frac{+10}{S + 2} \tag{6.86}$$

Equations (6.84) to (6.86) define the isoclines for the phase-plane trajectory of the error response, and these are shown in Figure 6.143. From any given initial conditions the phase-plane trajectory may be patched in by smoothing a curve through the known values of S for each of the isoclines shown in the figure. For the particular system constants given a step input of 1 radian, the phase-plane trajectory of the error response shows that the system is stable and that the amplifier does not saturate on the first overshoot.

System with a Dead-Band Characteristic

Figure 6.144 shows a nonlinear system element which exhibits a dead-band characteristic. As in the previous example, there are three distinctive regions in which the behavior of the nonlinear element may be uniquely described:

1. $E > 0.1$, $U = 1(E - 0.1)$
2. $-0.1 < E < +0.1$, $U = 0$, i.e., the dead band
3. $E < -0.1$, $U = 1(E + 0.1)$

Following a similar approach to that adopted earlier in all regions

$$\frac{d^2E}{dt^2} + 2\zeta\omega_n \frac{dE}{dt} + U\omega_n^2 = 0$$

Choosing the state variable as $X1 = dE/dt$,

$$X1 = \frac{-U\omega_n^2}{2\zeta\omega_n + S} \tag{6.87}$$

In region 1 the isoclinal equation becomes

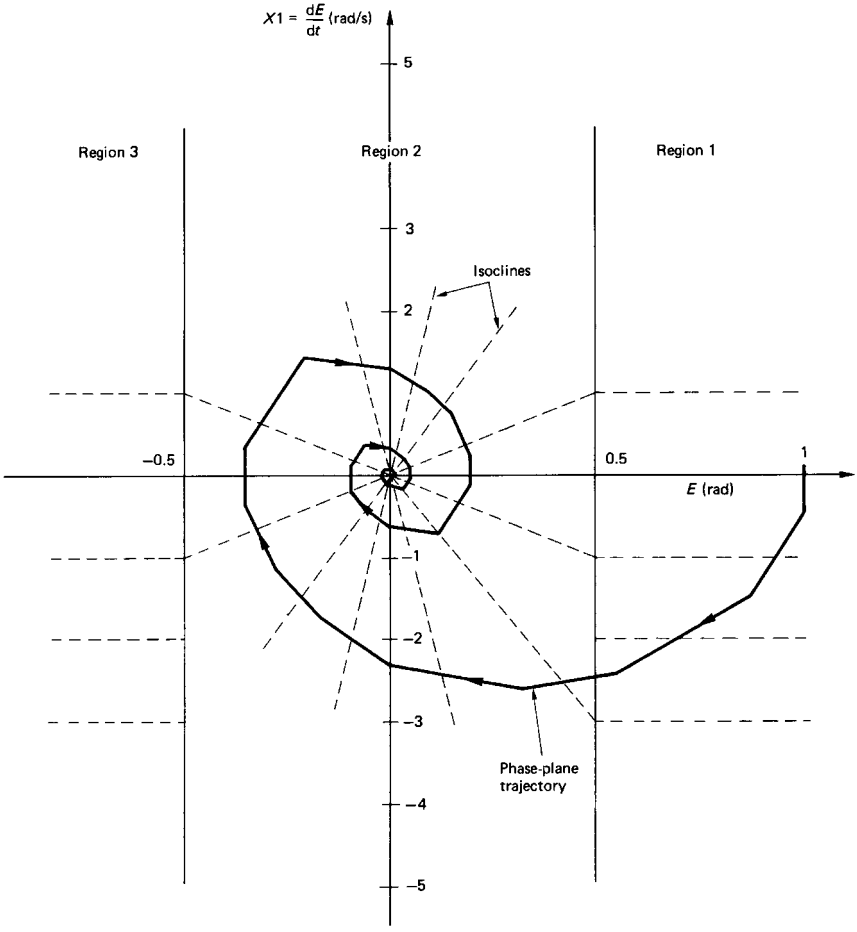


FIGURE 6.143 Phase-plane trajectory for saturation characteristic.

$$X1 = \frac{-\omega_n^2(E - 0.1)}{S + 2\zeta\omega_n} \tag{6.88}$$

Equation (6.88) defines a family of isoclines as a set of constant slope lines with origin (+0.1,0). Region 3 is similar to region 1 and the isoclinal equation defines a family of constant slope with origin (−0.1,0). In the dead-band region (2) $U = 0$ and the slope becomes

$$\text{Slope} = S = \frac{-2\zeta\omega_n X1}{X1} = -2\zeta\omega_n$$

The resulting phase-plane trajectory takes the form shown in Figure 6.145.

Other system nonlinearities may be handled similarly. These may include, for example, the on/off characteristics of a relay or hysteresis effects as typified by

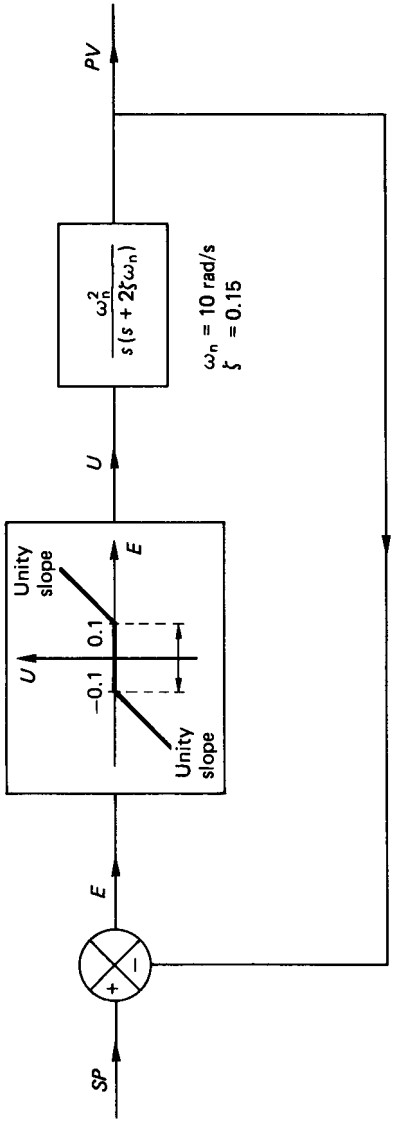


FIGURE 6.144 Control system with a nonlinear dead-band characteristic.

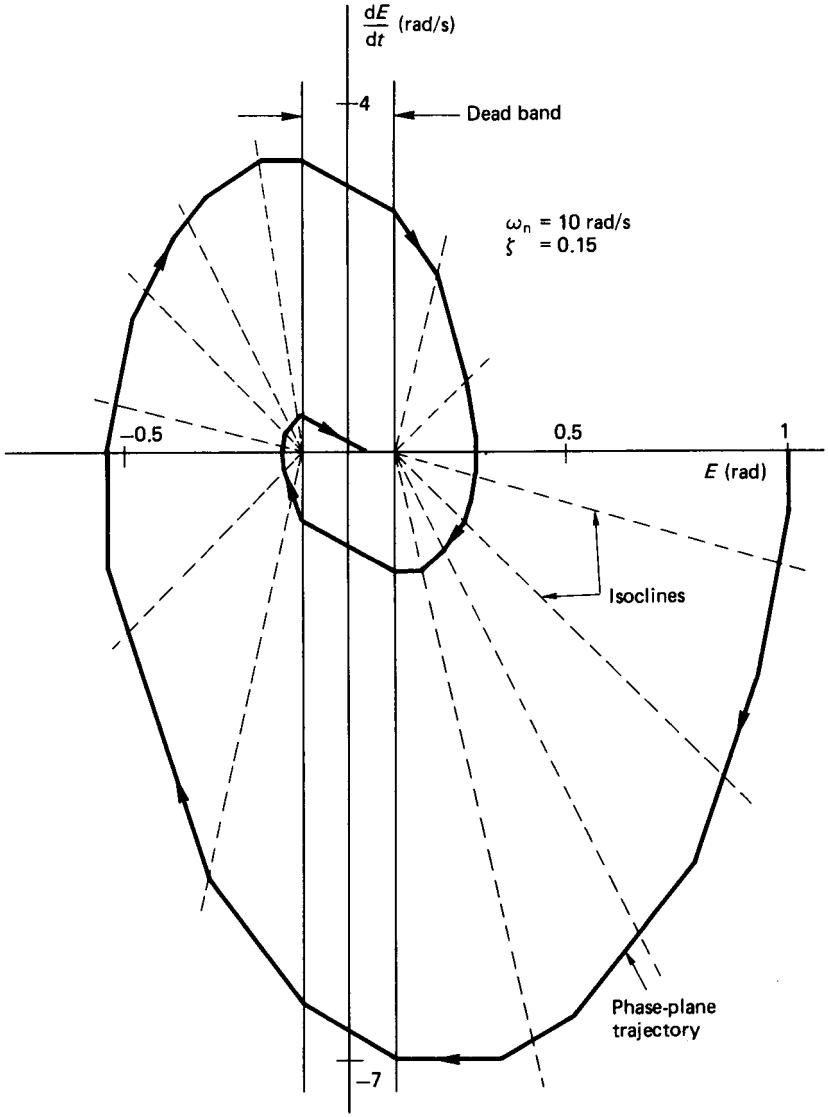


FIGURE 6.145 Phase-plane trajectory for a system with a dead-band characteristic.

backlash in geared systems. The general representation and typical phase-plane trajectories for some of the common types of system nonlinearities are shown in Figure 6.146.

The two examples which include hysteresis effects result in a phase-plane trajectory which traces out a limit cycle. Limit cycles are characterized by isolated closed curves in the phase plane and all stable near trajectories approach the limit

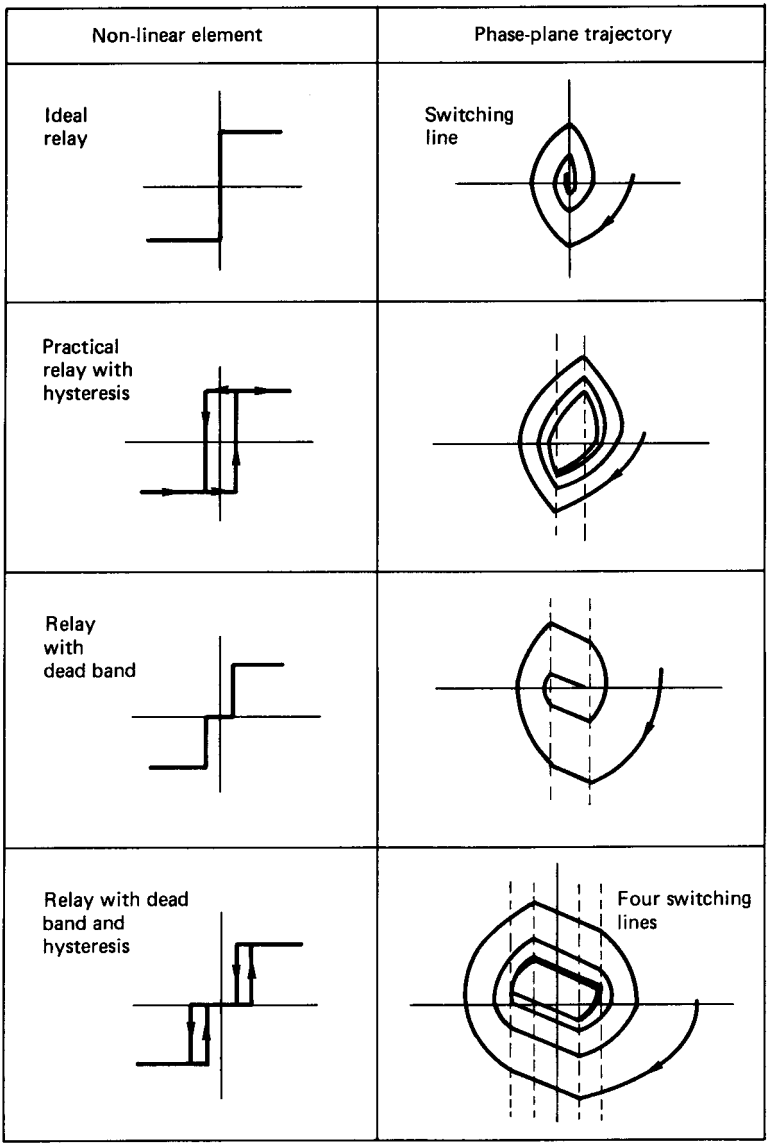


FIGURE 6.146 Common system nonlinearities and corresponding phase-plane trajectories.

cycle as time approaches infinity. The stable limit cycle corresponds to a stable periodic motion in the physical system. Limit cycles may also be unstable, in which case the near trajectories move away from the closed curve.

Unfortunately, the phase-plane approach is generally limited to unforced second-order systems. In theory, the phase-plane concepts may be extended to higher-order state space. This, however, has not been successful, due to the problems of determining, presenting, and visualizing trajectories in n-dimensional state space.

6.55 SAMPLED-DATA SYSTEMS

In a discretely sampled system the relationship between output and the input can be represented in terms of z-transforms or, ultimately, as a time-difference equation. In any system described by linear difference equations a block diagram can be similarly constructed using the basic building elements of the adder, the amplifier, and the unit time delay. The unit time delay, which is analogous to the integrator in continuous systems, is depicted in Figure 6.147. In the figure the input to the unit delay appears at the output one period later, or delayed by a time T.

If the difference equation has, for example, the form given in equation (6.89) the block diagram for the system is represented as that shown in Figure 6.148, i.e.,

$$y_i + ay_{i-1} + by_{i-1} = x_i \tag{6.89}$$

or

$$y_i = x_i - ay_{i-1} - by_{i-2}$$

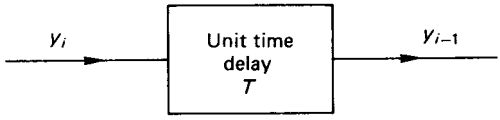


FIGURE 6.147 The unit time delay.

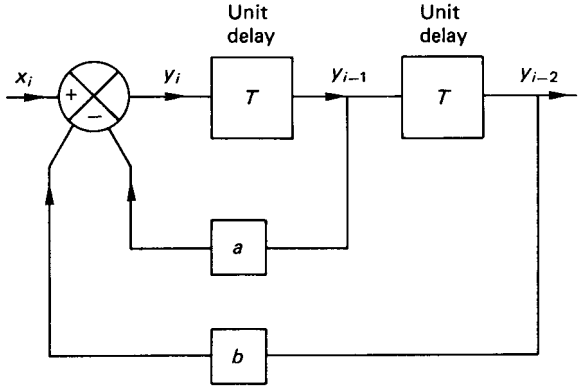


FIGURE 6.148 Block representation of equation (3.89).

The system may also be handled in z-transform notation and a block diagram constructed to represent the system in terms of the appropriate operators.

Following the procedures outlined in Part 9, a closed- loop digital control system can be represented by the appropriate z-transform operators. The closed-loop transfer operator may subsequently be manipulated and reduced to a finite-difference or time-sequence equation. For the $P + I$ control strategy used in conjunction with a process which can be represented as a first-order differential equation the resulting time-sequence equation (6.61) rewritten as

$$PV_i = aSP_{i-1} - bSP_{i-2} + cPV_{i-1} + dPV_{i-2} \tag{6.61}$$

where a, b, c, and d are as defined and i is an integer variable used to denote the time periods.

The block diagram to represent the difference equation is shown in Figure 6.149. The appropriate state variables are chosen as

$$X1 = PV_{i-2}: X2 = PV_{i-1}$$

The state variable equations can be written in the form

$$X1_{k+1} = X2_k \tag{6.90}$$

$$X2_{k+1} = aSP_k - bSP_{k-1} + cX2_k + dX1_k \tag{6.91}$$

where k is an integer variable used to denote the time periods. In matrix notation the state variable representation is

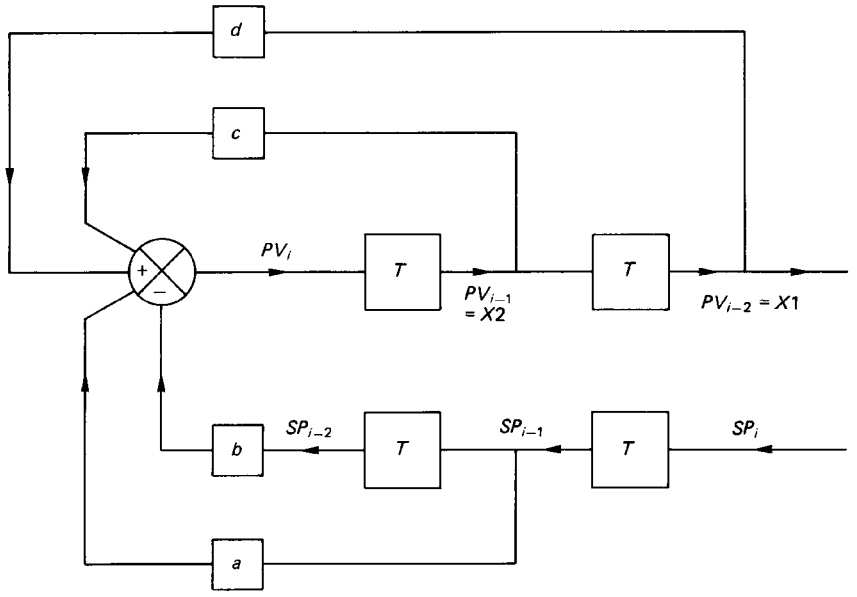


FIGURE 6.149 Block diagram representation of equation (6.61).

$$\begin{bmatrix} X1_{k+1} \\ X2_{k+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ d & c \end{bmatrix} \begin{bmatrix} X1_k \\ X2_k \end{bmatrix} + (aSP_k - bSP_{k-1}) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{6.92}$$

$$PV_i = [0 \quad 1] \begin{bmatrix} X1_{k+1} \\ X2_{k+1} \end{bmatrix} + (aSP_k - bSP_{k-1})[0] \tag{6.93}$$

With any given initial conditions, equations (6.90) and (6.91) may be solved simultaneously for discrete time steps in T. Figure 6.150 shows the response to a step input when $K = 2$, $T_i = 0.1$, and T is set at 0.2 s. The phase-plane trajectory is shown in two forms in the figure. In one form the plot is given of $X2$ against $X1$, these being the chosen state variables. This trajectory shows that the system is stable under the imposed conditions. An alternative representation is shown where the parameter $(X2 - X1)/T$ is plotted against $X2$. In terms of the actual control system, this second plot is equivalent to a graph of the discrete time rate of change of the process variable against the process variable itself. In either plot, however, the stability of the system is adequately depicted in graphical form.

6.56 STATE VARIABLE TRANSFORMATIONS

The concept of state variables, introduced earlier, was applied to the solution of closed-loop feedback control systems. Writing the state variable equations in matrix form for the example given there gives

$$\begin{bmatrix} \dot{X}1 \\ \dot{X}2 \\ \dot{X}3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \dfrac{-kK}{T_i} & -(\omega_n^2 + kK) & -2\zeta\omega_n \end{bmatrix} \begin{bmatrix} X1 \\ X2 \\ X3 \end{bmatrix} + kK\left(\dot{SP} + \frac{SP}{T_i}\right) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{6.94}$$

In addition, the system output PV, or $X1$, can be written in matrix form as

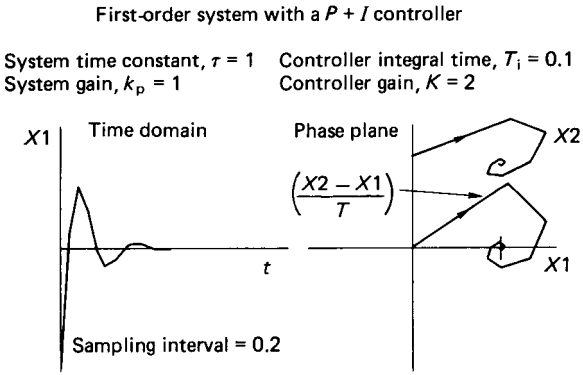


FIGURE 6.150 System response to a step input.

$$PV = [1 \quad 0 \quad 0] \begin{bmatrix} X1 \\ X2 \\ X3 \end{bmatrix} + kK \left(\dot{S}P + \frac{SP}{T_i} \right) [0] \tag{6.95}$$

These matrix relations fall into a generalized form which is also apparent in the pairs of equations given as (6.67) and (6.68). In shorthand notation, the equations can be written as

$$\dot{X} = AX + Bu \tag{6.96}$$

$$y = CX + Du \tag{6.97}$$

The matrix A is said to be in companion form and the components of X are the state variables $X1$, $X2$, $X3$, . . . X_n. It can be seen that in this form the highest subscripted X_n will always be a function of all the state variables. In many applications it is more useful to have the state variables decoupled, in which case each $\dot{X}_i$ is a function of X_i and the input only. This implies that the A matrix will then be in a convenient diagonal form.

Matrix B is the input matrix and C is the output matrix. For a single-input–single-output system, the matrix D is always zero. If, however, the system has multiple inputs and multiple outputs, then the input u and the output y are vectors and the matrix D then has a finite value.

The conversion of the state equations to form a diagonal A matrix can be accomplished using a variety of techniques. One method is included here for an example.

Consider the transformation where

$$X = Mq \tag{6.98}$$

Substituting this transformation into equations (6.96) and (6.97) gives

$$M\dot{q} = AMq + Bu$$

and

$$y = CMq$$

or

$$\dot{q} = M^{-1}AMq + M^{-1}Bu$$

i.e.,

$$\dot{q} = A_1 q + B_1 u \tag{6.99}$$

and

$$y = C_1 q \tag{6.100}$$

where $A_1 = M^{-1}AM$, $B_1 = M^{-1}B$, and $C_1 = CM$.

The procedure involved in transforming the state equations are best illustrated in a simple example. Suppose a particular dynamical system can be represented by the differential equation

$$\dddot{y} - 2\ddot{y} + \dot{y} - 2y = u \tag{6.101}$$

The state equations in companion form are

$$\begin{bmatrix} \dot{X}1 \\ \dot{X}2 \\ \dot{X}3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -1 & 2 \end{bmatrix} \begin{bmatrix} X1 \\ X2 \\ X3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u \tag{6.102}$$

and

$$y = [1 \quad 0 \quad 0] \begin{bmatrix} X1 \\ X2 \\ X3 \end{bmatrix} \tag{6.103}$$

The characteristic equation of A is

$$[\lambda I - A] = \lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + a_1\lambda^1 + a_0 = 0 \tag{6.104}$$

where I is the unit matrix, λ_i are the eigenvalues of A, and a_i are the coefficients associated with the left-hand side of equation (6.101). Thus,

$$\lambda^3 - 2\lambda^2 + \lambda - 2 = 0$$

i.e.,
$$(\lambda - 2)(\lambda^2 + 1) = 0$$

Hence,
$$\lambda_1 = 2, \lambda_2 = i \text{ and } \lambda_3 = -i.$$

where $i = \sqrt{-1}$.

Defining the modal matrix, M, such that

$$AM = A_1M$$

where

$$A_1 = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

Then
$$A_1 = M^{-1}AM \tag{6.105}$$

It can be shown that

$$M = \begin{bmatrix} 1 & 1 & 1 & — & — & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 & & & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 & — & — & \lambda_n^2 \\ — & — & — & — & — & — \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \lambda_3^{n-1} & — & — & \lambda_n^{n-1} \end{bmatrix} \tag{6.106}$$

Thus, for example,

$$M = \begin{bmatrix} 1 & 1 & 1 \\ 2 & i & -i \\ 4 & -1 & -1 \end{bmatrix}$$

and

$$M^{-1} = \frac{1}{10}\begin{bmatrix} 2 & 0 & 2 \\ 4 + 2i & -5i & -1 + 2i \\ 4 - 2i & +5i & -1 - 2i \end{bmatrix}$$

$$\therefore B_1 = M^{-1}B = \frac{1}{10}\begin{bmatrix} 2 & 0 & 2 \\ 4 + 2i & -5i & -1 + 2i \\ 4 - 2i & +5i & -1 - 2i \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\therefore B_1 = \frac{1}{10}\begin{bmatrix} 2 \\ -1 + 2i \\ -1 - 2i \end{bmatrix}$$

And $C_1 = CM$,

$$\therefore C_1 = [1 \quad 0 \quad 0]\begin{bmatrix} 1 & 1 & 1 \\ 2 & i & -i \\ 4 & -1 & -1 \end{bmatrix}$$

$$= [1 \quad 1 \quad 1]$$

The transformed state equations become

$$\begin{bmatrix} \dot{q}1 \\ \dot{q}2 \\ \dot{q}3 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{bmatrix}\begin{bmatrix} q1 \\ q2 \\ q3 \end{bmatrix} + \frac{1}{10}\begin{bmatrix} 2 \\ -1 + 2i \\ -1 - 2i \end{bmatrix} u \tag{6.107}$$

and

$$y = [1 \quad 1 \quad 1]\begin{bmatrix} q1 \\ q2 \\ q3 \end{bmatrix} \tag{6.108}$$

Equations (6.107) and (6.108) represent a decoupled system of state equations due to the fact that the A matrix is diagonal. These are referred to as the normal form for the state equations, and they enable full computational use to be made of the diagonality of A.

6.57 THE STATE OF TRANSITION MATRIX

Given the vector matrix form for a linear system of equations (i.e., equation (6.96)), this can be rearranged and premultiplied by the matrix e^{-At}, i.e.,

$$e^{-At}(\dot{X} - AX) = e^{-At}Bu$$

$$\therefore \frac{d}{dt}(e^{-At} \cdot X) = e^{-At}Bu$$

On integration this becomes

$$e^{-At}X(t) - X(0) = \int_0^t e^{-A\tau}Bu(\tau)d\tau \tag{6.109}$$

The matrix e^{At} is termed the state transition matrix and is denoted by $\Phi(t)$, i.e.,

$$\Phi(t) = e^{At} \tag{6.110}$$

For the unforced system, when $u(t) = 0$, the solution of equation (6.109) becomes

$$X(t) = \Phi(t)X(0) \tag{6.111}$$

Equation (6.111) shows that $\Phi(t)$ transforms the system from its state $X(0)$ at some initial time, $t = 0$, to the state $X(t)$ at some subsequent time, $t > 0$.

The calculation of the state transition matrix may be performed in several different ways. The most convenient method, for many problems, involves the use of Laplace and inverse Laplace transforms. It may be shown that the state transition matrix can be written as

$$\Phi(t) = e^{At} = \mathcal{L}^{-1}\{[sI - A]^{-1} \tag{6.112}$$

Suppose that a system is characterized by the state equation

$$\begin{bmatrix} \dot{X}1 \\ \dot{X}2 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ -1 & -3 \end{bmatrix} \begin{bmatrix} X1 \\ X2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \tag{6.113}$$

If the forcing function $u(t) = 1$ for $t > 0$ and

$$X(0) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

then the state X of the system at any time t is determined as follows:

$$[sI - A] = \begin{bmatrix} s & -2 \\ 1 & s + 3 \end{bmatrix}$$

$$[sI - A]^{-1} = \frac{1}{s(s + 3) + 2} \begin{bmatrix} s + 3 & 2 \\ -1 & s \end{bmatrix}$$

$$= \frac{1}{(s + 1)(s + 2)} \begin{bmatrix} s + 3 & 2 \\ -1 & s \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{s + 3}{(s + 1)(s + 2)} & \dfrac{2}{(s + 1)(s + 2)} \\ \dfrac{-1}{(s + 1)(s + 1)} & \dfrac{s}{(s + 1)(s + 2)} \end{bmatrix}$$

The matrix $[sI - A]^{-1}$ is called the resolvent matrix. To evaluate the inverse transforms it is necessary to express each element of the resolvent matrix in partial fraction form:

$$\therefore [sI - A]^{-1} = \begin{bmatrix} \dfrac{2}{s + 1} - \dfrac{1}{s + 2} & \dfrac{2}{s + 1} - \dfrac{2}{s + 2} \\ -\dfrac{1}{s + 1} + \dfrac{1}{s + 2} & -\dfrac{1}{s + 1} + \dfrac{2}{s + 2} \end{bmatrix}$$

$$\mathcal{L}^{-1}\{[sI - A]^{-1}\} = \begin{bmatrix} 2e^{-t} - e^{-2t} & 2e^{-t} - 2e^{-2t} \\ -e^{-t} + e^{-2t} & -e^{-t} + 2e^{-2t} \end{bmatrix}$$

This gives

$$\Phi(t)X(0) = \begin{bmatrix} 2e^{-t} - 2e^{-2t} & 2e^{-t} - 2e^{-2t} \\ -e^{-t} + e^{-2t} & -e^{-t} + 2e^{-2t} \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$= \begin{bmatrix} e^{-2} \\ -e^{-2t} \end{bmatrix} \tag{6.114}$$

The above is the complementary function which applies to the unforced system. On consideration of the particular integral it can be noted that

$$\mathcal{L}[u(t)] = \frac{1}{s}$$

Thus,

$$[sI - A]^{-1}Bu(s) = \frac{1}{(s+1)(s+2)} \cdot \frac{1}{s} \begin{bmatrix} s+3 & 2 \\ -1 & s \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{2}{s(s+1)(s+2)} \\ \dfrac{1}{(s+1)(s+2)} \end{bmatrix}$$

Expanding the above into partial fractions gives

$$\begin{bmatrix} \dfrac{1}{s} - \dfrac{2}{s+1} + \dfrac{1}{s+2} \\ \dfrac{1}{s+1} - \dfrac{1}{s+2} \end{bmatrix}$$

Taking the inverse Laplace transform gives

$$\begin{bmatrix} 1 - 2e^{-t} + e^{-2t} \\ e^{-t} - e^{-2t} \end{bmatrix}$$

The complete solution is given as the sum of the complementary function and the particular integral, i.e.,

$$X(t) = \begin{bmatrix} e^{-2t} \\ -e^{-2t} \end{bmatrix} + \begin{bmatrix} 1 - 2e^{-1} + e^{-2t} \\ e^{-t} - e^{-2t} \end{bmatrix}$$

$$\begin{bmatrix} X1(t) \\ x2(t) \end{bmatrix} = \begin{bmatrix} 1 - 2e^{-t} + 2e^{-2t} \\ e^{-t} - 2e^{-2t} \end{bmatrix} \tag{6.115}$$

Equation (6.115) defines the state of the system at any time t and provides a direct solution to the state equations.

Results obtained from the solution of equation (6.115) and that obtained from a finite-difference approximation of the state equation (6.113) show no significant difference in the numerical results generated.

An alternative approach to the evaluation of $\Phi(t)$ is to expand e^{At} as a series, and to truncate the series at an appropriate point. This is well suited to a computer.

Further consideration of the state-variable approach is essentially beyond the scope of this chapter. The reader is therefore referred to DeRusso et al. (1965) and Burghes and Graham (1980) for comprehensive, in-depth coverage.

REFERENCES

Burghes, D. and Graham, A. 1980. *Introduction to Control Theory, Including Optimal Control,* Mathematics and Its Applications, Ellis Horwood, Chichester, U.K.

DeRusso, M. P., Roy, R. J., and Close, C. M. 1965. *Variables for Engineers,* John Wiley, Chichester, U.K.

FURTHER READING

Bannister, B. R. and Whitehead, D. G., *Transducers and Interfacing,* Van Nostrand Reinhold, Wokingham, U.K. (1986).

Barney, G. C., *Intelligent Instrumentation,* Prentice-Hall, Englewood Cliffs, NJ (1985).

Bollinger, J. G. and Duffie, N. A., *Computer Control of Machines and Processes,* Addison-Wesley, Reading, MA (1988).

Burr Brown, *The Handbook of Personal Computer Instrumentation—for Data Acquisition, Test Measurement and Control,* Burr Brown Corporation, January (1988).

Cassel, D. A., *Microcomputers and Modern Control Engineering,* Reston, New York (1983).

Cluley, J. C., *Transducers for Microprocessor Systems,* Macmillan, London (1985).

Doebelin, E. O., *Control System Principles and Design,* John Wiley, Chichester, U.K. (1985).

Dorf, R. C., *Modern Control Systems,* 3d ed., Addison-Wesley, Reading, MA (1980).

Gayakwad, R. and Sokoloff, L., *Analog and Digital Control Systems,* Prentice-Hall, Englewood Cliffs, NJ (1988).

Hunt, V. D., *Mechatronics: Japan's Newest Threat,* Chapman & Hall, London (1988).

Ismail, A. R. and Rooney, V. M., *Microprocessor Hardware and Software Concepts,* Collier-Macmillan, New York (1987).

Kafrissen, E. and Stephans, M., *Industrial Robots and Robotics,* Reston, New York (1984).

Kief, H. B, Olling, G., and Waters, T. F., *Flexible Automation—The International CNC Reference Book,* Becker (U.K.) Ltd (1986).

Meadows, R. and Parson, A. J., *Microprocessors: Essentials, Components and Systems,* Pitman, London (1985).

Scott, P. B., *The Robotics Revolution,* Blackwell, Oxford (1984) Sharon, D., Harstein, J. and Yantian, G., *Robotics and Automated Manufacturing,* Pitman, London (1987).

Snyder, W. E., *Industrial Robots, Computer Interfacing and Control,* Prentice-Hall, Englewood Cliffs, NJ (1985).

Warnock, I. G., *Programmable Controllers: Operation and Application,* Prentice-Hall, Englewood Cliffs, NJ (1988).

SECTION 7
AERONAUTICAL PROPULSION

Scott Eberhardt

7.1 NEWTON AND PROPULSION

Aircraft propulsion systems are applications of Newton's second law. The most familiar form of the second law is $F = ma$ (Force equals mass times acceleration). In the case of propulsion systems, the derivative form of Newton's second law is more appropriate:

$$F = \frac{dm}{dt} v$$

Here the force (or thrust), F, is equal to the rate that mass (in the form of propwash or jet exhaust) is expelled from the propulsion system, $dm/dt = \dot{m}$, times the velocity of that mass, v. To increase the thrust of an engine one can increase $\dot{m}$, v, or both.

The objective of a propulsion system is to produce the needed thrust or propulsive power as efficiently as possible. As will be discussed in the following section, thrust is the appropriate parameter for jets and propulsive power (thrust times speed) is the appropriate parameter for propellers and rotors. In the production of propulsion, kinetic energy ($\frac{1}{2} mv^2$) is given to the exhaust or propwash. This is wasted energy, and it is desirable to minimize it. The rate that energy is wasted to kinetic energy is the wasted power:

$$P_w = \frac{1}{2} \dot{m} v^2$$

Therefore, in order to maximize the efficiency of a propulsion system, one must maximize $\dot{m}$ and minimize v. Ideally, one would accelerate an almost infinite amount of gas at almost zero velocity to produce the desired propulsion.

An understanding of the physics of propulsion efficiency makes it easy to understand why fixed-wing aircraft are more efficient than helicopters and why large jet aircraft are equipped with such large turbofan engines. The wing is efficient because it diverts a large amount of air at a low velocity to produce lift, while a helicopter's rotors divert a relatively small amount of air at a high velocity. Likewise, the modern turbofan engines accelerate a large amount of air with the fan

(which bypasses the turbine core of the engine) to produce as much as 90% of the thrust. This greatly improves the engine's efficiency.

7.2 TURBOJETS AND PROPELLERS

Jet engines and propeller produce the necessary propulsion in slightly different ways. Jet engines produce almost constant thrust at all speed. Constant-speed propellers produce almost constant propulsive power over much of the operating range of the aircraft.

To first order, the thrust of a turbojet is proportional to the rate that air is processed by the engine, $\dot{m}$, times the change in velocity of that air, Δv. As the airplane's speed increase, $\dot{m}$ increases, while Δv decreases in such a way that thrust remains almost constant, as shown in Figure 7.1.

Since propulsive power is equal to thrust times speed, the propulsive power of a jet engine increases linearly with speed. The engine power is approximately constant as a function of speed for a turbojet. Since the propulsive power increases linearly with speed, the engine's efficiency also increases with speed. Because of the effect of the nacelle around the fan in a turbofan, the above discussion of the performance to a turbojet also holds for a turbofan.

The situation is somewhat different for a propeller-driven propulsion system with a constant-speed propeller. The engine, whether turbojet or piston engine, produces a constant power, independent of the speed of the aircraft. The propeller delivers a fairly constant propulsive power as a function of aircraft speed, as shown in Figure 7.2. In this case, the thrust decreases with speed. Thus the efficiency of a propeller-driven propulsion system is fairly constant with speed and can be as high as 85%.

The following sections will cover engine cycles, cycle analysis, and basic component efficiencies for turbine engines. Then, ramjets and scramjets will be discussed followed by engines and propellers and then fuels and noise issues.

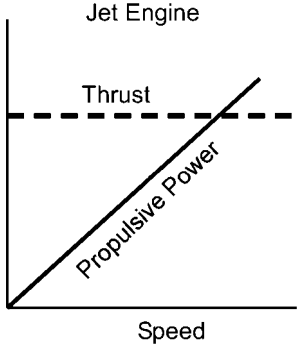


FIGURE 7.1 Thrust and power versus speed for a jet engine.

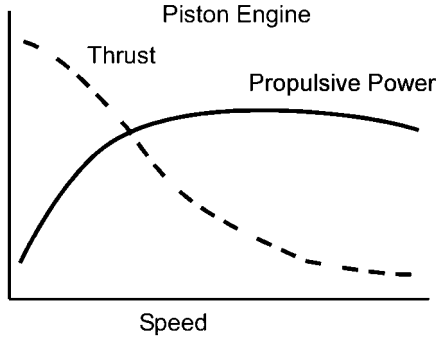


FIGURE 7.2 Thrust and power versus speed for a piston engine/propeller.

7.3 THRUST EQUATION

An airbreathing aircraft engine is basically a reaction motor, much like a rocket engine. Net thrust is obtained by accelerating gas out the exhaust. The primary difference between an airbreathing engine and a non-airbreathing engine is that the airbreathing engine uses oxygen from the environment, rather than carrying it along. Newton's second law can be rewritten

$$F = (\dot{m} + \dot{m}_f)v_e - \dot{m}v_i$$

where subscripts f, e, and i stand for fuel, exit, and inlet, respectively. A figure of performance is the specific thrust, which is the thrust divided by the mass flow rate of the air, or

$$\frac{F}{\dot{m}} = \left(1 + \frac{\dot{m}_f}{\dot{m}}\right) v_e - v_i$$

This definition ignores pressure forces on the inlet and exit. An ideal engine will exhaust at freestream pressure, so that $p_e = p_i$. If they are not the same, the equation for specific thrust of an uninstalled engine is given by

$$\frac{F}{\dot{m}} = \left[\left(1 + \frac{\dot{m}_f}{\dot{m}}\right) v_e - v_i\right] + (p_e - p_i)\frac{A_e}{\dot{m}}$$

The performance of an aircraft engine is dependent on its installation. An engine can be mounted in nacelles (pods) under the wings, as shown in Figure 7.3, or inside the fuselage, as shown in Figure 7.4. When discussing the thrust of an engine, it is important to distinguish between the *installed thrust* and the *uninstalled thrust*. The uninstalled thrust was developed above. The installed thrust includes drag on the housing. This can be generically broken up into two parts, forward and aft, which we shall call the inlet and nozzle drag. So the installed thrust is simply

$$T_{\text{installed}} = F_{\text{uninstalled}} - D_{\text{inlet}} - D_{\text{nozzle}}$$

where T is the installed thrust and F is the uninstalled thrust.

Another figure of merit can be introduced that relates to fuel consumption. The thrust specific fuel consumption, or TSFC, is the ratio of the fuel flow rate to the thrust, or

$$\text{TSFC} = \frac{\dot{m}_f}{T}$$

Engine performance is usually given in uninstalled form, since the actual performance will depend on installation. So the uninstalled thrust specific fuel consumption, S, is

$$S = \frac{\dot{m}_f}{F}$$

F and S are usually given at sea-level, standard conditions. With less oxygen at

FIGURE 7.3 C-17 Globemaster with wing-mounted nacelles (photo courtesy of the United States Air Force).

altitude, both F and S will decrease as altitude increases, up to the isothermal layer at about 11,000 m (36,000 ft). In the next section, ideal engine cycle analysis will be outlined to predict the specific thrust and the uninstalled specific fuel consumption.

Propulsive efficiency is a parameter that relates the aircraft power to the power out of the engine. The aircraft power is the thrust times the velocity, or Tv_i. The work done by the engine is, $W_{\text{out}} = \frac{1}{2}((\dot{m} + \dot{m}_f)v_e^2 - \dot{m}v_i^2)$. Therefore, propulsive efficiency is

$$\eta_p = \frac{Tv_i}{W_{\text{out}}} = \frac{2((\dot{m} + \dot{m}_f)v_e - \dot{m}v_i)v_i}{(\dot{m} + \dot{m}_f)v_e^2 - \dot{m}v_i^2}$$

In this expression the installation drag has been left out. Usually it is small and so we have ignored it. In addition, we can assume that $\dot{m}_f << \dot{m}$, and so η_p reduces to

$$\eta_p = \frac{2}{v_e/v_i + 1}$$

7.4 ENGINE CYCLES

Thermodynamic analysis of engines starts with the engine cycle. The three most common engine cycles are the Otto, Diesel, and Brayton cycles. The Otto cycle is

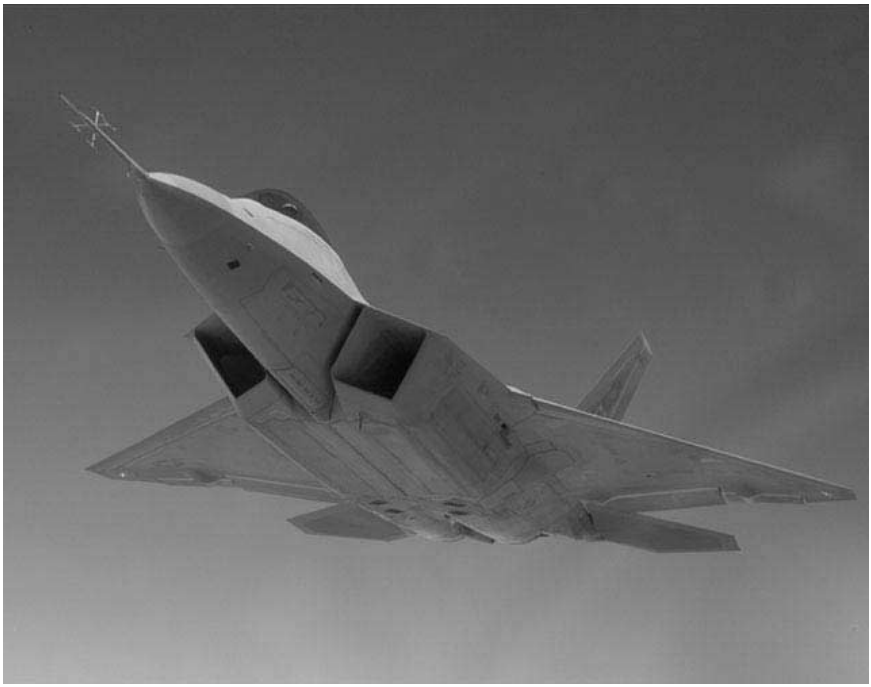


FIGURE 7.4 F-22 Raptor with internal engine (photo courtesy of the United States Air Force).

the thermodynamic description of the common internal combustion engine. The Diesel cycle, as the name suggests, is the thermodynamic description of the Diesel engine. The Brayton cycle is used to describe the thermodynamics of the jet engine.

Engine cycles describe where work is done on the gas, where the gas does work on the air, and where energy is added to the system. It describes the basic thermodynamic process of each part of the cycle. The Brayton cycle is the easiest to illustrate and so will be discussed first.

The Brayton cycle is shown in Figures 7.5(a) and (b). At point 2 air is isentropically compressed. The volume of the air is decreased as the pressure is increased. This isentropic process continues until heat addition begins at point 3. From points

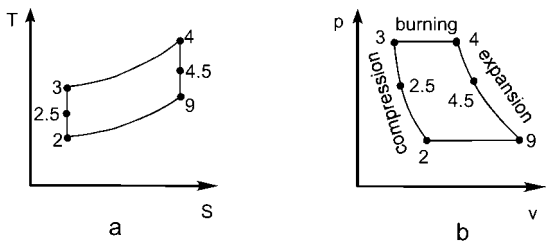


FIGURE 7.5 Brayton cycle.

3 to 4 heat is added, usually through combustion, at constant pressure. From point 4 to point 9 the air is expanded isentropically, and from 9 back to 2 the air is compressed at constant pressure (note that points on the figures not explicitly discussed in this section will be described later).

In reality, the cycle of a jet engine is not closed, as indicated in the figure. Point 2 represents the inlet and point 9 the exit of the engine. At point 2 gas from the environment is used, and at point 9 it is exhausted back into the atmosphere. The installation of a jet engine does not reroute the exhaust back to the inlet, as implied by the figures.

The efficiency of the Brayton cycle is given by

$$\eta_T = 1 - \left(\frac{p_2}{p_3}\right)^{(\gamma-1/\gamma)}$$

Note that if the process from points 2 to 3 is isentropic, then

$$\left(\frac{p_2}{p_3}\right)^{(\gamma-1/\gamma)} = \frac{T_2}{T_3}$$

and thus

$$\eta_T = 1 - \left(\frac{T_2}{T_3}\right)$$

For a given p_3/p_2 and T_3/T_2, it can be show that the maximum work for a given mass flow is given by

$$\frac{T_3}{T_2} = \sqrt{\frac{T_4}{T_2}}$$

A key point to recognize is that better efficiency occurs when the temperature at point 4 is made as high as possible. This gives the maximum area inside the T-S diagram. Unfortunately, the maximum temperature at point 4 is limited by material temperature limits. As will be shown later, point 4 represents what is known as the *turbine inlet temperature*.

The Otto cycle is almost exclusively used for internal combustion, propeller-driven aircraft, although some are looking into Diesel aircraft engines. Figure 7.2 illustrates the Otto cycle on T–S (temperature–entropy), Figure 7.6(a), and P–V (pressure–volume), Figure 7.6(b), plots. The Otto cycle is a *closed cycle*, which means that the gas is contained during the cycle. The Otto cycle is also referred to

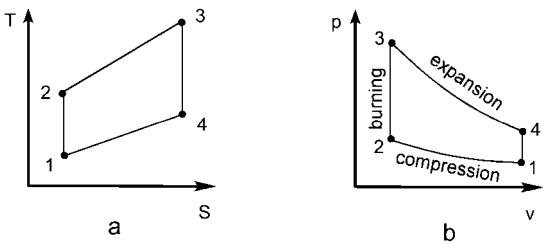


FIGURE 7.6 Otto cycle.

as a *constant volume cycle* because combustion occurs at a constant volume. The expansion phase of the cycle is the only place where power can be extracted.

The internal combustion engine uses pistons inside cylinders for compression and expansion. Compression occurs between point 1 and 2. Like the Brayton cycle, this process is isentropic. However, during compression, pressure and volume change together. Expansion is between points 3 and 4 and is also isentropic. Combustion occurs from 2 to 3 at a constant volume. In a single cylinder, each phase of the cycle occurs one at a time. The piston and cylinder are where compression, combustion, and expansion all occur. This contrasts with the jet engine where specific components are responsible for each part of the cycle and all occur simultaneously. The compressor compresses the gas, the burner combusts, and the turbine and nozzle expand the gas.

The thermal efficiency of the Otto cycle is given by the temperature ratio through the compression stage. For the Otto cycle the thermal efficiency is expressed in terms of static temperatures, since the changes in static temperature are proportional to the addition of heat. The efficiency is, therefore,

$$\eta_T = 1 - \frac{T_4 - T_1}{T_3 - T_2}$$

7.5 GAS TURBINE ENGINES

A gas turbine engine is known as an *open cycle* engine. Air continually flows through the engine along an uninterrupted path from inlet to outlet. In this section we will introduce the different applications of the gas turbine engine in aircraft.

The distinction between the gas turbine core and the housing it goes into must be made clear. In many applications the engine is stored in a nacelle as shown in Figure 7.7. Sometimes, most notably in military fighters and trainers, the engine is stored inside the body of the aircraft. The engine nacelle, or housing, conditions the gas before and after the core. The engine components that are a part of this housing are the diffuser and, sometimes, the nozzle.

The core of the engine, called the gas generator, consists of three primary parts, the compressor, burner, and turbine. Returning to the Brayton cycle of Figure 7.5, these three components describe the isentropic compression from points 2 to 3, the heat addition from points 3 to 4, and the isentropic expansion through the turbine from points 4 to 9. The traditional numbering system for gas turbine engines will be used. Figure 7.8 illustrates an example of the numbering system. Between points 2 and 3 is the compressor, 3 and 4, the burner, and 4 and 5, the turbine. Points 2.5 and 4.5 represent the situation when dual compressors and turbines are used. The numbering system is useful for associating properties with location in the engine. For example, T_4 is the turbine inlet temperature at station 4. Note that station 4 is both the turbine inlet location and the burner exit location. From this point on, values with numeric subscripts will refer to the points in Figure 7.8.

7.6 IDEAL ENGINE CYCLE ANALYSIS

An overview analysis of a jet engine from inlet to exhaust is called a *cycle analysis*. The idea of cycle analysis is to perform parametric studies to determine the opti-

FIGURE 7.7 Engine nacelle.

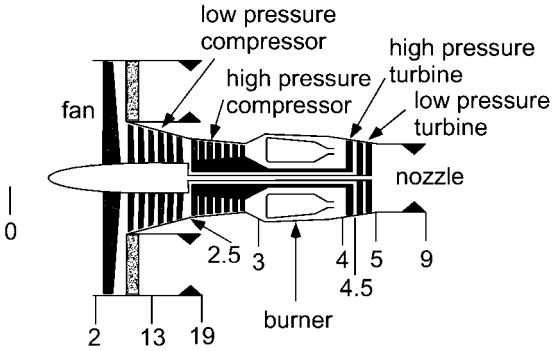


FIGURE 7.8 Turbofan showing number system.

mum engine design for a specific thrust and fuel consumption requirement. Later we will look at off-design engine analysis, which will show how to examine a particular engine under many conditions.

The specific engine parameters that cycle analysis will give are specific thrust, thrust-specific fuel consumption, and thermal and propulsive efficiency. For cycle analysis, we will need to characterize changes in conditions through engine components. So, for example, we will need the pressure rise across the compressor. Once again, conventions are used to define these changes in the term of ratios. Ratios of stagnation, or total, pressure use the symbol π and ratios of total temperature, τ. Table 7.1 identifies the application of each subscript.

TABLE 7.1

$\pi_d = \frac{p_{t2}}{p_{t0}}$	$\tau_d = \frac{T_{t2}}{T_{t0}}$	Diffuser
$\pi_c = \frac{p_{t3}}{p_{t2}}$	$\tau_c = \frac{T_{t3}}{T_{t2}}$	Compressor
$\pi_b = \frac{p_{t4}}{p_{t3}}$	$\tau_b = \frac{T_{t4}}{T_{t3}}$	Burner
$\pi_t = \frac{p_{t5}}{p_{t4}}$	$\tau_t = \frac{T_{t5}}{T_{t4}}$	Turbine
$\pi_n = \frac{p_{t9}}{p_{t7}}$	$\tau_n = \frac{T_{t9}}{T_{t7}}$	Nozzle

Other subscripts include (AB) for an afterburner, points 5 to 7, λ from freestream to turbine inlet, and λ_{AB}, from freestream to afterburner exit. The terms π_λ and τ_λ can be expressed as follows:

$$\pi_\lambda = \pi_d \pi_c \pi_d, \quad \tau_\lambda = \tau_d \tau_c \tau_d$$

One other set of parameters is the freestream stagnation to static ratios. Deviating from convention, which uses only total pressure and temperature, these also use the symbols π and τ and are defined as follows:

$$\pi_r = \frac{p_{t0}}{p_0} = \left(1 - \frac{\gamma - 1}{2} M_0^2\right)^{\gamma/\gamma - 1}$$

$$\tau_r = \frac{T_{t0}}{T_0} = 1 - \frac{\gamma - 1}{2} M_0^2$$

where M_0 is the flight Mach number and γ is the ratio of specific heats.

A loose association can be made with π and τ. Any deviation of π from the ideal can be associated with losses due to skin friction or shock waves. The changes in τ can be associated with work done or extracted from the gas, or energy added from combustion.

7.7 GOALS OF CYCLE ANALYSIS

Cycle analysis is used determine the specific thrust and the specific thrust fuel consumption in a parametric study. The specific thrust and specific thrust fuel consumption are defined as

$$\text{specific thrust} = \frac{F}{\dot{m}} \quad \text{and} \quad \text{specific thrust fuel consumption} = S = \frac{\dot{m}_f / \dot{m}}{F/\dot{m}}$$

In the above expressions, F is the uninstalled thrust, $\dot{m}$ is the mass flow rate through

the engine, and $\dot{m}_f$ is the fuel mass flow rate through the engine. It is convenient to look at the specific values of thrust and fuel consumption because the results can be scaled for any size engine. Variable inputs include Mach number, component (e.g., compressor, turbine, etc.) efficiencies, component pressure ratios, and the turbine inlet temperature. Thus, cycle analysis is really a trade-off study of how flight parameters and component efficiencies affect the specific thrust and fuel consumption, without regard to the actual thrust required for the application.

In ideal cycle analysis several assumptions are made. The first is that the gas is frictionless, so that no total pressure losses occur due to friction. This results in the following for the diffuser and nozzle:

$$\pi_d = \pi_n = 1, \quad \tau_d = \tau_n = 1$$

The compressor and turbine we assume isentropic, so

$$\tau_c = \pi_c^{(\gamma-1)/\gamma} \quad \text{and} \quad \tau_t = \pi_t^{(\gamma-1)/\gamma}$$

In the burner $\pi_b = 1$ but $\tau_b \neq 1$, since the heat addition due to combustion raises the total temperature.

It is also assumed that the mass flow rate of fuel, $\dot{m}_f$, is much less than the mass flow rate of the gas flowing through the engine, $\dot{m}$. Thus,

$$\dot{m} + \dot{m}_f \cong \dot{m}$$

Also, the gas is considered to be a perfect gas with constant specific heat, c_p. Finally, the exit pressure is assumed to be equal to the inlet, or ambient, pressure.

7.8 GENERAL PROCEDURE FOR CYCLE ANALYSIS

The procedure for analyzing an engine cycle is the same whether we are analyzing a simple turbojet or an afterburning fan jet. Starting with the thrust equation, we can determine what variables we will need. If we introduce the sound speed, a_0, we have

$$\frac{F}{\dot{m}} = \left[\left(1 + \frac{\dot{m}_f}{\dot{m}}\right) v_9 - v_0\right] + A_9 (p_9 - p_0) \cong v_9 - v_0 = a_0 \left(\frac{v_9}{a_0} - M_0\right)$$

M_0 and a_0 come from flight conditions, but we will have to determine v_9. Now we introduce a_9 and use the definition of sound speed for an ideal gas to get

$$\left(\frac{v_9}{a_0}\right)^2 = \left(\frac{a_9}{a_0}\right)^2 M_9^2 = \frac{T_9}{T_0} M_9^2$$

The next step is to find the exit Mach number and temperature. For the exit Mach number, we use the expression for total pressure in terms of Mach number:

$$p_{t9} = p_9 \left(1 + \frac{\gamma - 1}{2} M_9^2\right)^{\gamma/(\gamma-1)}$$

which gives

$$M_9^2 = \frac{2}{\gamma - 1}\left[\left(\frac{p_{t9}}{p_9}\right)^{(\gamma-1)/\gamma} - 1\right]$$

The pressure ratio is found as follows:

$$\frac{p_{t9}}{p_9} = \frac{p_0}{p_9}\frac{p_{t0}}{p_0}\frac{p_{t2}}{p_{t0}}\frac{p_{t3}}{p_{t2}}\frac{p_{t4}}{p_{t3}}\frac{p_{t5}}{p_{t4}}\frac{p_{t7}}{p_{t5}}\frac{p_{t9}}{p_{t7}} = \frac{p_0}{p_9}\,\pi_r\pi_d\pi_c\pi_b\pi_t\pi_{AB}\pi_n$$

Using our ideal engine assumptions with $\pi_d = \pi_b = \pi_{AB} = \pi_n = i$ and $p_9 = p_0$, we get the simplified, ideal result,

$$\frac{p_{t9}}{p_9} = \pi_r\pi_c\pi_t$$

Similarly, for T_{t9}/T_0 we have

$$\frac{T_{t9}}{T_0} = \frac{T_{t0}}{T_0}\frac{T_{t2}}{T_{t0}}\frac{T_{t3}}{T_{t2}}\frac{T_{t4}}{T_{t3}}\frac{T_{t5}}{T_{t4}}\frac{T_{t7}}{T_{t5}}\frac{T_{t9}}{p_{t7}} = \tau_r\tau_d\tau_c\tau_b\tau_t\tau_{AB}\tau_n$$

and for T_{t9}/T_9 we have

$$\frac{T_{t9}}{T_9} = \left(\frac{p_{t9}}{p_9}\right)^{(\gamma-1)/\gamma}$$

Combining the two temperature and Mach number relationships gives

$$\left(\frac{v_9}{a_0}\right)^2 = \frac{T_9}{T_0}M_9^2 = \left[\frac{\tau_r\tau_d\tau_c\tau_b\tau_t\tau_{AB}\tau_n}{\left(\frac{p_{t9}}{p_9}\right)^{(\gamma-1)/\gamma}}\right]\left[\frac{2}{\gamma-1}\left[\left(\frac{p_{t9}}{p_9}\right)^{(\gamma-1)/\gamma} - 1\right]\right]$$

This expression, with the flight conditions, which give a_0 and M_0, gives the specific thrust, $F/\dot{m}$.

To obtain the thrust-specific fuel consumption we need to perform and energy balance across the burner. Generically, this can be written

$$\dot{m}h_{t3} + \dot{m}_f h_f^0 = \dot{m}h_{t4}$$

where h_{t3} and h_{t4} are the total specific enthalpies at points 3 and 4 and h_f^0 is the heat content of the fuel at standard conditions. Making use of our ideal engine assumption of a perfect gas,

$$\dot{m}c_pT_{t3} + \dot{m}_f h_f^0 = \dot{m}c_pT_{t4}$$

Recall that the uninstalled thrust-specific fuel consumption is

$$S = \frac{\dot{m}_f}{F} = \frac{\dot{m}_f/\dot{m}}{F/\dot{m}}$$

From the energy balance equation, we find that, for constant c_p,

$$\frac{\dot{m}_f}{\dot{m}} = \frac{c_p}{h_f^0}(T_{t4} - T_{t3}) = \frac{c_p T_0}{h_f^0}\left(\frac{T_{t4}}{T_0} - \frac{T_{t3}}{T_0}\right) = \frac{c_p T_0}{h_f^0}(\tau_\gamma - \tau_r \tau_d \tau_c)$$

Now we will apply this general procedure to various types of ideal gas turbine engines.

7.9 THE TURBOJET

The turbojet consists of the gas generator, with a compressor, burner and turbine, and a diffuser and nozzle, as shown in Figure 7.9. As we saw in the introductory section of this chapter, the compressor is used to compress the air and increase the cycle, or thermal, efficiency. The burner is where energy is added at constant pressure. Air is then expanded in the turbine, which powers the compressor. The work done by turbine is equal to the work necessary to run the compressor. There is excess energy after the turbine, which can be used to provide thrust.

Reexamining the thrust equation, for an ideal engine, we recall that

$$F = \dot{m}(v_9 - v_0)$$

So, as was discussed earlier, one can increase thrust by increasing the mass flow through the engine, $\dot{m}$, or the exit velocity, v_9. It is the excess kinetic energy left after the turbine that contributes to increasing v_9. Unfortunately, this same excess energy reduces propulsive efficiency, η_p.

Now we will look at the cycle analysis of an ideal turbojet. Recall, the following assumptions are made: $\pi_d = \tau_d = \pi_n = \tau_n = 1$, since the diffuser and nozzle have no energy transfer total, or stagnation, conditions remain constant. In addition, through the burner, $p_{t4} = p_{t3}$ so $\pi_b = 1$. Thus,

$$\frac{p_{t9}}{p_{t0}} = \pi_d \pi_c \pi_b \pi_t \pi_n = \pi_c \pi_t$$

and

$$\frac{T_{t9}}{T_{t0}} = \tau_d \tau_c \tau_b \tau_t \tau_n = \tau_c \tau_b \tau_t$$

With a little manipulation, and making use of ideal, isentropic components, we can write

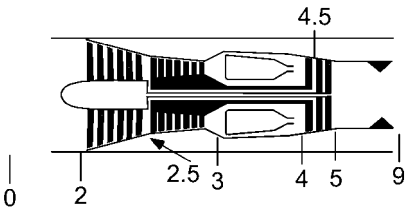


FIGURE 7.9 Turbojet.

$$\frac{T_9}{T_0} = \left(\frac{T_9}{T_{t9}}\right)\left(\frac{T_{t9}}{T_{t0}}\right)\left(\frac{T_{t0}}{T_0}\right) = \left[\left(\frac{p_9}{p_{t9}}\right)^{\gamma-1/\gamma}\right][\tau_c\tau_b\tau_t][\tau_r]$$

and

$$\frac{p_{t9}}{p_9} = \left(\frac{p_{t9}}{p_{t0}}\right)\left(\frac{p_{t0}}{p_0}\right)\left(\frac{p_0}{p_9}\right) = [\pi_c\pi_t][\pi_r][1] = (\tau_r\tau_c\tau_t)^{\gamma/\gamma-1}$$

Plugging into the equation for T_9/T_0, we find that

$$\frac{T_9}{T_0} = \frac{\tau_r\tau_c\tau_b\tau_t}{\tau_r\tau_c\tau_t} = \frac{\tau_\lambda}{\tau_c\tau_t} = \tau_b$$

This result should not be too surprising. In an ideal turbojet the compressor and turbine are isentropic, and the work done by the compressor is equal to the work extracted from the turbine. Thus, the added net energy is from the burner only. We can examine the work done by the compressor and turbine, using total temperatures, which assumes a thermally perfect gas.

$$\dot{W}_c = \dot{m}c_p(T_{t3} - T_{t2}) = \dot{m}T_{t2}c_p(\tau_c - 1)$$

$$\dot{W}_t = \dot{m}c_p(T_{t4} - T_{t5}) = \dot{m}T_{t4}c_p(1 - \tau_t)$$

Making these two equal for an ideal turbojet gives

$$\tau_t = 1 - \frac{T_{t2}}{T_{t4}}(\tau_c - 1) = 1 - \frac{\tau_r}{\lambda_\gamma}(\tau_c - 1)$$

Now we can solve for the specific thrust. Recall,

$$\frac{F}{\dot{m}} = v_9 - v_0 = a_0\left(\frac{v_9}{a_0} - M_0\right)$$

with

$$\frac{v_9}{a_0} = \sqrt{\frac{T_9}{T_0}M_9^2} \quad \text{and} \quad M_9^2 = \frac{2}{\gamma-1}\left(\frac{T_{t9}}{T_9} - 1\right) = \frac{2}{\gamma-1}(\tau_r\tau_c\tau_t - 1)$$

So, finally,

$$\frac{F}{\dot{m}} = a_0\left[\sqrt{\frac{2}{\gamma-1}\frac{\tau_\gamma}{\tau_r\tau_c}(\tau_r\tau_c\tau_t - 1)} - M_0\right]$$

The next step will be to find the thrust-specific fuel consumption,

$$S = \frac{\dot{m}_f/\dot{m}}{F/\dot{m}}$$

For, $\dot{m}_f/\dot{m}$ we have

$$\frac{\dot{m}_f}{\dot{m}} = \frac{c_p T_0}{h_f^0}(\tau_\lambda - \tau_r \tau_c)$$

so

$$S = \frac{c_p T_0(\tau_\lambda - \tau_r \tau_c)}{a_0 h_f^0 \left[\sqrt{\dfrac{2}{\gamma - 1}\dfrac{\tau_\lambda}{\tau_r \tau_c}(\tau_r \tau_c \tau_t - 1)} - M_0\right]}$$

Finally, we can find the thermal efficiency, which is the net work done divided by the energy input, and the propulsive efficiency. These are

$$\eta_T = 1 - \frac{T_0}{T_3} = 1 - \frac{1}{\tau_r \tau_c}$$

$$\eta_p = \frac{2v_0}{v_9 - v_0} = \frac{M_0}{v_9/a_0 + M_0}.$$

The two important things to note are that the thermal efficiency increases with increasing T_3, and the propulsive efficiency increases with decreasing v_9. The former implies that we want to compress the gas as much as possible, and the latter implies that we want as little kinetic energy left in the jet as possible. There are two additional results not explicitly stated, however. Recall that

$$\left(\frac{v_9}{a_0}\right)^2 = \frac{T_9}{T_0} M_9^2$$

Since $T_9/T_0 = \tau_b$, and we have to limit T_{t4} due to material constraints, the higher the compressor temperature rise, τ_c, the less energy we can add in the burner. The result will be less thrust. Thus, thermal efficiency and thrust are at odds with each other. The other implicit statement is simply that thrust increases with v_9, so propulsive efficiency and thrust are at odds with each other. The resulting engine must compromise efficiency with thrust.

The next step will be to look at a turbojet with an afterburner. The afterburner adds a second burner after the turbine, as shown in Figure 7.10. The purpose is to add energy back into the gas after the turbine has taken some out. This additional burner changes the cycle, as shown in Figure 7.11.

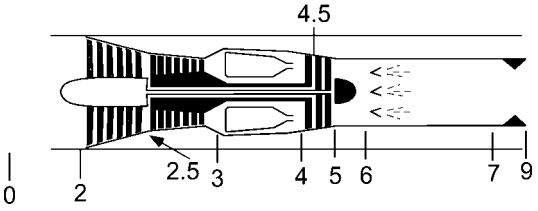


FIGURE 7.10 Turbojet with afterburner.

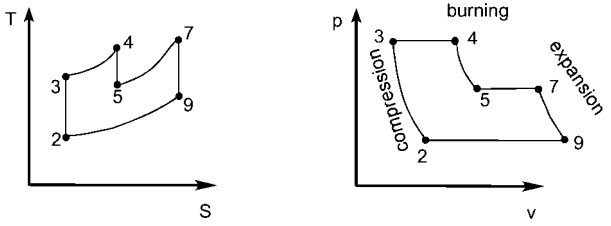


FIGURE 7.11 Brayton cycle with afterburner.

Let's look at the changes that occur from the non-afterburning engine. We will use the station 9′ to denote the normal turbojet exit condition. The exit velocity divided by free-stream sound speed is now

$$\left(\frac{v_9}{a_0}\right)^2 = \left(\frac{v_9}{v'_9}\right)^2 \left(\frac{v'_9}{a_0}\right)^2 = \left(\frac{v_9}{v'_9}\right)^2 \left[\frac{2}{\gamma - 1}\frac{\tau_\lambda}{\tau_r\tau_c}(\tau_r\tau_c\tau_t - 1)\right]$$

Note that with the exception of τ_λ, all of the component total-temperature changes are the same as the ideal turbojet, since we are assuming they use the same components. The trick now is to determine the ratio of exit velocities for the afterburning and non-afterburning engines. Using energy relations, i.e.,

$$\frac{v_9^2}{2} = c_p T_{t9}\left[1 - \left(\frac{p_9}{p_{t9}}\right)^{(\gamma-1)/\gamma}\right]$$

Also, the process through the afterburner is constant pressure, so

$$\frac{p'_9}{p'_{t5}} = \frac{p_9}{p_{t9}}$$

and so,

$$\left(\frac{v_9}{v'_9}\right)^2 = \frac{T_{t9}}{T'_{t5}}$$

Introducing $\tau_{\lambda AB}$ as T_{t7}/T_0, which follows the definition of τ_λ, we can solve for T_{t9}/T'_{t5} as follows:

$$\frac{T_{t9}}{T'_{t5}} = \frac{\left(\frac{T_{t9}}{T_{t7}}\right)\left(\frac{T_{t7}}{T_0}\right)}{\left(\frac{T'_{t5}}{T'_{t4}}\right)\left(\frac{T'_{t4}}{T_0}\right)} = \frac{\tau_{\lambda AB}}{\tau_\lambda\tau_t}$$

Plugging this result into the earlier equation for v_9/a_0 gives

$$\left(\frac{v_9}{a_0}\right)^2 = \frac{\tau_{\lambda AB}}{\tau_\lambda\tau_t}\left[\frac{2}{\gamma - 1}\frac{\tau_\lambda}{\tau_r\tau_c}(\tau_r\tau_c\tau_t - 1)\right] = \tau_{\lambda AB}\left(\frac{2}{\gamma - 1}\right)\left(1 - \frac{1}{\tau_r\tau_c\tau_t}\right)$$

Finally,

$$\frac{F}{\dot{m}} = a_0 \left[\sqrt{\tau_{\lambda AB} \left(\frac{2}{\gamma - 1} \right) \left(1 - \frac{2}{\tau_r \tau_c \tau_t} \right)} - M_0 \right]$$

The total fuel flow mass ratio is determined by energy balance through the entire engine, from station 0 to 9. Energy balance gives

$$\dot{m}_f h_f^0 = \dot{m} c_p (T_{t9} - T_{t0})$$

or

$$\frac{\dot{m}_f}{\dot{m}} = \frac{c_p T_0}{h_f^0} \left(\frac{T_{t9}}{T_0} - \frac{T_{t0}}{T_0} \right) = \frac{c_p T_0}{h_f^0} (\tau_{\lambda AB} - \tau_r)$$

This equation is used in the expression for fuel consumption, which gives

$$S = \frac{\dfrac{c_p T_0}{h_f^0} (\tau_{\lambda AB} - \tau_t)}{a_0 \left[\sqrt{\tau_{\lambda AB} \left(\dfrac{2}{\gamma - 1} \right) \left(1 - \dfrac{1}{\tau_r \tau_c \tau_t} \right)} - M_0 \right]}$$

7.10 THE TURBOFAN

The turbofan, shown in Figure 7.8, adds a fan to the core gas turbine engine. The driving reason for using a fan is that the excess energy from the turbine drives a fan, which increases $\dot{m}$ in the thrust equation. Since either $\dot{m}$ or v_9 can be increased to increase thrust, but increasing v_9 decreases efficiency, the turbofan increase thrust by increasing $\dot{m}$. Thrust is then increased without loss of efficiency.

The mass flow through the engine is divided into two parts, the core, $\dot{m}_C$, and the fan, $\dot{m}_F$. The ratio of these two is known as the bypass ratio and is given by

$$\alpha = \frac{\dot{m}_F}{\dot{m}_C}$$

So,

$$\dot{m} = \dot{m}_C + \dot{m}_F = (1 + \alpha)\dot{m}_C$$

Through the core, the cycle follows the turbojet. The fan, however, has no burner or turbine. Thus, the work done to the air by the fan must come from the core's turbine. This leaves less energy in the core's jet, but adds energy to the fan's exhaust. Since the fan moves a large amount of mass, the turbofan is more efficient than the turbojet.

Cycle analysis of an ideal turbofan is straightforward after the turbojet. First we introduce new stations, 13 and 19, which represent the station after the fan and the station after the nozzle, respectively. Thus, we have the new variables π_F, τ_F, π_{nF}, and τ_{nF}, which represent the pressure and temperature rises across the fan and fan's nozzle, respectively. Again, $\pi_d = \tau_d = \pi_n = \tau_n = \pi_b = 1$. Also, for the ideal turbofan, $\pi_{nF} = \tau_{nF} = 1$.

First, we must look at how to modify the thrust equation to include the fan. The specific thrust is

$$\frac{F}{\dot{m}} = \frac{a_0}{(1 + \alpha)} \left[\left(\frac{v_9}{a_0} - M_0 \right) + \alpha \left(\frac{v_{19}}{a_0} - M_0 \right) \right]$$

where the first term in the parenthesis is due to the core, and the second is due to the fan.

The core follows exactly as the turbojet, so we have

$$\frac{v_9}{a_0} = \sqrt{\frac{2}{\gamma - 1} \frac{\tau_\lambda}{\tau_r \tau_c} (\tau_r \tau_c \tau_t - 1)}$$

Determining v_{19}/a_0 is simple when you recognize that $a_{19} = a_0$, $T_{19} = T_0$, and $p_{19} = p_0$. We follow the steps for determining M_9 to find M_{19} as follows:

$$M_{19}^2 = \frac{2}{\gamma - 1} \left(\frac{T_{t19}}{T_{19}} - 1 \right) = \frac{2}{\gamma - 1} (\tau_r \tau_F - 1)$$

Thus,

$$\frac{v_{19}}{a_0} = M_{19} = \sqrt{\frac{2}{\gamma - 1} (\tau_r \tau_F - 1)}$$

For specific thrust we have

$$\frac{F}{\dot{m}} = \frac{a_0}{(1 + \alpha)} \left[\left(\sqrt{\frac{2}{\gamma - 1} \frac{\tau_\lambda}{\tau_r \tau_c} (\tau_r \tau_c \tau_t - 1)} - M_0 \right) + \alpha \left(\sqrt{\frac{2}{\gamma - 1} (\tau_r \tau_F - 1)} - M_0 \right) \right]$$

Next we need to find out how much power the turbine must produce. The total work extracted from the gas by the turbine must be equal to the work added by the compressor and the fan. The work done by the compressor is

$$\dot{W}_c = \dot{m}_C c_p (T_{t3} - T_{t2}) = \dot{m}_C T_{t2} c_p (\tau_c - 1)$$

The work done by the fan is

$$\dot{W}_F = \dot{m}_F c_p (T_{t13} - T_{t2}) = \dot{m}_F T_{t2} c_p (\tau_F - 1)$$

The word done on the turbine is

$$\dot{W}_t = \dot{m}_C c_p (T_{t4} - T_{t5}) = \dot{m}_c T_{t4} c_p (1 - \tau_t)$$

which makes use of the assumption that $\dot{m}_f << \dot{m}_C$.

Equating the work done by the fan and compressor by the work output of the turbine gives

$$T_{t4}(1 - \tau_t) = T_{t2}(\tau_c - 1) + \alpha T_{t2}(\tau_F - 1)$$

With $\tau_\lambda = T_{t4}/T_0$ we get

$$\frac{T_{t4}}{T_{t2}}(1 - \tau_t) = \frac{\tau_\lambda}{\tau_r}(1 - \tau_t) = (\tau_c - 1) + \alpha(\tau_F - 1)$$

Rearranging gives

$$\tau_t = 1 - \frac{\tau_r}{\tau_\lambda}[(\tau_c - 1) + \alpha(\tau_F - 1)]$$

Now let's look at thrust-specific fuel consumption. Energy balance through the core gives the same result as for the turbojet, namely,

$$\frac{\dot{m}_f}{\dot{m}_C} = \frac{c_p T_0}{h_f^0}(\tau_\lambda - \tau_r\tau_c)$$

The thrust-specific fuel consumption is

$$S = \frac{\dot{m}_f/\dot{m}}{F/\dot{m}} = \frac{\dot{m}_f/\dot{m}_C}{(1 + \alpha)F/\dot{m}}$$

Thus,

$$S = \frac{c_p T_0(\tau_\lambda - \tau_r\tau_c)}{a_0 h_f^0\left[\left(\sqrt{\frac{2}{\gamma - 1}\frac{\tau_\lambda}{\tau_r\tau_c}(\tau_r\tau_c\tau_t - 1)} - M_0\right) + \alpha\left(\sqrt{\frac{2}{\gamma - 1}(\tau_r\tau_F - 1)} - M_0\right)\right]}$$

The thermal efficiency of the turbofan is the same as for the turbojet, namely,

$$\eta_T = 1 - \frac{1}{\tau_r\tau_c}$$

The propulsive efficiency is

$$v_p = 2\frac{\left(\frac{v_9}{v_0} - 1\right) + \alpha\left(\frac{v_{10}}{v_0} - 1\right)}{\left[\left(\frac{v_9}{v_0}\right)^2 - 1\right] + \alpha\left[\left(\frac{v_{19}}{v_0}\right)^2 - 1\right]}$$

7.11 THE TURBOPROP

The turboprop, shown in Figure 7.12, adds a propeller shaft to the a core gas turbine engine. The motivation follows that of the turbofan, but now the excess energy from the turbine drives a propeller. Once again, $\dot{m}$ is increased to increase thrust, instead of increasing v_9. An important point to note about turboprops is that they must be geared. The gas turbine core turns much faster than a propeller can. Therefore, a substantial reduction gear is required. For the ideal turboprop we will assume there are no mechanical losses through the gears.

Most turboprops leave little energy left for the core exhaust. Most of the excess power from the turbine is used to drive the propeller. What power is left over gives thrust from the core. Thus, the thrust developed with a turboprop is from the pro-

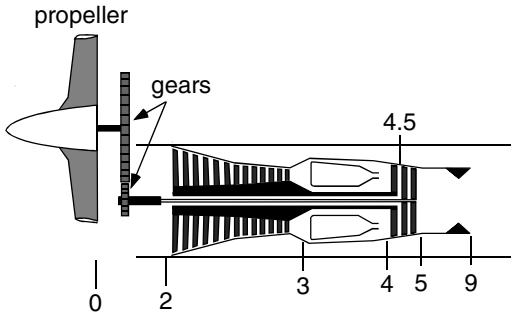


FIGURE 7.12 Turboshaft engine.

peller, designated F_{prop}, and the core. Because the second stage turbine is not being used to provide power to the compressor, it is sometimes referred to as a *free turbine*.

The propeller efficiency is the power from the propeller over the power to the propeller, which is written

$$\eta_p = \frac{F_p v_0}{\dot{W}_{\text{prop}}}$$

The thrust of the propeller is given by

$$F_p = \dot{m}_p(v_p - v_0)$$

where $\dot{m}_p$ is the mass flow through the propeller and v_p is the velocity of the air leaving the propeller. It is possible to write the total force as

$$F = F_p + F_C = \dot{m}_p(v_p - v_0) + \dot{m}(v_9 - v_0)$$

The ratio of $\dot{m}_p$ to $\dot{m}$ would be analogous to the bypass ratio, but without any flow constraints at the tip of the propeller this ratio is not as simple as the bypass ratio. Instead, we define a parameter to replace the specific thrust with a work output coefficient, C. This coefficient starts with the ratio of power to the mass flow through the core, i.e., $F_p v_0/\dot{m}$ and $F_C v_0/\dot{m}$. This ratio is then divided by the total specific enthalpy to give the dimensionless coefficients

$$C_C = \frac{F_C v_0/\dot{m}}{h_0} \quad \text{and} \quad C_p = \frac{F_p v_0/\dot{m}}{h_0}$$

These definitions give the following results:

$$C_{\text{total}} = C_p + C_C$$

and

$$\frac{F_{\text{total}}}{\dot{m}} = \frac{h_0 C_{\text{tot}}}{v_0} = \frac{h_0}{v_0}(C_p + C_C)$$

This now gives an explicit expression for specific thrust by relating the thrust to only the core mass flow.

Figures 7.9 and 7.8, which show the turbojet and turbofan, illustrate two spools. A high-pressure compressor and turbine combination are used inside a low compressor and turbine combination. The points that represent the partition between the two are 2.5 between the low and high-pressure compressors and 4.5 between the high and low-pressure turbines. In a turboprop the high-pressure compressor is turned by the high-pressure turbine, but there is no low-pressure compressor; station 2.5 does not exist. But there is a low-pressure turbine which turns the propeller, so station 4.5 exists. Thus, we define new turbine properties

$$\pi_{tH} = \frac{p_{t4.5}}{p_{t4}}, \quad \pi_{tL} = \frac{p_{t5}}{p_{t4.5}}, \quad \tau_{tH} = \frac{T_{t4.5}}{T_{t4}}, \quad \tau_{tL} = \frac{T_{t5}}{T_{t4.5}}$$

The core engine analysis is identical to the basic turbojet. Specific thrust is

$$\frac{F}{\dot{m}} = a_0 \left[\sqrt{\frac{2}{\gamma - 1} \frac{\tau_\lambda}{\tau_r \tau_c} (\tau_r \tau_c \tau_{tH} \tau_{tL} - 1)} - M_0 \right]$$

so

$$C_C = \frac{a_0}{c_p T} \left[\sqrt{\frac{2}{\gamma - 1} \frac{\tau_\lambda}{\tau_r \tau_c} (\tau_r \tau_c \tau_{tH} \tau_{tL} - 1)} - M_0 \right]$$

Also, the fuel flow is also the same, namely,

$$\frac{\dot{m}_f}{\dot{m}} = \frac{c_p T_0}{h_{8f}^0} (\tau_\lambda - \tau_r \tau_c)$$

Now we will return to the propeller to find C_p. The power output from the low-pressure turbine must equal the power input to the propeller. In other words,

$$\dot{W}_p = \dot{m} c_p (T_{t4.5} - T_{t5})$$

C_p can be rewritten

$$C_p = \frac{\eta_p \dot{W}_p}{\dot{m} h_0} = \eta_p \tau_\lambda \tau_{tH} (1 - \tau_{tL})$$

Combining the core and propeller parts gives us the total specific thrust,

$$\frac{F_{\text{total}}}{\dot{m}} = \frac{(C_C + C_p)}{v_0} c_p T$$

This can then be used to find the thrust-specific fuel consumption:

$$S = \frac{v_0 (\tau_\lambda - \tau_r \tau_c)}{h_f^0 (C_C + C_p)}$$

Piston engine/propeller combinations are usually measured with power. Therefore, rather than use the specific thrust to determine the thrust-specific fuel consumption, we will use power-specific fuel consumption, S_p, where S_p is given by

$$S_p = \frac{\dot{m}_f/\dot{m}}{W_p/\dot{m}} = \frac{\tau_\lambda - \tau_r\tau_c}{h_f^0(C_C + C_p)}$$

The high-pressure turbine total temperature ratio, τ_{tH}, is the same as in ideal turbojet analysis, i.e.,

$$\tau_{tH} = 1 - \frac{\tau_r}{\tau_\lambda}(\tau_c - 1)$$

The low-pressure turbine, or free turbine, that drives the propeller, must be determined using energy balance. The energy balance between the free turbine and the propeller is

$$\dot{W}_p = \dot{m}c_p(T_{t4.5} - T_{t5})$$

Using our coefficient C_p, we get

$$C_p = \frac{\eta_p\dot{W}_p}{\dot{m}c_pT_0} = \eta_p\tau_\lambda\tau_{tH}(1 - \tau_{tL})$$

A turboprop is one type of the generic gas turbine engine called the turboshaft engine. The turboshaft engine uses the free turbine to power a shaft to power many things. For example, power stations can use a turboshaft engine for electricity generation. Helicopters use turboshaft engines, which, through gearboxes, power the rotor.

7.12 GAS TURBINE COMPONENT TECHNOLOGY

Diffusers

A real engine has aerodynamic skin friction, component friction, non-isentropic flows, and real gas properties. In this section the effect of these will be discussed with regard to individual components, except the last, which will be covered separately.

The diffuser is a component that involves careful integration of the engine with its housing. Air that enters the inlet of the engine is conditioned so that the first stage of the compressor receives gas at the proper conditions, namely Mach number and pressure. The diffuser has to function at all flight speeds of the aircraft without large performance losses. Therefore, the same diffuser that conditions air when standing still must work at cruise speeds as well.

An ideal diffuser has no losses, so total pressure and temperature remain constant. Unfortunately, frictional losses will serve to decrease the total pressure so that π_d is not unity. Therefore, the figure of merit for the inlet is just π_d. It is also convenient to introduce the isentropic efficiency, η_d, which is defined as

$$\eta_d = \frac{h_{t2} - h_0}{h_{t0} - h_0} = \frac{T_{t2} - T_0}{T_{t0} - T_0}$$

Under ideal conditions, and $h_{t2} = h_{t0}$ and $T_{t2} = T_{t0}$, so that $\eta_d = 1$. Relating the diffuser pressure ratio, π_d, to the isentropic efficiency, we get

$$\eta_d = \frac{\pi_d^{(\gamma-1)/\gamma} - 1}{\tau_r - 1}$$

Typical values of π_d and η_d for a subsonic inlet are shown in Figure 7.13.

Supersonic inlets have shock waves, which make them less efficient due to the total pressure drop across the shocks. Therefore, in addition to frictional losses discussed above, there will be shock, or "ram" recovery, losses. An empirical formula for dealing with the ram recovery is given in Military Specification 5008B, which first breaks the diffuser pressure ratio into two parts as follows:

$$\pi_d = \pi_{d\,\text{friction}}\eta_r$$

where $\pi_{d\,\text{friction}}$ is the pressure loss due to friction discussed above and η_r is the ram recovery factor. The empirical formula for η_r is

$$\eta_r = \begin{cases} 1 & M_0 \le 1 \\ 1 - 0.075(M_0 - 1)^{1.35} & 1 < M_0 < 5 \\ \dfrac{800}{M_0^4 + 935} & 5 \le M_0 \end{cases}$$

and is illustrated in Figure 7.14.

Supersonic inlets are an integral part of ramjet and scramjet engines and will be discussed in more detail in that context below.

Compressors

The job of the compressor is to compress the air to increase engine thermal efficiency. The compressor does work on the air by increasing the enthalpy of the gas.

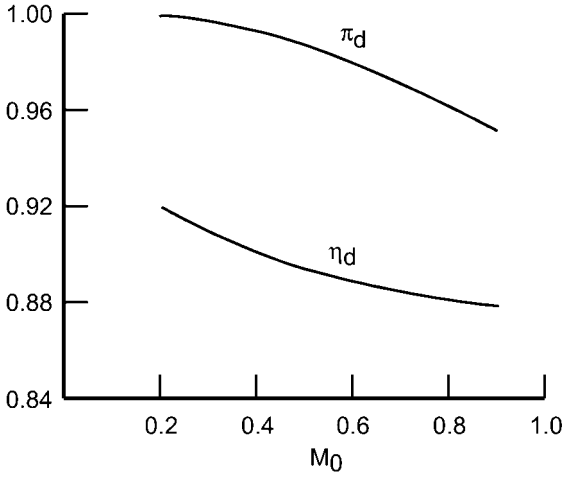


FIGURE 7.13 Diffuser efficiency and pressure ratio versus Mach number.

AERONAUTICAL PROPULSION

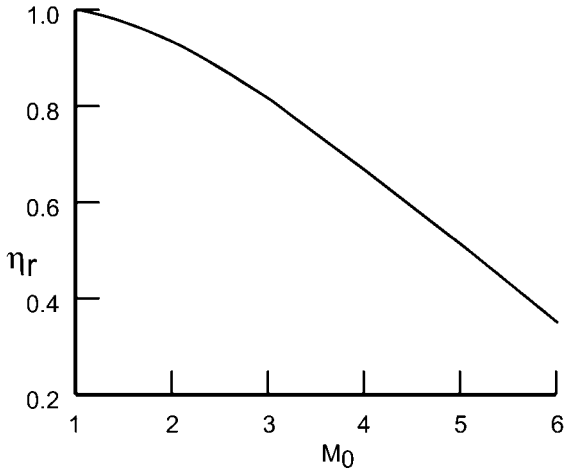


FIGURE 7.14 Recovery factor versus Mach number.

Without a compressor, the thermal efficiency of an engine would be due to ram compression alone. Reviewing thermal efficiency, we see that

$$\eta_T = \frac{\text{net } \dot{W}_{\text{out}}}{\dot{Q}_{\text{in}}} = \frac{\dot{m}(\Delta h_t - \Delta h_c)}{\dot{m}\Delta h_b} = \frac{\dot{m}c_p[T_4 - T_9 - (T_3 - T_2)]}{\dot{m}c_p(T_4 - T_3)}$$

for an ideal Brayton cycle, $T_3/T_2 = T_4/T_9 = (p_3/p_2)^{(\gamma-1)/\gamma}$. Thus,

$$\eta_T = 1 - \left(\frac{1}{p_3/p_2}\right)^{(\gamma-1)/\gamma}$$

and so we see that greater thermal efficiency is obtained by increasing the compressor pressure ratio, p_3/p_2.

In order to compress the air, the compressor must do work. This is accomplished by either an axial or a centrifugal compressor. The axial compressor works using stages of rotating and stationary blades, called rotors and stators, respectively. The compressor blades create channels that add energy to the air, as shown in Figure 7.15. As the name implies, the rotors rotate and the stators are stationary. A rotor–stator combination is called a *stage*. There are typically six to eight stages per compressor. Compression must be performed in stages because, as with a wing, too great an adverse pressure gradient will result in separated flow. In a compressor this is known as compressor stall. Thus, smaller increments of pressure increases are performed by each of the stages.

Centrifugal compressors add energy by accelerating the gas outwardly. The advantage of a centrifugal compressor is that it is cheaper to manufacture since a single unit can be cast to form the entire compressor. The disadvantage is that the rerouting of the air from axial to radial and then ultimately back to axial creates high losses.

Losses occur in compressors just as they do diffusers. There are frictional losses, leakage around the blade tips, and wakes from each blade. The result of these losses is a modification to the ideal Brayton cycle as shown in Figure 7.16. The compressor ends with a higher temperature, at the same pressure, and higher entropy.

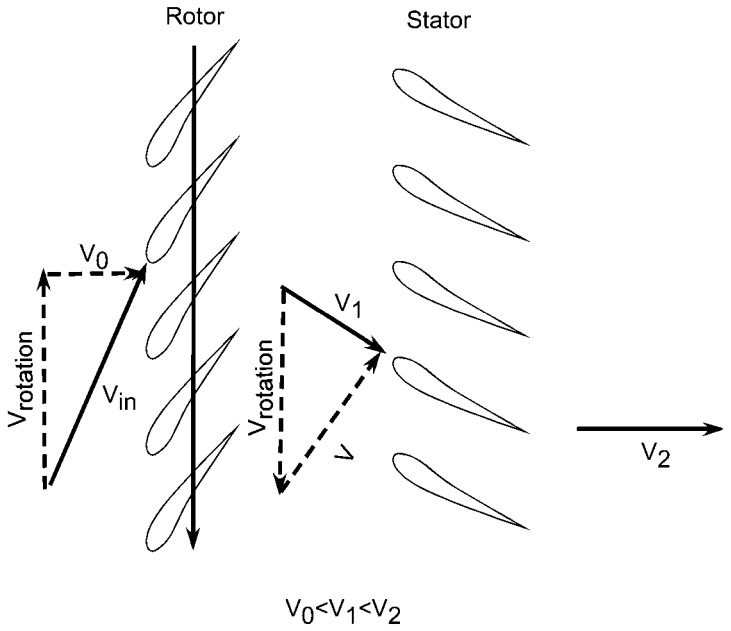


FIGURE 7.15 Compressor blades.

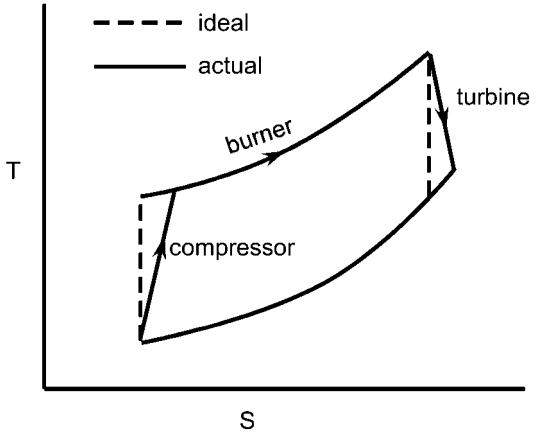


FIGURE 7.16 Modified Brayton cycle.

Studying a compressor requires looking at each stage. In the following, we will denote each stage by the subscript i. π_{ci} and τ_{ci} represent the total pressure and temperature rise across compressor stage i. The efficiency is for a single stage is thus written

$$\eta_{ci} = \frac{\pi_{ci}^{(\gamma-1)/\gamma} - 1}{\tau_{ci} - 1}$$

We can rearrange the above equation to find τ_{ci}, which gives

$$\tau_{ci} = \frac{T_{ti}}{T_{t(i-1)}} = 1 + \frac{1}{\eta_{ci}}\left[\left(\frac{p_{ti}}{p_{t(i-1)}}\right)^{(\gamma-1)/\gamma} - 1\right]$$

The total temperature rise over all stages of an N-stage compressor, τ_c, is the product of the total temperature rise of the stages. So,

$$\tau_c = \prod_{i=1}^{N}\left\{1 + \frac{1}{\eta_{ci}}\left[\left(\frac{p_{ti}}{p_{t(i-1)}}\right)^{(\gamma-1)/\gamma} - 1\right]\right\}$$

and similarly,

$$\pi_c = \prod_{i=1}^{N}\frac{p_{ti}}{p_{t(i-1)}} = \frac{p_{tN}}{p_{t0}}$$

The total compressor efficiency can then be written

$$\eta_c = \frac{\left(\prod_{i=1}^{N}\frac{p_{ti}}{p_{t(i-1)}}\right)^{(\gamma-1)/\gamma} - 1}{\left(\prod_{i=1}^{N}\left\{1 + \frac{1}{\eta_{ci}}\left[\left(\frac{p_{ti}}{p_{t(i-1)}}\right)^{(\gamma-1)/\gamma} - 1\right]\right\}\right) - 1}$$

Introducing the polytropic efficiency, e_c, which is the ratio of ideal work of compression for a differential pressure change to the actual work, simplifies the above expression. The polytropic efficiency is written

$$e_c = \frac{dw_{\text{ideal}}}{dw_{\text{actual}}} = \frac{dh_{t-\text{ideal}}}{dh_{t-\text{actual}}} = \frac{dT_{t-\text{ideal}}}{dT_{t-\text{actual}}}$$

This, with the isentropic relation

$$\frac{dT_{t-\text{ideal}}}{T_t} = \frac{\gamma - 1}{\gamma}\frac{dp_t}{p_t}$$

gives

$$e_c = \frac{dT_{t-\text{ideal}}/T_t}{dT_{t-\text{actual}}/T_t} = \frac{\gamma - 1}{\gamma}\frac{dp_t/p_t}{dT_{t-\text{actual}}/T_t}$$

Rewriting this equation gives the simple relationship

$$\frac{dT_{t-\text{actual}}}{T_t} = \frac{\gamma - 1}{\gamma e_c} \frac{dp_t}{p_t}$$

and then, by integrating from points 2 to 3,

$$\frac{T_{t3}}{T_{t2}} = \left(\frac{p_{t3}}{p_{t2}}\right)^{(\gamma-1)/\gamma e_c}$$

or, finally,

$$\tau_c = \pi_c^{(\gamma-1)/\gamma e_c}$$

Plugging this result into the equation for efficiency gives the compressor efficiency as

$$\eta_c = \frac{\pi_c^{(\gamma-1)/\gamma} - 1}{\pi_c^{(\gamma-1)/\gamma e_c} - 1}$$

The key parameters to determine compressor performance, namely π_c and η_c, are the corrected mass flow and corrected engine speed. A four-parameter map, called a compressor map, is normally used to determine performance. In engine design groups software replaces this role, but a compressor map gives an overall picture of the compressor performance. Figure 7.17 illustrates a compressor map.

Note the addition of the "stall surge line" in the figure. The stall surge line is an operational limiting line of the compressor. Above this line, the compressor will stall and the burner will force hot gas out through the compressor, instead of just the turbine. Since the compressor is not designed for such temperatures, a compressor stall can be catastrophic.

Burner

Incomplete combustion and total pressure losses occur in the burner. The combustion efficiency, η_b, can be difficult to determine because the properties of the gas are different before and after combustion. For convenience, we'll label the two specific heats, c_p, as c_{pc} and c_{pt}, for the compressor and turbine values, respectively. Determining the amount of energy added to that available gives

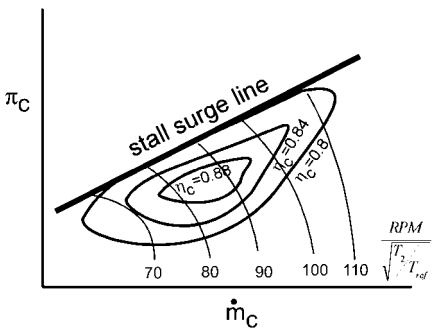


FIGURE 7.17 Compressor map.

$$\eta_b = \frac{(\dot{m} + \dot{m}_f)c_{pt}T_{t4} + \dot{m}c_{pc}T_{t3}}{\dot{m}_f h_{pr}}$$

The total pressure loss means that π_b is not unity, i.e.,

$$\pi_b = \frac{p_{t4}}{p_{t3}} < 1$$

Turbines

The turbine acts just like the compressor except that it expands the gas and extracts work. Developing equations for η_t and π_t follow similar procedures to the development of η_c and π_c. See Figure 7.16 for an illustration of the effect the nonideal turbine has on the Brayton cycle. If we define η_t as the ratio of actual to ideal turbine work at a given π_t, we get

$$\eta_t = \frac{h_{t4} - h_{t5}}{h_{t4} - h_{t5i}} = \frac{T_{t4} - T_{t5}}{T_{t4} - T_{t5i}}$$

or,

$$\eta_t = \frac{1 - \tau_t}{1 - \pi_t^{(\gamma-1)/\gamma}}$$

Turbines also work in stages. However, since the pressure gradient is favorable for a turbine (i.e., the pressure downstream is lower than upstream), fewer stages are required than were needed for the compressor. A typical turbine may only have two stages of rotor/stator combinations. Following the development for a multistage compressor, we find that for an N-stage turbine,

$$\eta_t = \frac{1 - \prod_{i=1}^{N}\left[1 - \frac{1}{\eta_{si}}(\pi^{(\gamma-1)/\gamma} - 1)\right]}{1 - \pi_t^{(\gamma-1)/\gamma}}$$

where η_{si} is the stage efficiency of the i^{th} stage.

Along similar lines as with the compressor, we can define a polytropic turbine efficiency, e_t, defined as the actual turbine work for a differential pressure change over the ideal work. Thus,

$$\pi_t = \tau_t^{\gamma/[(\gamma-1)e_t]}$$

This gives, for the turbine,

$$\eta_t = \frac{1 - \tau_t}{1 - \tau_t^{1/e_t}} = \frac{1 - \pi_t^{(\gamma-1)/\gamma e_t}}{1 - \pi_t^{(\gamma-1)/\gamma}}$$

Nozzle

The exhaust nozzle is very nearly adiabatic. Therefore, τ_n can be considered to equal unity. However, π_n is less than one due to frictional losses. Thus,

$$\pi_n = \frac{p_{t9}}{p_{t8}} < 1$$

7.13 REAL GAS PROPERTIES

In the ideal engine, we have assumed that the gas properties are constant and ideal through the engine. Thus, the properties of the fuel and the combustion products are assumed the same as the freestream air. These properties include the gas molecular mass and specific heats.

We have also only considered one state of the engine, the design condition. An engine must operate with a wide range of operating conditions, and only one is considered the on-design condition.

In the previous analyses the ratio of specific heats used was constant, γ. The air flowing through the compressor will not change significantly, but the combustion produced will have a different γ. Therefore, the analysis of the compressor and the turbine must include the appropriate values. Also, at high temperatures γ is no longer a constant.

Engine cycle analysis determines the performance given specific component performances. Once the optimum engine cycle is determined, the performance losses for off-condition states must be analyzed. If an engine is designed for a cruise condition, how does it behave during climb, during idle, etc.? To look at off-design conditions, each component must be analyzed individually. Unfortunately, there is not sufficient space in this handbook to consider these details.

7.14 RAMJETS AND SCRAMJETS

Ramjets

In Section 7.2 we saw that engine thermal efficiency increased at higher burner inlet pressures. Therefore, compressors where added to compress the air. It is possible, however, to simply use the pressure rise from ram air to compress the air. Figure 7.18 illustrates the ideal ramjet, which has no compressor or turbine.

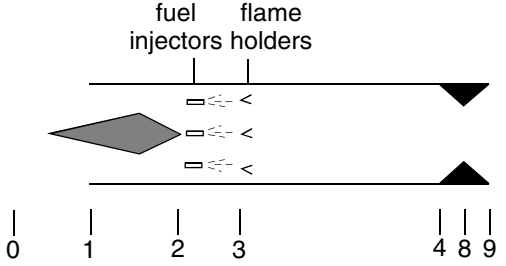


FIGURE 7.18 Ideal ramjet.

If an ideal engine were to bring the ram air to a complete rest before combustion, we could use the isentropic relations to determine the pressure at the burner inlet. Then the burner inlet pressure and temperatures would be

$$p_{t2} = p_0 \left(1 + \frac{\gamma - 1}{2} M_0^2\right)^{\gamma/(\gamma-1)}$$

$$T_{t2} = T_0 \left(1 + \frac{\gamma - 1}{2} M_0^2\right)$$

One thing we can see here is that no compression can occur when there is no forward motion. Now we will look at an ideal ramjet cycle analysis in more detail. The thrust from an ideal ramjet engine, as with the turbojet engine, is given by

$$\frac{F}{\dot{m}} = v_9 - v_0 = a_0 \left(\frac{v_9}{a_0} - 1\right)$$

Again, as with the turbojet, we can use our assumptions and the definition of the sound speed to find v_9/a_0, which is

$$\left(\frac{v_9}{a_0}\right)^2 = \frac{T_9}{T_0} M_9^2$$

With no compressor or turbine, there is no total pressure change, in an ideal ramjet. Thus,

$$\frac{p_{t9}}{p_0} = \pi_r$$

We also have the relationship that

$$\frac{p_{t9}}{p_9} = \frac{p_{t9}}{p_0}\frac{p_0}{p_9} = \pi_r$$

The exit Mach number, M_9, can be found from the above relationship and is

$$M_9^2 = \frac{2}{\gamma - 1}\left[\left(\frac{p_{t9}}{p_9}\right)^{(\gamma-1)/\gamma} - 1\right] = \frac{2}{\gamma - 1}\left[\pi_r^{(\gamma-1)/\gamma} - 1\right]$$

Adding in the isentropic relationship between total temperature and total pressure changes, we have

$$M_9^2 = \frac{2}{\gamma - 1}\left[\tau_r - 1\right] = M_0^2$$

Thus, for an ideal ramjet, $M_9 = M_0$. Next we need to find T_9/T_0. We start with

$$\frac{T_{t9}}{T_0} = \tau_r \tau_d \tau_b \tau_n = \tau_r \tau_b$$

Using isentropic relations, we have

$$\frac{T_{t9}}{T_9} = \left(\frac{p_{t9}}{p_0}\right)^{(\gamma-1)/\gamma} = \pi_r^{(\gamma-1)/\gamma} = \tau_r$$

Putting these last two expressions together gives

$$\frac{T_9}{T_0} = \frac{T_{t9}/T_0}{T_{t9}/T_9} = \frac{\tau_r \tau_b}{\tau_r} = \tau_b$$

The results for T_9/T_0 and M_9 give us the specific thrust, namely,

$$\frac{F}{\dot{m}} = a_0\left(\frac{v_9}{a_0} - M_0\right) = a_0\left(\sqrt{\frac{T_9}{T_0}M_9^2} - M_0\right) = a_0 M_0(\sqrt{\tau_b} - 1)$$

This result shows that with no forward speed, $M_0 = 0$, there is no thrust. Thus, the biggest problem with using a ramjet is establishing a forward velocity so the engine can begin to work. Another problem is that high speeds are needed to achieve a reasonable thermal efficiency. The thermal efficiency is given by

$$\eta_T = 1 - \frac{1}{\tau_r}$$

which means that τ_r must be large and so M_0 must be large. To put this in perspective, a Mach 1 ideal ramjet has a thermal efficiency of 17%, as shown in Figure 7.19.

If the ramjet engine is going supersonic, shock compression can be used to compress the air before it reaches the burner. Due to the need to run at high speeds, ramjets are most often used at speeds above Mach 2. This will result in strong shock waves, with their associated compression and losses. The next step will be to look at nonideal ramjets and shock compression.

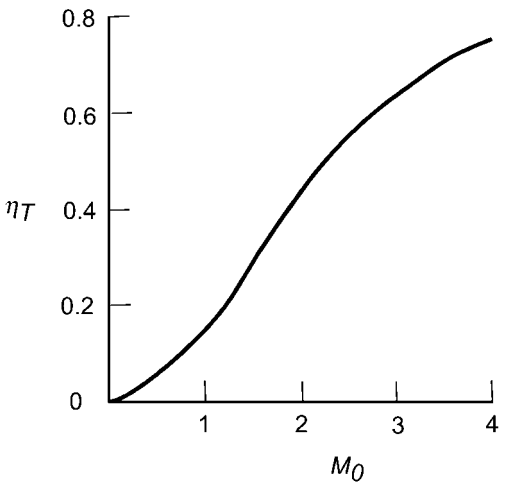


FIGURE 7.19 Thermal efficiency of ideal ramjet versus Mach number.

A supersonic ramjet has inlet shocks, as shown in Figure 7.20. A normal shock results in the largest pressure rise, but also the largest total pressure loss, i.e., $\pi_d < 1$. Unfortunately, the total pressure loss is the dominating factor since thermal efficiency will decrease. The result is that an efficient inlet, using a series of oblique shocks, must be used.

Oblique shock waves will result in an increase in pressure and total pressure loss, but the end result is much less than a normal shock. Let's look at an example of a Mach 2 inlet with a single normal shock verses two oblique shocks terminated with a normal shock. A normal shock wave at Mach 2 has the following properties:

$$\frac{p_2}{p_1} = 4.5 \quad \text{and} \quad \frac{p_{t2}}{p_{t1}} = 0.72$$

(note that subscripts 1 and 2 signify the upstream and downstream conditions of the shock wave, respectively). From this result, we find that the recovery efficiency, η_r, is 72%. For the oblique shocks we have to decide on a diffuser design. Let's select a leading wedge angle of 10° and put position 1 where the shock impinges on the cowl. Note that for this simplified example we will use a two-dimensional engine, which gives us simple two-dimensional oblique shocks instead of conical shock waves. The resulting Figure 7.21 illustrates the shock waves used in this example, with numbers in Table 7.2. The net result is that the static pressure rises by a factor of 4.9 while the total pressure loss is only 4.5%, for a recover efficiency of 95.5%. Thus, in this example we clearly see the advantage of using oblique shock waves.

Scramjets

The cycle analysis for ramjets shows that the specific thrust is proportional to the flight Mach number and the burner total temperature ratio. The dependency on

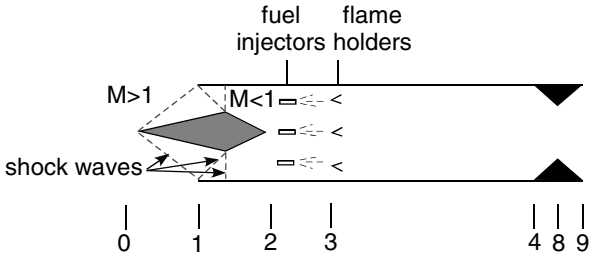


FIGURE 7.20 Ramjet engine with inlet shock waves.

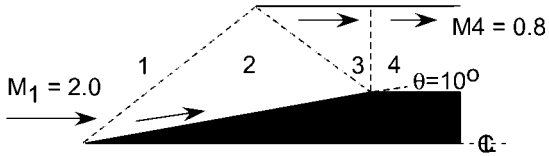


FIGURE 7.21 Inlet shock waves.

TABLE 7.2

Shock wave	p_2/p_1	p_{t2}/p_{t1}	β	θ	M_2
1–2	1.72	0.984	39.3	10	1.64
2–3	1.64	0.988	49.4	10	1.28
3–4	1.75	0.982	90	0	0.80
Total	4.9	0.955			

Mach number suggests that ramjets develop more thrust the faster they go, with a corresponding increase in fuel consumption. However, this ignores thermal properties of materials used to build the engine. Assuming there is a limit on the burner exit temperature, T_t, we find that τ_b will be limited. This is due to an increasing burner inlet total temperature, T_{t3}, as the Mach number increases. Recall that, for an ideal ramjet engine,

$$\frac{T_{t3}}{T_0} = \frac{T_{t2}}{T_0} = \frac{T_{t0}}{T_0} = \tau_r = 1 + \frac{\gamma - 1}{2} M_0^2$$

Eventually T_{t3} will approach T_{t4}, which T_4 depends on, so that at some point τ_b will go to zero. Thus, the specific thrust will go to zero.

The key problem is that T_{t3} approaches T_{t4}. Thus, the solution is to find a way to reduce T_{t3}. The burner inlet temperature can be reduced by exchanging thermal energy of the air at that point with kinetic energy. If we look at the thermal energy/kinetic energy tradeoff, we know that

$$\frac{T_3}{T_0} = \frac{\dfrac{T_{t3}}{T_2}}{\dfrac{T_{t0}}{T_0}} = \frac{1 + \dfrac{\gamma - 1}{2} M_0^2}{1 + \dfrac{\gamma - 1}{2} M_3^2}$$

If we then assume that the air flowing into the burner is subsonic ($M_3 < 1$), the minimum value for T_3/T_0 will be

$$\frac{T_3}{T_0} = \frac{1 + \dfrac{\gamma - 1}{2} M_0^2}{1 + \dfrac{\gamma - 1}{2}}$$

The trade-off between thermal and kinetic energy shows how T_3 is dependent not only on M_0 but also on the Mach number at the burner inlet. Suppose that, rather than assuming the flow through the engine is subsonic, we allow the air to flow supersonically through the burner. Then the result, derived above, shows that the higher the Mach number of the burner entrance, the lower T_3. See Table 7.3 for ideal gas solutions for various values of M_0 and M_3.

Looking at a flight altitude of 6000 m, where the outside temperature is approximately −20°C, we find that unrealistically high values of T_3 result. For reference, water dissociates at roughly 4000°C, so hydrogen will not burn in oxygen.

Thus, at hypersonic speeds we are motivated to have the flow pass through the burner supersonically. The idea is the same as a ramjet, except that the air never

AERONAUTICAL PROPULSION

TABLE 7.3

M_0	M_3	T_3/T_0	T_3 at 6000 m (°C)
4	0.4	4.07	760
4	2	2.33	320
6	0.4	7.95	1700
6	2	4.55	880
6	4	1.95	220
8	0.4	13.4	3100
8	2	7.67	1700
8	4	3.29	560
10	0.4	20.3	4900
10	2	11.7	2700
10	4	5.00	990

becomes subsonic as shown in Figure 7.22. This engine is called a scramjet (supersonic combustion ramjet).

New problems enter the design of a scramjet engine. The engine inlet becomes a dominant part of the design. The engine and airframe must be highly integrated. Figure 7.22 shows a conceptual example of a hypersonic vehicle design. The inlet, with its complex, interacting shock waves, must work over a range of flight Mach numbers. Normal shocks must be avoided. At high speeds through the burner the mixing of fuel and air can be a problem. These problems are slowly being conquered, but the details are beyond the scope of this handbook.

7.15 RECIPROCATING ENGINES

Piston Engines

Virtually all internal combustion aircraft engines are based on the Otto cycle. A few engines in development are exploring using the diesel cycle.

The efficiency of the Otto cycle is given by

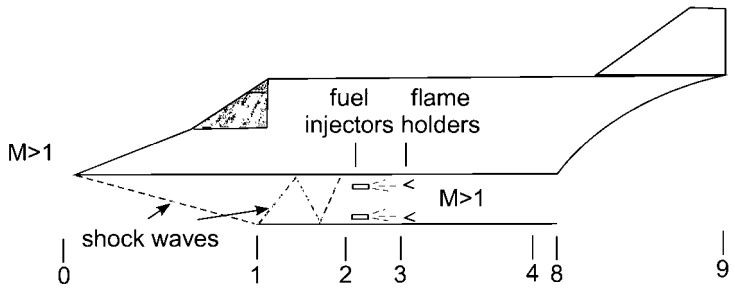


FIGURE 7.22 Scramjet application.

$$\eta_{\text{Otto}} = 1 - \frac{T_4 - T_1}{T_3 - T_2}$$

For an ideal piston engine, the expansion and compression occur with no heat addition and are reversible and therefore isentropic. Thus, the relation $TV^{\gamma-1}$ holds from 1 to 2 and from 3 to 4. Thus, we have

$$T_3 = T_4 * \left(\frac{V_4}{V_3}\right)^{\gamma-1} \quad \text{and} \quad T_2 = T_1 * \left(\frac{V_1}{V_2}\right)^{\gamma-1}$$

In a piston engine, the compression and expansion strokes have identical volume changes, so we can define the compression ratio:

$$r_c = \frac{V_2}{V_1} = \frac{V_3}{V_4}$$

Which gives

$$\eta_{\text{Otto}} = 1 - \frac{1}{r_c^{\gamma-1}}$$

Therefore, the cycle efficiency of the Otto cycle improves with greater compression. In a normal engine, there is heat transfer to the piston and the piston does not remain fixed, to fix the volume, during combustion. These leads to inefficiencies in the engine cycle. In addition, mechanical losses enter from pumping the pistons during the exhaust/intake cycle. There can also be leakage around the pistons and imperfect fuel/air mixing. In a ''real'' engine we must adjust the efficiency to take into account these engine cycle and mechanical losses, which gives a definition for overall engine efficiency:

$$\eta_0 = \eta_i \eta_m \eta_{\text{Otto}}$$

with representing the thermal losses due to a imperfect engine cycle and representing the mechanical losses. Typical engines have overall efficiencies around 30%.

Propellers

The tools to design propellers have improved greatly since the maturing of computer modeling. However, the classical methods, developed before World War II, are still useful for estimates. The first such theory was the actuator disk theory which replaces the propeller with a disk that accelerates the fluid. The second is blade element theory, which sums the forces on cross sections, or elements, of the propeller blade. After World War II, lifting line and lifting surface methods were developed. These methods are much more complex and require computers to use.

In actuator disk theory, the thrust is calculated by considering the change in momentum for air flowing through a disk. Figure 7.23 illustrates how air behaves as it passes through a propeller. A notable feature is that there is a contraction of the stream-tube that flows through the propeller.

Actuator disk theory simplifies this picture by assuming a discontinuous change in velocity and pressure and no contraction of the stream tube. The simplified picture is shown in Figure 7.24. In the figure, $u_1 = V$ and $u_2 = V + \Delta V$. This figure

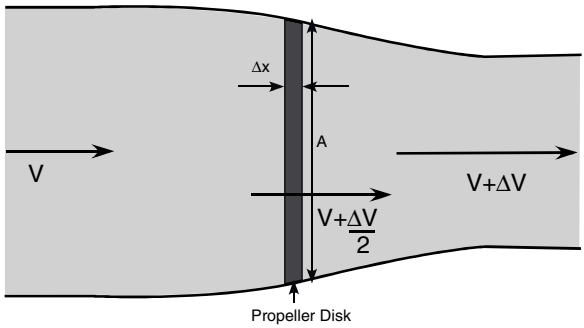


FIGURE 7.23 Propeller disk.

is convenient for using in analysis of flow behind the propeller. From the figures we can determine that the change in momentum in time Δt, or thrust, is given by:

$$T = \frac{\rho A \Delta x(V + \Delta V) - \rho A \Delta x V}{\Delta t} = \rho A \frac{\Delta x}{\Delta t} \Delta V = \rho A \left(V + \frac{\Delta V}{2}\right) \Delta V$$

where ρ is the air density and A is the disk area, which is equal to $(\pi/4)D^2$. The velocity through the propeller disk, $\Delta x/\Delta t$, is taken to be the average of the velocities before and after the propeller

$$\left(\frac{V + (V + \Delta V)}{2}\right) = V + \frac{\Delta V}{2}$$

Next, we would like to find the power output of the propeller and the propulsive efficiency. The useful power output of the propeller, the power that is going to create thrust, is given by $P_p = TV$. The power that is imparted by the propeller is the difference kinetic energy before and after the propeller disk per second. The propulsive efficiency is the useful propeller power output divided by the power imparted to the fluid, or $\eta_p = (TV/P_{\text{fluid}})$. The power imparted to fluid is

$$\begin{aligned} P_{\text{fluid}} &= \frac{1}{2} \rho A \left(V + \frac{\Delta V}{2}\right) (V + \Delta V)^2 - \frac{1}{2} \rho \left(V + \frac{\Delta V}{2}\right) V^2 \\ &= \frac{1}{2} \rho A \left(V + \frac{\Delta V}{2}\right) (2V\Delta V + \Delta V^2) \end{aligned}$$

The propeller efficiency is thus

$$\eta_p = \frac{\rho A \left(V + \frac{\Delta V}{2}\right) \Delta V * V}{\frac{1}{2} \rho A \left(V + \frac{\Delta V}{2}\right) (2V\Delta V + \Delta V)^2} = \frac{V}{V + \frac{\Delta V}{2}} = \frac{1}{1 + \frac{\Delta V}{2V}}$$

Two obvious limits are that the propulsive efficiency goes to one as ΔV goes to zero and goes to zero as V goes to zero.

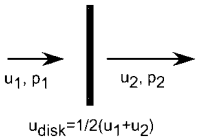


FIGURE 7.24 Actuator disk model.

Another method for calculating the thrust and torque on a propeller is blade-element theory. In blade-element theory, the propeller is divided into sections as shown in Figure 7.25. Each section is solved as a quasi-3D problem and then summed to give the overall thrust and torque.

The section between r and $r + \Delta r$ is solved as an airfoil as shown in Figure 7.26.

The thrust and torque of the element are

$$\Delta T = L \cos(\theta - \alpha) - D \sin(\theta - \alpha)$$

and

$$\Delta Q = [L \sin(\theta - \alpha) + D \cos(\theta - \alpha)] * r$$

or using coefficients,

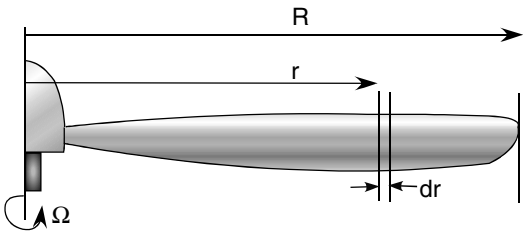


FIGURE 7.25 Blade elements.

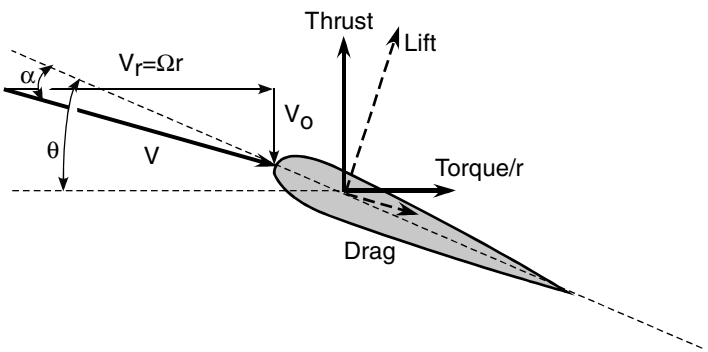


FIGURE 7.26 Propeller section.

$$\Delta T = \frac{1}{2}\rho V^2[C_L \cos(\theta - \alpha) - C_D \sin(\theta - \alpha)] * cdr$$

and

$$\Delta Q = \frac{1}{2}\rho V^2[C_L \sin(\theta - \alpha) + C_D \cos(\theta - \alpha)] * r * cdr$$

with the local propeller chord, c, and $V^2 = V_0^2 + V_r^2$. The total thrust and torque are summed over all elements, including all blades.

An additional complexity is that the velocity, V, seen by the propeller section includes the effect of all other propeller sections in the form of swirl. In other words, V_0 is not the same as the forward airspeed and V_r is not really Ωr. A modification to the above equations is necessary to correct for the swirl of the incoming airflow.

It is convenient to consider thrust and torque coefficients to allow for generalizations. Convenient variables to obtain nondimensional coefficients are blade diameter, D, and rotations per second (or minute), n (rpm). The dimensions of velocity, m/s, can be obtained using the product of D and n. Thus, we see that

$$T \propto \rho(nD)^2 * D^2 \quad \text{and} \quad Q \propto \rho(nD)^2 * D^3$$

so

$$C_T = \frac{T}{\rho n^2 D^4} \quad \text{and} \quad C_Q = \frac{Q}{\rho n^2 D^5}$$

Another useful parameter for describing propellers is the advance ratio, J. The advance ratio is the nondimensional ratio of aircraft forward velocity, V_0 and twice the propeller tip rotational velocity, nD, or

$$J = \frac{V_0}{nD}$$

The power supplied to the propeller is $P = 2\pi nQ$, so the efficiency is given by

$$\eta = \frac{TV_0}{2\pi nQ} = \frac{C_T V_0}{2\pi nDC_Q} = \frac{J}{2\pi}\left(\frac{C_T}{C_Q}\right)$$

In most literature the torque coefficient is replaced by a power coefficient, which is simply $C_p = 2\pi C_Q$, which gives an efficiency of

$$\eta = \frac{C_T}{C_p} J$$

7.16 AIRCRAFT ENGINE EMISSIONS AND FUELS

Engine Fuels

Gas turbine engines, in principle, can run on any type of fuel. But the stringent requirements and high performance of aircraft gas turbines limits the fuels used

due to the need for specific characteristics. Jet fuels are typically kerosene based. Kerosene, a petroleum product, actually designates a type of fuel, not a specific fuel. Fuels for jet engines must meet stringent requirements for volatility, specific gravity, and addition of additives. Specifications for jet fuels are outlined by the American Society for Testing Materials.

Volatility represents the ability of the fuel to vaporize. At cool temperatures, a particular fuel may not vaporize enough for efficient combustion. Similarly, some fuels may produce vapor while in storage, resulting in a highly combustible and dangerous situation.

Fuel density changes with temperature. Changes in the fuel's specific gravity can result in incorrect fuel flow to the burner.

Additives are required to protect fuel lines from corroding, prevent bacterial growth, lower the freezing temperature of the fuel, and help emissions.

The military and commercial fuel markets use different fuels. The U.S. Air Force uses JP-4, while the Navy uses JP-5. The primary difference between the two is that JP-5 is less volatile, which makes it more convenient for storage on ships. JP-4 is a "wide-cut" fuel, which means that its boiling point is between that of kerosene and gasoline. JP-5 is a kerosene-based fuel. Commercial aircraft use Jet A or Jet A-1 fuels. These fuels are denser and release more heat per volume then their military counterparts. The specifications for these fuels are given in ASTM D1655 Standard Specification for Aviation Turbine Fuels. The difference between Jet A and Jet A-1 is that Jet A-1 has a lower freezing point. (Refer to Table 7.4.)

Aviation piston engines are powered by leaded gasoline that meets ASTM D910, Standard Specification for Aviation Gasolines. The fuels used for aviation piston engines are referred to as "avgas." Three varieties of avgas are being phased out, 80 octane, 100LL, and 100 octane, while an 82 octane, unleaded fuel is being developed. 100LL is a low-lead version of the 100 octane gas, which is no longer manufactured. Manufacturers have stopped making 80 octane avgas and it is now virtually nonexistent. The "octane" in the fuel is a measure of the percentage of "iso-octane" in the fuel. The octane gives a measure of the antiknock performance of the fuel. Knock occurs when the fuel preignites and causes a high-pressure knock. High-octane fuels slow ignition to avoid the large pressure spikes.

Emissions

The combustion process of hydrocarbons and air results in the combustion products of H_2O and CO_2. In addition to these two gases, unburned hydrocarbons, CO, NO_x, and SO_x, are byproducts of the process and incomplete combustion. Carbon mon-

TABLE 7.4

Fuel	Typical density at 15°C (g/mL)	Typical energy content	
		Specific MJ/kg	Volumetric MJ/L
Aviation Gasoline	0.715	43.71	31.00
JP-4	0.751	43.46	32.64
JP-5	0.818	43.00	35.17
Jet A-1	0.710	43.23	30.70

oxide, CO, results from incomplete combustion and so will be greater for richer mixtures. CO is more of a problem for piston engines as better gas turbine combustor designs limit CO production. (Catalytic converters are used in the auto industry to treat CO.) High burner temperatures will increase CO emissions, conflicting with engine efficiency.

Nitrogen oxides, NO_x, are affected most by flame temperature. Emissions are far worse during takeoff when maximum power, and thus high flame temperature, are required. During cruise, when the engine is throttled back significantly, the emissions of NO_x can be one-third of those during maximum power output. Therefore, most work on improving NO_x emissions focuses on the takeoff phase.

The production of NO_x is maximum at stoichiometric mixtures and less when the fuel/air mixture is either rich or lean. Unfortunately, CO emissions behave exactly the opposite. Methods to decrease NO_x emissions include water injection, catalytic converters, and burner design. Water injection decreases flame temperature and increases power output. So adding water, particularly during the takeoff phase, has the beneficial effect of both increasing power output and reducing NO_x. Water injection is now out of favor because of operating problems (water injection pushes the compressor closer to the surge line for one). Catalytic reduction of NO_x is not practical for aircraft gas turbines. The most effective method of reducing NO_x is through improved burner design.

Decreasing flame temperature, while allowing for complete combustion to eliminate incomplete combustion, can be done by running lean and providing staged combustors, one for idle and one for primary combustion, as illustrated in Figure 7.27. The two stages, or zones, can be adjusted to run lean at high power settings, as shown in Figure 7.27. Only one of the burners will be lit during certain phases of the flight. Premixing the fuel and air can help thwart incomplete combustion.

7.17 ENGINE NOISE

Noise has become one of the primary constraints placed on commercial airlines over the last few decades. Starting with the Noise Control Act of 1972, the EPA set guidelines for allowable jet noise. The regulations regarding noise set by the International Civil Aviation Organization (ICAO) have been phased in over years. The current level of noise is known as Chapter 3, or Stage 3, as referred to by the FAA.

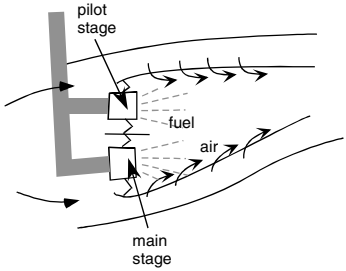


FIGURE 7.27 Dual burner for reduced emissions.

In the United States, the Airport Noise and Capacity Act of 1990 stipulated that 100% of the aircraft fleet be Stage 3 by the year 2000. Stage 3 refers to the level of noise allowable for takeoff and approach as well as sideline noise. As of June 2001 a new set of ICAO standards, Chapter 4, has been outlined, with plans to be phased in by 2006. The goal noise reduction is to reduce the "decibels" of the sound. Decibels are a logarithmic measure of sound energy, with levels spanning a range of 10^{12} in amplitude, in 1/log(10) increments. So, the decibel scale goes from 1 to 120, with each decibel a factor of 1.259 greater than the previous. At the low end, a clock ticking and the noise in a library are about 30 decibels. At the other end, pushing a gas-powered lawn mower or standing three meters from a pneumatic drill exposes you to 100 decibels. 100 decibels is 10^7 more energy than 30 decibels. From 50 meters, a jet engine will range from 100 to 135 decibels, depending on the particular engine and its thrust level. To get a feel for decibels in terms of power, the relationship between noise power (in watts) versus decibels is:

$$\text{dB} = 10 \log_{10} \frac{\text{Power}}{10^{-12}}$$

Tables 7.5 to 7.7 show noise requirements for aircraft of different weights and numbers of engines. The required drop in noise at takeoff from Stage 2 to Stage 3 aircraft is 60% for light aircraft and 50% for heavy aircraft. For sideline noise, lighter two-engine aircraft must decrease noise by over 80%. Chapter 4 is looking at another 10-db drop, or 10% of the noise associated with current aircraft (Chapter 3). These can be shown graphically as illustrated in Figures 7.28 and 7.29.

The noise energy is produced from a variety of sources, as illustrated in Figure 7.30. The most significant used to be jet exhaust, due to turbulent mixing of the jet exhaust with the surrounding air. The higher the velocity of the jet exhaust, the more noise generated. The introduction of the fanjet increased thrust by favoring mass flow over exhaust velocity. This gave the added benefit of slowing down the jet velocity and reducing noise. Today, engines with bypass ratios of 8 produce less than one-quarter the noise of older engines, with bypass ratios on the order of 2. Because higher-bypass ratio engines are the norm, fan noise has become an important aspect of noise reduction.

Fan noise comes from several sources. First, the fan exhaust is now a major noise consideration. The shear layer between the fan and the outer airflow and the fan and the primary jet both result in significant noise. Second, fan blades produce

TABLE 7.5 Stage 2: Noise Requirements

	Takeoff limits (EPNdB)	Sideline limits (EPNdB)	Approach limits (EPNdB)
Up to and including 75,000 lb	93	102	102
Over 75,000 to 600,000 lb	$93 + 5\left\{\dfrac{\log \frac{W}{75{,}000}}{\log 2}\right\}$	$102 + 2\left\{\dfrac{\log \frac{W}{75{,}000}}{\log 2}\right\}$	$102 + 2\left\{\dfrac{\log \frac{W}{75{,}000}}{\log 2}\right\}$
Over 600,000 lb	108	108	108

TABLE 7.6 Stage 3: Takeoff Limits

2-Engine		4-Engine	
Weight	EPNdB	Weight	EPNdB
Up to and including 106,250 lb	89	Up to and including 44,673 lb	89
Over 106,250 to 850,000 lb	$89 + 9\left\{\dfrac{\log \frac{W}{106{,}250}}{\log 2}\right\}$	Over 44,673 to 850,000 lb	$89 + 9\left\{\dfrac{\log \frac{W}{44{,}673}}{\log 2}\right\}$
Over 850,000 lb	101	Over 850,000 lb	106

TABLE 7.7 Sideline and Approach Limits for Two-Engine Aircraft

Sideline limits		Approach limits	
Weight	EPNdB	Weight	EPNdB
Up to and including 77,200 lb	94	Up to and including 77,200 lb	98
Over 77,200 to 882,000 lb	$94 + 2.56\left\{\dfrac{\log \frac{W}{77{,}200}}{\log 2}\right\}$	Over 77,200 to 617,300 lb	$98 + 2.33\left\{\dfrac{\log \frac{W}{77{,}200}}{\log 2}\right\}$
Over 882,000 lb	103	Over 617,300 lb	105

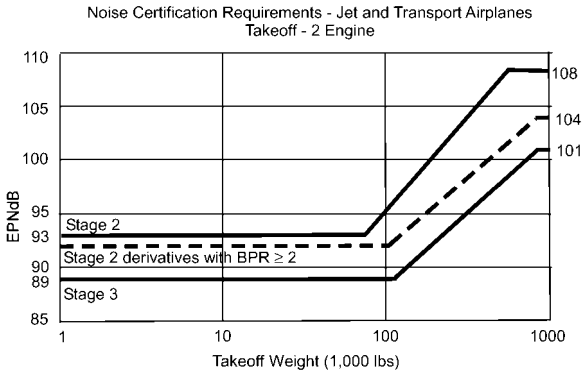


FIGURE 7.28 Two-engine takeoff noise limits.

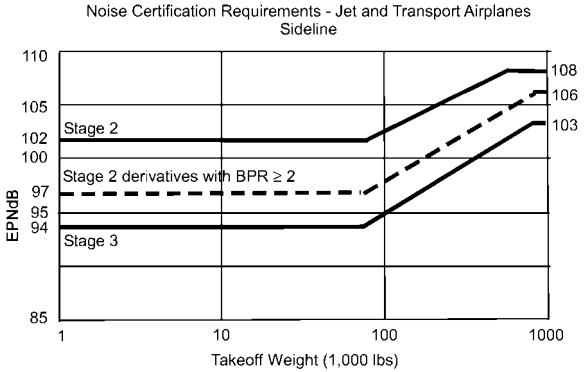


FIGURE 7.29 Sideline noise limits for jet transports.

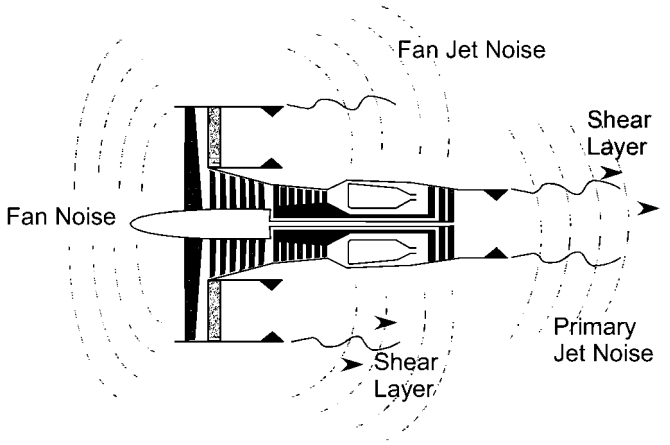


FIGURE 7.30 Where jet engine noise comes from.

noise if turning behind the wake of turning vanes. This will also be true for the compressor and turbine sections. This noise will be of a discrete frequency, relating to the blade passage behind the wake. Third, fan blade tips may be traveling supersonically, generating shock waves. The shock wave noise propagates forward through the inlet and is only heard forward. Because of the frequencies involved, this is usually called "engine buzz," after the noise from a buzz saw.

Analysis of noise in a jet engine involves the solution of linearized pressure waves. A useful tool is to use monopoles, dipoles, and quadrupoles to simplify the solution. Solutions by summing various poles are beyond the scope of this handbook. However, a result of these solutions is that the acoustic power of a subsonic jet is approximately

$$P_j \approx \frac{2\pi\rho_0 u_j^8 D^2}{a_0^5}$$

where subscript j refer to the jet properties (e.g., ρ_j is the density of the gas in the jet), D is the diameter of the jet, subscript o refers to ambient conditions, and a refers to the sound speed. From this result, we see the u^8 relation found by Lighthill. The power of the generated noise goes as the 8th power of the jet velocity. It is clear why it is so important to reduce jet velocity and instead increase mass flow through high bypass fanjets.

The other piece of the solution shows that the noise intensity is largest 45° behind the jet. The noise footprint is shown in Figure 7.31.

The most intense noise from a jet hits the ground 45° after the jet has passed. It continues for a while before it finally weakens.

Noise Suppression

There are many different techniques used for noise suppression. Modern jet engines use some of each of these methods. The technique that has had the largest impact has been to decrease the shear in the jet. The shear is due to high jet velocities. Going to high bypass ratios reduces the shear, resulting in lower noise as well as increased efficiency.

Following the same idea of decreasing noise through decreasing the shear, work is done at increasing the jet mixing with the surrounding air. The objective is to increase the perimeter of the nozzle. Figure 7.32 shows a photo of a corrugated perimeter nozzle. In this example, the lobes help break up the main, noise-producing jet.

Careful placing or removal of inlet guide vanes is used to reduce noise. Most modern fanjets do not have turning vanes in front of the van, so that there is no

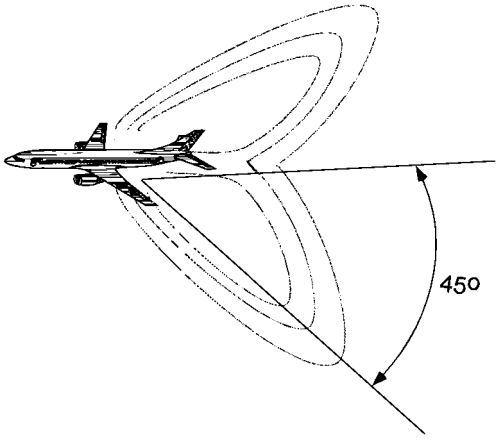


FIGURE 7.31 Noise footprint for typical jet engine.

FIGURE 7.32 Corrugated perimeter nozzle.

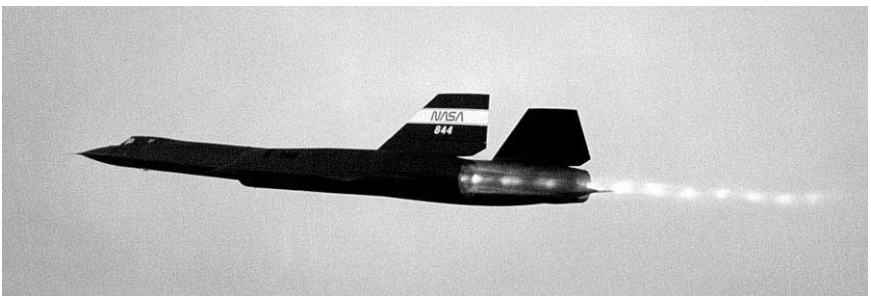


FIGURE 7.33 Shock diamonds in SR-71 exhaust (photo courtesy of NASA).

vane wake-fan interaction. Similarly, the number and spacing of rotor and stator blades in a compressor can be altered to reduce noise.

Noise-attenuating material is used to reduce fan noise. A porous material in the engine cowl can help reduce noise reflected off the inside of the cowl. The same is true for exhaust ducts.

AERONAUTICAL PROPULSION

Supersonic Engines

Supersonic airplanes do not have the option of using high-bypass engines. The frontal area of a large fan produces too much drag. Supersonic military jets typically have bypass ratios of about 2. In addition, shock diamonds in the jet can form, as shown in Figure 7.33. These shock waves can radiate a lot of noise. One solution to the military noise problem is to adapt procedures to reduce noise. For example, climb profiles and airport locations may be changed to reduce the impact of noise.

SECTION 8
ROCKETS AND LAUNCH VEHICLES

Jerry Jon Sellers

Rockets take us where we us need to go in space. They form the core of the propulsion systems found on everything from fireworks to Space Shuttles to the Starship *Enterprise*. Propulsion systems:

- Get us into space,
- Move us around once we get there, and
- Change our attitude (the direction we're pointing)

Once a spacecraft is delivered to space by a launch vehicle, its propulsion subsystem provides the ΔV needed to take it to the final mission orbit and then make orbital corrections and maneuvers throughout the mission lifetime.

Propulsion is also essential for attitude control. These thrusters are either the sole method of attitude control (like the Shuttle's reaction-control system) or they complement the primary system for large slewing maneuvers or to provide a means for momentum dumping. In this chapter we will peel back the mysteries of rocket science to see how rockets work and how propulsion systems are put together for both spacecraft and launch vehicles.

8.1 ROCKET SCIENCE

A rocket is simply a system that takes mass plus energy and converts them into a force to move a vehicle. The input mass for a rocket is generally called *propellant*. The force produced by a rocket we call *thrust*.

Our examination of rocket systems begins by looking at the output—thrust. This requires us to dust off Newton's laws to see how high-speed exhaust going in one direction pushes a vehicle in another. Next we will see how this thrust, over time, produces a velocity change for the vehicle, and—most important for mission planning—how to calculate this effect and ensure we have enough propellant to get our vehicle where we want it to go. We will then turn our attention to the process at the heart of a rocket: how it converts mass plus energy into the high-speed exhaust.

We will tie all these concepts together by looking at the simplest example of a rocket—cold-gas thrusters—to see how varying some of the inputs and design variables changes their thrust and overall system efficiency.

Thrust

A rocket ejects mass at high speed in one direction so a vehicle can go in the other. The simplest example of this is a balloon. All of us have blown up a toy balloon and let go of the stem to watch it fly wildly around the room. What makes the balloon go? Recall Newton's third law: *For every action there is an equal but opposite reaction.*

When you blow into a balloon, you force air into it, making the rubber skin tighten, increasing the internal air pressure, and storing mechanical energy like a spring. When you let go of the stem, the air pressure has an escape route, so the skin releases, forcing the air out under pressure. Following Newton's third law, as the air, which has mass, is forced out in one direction (the action), an equal force pushes the balloon in the opposite direction (the reaction).

Let's look at this action/reaction situation in a bit more detail to see where the force comes from. Consider an astronaut perched in a wagon armed with a load of rocks. If he is initially at rest and begins to throw the rocks in one direction, because of Newton's third law, an equal but opposite force will move him (and the wagon load of rocks) in the opposite direction.

To throw the rocks, the astronaut has to apply a force to them. This force is identical in magnitude, but opposite in direction, to the force applied to the astronaut and thus, the wagon. However, remember the concept of conservation of linear momentum. It tells us the change in speed of the rock (because it has less mass) will be greater than the change in speed of the wagon.

We will say the rocks leave the wagon at a rate called the *mass flow rate*, $\Delta m/\Delta t = \dot{m}$ (kg/s). Recall that linear momentum is mass times velocity. If the exit speed of the rocks is V_{exit}, the ejected mass has a momentum of $\dot{m}V_{exit}$. But remember, momentum is always conserved! So as the momentum of the rocks goes in one direction, the momentum of the wagon goes in the other direction. This is the basic principle that produces rocket thrust. A rocket expends energy to eject mass out one end at high velocity, pushing it (and the attached vehicle) in the opposite direction.

Dropping the vector notation to look at magnitudes only, we can say

$$\dot{P}_{rocket} = \dot{P}_{exhaust}$$
$$\dot{P}_{rocket} = \dot{m}V_{exit} \tag{8.1}$$

where $\dot{P}_{rocket}$ = time rate of change of the rocket's momentum (N)
$\dot{P}_{exhaust}$ = time rate of change of the exhausted mass' momentum (N)
$\dot{m}$ = mass flow rate of the exhaust products (kg/s)
V_{exit} = exit velocity of the exhaust (m/s)

Notice this momentum change has the same units as force. This is the force on the rocket we defined to be the thrust. As we will see later, depending on the type of rocket used, the effective thrust delivered may be slightly different than $\dot{m}V_{exit}$. For this reason, we define a more comprehensive term called *effective exhaust velocity, C*, so we can express the rocket thrust as:

$$F_{\text{thrust}} = \dot{m}C \tag{8.2a}$$

where F_{thrust} = rocket's total thrust (N)
C = effective exhaust velocity (m/s)
$\dot{m}$ = mass flow rate (kg/s)

This relationship should make sense from our astronaut example. He can increase the thrust on the wagon either by increasing the rate at which he throws the rocks (higher $\dot{m}$) or by throwing the rocks faster (higher C). Or he can do both. For example, if he were throwing bowling balls he could achieve high $\dot{m}$ but with lower velocity than if he were throwing small pebbles.

Of course, exhaust velocities for typical rockets are much, much higher than anyone can achieve by throwing rocks. For typical chemical rockets, like the ones used by the Space Shuttle, the exhaust velocity can be as high as 3 km/s. Because these high velocities are hard to visualize, it is useful to think about the raw power involved in a rocket engine. Recall that kinetic energy is:

$$KE = \tfrac{1}{2}mV^2$$

We define *power* as energy expended per unit time. Thus, the power in the jet exhaust of a rocket is

$$P_J = \tfrac{1}{2}\dot{m}C^2 \tag{8.2b}$$

where P_J = Jet power in a rocket ($J/s = W$)

At lift-off, the Space Shuttle's three main engines plus its solid rocket boosters produce 26.6 billion watts of power. That is equivalent to over 13 Hoover Dams! We will see the effect of all that power next.

The Rocket Equation

To understand better how we use the thrust produced by rockets to get a vehicle where we want it to go, we must first introduce a new concept—impulse. Impulse will help us understand the total velocity change rockets deliver.

Impulse. So a rocket produces thrust that pushes on a vehicle. Then what happens? If you push on a door, it opens. If you hit a ball with a bat, it flies out to left field. Returning to our astronaut in the wagon, realize that to give the rocks their velocity, he has to apply a force to them over some length of time. Force applied to an object over time produces an *impulse*. Dropping vector notation, we express this law as

$$F = \frac{\Delta p}{\Delta t}$$

If we multiply both sides by Δt, we get

$$F\Delta t = \Delta p$$

The left side of this equation represents a force, F, such as the force we would generate if we hit a baseball with a bat, over some time, Δt. When our bat hits that

fastball, it seems like the impact is instantaneous, but the bat actually stays in contact with the ball for a fraction of a second, applying its force to the ball during that time. The result of that force acting over time is seen on the right side of the equation, where Δp represents the resulting change in momentum.

From this relationship, realize that to change momentum you can either apply a large force acting over a short time (like a bat hitting a ball) or a smaller force acting over a longer time (like an ant slowly moving a bread crumb). We define total impulse, I, to be force times time, or change in momentum

$$I \equiv F\Delta t = \Delta p \tag{8.3}$$

where I = total impulse (N/s)
F = force (N)
Δt = time (s)
Δp = momentum change (N/s)

Impulse works the same way for rockets as it does for baseballs. We want to change the velocity and hence the momentum of our rocket, so we must apply some impulse. This impulse comes from the rocket thrust acting over some time interval. But as we saw, we can produce the same impulse from a rocket by applying a small thrust over a long time or a large thrust over a short time.

Although total impulse is useful for telling us the total effect of rocket thrust, it does not give us much insight into the rocket's efficiency. To compare the performance of different types of rockets, we need a new parameter we call *specific impulse.* Specific impulse, I_{sp}, is the ratio of the total impulse to the propellant weight required to produce that impulse (how much "bang" you get for your buck).

$$I_{sp} \equiv \frac{I}{\Delta W} \tag{8.4}$$

where I_{sp} = specific impulse (s)
I = total impulse (N/s)
ΔW = change in propellant weight (N) (using weight rather than mass to calculate I_{sp} is simply a convention established by the founders of rocket science and, as a result, I_{sp} has the unusual units of seconds.)
Substituting for total impulse

$$I_{sp} = \frac{F_{thrust}\Delta t}{\Delta W}$$

$$I_{sp} = \frac{F_{thrust}}{\dot{W}} \tag{8.5}$$

where I_{sp} = specific impulse (s)
F_{thrust} = force of thrust (N)
$\dot{W}$ = propellant weight flow rate (N/s)

I_{sp} represents the ratio of what we receive (momentum change) to what we spend (propellant). So the higher the I_{sp}, the more efficient the rocket. To get a more useful expression for I_{sp}, we replace the weight flow rate with the mass flow rate, $\dot{m}$, times the gravitational acceleration at sea level (a constant).

ROCKETS AND LAUNCH VEHICLES

$$I_{sp} = \frac{F_{thrust}}{\dot{m} g_o} \tag{8.6}$$

where $\dot{m}$ = mass flow rate
g_o = gravitational acceleration constant = 9.81 m/s^2 (sea level)

Earlier, we found the force of thrust in terms of the mass flow rate and the effective exhaust velocity. By substituting equation (8.2) into equation (8.6), we get another useful expression for I_{sp}.

$$I_{sp} = \frac{C}{g_o} \tag{8.7}$$

where C = effective exhaust velocity (m/s)

Notice g_o is a constant value representing the acceleration due to gravity at sea level. This means no matter where we go in the universe, we humans will use the same value of g_o to measure engine performance.

As a measure of engine performance, I_{sp} is the same as the miles per gallon (mpg) rating given for cars. The higher the I_{sp} is for a rocket, the more ΔV it will deliver for a given mass of propellant. Another way to think about I_{sp} is that the faster a propulsion system can expel propellant, the more efficient it is.

Realize that we express I_{sp} in terms of propellant mass. For some space missions, especially those involving small satellites, conserving volume can be just as important, or even more important, than conserving mass. Therefore, comparing only the I_{sp} of two systems may not tell us the whole story. For this reason, we define another useful term called *density specific impulse,* I_{dsp}, which we find by multiplying the rocket's I_{sp} times the average density of the propellants, δ_{av}.

$$I_{dsp} = \delta_{av} I_{sp} \tag{8.8}$$

where δ_{av} = average density of propellants (kg/m^3)

By comparing the mass and volume between different system options, mission planners can do more realistic tradeoffs.

Velocity Change. When you take a long trip in your car, you have to make sure you will have enough gas in the tank to get there. This concern is equally important for a trip into space. But how do we determine how much propellant, we need for a given mission?

Naturally, some rockets are more efficient than others. For example, one rocket may need 100 kg of propellant to change velocity by 100 m/s while another may need only 50 kg. To figure out how much propellant we need for a given trip, we need a relationship between the velocity change and the amount of propellant used. We find this relationship by setting thrust equal to momentum change.

$$F_{thrust} = \dot{m} c = \frac{\Delta p_{rocket}}{\Delta t}$$

where F_{thrust} = effective thrust from the rocket (N)
$\dot{m}$ = propellant mass flow rate (kg/s)
c = effective exhaust velocity (m/s)

$\frac{\Delta p_{\text{rocket}}}{\Delta t}$ = the rocket's time rate of change of momentum (N)

From this relationship we can derive the *ideal rocket equation*. It tells us how much propellant we need for a given ΔV.

$$\Delta V = C \ln \left(\frac{m_{\text{initial}}}{m_{\text{final}}} \right) \tag{8.9}$$

where ΔV = velocity change (m/s)
C = effective exhaust velocity (m/s)
ln = natural logarithm of the quantity in the parenthesis
m_{initial} = vehicle's initial mass, before firing the engine (kg)
m_{final} = vehicle's final mass, after firing the engine (kg)

Equation (8.9) is one of the most useful relationships of rocket propulsion. Armed with this equation, we can determine how much propellant we need to do anything, from stopping the spin of a spacecraft in orbit, to launching a probe to another solar system. Notice that we are taking the natural logarithm of the ratio of initial to final mass. *The difference between initial and final mass represents the amount of propellant used.* ΔV is also a function of the effective exhaust velocity. This relationship should make sense because, as the propellant leaves the rocket faster, momentum changes more, and the rocket goes faster.

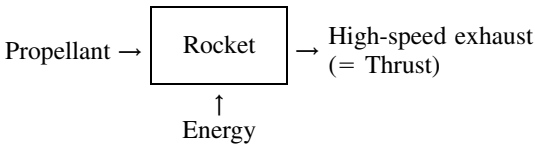
We can substitute the definition of I_{sp} into the rocket equation (8.9) to compute the ΔV for a rocket if we know the I_{sp}, as well as the initial and final rocket mass.

$$\Delta V = I_{\text{sp}} g_o \ln \left(\frac{m_{\text{initial}}}{m_{\text{final}}} \right) \tag{8.10}$$

where I_{sp} = propellant's specific impulse (s)
g_o = gravitational acceleration at sea level (9.81 m/s^2)

Rockets. Now that we've seen what rockets do—expel high-speed exhaust in one direction so a space vehicle can go in the other—let's take a closer look at how they do it. For purposes of discussion, we can break this process into two steps. First, energy must be *transferred* to the propellant in some form. Second, the energized propellant must be *converted* into high-speed exhaust. There are only two basic types of rockets currently in use. Their classification depends on the form of energy that is transferred to the propellant and converted to high speed exhaust. These are:

- Thermodynamic rockets—rely on thermodynamic energy (heat and pressure)
- Electrodynamic rockets—rely on electrodynamic energy (charge, electric and magnetic fields)



Thermodynamic energy is in the form of heat and pressure. We are all familiar

with this. A covered pot of water on the stove reaches high temperature and produces high-pressure steam. We have all seen how the thermodynamic energy in steam can be used to drive trains or in power plants to produce electricity. In a *thermodynamic rocket*, thermodynamic energy is transferred to the propellant in the form of heat and pressure. Heat is either liberated from the propellant by a chemical reaction or supplied directly from electrical, solar, or nuclear sources. Gaseous or liquid propellants are delivered to the rocket under pressure, supplying additional thermodynamic energy. However, for now the result is the most the important thing. Once energy is transferred to the propellant, we end up with a high-temperature, high-pressure gas, in other words, a gas with lots of thermodynamic energy. Air in a toy balloon or the result of burning liquid hydrogen and liquid oxygen inside the Shuttle main engines are two extreme examples.

Later in this section we will look at the simplest type of rocket, a cold gas thruster, that relies on gas under pressure alone as its source of thermodynamic energy. In subsection 8.2 we will look at other, more complex and efficient types of thermodynamic rockets.

Electrodynamic rockets rely on electrodynamic energy. *Electrodynamic energy* relates to the energy available from charged particles moving in electric and magnetic fields.To understand this, we need to use the concept of *charge*.

Charge is a fundamental property of matter, like mass, and can be either positive or negative. Like charges repel each other and opposite charges attract. Typically, a molecule of propellant has the same number of protons and electrons, making it electrically neutral. However, if one or more electrons can be "stripped off," the resulting molecule will be left with a net positive charge, making it an *ion*, as shown in Figure 8.1. To create the ion, some electrodynamic energy must be supplied by the electrical power subsystem (EPS). In a thermodynamic rocket the inherent energy of the energized propellant is quite high; however, the inherent energy of an ion is relatively low. However, once a particle is charged, it can be easily accelerated to very high velocities using additional electrodynamic energy. Later in this section we will look at how this is done inside an electrodynamic rocket.

The form of energy transferred to the propellant determines how it can be converted to high-speed exhaust. In the rest of this section we will look at the two ways of doing this:

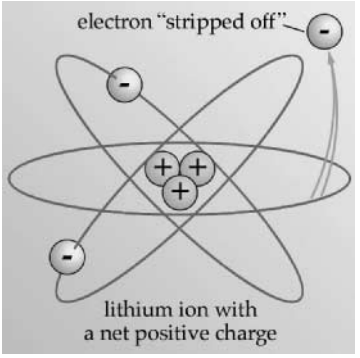


FIGURE 8.1 The ion.

- Thermodynamic expansion—using nozzles
- Electrodynamic acceleration—using electric and magnetic fields

Thermodynamic Expansion—Nozzles. By far the most commonly used types of rockets rely on nozzles. *Nozzles* convert the thermal energy produced by chemical, nuclear or electrical sources into kinetic energy through thermodynamic expansion. To understand how nozzles do this, we must first understand fluid mechanics. Once we understand how fluids behave, we can look at nozzles to see how they convert low-speed, high-temperature gases into high-velocity exhaust. We will then look at how we can predict and measure the performance of thermodynamic rockets and look at a simple example of a cold-gas rocket to see how all these principles fit together.

Fluid Mechanics. Let's start by looking at one of the simplest examples, the air in a balloon. Assuming the air in the balloon behaves as a perfect gas, we can relate the pressure, density, and temperature of the gas using the perfect-gas law:

$$P = \rho RT \tag{8.11}$$

where P = pressure (N/m^2)
ρ = density (kg/m^3)
T = temperature (K)
R = specific gas constant (J/kgK)
= Ru/M
Ru = universal gas constant (= 8314.41 J/kmol K)
M = molecular weight of the gas (kg/kmole)

It is important to review some basic assumptions that will make the our discussion valid. A perfect gas flow has the following properties:

- No heat transfer into/out of the fluid—this is known as an *adiabatic flow*.
- The flow is reversible, meaning total energy is conserved. Together, these two assumptions are called *isentropic flow*.
- Flow in one dimension—two and three dimensions get more complicated
- "Frozen flow"—meaning all chemical reactions are completed inside the combustion chamber.
- Steady flow—meaning mass flow rate, energy, and momentum are constant.

Earlier we discussed mechanisms for transferring energy to the propellant. In most cases the propellant arrives in the combustion chamber under pressure, so it already has some mechanical energy. As we said earlier, the most common means of adding energy is through heat, such as that produced in chemical reactions (combustion). We describe the total energy in fluid systems in terms of their *specific enthalpy, h.*

$$h = u + Pv \tag{8.12}$$

where h = specific enthalpy (J/kg)
u = internal energy (J/kg)
P = pressure (N/m^2)
v = specific volume (m^3/kg)

Computing enthalpy allows us to separate the energy in the exhaust due to the

internal energy (heat) from the mechanical energy (pressure). To see the usefulness of this concept, we need to put the fluid in motion. Let's connect our balloon to a pipe and watch the gas flow as shown in Figure 8.2.

Assuming mass flow rate is constant, we can compute that rate by multiplying the density of the exhaust by the flow velocity and the cross-sectional area of the pipe:

$$\dot{m} = \rho V A \tag{8.13}$$

where $\dot{m}$ = fluid's mass flow rate (kg/s)
ρ = fluid's density (kg/m^3)
V = velocity of flow (m/s)
A = pipe's cross-sectional area (m^2)

Now what happens when we vary the pipe's cross-sectional area? If we reduce the pipe's area, the flow velocity *increases* to maintain a constant mass flow rate. This increase in flow velocity to maintain constant mass flow rate due to a constriction in area is called the *Venturi effect*. Looking at it the other way, if we *increase* the pipe's cross-sectional area, the flow velocity decreases to maintain the same mass flow rate. This is illustrated in Figure 8.3.

These effects are common ones, so they are fairly intuitive. However, it only works for low speed flows. For very high-speed flows, the opposite effect takes place. As the area increases, the flow speeds up. For the case of steady flow, with no heat transfer to or from the system, the sum of the specific enthalpy, h, and one-half the square of the flow velocity, $\frac{1}{2}V^2$, is constant.

$$h + \tfrac{1}{2}V^2 = \text{constant} \tag{8.14}$$

where

$$h = u + Pv$$

Note that the second term is the same as the specific kinetic energy (independent of the mass) of the flow. Thus, the internal energy and gas pressure can be traded for kinetic energy. In other words, the *velocity can increase at the expense of*

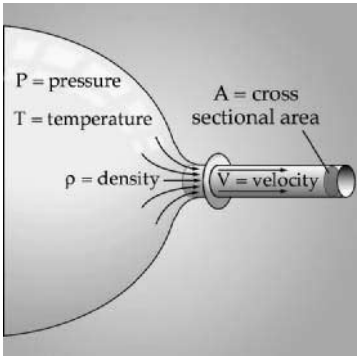


FIGURE 8.2 Fluid in motion.

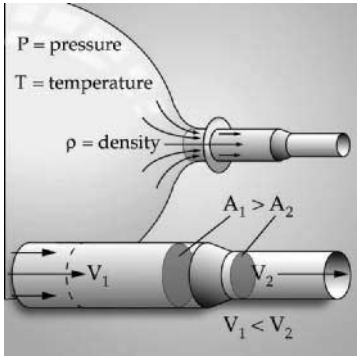


FIGURE 8.3 The Venturi effect.

enthalpy and vice versa. This relationship is the *Bernoulli principle*, named after its discoverer, Swiss mathematician Daniel Bernoulli (1700–1782). This is one of the most important concepts in science. It helps us explain the dynamics of weather and how birds and planes fly.

But this still does not immediately explain how the flow speed increases when the cross-sectional area expands for very high-speed flow. To understand this, we need to delve a little deeper into the behavior of high-speed gases by looking at the speed of sound.

In 1947, Chuck Yeager, piloting the Bell X-1 rocket plane, became the first person to break the sound barrier by traveling faster than the speed of sound. The speed of sound represents the velocity at which pressure disturbances move through a medium. In other words, if this book falls on the floor, it will create a pressure disturbance (sound) that will travel out from the source at a specific speed that depends on characteristics of the air in the room. You will hear the sound a fraction of a second before someone on the other side of the room hears it. For short distances like these, it seems almost instantaneous. But if you have ever witnessed a thunderstorm, you have seen flashes of lightning in the distance and heard the thunder a few seconds later. By counting the seconds between seeing the lightning and hearing the thunder, and by knowing the speed of sound, you can get a good estimate of the distance to the storm. We can find the speed of sound, a_o, from:

$$a_o = \sqrt{\gamma RT} \tag{8.15}$$

where a_o = speed of sound (m/s)
γ = ratio of specific heats (dimensionless)
R = specific gas constant (J/kgK)
T = temperature (K)

Notice in equation (8.15) we introduce the ratio of specific heats, γ. This parameter is constant for a particular gas (or gas mixture) at a given temperature and pressure. We can compute it using various gas-modeling techniques that are beyond the scope of our discussion here, or we can measure it experimentally. As we will see later, γ is an extremely useful parameter for calculating rocket efficiency and other characteristics. Applying equation (8.15) to air at 20°C, we get a speed of sound of 346 m/s.

The ratio of the velocity of the flow (or the velocity of the Bell X-1) to the speed of sound is called the Mach number, M_a.

$$M_a = V/a_o \tag{8.16}$$

From equations (8.14) and (8.16), we can derive a relationship between Mach number, M_a, the change in area, dA, and change in flow velocity, dV.

$$dA/A = (M_a^2 - 1)dV/V \tag{8.17}$$

where M_a = Mach number
dA = infinitesimal area change (m^2)
A = pipe's cross-sectional area (m^2)
dV = infinitesimal velocity change (m/s)
V = velocity of the flow (m/s)

Looking closely at this relationship, we can see that for subsonic flow ($M_a < 1$) a positive change in area (expansion) leads to a negative change in velocity (de-

ROCKETS AND LAUNCH VEHICLES

crease). However, for supersonic flow ($M_a > 1$) the opposite must occur, a positive area change (expansion) leads to a positive velocity change (increase). Figure 8.4 illustrates these relationships. This flip-flop in relationships between area and velocity change above and below the speed of sound is due to conservation of energy (the trade-off between enthalpy and velocity) and conservation of mass (constant flow rate).

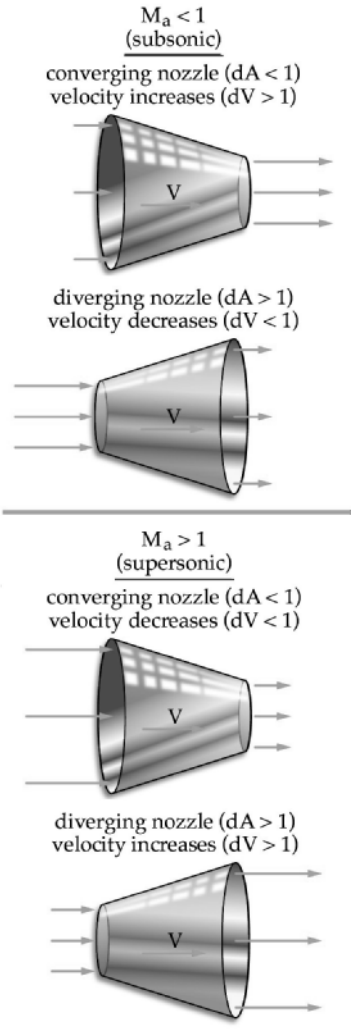


FIGURE 8.4 Changing Mach number versus changing area. For subsonic flow ($M_a < 1$) the flow velocity increases when area decreases. For supersonic flow ($M_a > 1$) the flow speed increases when area increases.

Without this effect, rockets as we know them would not be possible. We take advantage of this principle to convert the enthalpy of gases, in the form of heat and pressure in the combustion chamber, into kinetic energy using nozzles. Just as with our balloon example, the process begins in the combustion chamber where high-pressure (and usually high-temperature) exhaust products are created with little or no velocity ($M_a << 1$). From the combustion chamber, the products flow first into a converging section of the nozzle where we know the velocity of the flow increases. The narrowest portion of the nozzle is called the throat. In the throat, the flow velocity reaches the speed of sound ($M_a = 1$) and we say the flow is "choked." As the nozzle expands beyond the throat the velocity again increases to supersonic before being exhausted at the nozzle exit. Figure 8.5 shows a simplified combustion chamber and nozzle diagram.

Nozzles. The more we expand the exhaust, the higher the velocity. But there are practical limits. Equation (8.17) showed us that the increase in velocity comes at the expense of the enthalpy of the gases. Recall, enthalpy is a measure of the gases' internal energy (heat) plus mechanical energy (pressure). Therefore, as the gases gain velocity through expansion in the nozzle, they lose enthalpy, both temperature and pressure. Theoretically, we would need an infinitely long nozzle to expand the exhaust to zero exit pressure (vacuum). In practice, of course, this is not possible. Instead, rocket scientists design nozzles that are long enough for the conditions in which they operate. As we will see next, the most important condition to consider is the outside air pressure.

Earlier, we used our astronaut in the wagon example to show that rocket thrust equals the mass flow rate, $\dot{m}$, times the effective exhaust velocity, C.

$$F = \dot{m}C \tag{8.18}$$

However, this represents only the thrust produced from momentum change, called the *momentum thrust*. But for rockets using nozzles to convert thermal energy into kinetic energy, the thrust due to the momentum of the exhaust is only part of the story.

As we learned, unless a nozzle is infinitely long, the exhaust will have some exit pressure, P_{exit}. This pressure also contributes to the rocket's thrust. To see this, consider an imaginary "control volume" drawn around a rocket as shown in Figure 8.6. Acting on the boundaries of this volume, we've drawn the atmospheric pressure

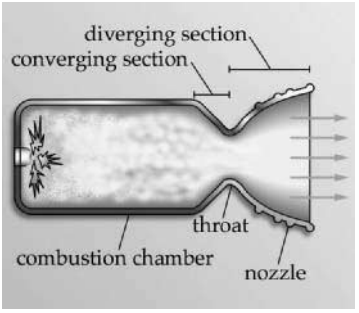


FIGURE 8.5 Standard combustion chamber and nozzle configuration.

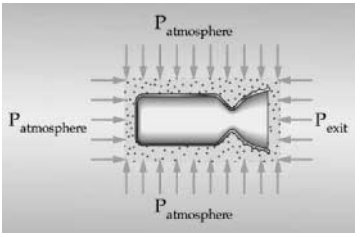


FIGURE 8.6 Pressure thrust.

$P_{atmosphere}$ acting on all sides except at the nozzle exit. At the nozzle exit, the pressure is P_{exit}, drawn inward for consistency. Notice that due to symmetry, $P_{atmosphere}$ cancels everywhere except in the direction parallel to momentum thrust over an area equal to the nozzle exit area, A_{exit}. The net force exerted on the rocket from this pressure differential is called the *pressure thrust*. It equals the difference between exit pressure, P_{exit}, and atmospheric pressure, $P_{atmosphere}$, times the exit area, A_{exit}. Its magnitude is

$$F_{\text{pressure thrust}} = A_{exit}(P_{exit} - P_{atmosphere}) \tag{8.19}$$

where $F_{\text{pressure thrust}}$ = pressure thrust (N)
A_{exit} = nozzle's exit area (m^2)
P_{exit} = exit pressure (N/m^2)
$P_{atmosphere}$ = atmospheric pressure (N/m^2)

By adding together momentum thrust and pressure thrust, we express the magnitude of the total thrust on the rocket, using the total rocket thrust equation

$$F_{thrust} = \dot{m}V_{exit} + A_{exit}(P_{exit} - P_{atmosphere}) \tag{8.20}$$

where F_{thrust} = rocket's total thrust (N)
V_{exit} = exit velocity of exhaust products (m/s)

To further simplify this equation, we can now fully define the *effective exhaust velocity, C*, to be

$$C = V_{exit} + \frac{A_{exit}}{\dot{m}}(P_{exit} - P_{atmosphere}) \tag{8.21}$$

So we are back to where we started with

$$F = \dot{m}C$$

Only now we know a lot more about where C comes from and how nozzle expansion effects it.

A casual glance at the total rocket thrust in equation (8.2) may lead us to conclude we'd want to make $P_{exit} >> P_{atmosphere}$ to maximize total thrust. Although this would appear to increase the amount of thrust generated, a big loss in overall efficiency would actually reduce the effective thrust. Recall that for supersonic flow, as the gases expand they *increase* in velocity while, due to the Bernoulli principle, they decrease in pressure. Thus, the higher the V_{exit}, the lower the P_{exit}. For the ideal case, the pressure thrust should be zero ($P_{exit} - P_{atmosphere} = 0$), which means the exit pressure exactly equals the atmospheric pressure ($P_{exit} = P_{atmosphere}$). In this case the exit velocity and thus the momentum thrust is maximized.

But what happens when $P_{exit} \neq P_{atmosphere}$? When this happens, we have a rocket that is not as efficient as it could be. We can consider two possible situations:

- *Over expansion*: $P_{exit} < P_{atmosphere}$. This is often the case for a rocket at lift-off. Because most launch pads are near sea level, the atmospheric pressure is at a maximum. This atmospheric pressure can cause shock waves to form just inside the nozzle. These shock waves represent areas where kinetic energy turns back into enthalpy (heat and pressure). In other words, they rob kinetic energy from the flow, lowering the exhaust velocity and thus decreasing the overall thrust.

- *Under expansion*: $P_{exit} > P_{atmosphere}$. In this case, the exhaust gases have not expanded as much as they could have within the nozzle and thus there is a loss in the sense that we have not converted all the enthalpy we could have into velocity. This is the normal case for a rocket operating in a vacuum, because P_{exit} is always higher than $P_{atmosphere}$ ($P_{atmosphere} = 0$ in vacuum). Unfortunately, you would need an infinitely long nozzle to expand the flow to zero pressure, so in practice we must accept some loss in efficiency.

In Section 8.3 we will see how we deal with this problem for launch-vehicle rocket engines. Figure 8.7 illustrates these cases of expansion.

The total expansion in the nozzle depends, of course, on its design. We define the nozzle *expansion ratio,* ε, to be the ratio between the nozzle exit area, A_e, and the throat area, A_t:

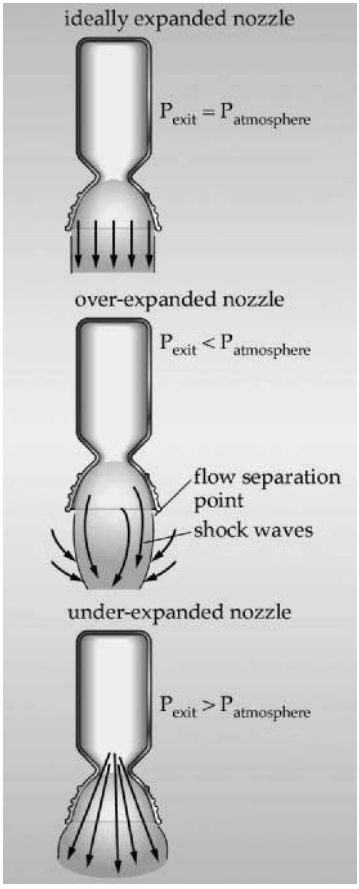


FIGURE 8.7 Nozzle expansion.

$$\varepsilon = A_e/A_t \tag{8.22}$$

Later in this section we will see how varying expansion ratio can effect engine performance.

Characteristic Exhaust Velocity. We know rocket thrust depends on effective exhaust velocity, C, the rocket's output. But how do you measure C? Unfortunately, you can't just stick a velocity transducer into superheated rocket exhaust. Therefore, when doing rocket experiments, we need to have some other, more measurable parameters available. We know we can vary rocket performance (thrust, I_{sp}, etc.) by changing the inputs: $\dot{m}$, of the propellant going into the combustion chamber, and the resulting pressure in the combustion chamber, P_c. In addition to these dynamic parameters, there are also important physical dimensions of the rocket that can be varied, such as the area of the nozzle throat, A_t, and the expansion ratio, ε. To understand the relationship between these design variables, we define another important rocket performance parameter called *characteristic exhaust velocity*, C^*, in terms of chamber pressure, P_c, mass flow rate, $\dot{m}$, and throat area, A_t.

$$C^* = P_c A_t/\dot{m} \tag{8.23}$$

The nice thing about C^* is that not only can we easily measure it experimentally, we can also compute it by modeling combustion and flow characteristics. Rocket scientists use a variety of computer codes to predict characteristic exhaust velocity. These techniques are beyond the scope of the discussion here, but for a complete description see Humble et al. (1995). One of the most important parameters to find using these modeling techniques is the ratio of specific heats, γ, for the products in the combustion chamber. Knowing this value, we can compute C^* using:

$$C^* = a_o/\Gamma \tag{8.24}$$

where a_o = speed of sound in the fluid (m/s)
$\Gamma = \gamma(2/(\gamma + 1))^{[(\gamma+1)/(2\gamma-2)]}$ (unitless)

By comparing C^* from experiments with C^* from our predictions, we can determine how efficiently the rocket transfers the available energy to the propellants relative to an ideal value. This is especially useful as it allows us to measure the performance of the combustion chamber (energy transfer) independent from the nozzle (energy conversion). No rocket is perfect. If the measured C^* is over 90% of predicted ideal C^*, it is considered good performance. As we learned earlier, absolute performance is measured by I_{sp} and I_{dsp}, both of which are a function of effective exhaust velocity, C. While it is fairly straight forward to determine I_{sp} experimentally by measuring delivered thrust and the total mass of propellant used, it is also important to be able to predict it for a given type of rocket. Fortunately, we can compute it directly from C^* and the ratio of specific heats for a given reaction.

As one would expect, one of the primary goals of rocket design is to maximize performance. We express rocket performance most often in terms of mass efficiency (I_{sp}). How do we maximize I_{sp}? We have just seen that effective exhaust velocity, and hence specific and density specific impulse are all functions of C^*. From equation (8.24), we know C^* depends on the speed of sound in the combustion chamber, a_o, which depends on the gas constant, R, and the combustion temperature, T. Looking at the equation, we can see that the higher the temperature, the higher the speed of sound and thus the higher the I_{sp}. What about R? Remember, R is

inversely proportional to the propellant's molecular weight. *Molecular weight* is a measure of the weight per molecule of propellant. Thus, to improve I_{sp} for thermodynamic rockets, we try to maximize the combustion temperature while minimizing the molecular weight of propellant. We can express this relationship more compactly as:

$$I_{sp} \propto \sqrt{\frac{T_{combustion}}{M}} \tag{8.25}$$

where I_{sp} = specific impulse (s)
$T_{combustion}$ = combustion temperature (K)
M = molecular weight (kg/mole)
[*Note:* the symbol "$\propto$" means proportional to]

As a result, the most efficient thermodynamic systems operate at the highest temperature with the propellants having the lowest molecular weight. For this reason, hydrogen is often the fuel of choice because it has the lowest possible molecular weight and achieves high temperatures during combustion. Unfortunately, the low molecular weight also means low density. Thus, while hydrogen systems achieve high I_{sp}, they often do it at the expense of I_{dsp}. Designers must trade off mass versus volume efficiency depending on the mission requirements.

Finally, we need to know what total thrust our rocket produces. We can relate characteristic exhaust velocity, C^*, to effective exhaust velocity, C, through yet another parameter called the *thrust coefficient*, C_F.

$$C_F = C/C^* \tag{8.26}$$

We can also relate the thrust coefficient to the thrust, F, chamber pressure, P_c, and throat area, A_t, through

$$C_F = F/(P_c A_t) \tag{8.27}$$

In this way, we can compare the measured rocket thrust to the ideal thrust from theoretical modeling to determine how well the nozzle converts enthalpy into kinetic energy. Like characteristic exhaust velocity, thrust coefficient gives us a way to determine the performance of the nozzle, independent from the combustion chamber. Again, no nozzle is perfect. However, a well-designed nozzle, for the correct expansion conditions, should achieve 95% or better.

Summary. Let's review what we have discussed about thermodynamic rockets. Figure 8.8 further expands our systems view of a thermodynamic rocket and summarizes important performance parameters. Recall, there are two important steps to the rocket propulsion process—energy transfer and mass acceleration. These two steps take place in the combustion chamber and nozzle, respectively. The most important output is the thrust that moves the vehicle from point *A* to point *B*.

Now let's put all these principles together by looking at one specific example—the simplest type of rocket in use, a cold-gas thruster.

Cold-Gas Rockets. A *cold-gas rocket* uses only mechanical energy in the form of pressurized propellant as its energy source, similar to the toy balloon example we talked about at the beginning of the chapter. While spacecraft designers do not send balloons into orbit, the basic principles of cold-gas rockets are not that different. A coiled spring stores mechanical energy that can be converted to work, such as running an old-fashioned wind-up watch. Similarly, any fluid under pressure

ROCKETS AND LAUNCH VEHICLES

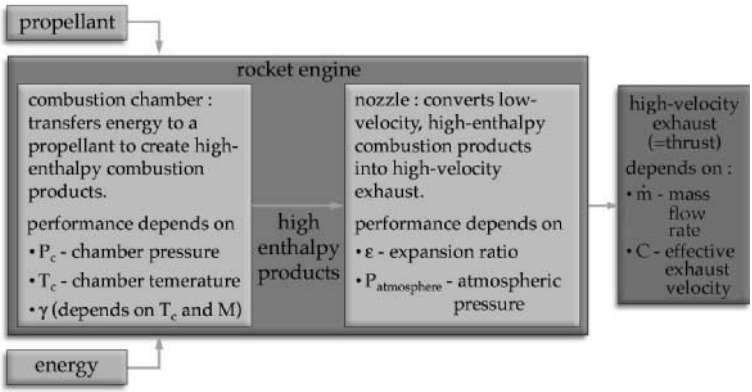


FIGURE 8.8 Expanded systems view of a thermodynamic rocket.

TABLE 8.1 Summary of Cold-Gas Rockets

Operating principle	Uses the mechanical energy contained in a compressed gas and thermodynamically expands the gas through a nozzle producing high-velocity exhaust
Propellants	Helium (He), nitrogen (N_2), carbon dioxide (CO_2), or virtually any compressed gas
Advantages	• Extremely simple • Reliable • Safe, low-temperature operation • Short impulse bit (thrust pulses)
Disadvantages	Low I_{sp}, I_{dsp} compared to chemical rockets
Example	UoSAT-12 cold-gas thrusters: Propellant = N_2, Pc = 4 bar, thrust = 0.1 N

has stored mechanical energy that can be used to do work. Any rocket system containing fluids under pressure (and virtually all do) uses this mechanical energy in some way. As we will see, usually this energy is a minor contribution to the overall energy of the propellant. However, for cold-gas rockets this is the only energy the propellant has.

Table 8.1 summarizes basic principles and propellants used by cold-gas rockets. There is a small volume upstream of the nozzle where the gas collects prior to expulsion. This is similar to the combustion chamber needed for chemical rockets.

Cold-gas rockets are very reliable and can be turned on and off repeatedly, producing very small, finely controlled thrust pulses (also called impulse bits)—a desirable characteristic for attitude control. A good example of them is on the manned maneuvering unit (MMU) used by Shuttle astronauts. The MMU uses compressed nitrogen and numerous small thrusters to give astronauts complete freedom to maneuver.

Unfortunately, due to their relatively low thrust and I_{sp}, we typically use cold-gas systems only for attitude control or limited orbital maneuvering on small space-

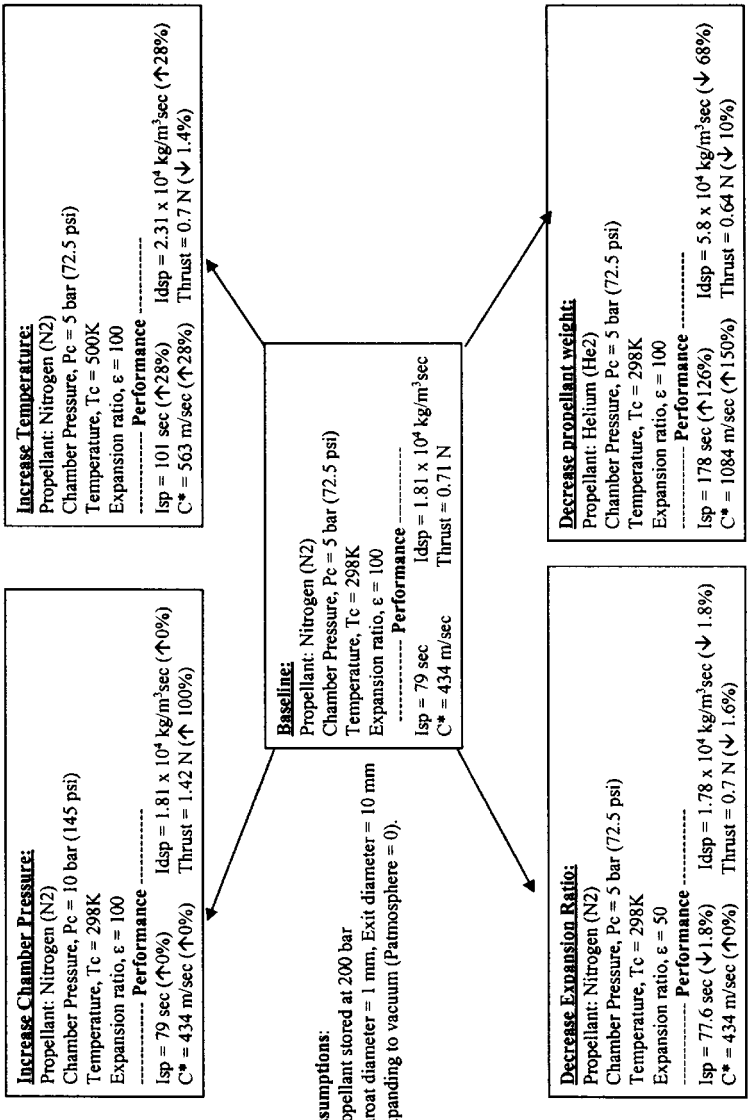


FIGURE 8.9 Cold-gas rocket trade-offs. This figure illustrates various trade-offs in rocket design by looking at the simplest type of rocket, a cold-gas thruster. From the baseline case using nitrogen as the propellant, the specific impulse, density specific impulse, characteristic exhaust velocity, C^*, and thrust are affected by changing chamber pressure, chamber temperature, expansion ratio, and the molecular weight of propellant. (Analysis courtesy of Johnson Rocket Company, Inc.)

craft. Even so, they can serve as a good example of trading off some of the basic rocket parameters we have talked about in this section.

Figure 8.9 presents the results of analyzing five variations of the same basic cold-gas rocket. From the baseline design using nitrogen as the propellant at room temperature (298K) and 5 bar (72.5 psi) chamber pressure, the figure shows the effects of increasing P_c or T_c and reducing ε or lowering the molecular weight (and γ) of the propellant by switching to helium. From this analysis, we can draw some general conclusions about basic trade-offs in rocket design:

- Increasing chamber pressure increases thrust (with little or no effect on specific impulse or density-specific impulse).
- Increasing chamber temperature increases specific impulse and density-specific impulse with a slight decrease in thrust.
- Decreasing expansion ratio (underexpanded condition) decreases specific impulse, density-specific impulse, and thrust.
- Decreasing propellant molecular weight increases specific impulse at the expense of decreasing density-specific impulse and thrust.

Electrodynamic Acceleration. We have spent some time discussing thermodynamic expansion and acceleration of exhaust using nozzles to convert propellant with thermodynamic energy into high-speed flow. But there is a second method for propellant acceleration currently gaining wider use in spacecraft—electrodynamic acceleration. To take advantage of this method, we must start with a charged propellant. Recall, the force of attraction (or repulsion) between charges depends on the strength of the charges involved and the distance between them. We expressed this as Coulomb's law:

$$F = KQq/R^2R \tag{8.28}$$

where F = electrostatic force on charge 1 (N)
K = constant (9×10^9)
Q = value of charge 1 (coulombs, C)
q = value of charge 2 (coulombs, C)
R = distance between charges (m)
$\boldsymbol{R}$ = unit vector in direction of force

An *electric field* exists when there is a difference in charge between two points—That is, there is an imbalance between positive and negative charges in a confined region. We call the energy an electric field can transmit to a unit charge the *electrical potential*, described in terms of volts/m. The resulting force on a unit charge is called *electrostatic force*.

A simple example of electrostatic force in action can be seen if you have ever rubbed a balloon through your hair and stuck it to a wall. As you rub the balloon it picks up a net positive charge. When it is placed against the wall, initially neutral, the positive charges on the surface are pushed away leaving a net negative charge. The opposite charges attract each other creating a force strong enough to keep the balloon in place. The strength of the electrostatic force is a function of the charge and the electric field.

$$\mathbf{F}_i = m_i\mathbf{a}_i = q_i\mathbf{E} \tag{8.29}$$

where $\mathbf{F}_i$ = electrostatic force vector (N)
m_i = mass of charged particle (kg)

$\mathbf{a}_i$ = acceleration of charged particle (m/s^2)
q_i = charge on particle (coulombs, C)
$\mathbf{E}$ = electric field vector (electric potential) (V/m)

This is illustrated in Figure 8.10. Notice that the direction of the force is parallel to the electric field.

Electrodynamic rockets take advantage of this principle to create thrust. In the simplest application, they only need some charged propellant and an electric field. As with any rocket, the two things we are most interested in are thrust, F, and specific impulse, I_{sp} From equation (8.2) we know thrust depends on the mass flow rate, $\dot{m}$, and effective exhaust velocity, C:

$$F = \dot{m}C \tag{8.30}$$

and from Equation (8.7) we know specific impulse, I_{sp}, is directly related to C by

$$C = I_{sp}g_o$$

In an electrodynamic rocket, high $\dot{m}$ is achieved by having a high density of charged propellant. High exhaust velocity comes from having a strong electric field and/or applying the electrostatic force for a longer time. We can summarize these effects on performance as follows:

- Higher charge density → higher $\dot{m}$ → higher thrust

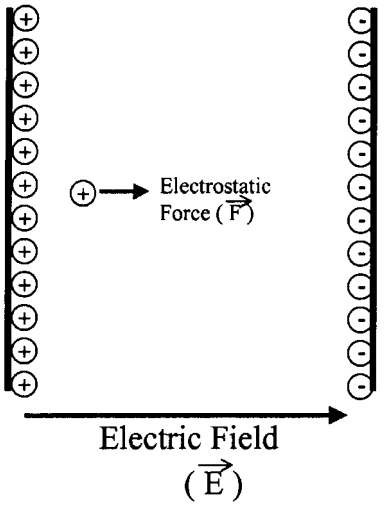


FIGURE 8.10 Electrostatic force. An electric field exists when there is an imbalance between positive and negative charges in a confined region. This field will impart an electrostatic force on a charged particle within the field, making it accelerate.

- Stronger electric field → stronger electrostatic force on propellant → higher acceleration → higher exhaust velocity → higher I_{sp}

Thus, by varying the charge density and the applied field, a wide range of thruster designs can be created. Naturally, there are practical design issues that limit how high you can increase each of these parameters. Let's start with charge density.

Charge density is limited by the nature of the propellant and how it is charged. Earlier, we defined an ion to be a positively charged propellant molecule that has had one or more electrons stripped off. Ions are handy in that they are simple to accelerate in an electric field. Unfortunately, when you try to pack lots of positive ions into a small confirmed space, they tend to repel each other. This creates a practical limit to the charge density you can achieve with ions.

One way around this is to create a *plasma* with the propellant instead. A plasma is an electrically neutral mixture of ions and free electrons. A plasma can be seen inside a common florescent lamp or neon light. When a gas, such as neon, is placed in a strong electric field, the electrons become only weakly bound to the molecules creating a "soup" of ions and free electrons. The glow that is given off is the result of electrons jumping back and forth between energy states within the molecule. Because it is electrically neutral, a plasma can contain a much higher charge density than simply an collection of ions alone.

So far we have only considered the acceleration effect from an applied electric field. However, whenever an electric field is applied to a plasma, a magnetic field is also created (or said to be *induced*). Charged particles are also accelerated by magnetic fields but at right angles to the field, instead of parallel to it. To determine the combined effect of electric and magnetic fields on a charged particle we must look at the cross-product of their interaction.

$$F_{em} = m_i a_{em} = q_i(\boldsymbol{E} + V_i \times \boldsymbol{B}) \tag{8.31}$$

where F_{em} = electromagnetic force on a particle (N)
a_{em} = electromagnetic acceleration (m/s2)
$\boldsymbol{E}$ = electric field vector (V/m)
V_i = velocity of particle (m/s)
$\boldsymbol{B}$ = Magnetic field vector (tesla)

Some types of electrodynamic rockets rely on this combined effect to produce thrust. However, for most cases the electrostatic force is dominant and we can ignore the effect of the magnetic field for simple analysis of performance. When we consider the second parameter that limits thruster performance, the strength of the electric field, we can focus mainly on the practical limits of applied power. From equation (8.2b) the jet power of a rocket is found from:

$$P_J = \tfrac{1}{2}\dot{m}C^2$$

Thus, for a given charge density, the exhaust velocity increases with the square root of the power. As you would expect, there are practical limits to the amount of power available in any spacecraft, limiting the ultimate performance of electrodynamic rockets.

While exhaust velocity (and I_{sp}) go up with power, there is a trade-off between thrust and exhaust velocity, as illustrated by the following relationship:

$$F = 2P/C \tag{8.32}$$

where F = thrust (N)
P = power (W)
C = effective exhaust velocity (m/s)

Therefore, when designing an electric thruster, you can have high exhaust velocity or high thrust, but not both at the same time. In subsection 8.2 we will look at some specific examples of electric thrusters and compare their performance.

8.2 PROPULSION SYSTEMS

Subsection 8.1 gave us a look at rockets as a system. But rockets, as important as they are, comprise only one part of an entire propulsion system. In this section, we will concentrate less on rocket theory and more on propulsion system technology to learn what essential components we need, and how they're put together.

Figure 8.11 shows a block diagram for an entire propulsion system. To design a specific system, we start with the desired thrust, usually at some very specific time. The propulsion system controller manages these inputs and formulates commands to send to the propellant management actuators to turn the flow of propellant on or off. For some systems, the controller also manages the energy input to the rocket. For example, in an electrodynamic rocket thermoelectric rocket, the system has to interface with the spacecraft's electrical power subsystem (EPS) to ensure it provides the correct power level. The controller uses sensors extensively to monitor the temperature and pressure of the propellant throughout.

Regardless of whether you are using an thermodynamic or an electrodynamic rocket, you need to store and handle some propellant. In this section, we will start by looking at propellant management, how to store liquid or gaseous propellants and supply them to the rocket as needed. We will then review in detail most of the various thermodynamic and electrodynamic rocket technologies currently in use or

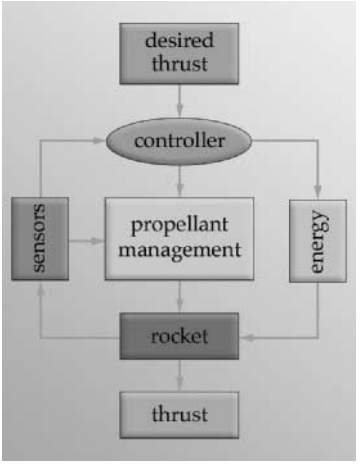


FIGURE 8.11 Block diagram of a complete propulsion system.

on the drawing boards. Following this discussion, we will look briefly at important factors for selecting and testing propulsion systems. Finally, since rocket scientists are always striving to improve propulsion system performance, we will look at what exotic concepts may one day take spacecraft to the stars.

Propellant Management

All rockets need propellant. The job of storing propellant and getting it where it needs to go at the right time is called *propellant management*. The propellant management portion of a propulsion system has four main tasks:

1. Propellant storage
2. Pressure control
3. Temperature control
4. Flow control

We will look briefly at the requirements and hardware for each of these tasks.

Gaseous propellants, such as nitrogen for cold-gas rockets, are normally stored in tanks under high pressure to minimize their volume. Typical gas storage pressures are 200 bar (3000 psi) or more. Unfortunately, we cannot make a liquid propellant denser by storing it under pressure. However, depending on how we pressurize the liquid propellant for delivery to the combustion chamber, the storage tanks may have to be designed to take high pressure as well. In any case, propellant tanks are typically made from aluminum, steel, or titanium and designed to withstand whatever pressure the delivery system requires.

As we learned in subsection 8.1, combustion chamber pressure is an important factor in determining rocket thrust. This pressure depends on the delivery pressure of the propellants. Pressurizing the flow correctly is another function of propellant management. There are two approaches to achieving high pressure flow: pressure-fed systems and pump-fed systems.

Figure 8.12 shows how a *pressure-fed propellant system* relies on either a gaseous propellant stored under pressure, or a separate tank attached to the main tank, filled with an inert, pressurized gas such as nitrogen or helium to pressurize and expel a liquid propellant. The high-pressure gas squeezes the liquid propellant out of the storage tank at the same pressure as the gas.

To minimize volume, the storage pressure of the gas is typically much higher than the pressure needed in the combustion chamber. To reduce or regulate the high pressure in the gas storage tank to the lower pressure for propellant delivery, mechanical regulators are typically used. As high-pressure gas flows into a *regulator*, the gas pushes against a carefully designed diaphragm. The resulting balance of forces maintains a constant flow rate but at a greatly reduced output pressure. For example, a gas stored at 200 bar may pass through a regulator that reduces it to 20 bar before it goes into a liquid propellant tank. Pressure regulators are common devices found in most rocket plumbing systems. SCUBA tanks use regulators to reduce high-pressure air stored in the tank to a safe, lower-pressure for breathing.

The main drawback of pressure-fed systems is that the more liquid propellant in the tank, the more pressurizing gas is needed. For very large propulsion systems, such as on the Space Shuttle, enormous quantities of high-pressure propellant must be delivered to the combustion chamber each second. To do this using a pressure-fed system would require additional large, high-pressure gas tanks, making the

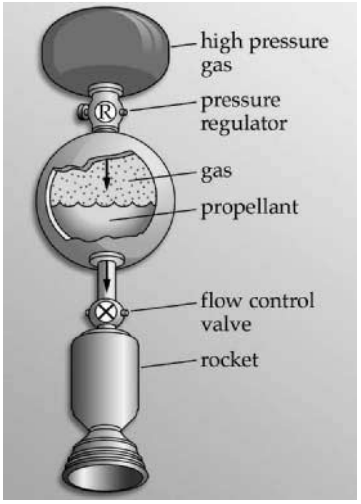


FIGURE 8.12 Pressure-fed propellant system.

entire launch vehicle larger and heavier. Instead, most launch vehicles use pump-fed delivery systems.

Pump-fed delivery systems rely on pumps to take low-pressure liquid and move it toward the combustion chamber at high pressure, as shown in Figure 8.13. Pumps impart kinetic energy to the propellant flow, increasing its pressure. On the Space Shuttle, massive turbo-pumps burn a small amount of H_2 and O_2 to produce mechanical energy. This energy is used to take the liquid propellants normally stored at a few bar and boost the feed pressure to over 480 bar (7000 psi) at a flow rate

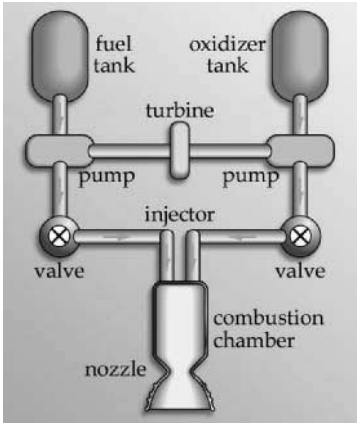


FIGURE 8.13 Pump-fed propellant management.

ROCKETS AND LAUNCH VEHICLES

of 2.45×10^5 liters/sec (6.5×10^4 gal/sec). Spinning at over 30,000 rpm, the Shuttle propellant pumps could empty an average-size swimming pool in only 1.5 seconds!

Regardless of the propellant-delivery system used, the pressure of propellants and pressurizing gases must be constantly monitored. *Pressure transducers* are small electromechanical devices used to measure the pressure at various points throughout the system. This information is fed back to the automatic propellant controller and sent to ground controllers via telemetry.

Temperature control for propellant and pressurant gases is another important propellant-management function. The ideal gas law tells us that a higher gas temperature causes a higher pressure and vice versa. The propellant-management subsystem must work with the spacecraft environmental control and life support subsystem (ECLSS) to maintain gases at the right temperature and prevent liquid propellants from freezing or boiling. In the deep cold of outer space, there is a danger of propellants freezing. For instance, hydrazine, a common spacecraft propellant, freezes at 0°C. Usually the spacecraft ECLSS maintains the spacecraft well above this temperature, but in some cases exposed propellant lines and tanks may need heaters to keep them warm.

On launch vehicles, propellant thermal management often has the opposite problem. It must maintain liquid oxygen (LOX) and liquid hydrogen at temperatures hundreds of degrees below zero Centigrade. Using insulation extensively helps control the temperature. However, some boil-off of propellants prior to launch is inevitable and must be planned for.

Finally, the propellant management system must control the flow of gases and liquids. It does this using valves. Valves come in all shapes and sizes to handle different propellants, pressures, and flow rates. Fill and drain valves are needed to fill the tanks prior to launch (and drain them in emergencies). Tiny, electrically controlled low-pressure valves pulse cold-gas thrusters on and off to deliver precise micro-amounts of thrust. Large pyrotechnic valves mounted below liquid-propellant tanks keep them sealed until ignition. When the command is sent, a pyrotechnic charge fires, literally blowing the valve open, allowing the propellant to flow. Of course, these valves are good for only one use. To protect against overpressure anywhere in the system, *pressure relief valves* automatically release gas if the pressure rises above a preset value. Check-valves are designed to allow liquid to flow in only one direction, preventing backflow in the wrong direction. Other valves throughout the system ensure propellant flows where it needs to when the system controller sends the command. Some of these other valves lead to redundant lines that ensure the propellant flows even when a main valve malfunctions.

Let's briefly review the components needed for propellant management. Propellants and pressurant gas are stored in tanks. Below the tanks, valves control the flow throughout the system and regulators reduce the pressure where needed. Pressure and temperature are measured at various points in the system using transducers and other sensors. Figure 8.14 gives an example schematic of a simple cold-gas rocket system showing the necessary components for propellant management.

Thermodynamic Rockets

As we learned in subsection 8.1, thermodynamic rockets convert thermodynamic energy (heat and pressure) transferred to a propellant into high-speed exhaust using

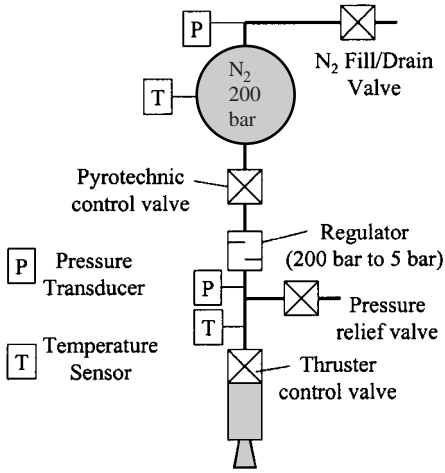


FIGURE 8.14 Example schematic for a cold-gas rocket system.

nozzles. There are a wide variety of other thermodynamic rockets currently available or being considered. We can classify these based on their source of energy:

- Cold-gas—use mechanical energy of a gas stored under pressure
- Chemical—rely on chemical energy (from catalytic decomposition or combustion) to produce heat
- Solar thermal—use concentrated solar energy to produce heat
- Thermoelectric—use the heat produced from electrical resistance
- Nuclear thermal—use the heat from a nuclear reaction

Because we examined simple cold-gas rockets in detail in the last section, here we will review the other four types and compare their relative performances.

Chemical Rockets. The vast majority of rockets in use today rely on chemical energy. When you strike a match, the flame represents a combustion process taking place. The fuel—the wood in the match—is chemically combining with the oxygen in the air to form various chemical by-products (CO, CO_2, water, etc.) and, most importantly, heat. In *chemical rockets*, the energy is stored in the chemical bonds of the propellants. In the Shuttle main engines, liquid hydrogen (H_2) and liquid oxygen (O_2) are combined in the most basic of all chemical reactions:

$$2H_2 + O_2 \rightarrow 2H_2O + \text{Heat}$$

All combustion reactions must have a *fuel* (such as hydrogen) plus an *oxidizer* (such as oxygen). These two combine, liberating a vast amount of heat and creating by-products that form the exhaust. The heat is transferred to the combustion products, raising their temperature. The place where this chemical reaction and energy transfer takes place is called the *combustion chamber*. Although the propellants arrive in the combustion chamber under pressure, delivered by the propellant man-

agement system, this mechanical energy is small compared to the thermal energy released by the chemical reaction.

Chemical rockets generally fall into one of three categories:

- Liquid
- Solid
- Hybrid

Let's briefly review the operating principles and performance parameters of each.

Liquid Chemical Rockets. Liquid chemical rockets are usually one of two types, bipropellant or monopropellant. As the name implies, *bipropellant* rockets use two liquid propellants. One is a fuel, such as liquid hydrogen (LH_2), and the other is an oxidizer, such as liquid oxygen (LOX). Brought together under pressure in the combustion chamber by the propellant management system, the two compounds chemically react (combust) releasing vast quantities of heat and producing combustion products (these vary depending on the propellants). To ensure complete, efficient combustion, the oxidizer and fuel must mix in the correct proportions. The *oxidizer/fuel ratio* (O/F) is the proportion, by mass, of oxidizer to fuel.

Some propellant combinations, such as hydrogen and oxygen, will not spontaneously combust on contact. They need an igniter, just as your car needs a spark plug, to get started. This need, of course, increases the complexity of the system somewhat. Propellant chemists thus strive to find combinations that react on contact. We call these propellants *hypergolic* because they do not need a separate means of ignition. Hydrazine (N_2H_4) plus nitrogen tetroxide (N_2O_4) is an example of a hypergolic fuel and oxidizer combination.

Another important feature in selecting a propellant is storability. Although the liquid hydrogen and liquid oxygen combination used in the Space Shuttle main engines offers high performance (specific impulse around 455 s), it must be supercooled to hundreds of degrees below zero Centigrade. Because of their low storage temperature, we call these propellants *cryogenic*. Unfortunately, it is difficult to maintain these extremely low temperatures for long periods (days or months). When the mission concept calls for long-term storage, designers turn to *storable propellants* such as hydrazine and nitrogen tetroxide that remain stable at room temperature for a very long time (months or even years).

The Titan, an early intercontinental ballistic missile (ICBM), used hypergolic, storable propellants because the missiles stayed deep in underground silos for many years. These propellants are also used in the Shuttle's orbital-maneuvering engines and reaction-control thrusters as well in virtually all spacecraft using liquid chemical rockets. The penalty paid for the extra convenience of spontaneous combustion and long-term storage is a much lower performance than the cryogenic option (specific impulse around 300 s). In addition, current hypergolic combinations are extremely toxic and require special handling procedures to prevent propellant release. Table 8.2 summarizes key points about bipropellant rockets.

As the name implies, monopropellant chemical rockets use only a single propellant. These propellants are relatively unstable and easily decompose through contact with the correct catalyst.

Hydrogen peroxide (H_2O_2) is one example of a monopropellant. You may have used a low-concentration (3%), drug store variety of this compound to disinfect a bad scrape or to bleach your hair. Rocket-grade hydrogen peroxide, also called high-test peroxide (HTP), has a concentration of 85% or more. It is relatively safe to handle at room temperatures, but when passed through an appropriate catalyst

TABLE 8.2 Bipropellant Rockets

Operating principle:	
A liquid oxidizer and a liquid fuel are react together in a combustion process liberating heat and creating exhaust products that are thermodynamically expanded through a nozzle.	
Advantages:	**Disadvantages:**
• High I_{sp}	• Must manage two propellants
• Can be throttled	• Intense heat of combustion creates thermal control problems for chamber and nozzle
• Can be re-started	

(such as silver) it readily decomposes into steam (H_2O) and oxygen, releasing significant heat in the process. Typical HTP reactions exceed 630°C. This relatively high temperature, combined with the molecular weight of the reaction products, gives HTP monopropellant rockets an I_{sp} of about 140 s. These types of thrusters were successful used on the X-15 rocket plane and the Scout launch vehicle.

By far the most widely used monopropellant today is hydrazine (N_2H_4). It readily decomposes when exposed to a suitable catalyst, such as iridium, producing an I_{sp} of about 180 sec. The main disadvantage of hydrazine is that it is higly toxic. This means specialized handling procedures and equipment are needed during all testing and launch operations.

The biggest advantage of monopropellant over bipropellant systems is simplicity. The propellant-management system needs to maintain only one set of tanks, lines, and valves. Unfortunately, there is a significant penalty in performance for this added simplicity (two-thirds the I_{sp} of a bipropellant system or less). However, for certain mission applications, especially station keeping and attitude control on large communication satellites, this trade-off is worth it. The benefit grows when we use hydrazine as the fuel with nitrogen tetroxide in a large bipropellant rocket for initial orbital insertion and then, by itself in a smaller, monopropellant rocket for station keeping. Such ''dual-mode'' systems take advantage of the flexibility offered by hydrazine to maximize overall system performance and simplicity. Table 8.3 summarizes key points about monopropellant rockets.

Solid Chemical Rockets. The fireworks we watch on the Fourth of July are a good example of solid rockets at work. Solid rockets date back thousands of years to the Chinese, who used them to confuse and frighten their enemies on the battlefield. In modern times these rockets create thrust for intercontinental ballistic missiles as well as space launch vehicles.

Just as liquid bipropellant rockets combine fuel and oxidizer to create combustion, solid rockets contain a mixture of fuel, oxidizer, and a binder, blended in the correct proportion and solidified into a single package called a *motor*. A typical

TABLE 8.3 Monopropellant Rockets

Operating principle:	
A single propellant is decomposed using a catalyst releasing heat and creating by-products that are thermodynamically expanded through a nozzle.	
Advantages:	**Disadvantages:**
• Simple, reliable	• Lower I_{sp} than bipropellant
• Single propellant to manage	

ROCKETS AND LAUNCH VEHICLES

composite solid rocket fuel is powdered aluminum. The most common oxidizer is ammonium perchlorate (AP). Together, the fuel and oxidizer comprise about 85–90% of the rocket motor mass, with an oxidizer/fuel ratio of about 8:1. The remaining mass of the motor consists of the binder, which holds the other ingredients together and provides overall structural integrity. Binders are usually hard, rubber-like compounds such as hydroxyl terminated polybutadiene (HTPB). During combustion, the binder also acts as additional fuel.

As we learned in subsection 8.1, rocket thrust depends on mass flow rate. In a solid rocket motor, this rate depends on the propellant burn rate (kg/s) and the burning surface area (m^2). The faster the propellant burns and the greater the burning surface area, the higher the mass flow rate and the higher the resulting thrust. Propellant burning rate depend on the type of fuel and oxidizer, their mixture ratio, and the binder material. The total burning area depends primarily on the inside shape of the solid propellant. During casting, designers can shape the hollow inner core of the solid propellant to adjust the surface area available for burning so they can control burning rate and thrust. The Space Shuttle's solid rocket motors, for example, have a star-shaped core, specifically tailored so the thrust decreases 55 seconds into the flight to reduce acceleration and the effects of aerodynamic forces.

Because solid motor combustion depends on exposed propellant surface area, manufacturers must carefully mold the propellant mixture to prevent cracks. Burning occurs on any exposed surface, even along undetected cracks in the propellant grain. Investigators linked the Space Shuttle *Challenger* accident to an improperly sealed joint between solid motor segments. This open seal exposed the motor case to hot gases, burning it through and causing the accident.

The *Challenger* disaster highlighted another drawback of solid motors—once they start, they are very difficult to stop. With a liquid rocket, we simply turn off the flow of propellant to shut off the engine. Solid motors burn until all the propellant is gone. To stop one prior to that would require us to blow off the top or split it open along its side, releasing internal pressure and thus, inhibiting combustion. That is not a very practical solution on the way to orbit!

Despite their drawbacks, solid motors are used on a variety of missions because they offer good, cost-effective performance in a simple, self-contained package that does not require a separate propellant management system. One important use of solid motors is to augment liquid engines on launch vehicles. Without the solid rocket boosters, the Space Shuttle could not get off the ground. Several expendable launch vehicles use various combinations of strap-on solid motors to give users a choice in payload-lifting capacity, without the need to redesign the entire vehicle. For example, three, six, or nine solid motors can be added to the Delta II launch vehicle, depending on the payload mass. Solid motors also provide thrust for strap-on upper-stages for spacecraft needing a well-defined ΔV to go from a parking orbit into a transfer orbit.

A solid rocket motor's I_{sp} performance depends on the fuel and oxidizer used. After the chemicals are mixed and the motor is cast, however, the I_{sp} and thrust profile are fixed. I_{sp} for typical solid motors currently in use ranges from 200–300 s, somewhat more than for a liquid, monopropellant rocket but slightly less than for a typical liquid bipropellant engine. Their big performance advantage is in terms of I_{dsp}. For example, the Shuttle's solid rocket boosters (SRBs) have an I_{dsp} 6% less than the I_{dsp} of the liquid main engines (SSMEs), even though the I_{sp} for the SSMEs is almost 70% higher. This makes solid motors ideal for volume-constrained missions needing a single, large ΔV. Table 8.4 summarizes key points about solid rocket motors.

TABLE 8.4 Solid Rockets

Operating principle:
An oxidizer and fuel are blended together in a single, solid grain. Combustion takes place along any exposed surface, producing heat and byproducts that are thermodynamically expanded through a nozzle.

Advantages:	**Disadvantages:**
• Simple, reliable	• Susceptible to cracks in grain
• No propellant management needed	• Cannot be restarted
• High I_{dsp} compared to bipropellant	• Difficult to stop
	• Modest I_{sp}

Hybrid Chemical Rockets. *Hybrid propulsion systems* combine aspects of both liquid and solid systems. A typical hybrid rocket uses a liquid oxidizer and a solid fuel. The molded fuel grain forms the combustion chamber and the oxidizer is injected into it as illustrated in Figure 8.15. A separate sparking system or a superheated oxidizer initiates combustion. The hybrid combustion process is similar to burning a log in the fireplace. Oxygen from the air combines with the log (fuel) in a fast oxidation process and burns. If we take away the air (throttle the oxidizer), the fire goes down. If we use a bellows or blow on the fire, we increase the flow of air and the fire grows.

A properly designed hybrid rocket can offer the flexibility of a liquid system with the simplicity and density of a solid motor. Hybrids are safe to handle and store, similar to a solid, but can be throttled and restarted, similar to a liquid engine. Their efficiencies and thrust levels are comparable to solids. For example, one interesting hybrid configuration uses high-test peroxide (HTP) oxidizer with polyethylene (plastic) fuel. At an O/F of 8:1 this system offers an I_{sp} of around 290 s and I_{dsp} of around 3.8×10^5 kg/m^{3}s. It has the added advantage that the HTP can be used alone as a monopropellant making it a dual-mode system. Unfortunately, at this time hybrid rocket research and applications lags far behind research on liquid and solid systems. Table 8.5 summarizes key points about hybrid rockets.

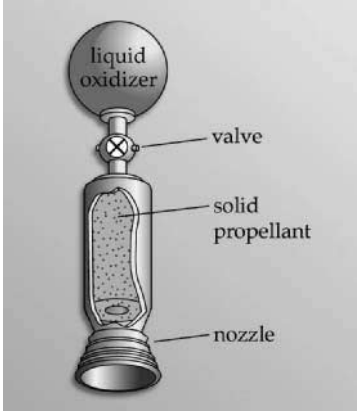


FIGURE 8.15 Hybrid rocket motor.

TABLE 8.5 Hybrid Rockets

Operating Principle:
Hybrid rockets typically use a liquid oxidizer with a solid fuel. The oxidizer is injected into a hollow port (or ports) within the fuel grain where combustion takes place along the boundary with the surface.

Advantages:
- Only one liquid propellant to manage
- Safe, cannot create an explosive combination of fuel/oxidizer
- Environmentally friendly products

Disadvantages:
- Limited heritage
- Modest I_{sp}

Chemical Rocket Summary. Table 8.6 compares the I_{sp} and I_{dsp} of the thermodynamic rockets we've discussed in this section and compares their performance and key features.

Solar Thermal Rockets. In chemical rockets, the heat is a by-product of a chemical reaction. But heat can also be produced in other ways, then transferred directly to the propellant using conduction and/or convection. One convenient source of heat is the sun. By concentrating solar energy using mirrors or lenses, we can create extremely high temperatures (2000K or more) on a focused point. By passing a propellant, such as hydrogen, through this point, it can directly absorb the heat, reaching very high temperatures before being expanded through a nozzle to achieve high velocity exhaust. In this way, *solar thermal rockets* use the limitless power of the sun to produce relatively high thrust with high I_{sp}.

The natural advantage of a solar thermal rocket is the abundant source of solar energy, saving the need to produce the energy on the spot or carry it along as chemical energy. It can use virtually any propellant. The best I_{sp}, of course, comes from using hydrogen. Theoretical and experimental results indicate a liquid-hydrogen, solar-thermal rocket could achieve a specific impulse of about 700 s. Thrust levels are limited by the basic engineering problems of efficiently transferring heat between the thermal mass that absorbs solar energy and the propellant. However, thrusts in the several-Newton range should be achievable. Another important operational challenge for solar thermal rockets is deploying and steering the large mirrors needed to collect and focus the solar energy. Several concepts for solar thermal rockets have been proposed, such as the Solar Orbital Transfer Vehicle, SOTV. However, up to now, none have been tested in orbit. Table 8.7 summarizes key features of solar thermal rockets.

Thermoelectric Rockets. Of course, the solar energy is only available when the sun is shining. A spacecraft in eclipse, or far from the sun, needs another heat source. If you hold your hand up next to a conventional light bulb, you will feel the heat given off by the resistance of the filament in the bulb. For space applications, the energy source is the electrical energy provided by the spacecraft's electrical power subsystem (EPS). By running electricity through a simple resistor, or by creating an arc discharge as you would with a spark plug, we can create heat. *Thermoelectric rockets* transfer this heat to the propellant by conduction and convection.

One of the simplest examples of a thermoelectric rocket is a *resisto-jet.* As we see in Figure 8.16, electrical current flows through a metal heating element inside a combustion chamber. The resistance (or electrical friction) in the metal causes it

TABLE 8.6 Thermodynamic Rockets

Type	Propellant combinations (O/F) [specific gravity]	I_{sp} (s)	I_{dsp} (s kg/m^3)	Advantages	Disadvantages
Liquid	—	—	—	• Can be re-started • Can be throttled	• Propellant management needed • Intense heat of combustion creates thermal control problems for chamber and nozzle
Bipropellant	LO_2/LH_2 (5:1) [1.15:0.07]	477	462	• High I_{sp} • "Green" propellants	• Cryogenic fuel/oxidizer difficult to store
	LO_2/Kerosene (RP-1) (2.25:1) [1.15:0.8]	370	385	• Storable fuel • Good I_{dsp}	• Cryogenic oxidizer
	N_2O_4/N_2H_4 (1.9:1) [1.43:1.0]	334	429	• Storable propellants • Good I_{sp}	• Toxic propellants
Monopropellant				• Simple • One set of propellant management • Moderate I_{sp} • Dual-mode capable	• Low I_{sp}
	N_2H_4 (hydrazine) [1.00]	245	246	• Large flight heritage	• Toxic
	H_2O_2 (hydrogen peroxide) [1.37]	181	247	• "Green" propellant	• Little flight heritage
Solid	NH_4ClO_4 (AP)/Al (includes a binder, e.g., HTPB) [1.95:1.26]	300	539	• Simple • High I_{dsp} • No propellant management needed	• Sensitive to cracks in propellant grain • Difficult to stop • Cannot restart
Hybrid	H_2O_2 (85%)/PE (8:1) [1.37:0.9]	333	437	• One set of propellant management • Safe operation (cannot explode) • "Green" propellants	• Limited heritage • Modest I_{sp}

LOX = Liquid oxygen; LH_2 = liquid hydrogen; RP-1 = rocket propellant-1, Kerosene; N_2O_4 = nitrogen tetroxide; N_2H_4 = hydrazine; HTP = high-test hydrogen peroxide (85% H_2O_2, 15%H_2O); HTPB = hydroxyl terminated polybutadiene (rubber); PE = polyethylene (plastic).

ROCKETS AND LAUNCH VEHICLES

TABLE 8.7 Solar Thermal Rockets

Operating principle:
Solar energy is concentrated using lenses or mirrors onto a heat transfer chamber. A propellant, such as liquid hydrogen, flows through the chamber, heating up, and is then expended through a nozzle.

Advantages:
- Limitless energy supply, can be refueled and reused
- Potentially very high I_{sp} (~700 s with H_2)
- Environmentally friendly products

Disadvantages:
- Less effective at the outer planets.
- Must carefully point a large mirror or lens
- No flight heritage

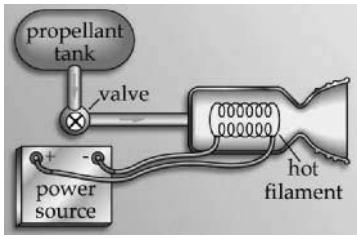


FIGURE 8.16 Resisto-jet.

to heat up. As propellant flows around the heating element, heat is transferred to it via convection, increasing its temperature before it expands through a nozzle.

This simple principle can be applied to virtually any propellant (NASA even investigated using urine on the Space Station as a propellant). The resisto-jet concept can significantly increase the specific impulse of a conventional cold-gas rocket, making it, in effect, a "hot-gas" rocket with twice the I_{sp} (from 60 s to more than 120 s for nitrogen). Resisto-jets improve the performance of a conventional hydrazine mono-propellant rockets by heating the exhaust products, thus boosting their I_{sp} by about 50% (from 200 s to over 300 s). The direct benefit of a resisto-jet rocket comes from adding heat to the propellant, so, the hotter it gets, the higher its I_{sp} and I_{dsp}. Hydrazine resisto-jets are gaining wide use as mission designers become increasingly able to trade extra electrical power for a savings in propellant mass. The Iridium spacecraft, for example, rely on hydrazine resisto-jets to take them from their initial parking orbit to their final mission orbit.

Another method for converting electrical energy into thermal energy is by using a spark or electric arc. To form an arc, we create a gap in an electrical circuit and charge it with a large amount of electricity. When the electrical potential between two points gets high enough, an arc forms (during a thunderstorm we see this as dazzling displays of lighting). An *arc-jet rocket* passes propellant through a sustained arc, increasing its temperature. Arc-jet systems can achieve relatively high I_{sp} (up to 1000 s) with small but significant thrust levels (up to 1 Newton). Like resisto-jets, arc-jet rockets can use almost any propellant. Current versions use hydrazine, liquid hydrogen, or ammonia. A schematic for an arc-jet system is shown in Figure 8.17. The ARGOS spacecraft, for example, was launched in 1999 to test

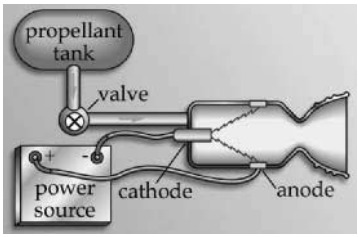


FIGURE 8.17 Arc-jet thruster.

a 25-kW ammonia arc-jet, producing a thrust of around 0.8 N with a specific impulse over 800 sec.

As you would expect, the primary limitation on thermoelectric rocket thrust and efficiency is the amount of power available. In subsection 8.1 we saw a simple relationship among input power, thrust, and specific impulse.

$$F = 2P/I_{sp}g_o$$

where F = thrust (N)
P = power (W)
I_{sp} = specific impulse (s)
g_o = gravitational acceleration = 9.81 m/s^2

Using this simple equation, we can fine-tune the design of a thermoelectric thruster, trading off thrust versus power versus I_{sp}. For example, if we double the power input we can increase thrust by a factor of 4 for the same I_{sp}. Table 8.8 summarizes key features of thermoelectric rockets.

Nuclear Thermal Rockets. Another potentially useful heat source in space is nuclear energy. On Earth, nuclear reactors harness the heat released by the fission of uranium to produce electricity. In much the same way, a *nuclear thermal rocket* uses its propellant, such as liquid hydrogen, to flow around the nuclear core, absorbing the thermal energy while cooling the reactor. As you can see in Figure 8.18, propellant enters the reaction chamber, where it absorbs the intense heat from the nuclear reaction. From there, thermodynamic expansion through a nozzle produces high thrust (up to 10^6 N) and high I_{sp} (up to 1000 s using hydrogen).

TABLE 8.8 Thermoelectric Rockets

Operating principle:
Heat is produced through electric resistance or spark discharge inside a heat transfer chamber. A propellant, such as water or ammonia (NH_3), flows through the chamber, heating up, and is then expended through a nozzle.

Advantages:
- Simple, reliable
- Can be used as an add-on to conventional monopropellant rocket to boost I_{sp} ~50%
- High power arc-jet offers very high I_{sp} (>800 sec with NH_3)

Disadvantages:
- Requires large amounts of onboard electrical power
- Relatively low thrust (<1 N)

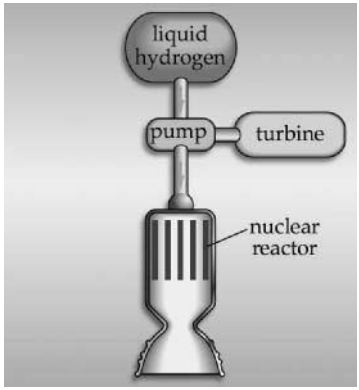


FIGURE 8.18 Nuclear thermal rocket.

Because of their relatively high thrust and better efficiencies, nuclear thermal rockets offer a distinct advantage over chemical systems, especially for manned planetary missions. These missions must minimize transit time to decrease the detrimental effects of free fall as well as exposure to solar and cosmic radiation on the human body. Ironically, future astronauts may escape the danger of space radiation by using the energy from a nuclear reactor to propel them to their destination faster. Extensive research into nuclear thermal rockets was done in the United States. in the 1960s as part of the NERVA program. Additional work was done in the 1980s when great theoretical advances occurred in heat transfer. Unfortunately, environmental and political concerns about safe ground testing of nuclear-thermal rockets (let alone the potential political problems of trying to launch a fully fueled nuclear reactor) have severely reduced research into this promising technology. Table 8.9 summarizes key features of nuclear thermal rockets.

Electrodynamic Rockets

While thermodynamic rockets offer relatively high thrust over a very wide range—10^{-1} to 10^{6} N—basic problems in heat transfer pose practical limits on the specific impulse that they can ultimately reach (up to 1000 s or so for nuclear rockets). To achieve the higher efficiencies demanded by future, more challenging interplanetary

TABLE 8.9 Nuclear Thermal Rockets

Operating principle:
Heat is produced through nuclear fission inside a reactor. A propellant, such as liquid hydrogen, flows through the reactor, heating up, and is then expended through a nozzle.

Advantages:
- Long-term energy supply, can be refueled and reused
- Potentially very high I_{sp} (~1000 s with H_2)
- High thrust (~10^6 N)

Disadvantages:
- Environmental and political problems with testing and launching nuclear reactors
- No flight heritage

and commercial missions, we need to take a different approach—electrodynamic rockets.

As discussed in subsection 8.1, electrodynamic rockets rely on electric and/or magnetic fields to accelerate a charged propellant to very high velocities (more than 10 times the exhaust velocity and I_{sp} of the Shuttle main engines). However, this high I_{sp} comes with a price tag—high power and low thrust. Recall from equation (8.2) the relationship among power, exhaust velocity, and thrust.

$$F = 2P/C$$

Power, of course, is always a limited commodity on a spacecraft—especially if you are trying to use an electrodynamic thruster. Given a finite amount of power, equation (8.2) tells us we can have high exhaust velocity only at the expense of thrust. As a result, practical limits on power availability make electrodynamic thrusters unsuitable for launch vehicles or when a spacecraft needs a quick, large impulse, such as when it brakes to enter a capture orbit. Even so, because of their very high I_{sp}, mission planners are increasingly willing to sacrifice power and thrust (and the extra time it will take to get where you need to go) in order to save large amounts of propellant mass.

As we indicated in subsection 8.1, there are many ways to use electric and/or magnetic fields to accelerate a charged propellant. Here we will focus on the two primary types of electrodynamic rockets currently in use operationally:

- Ion (or electrostatic) thrusters—use electric fields to accelerate ions
- Plasma thrusters—use electric and magnetic fields to accelerate a plasma

Ion Thrusters. An *ion thrusters* (also called an *electrostatic thruster*) uses an applied electric field to accelerate an ionized propellant. Figure 8.19 illustrates the basic operating principles of an ion thruster. First, the thruster ionizes propellant by stripping off the outer shell of electrons, making positive ions. It then accelerates these ions by applying a strong electric field. If the engine ejected the positive ions without neutralizing them, the spacecraft would eventually accumulate a negative charge over time due to the leftover electrons. To prevent this, as Figure 8.19 illustrates, it uses a neutralizer source at the exit plane to eject electrons into the exhaust, making it charge-neutral.

The ideal ion thruster propellant is easy to ionize, store, and handle. Early ion thruster research used mercury and cesium, both metals that are easy to ionize.

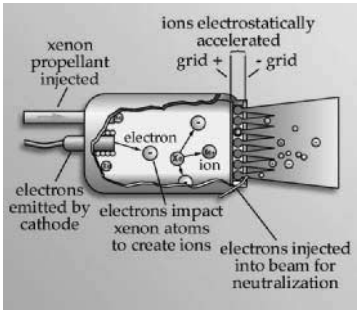


FIGURE 8.19 Simple ion thruster.

ROCKETS AND LAUNCH VEHICLES

Unfortunately, they are also toxic, making them difficult to store and handle. Currently, the most popular propellant for ion thrusters is xenon. Xenon is a safe, inert gas that stores as a dense liquid (4.56 times the density of water) under a moderate pressure of 58.4 bar at room temperature. This high-density propellant also gives ion thrusters excellent density-specific impulse.

Ion thrusters offer an electrically efficient (~90%, meaning 90% of the power goes to accelerating propellant) propulsion with very high specific impulse (as high as 10,000 s). As we discussed earlier, thrust and I_{sp} in electrodynamic rockets are limited only by available power. Ion thrusters have been used on a variety of space missions. Perhaps their most exciting application is on interplanetary missions. NASA's Deep Space 1 mission, which was the first to rely on an ion rocket for the primary propulsion system beyond Earth orbit.

Plasma Thrusters. As we discussed in subsection 8.1, there is a practical limit to the number of ions you can pack into a small space inside a thruster. However, a neutral plasma can have a much higher charge density. Plasma thrusters can take advantage of this fact to offer slightly higher thrust than ion thrusters for the same power input. Plasma thrusters use the combined effect of electric and magnetic fields to accelerate the positive ions within a plasma to high velocities.

Two types of plasma thrusters have seen some initial testing and applications in space:

- Stationary plasma thruster (SPT) (also called Hall effect thrusters, Figure 8.20)
- Pulsed-plasma thrusters (PPT) (Figure 8.21)

SPTs are currently the most widely used type of plasma thruster. They take advantage of a unique effect called a Hall current that is created when you apply a radial magnetic field to a conducting plasma. The interaction of the magnetic field with the resulting electric field creates the force that accelerates the positive ions in the plasma, as illustrated in Figure 8.20.

Russian scientists pioneered many of the modern advances in SPTs, having run them for several years for station-keeping applications. In general, SPTs are slightly less electrically efficient than ion thrusters (around 60%) but can deliver roughly the same thrust and specific impulse. As the propellant requirements for plasma

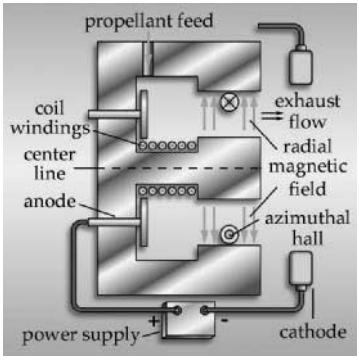


FIGURE 8.20 Hall effect thruster (HET).

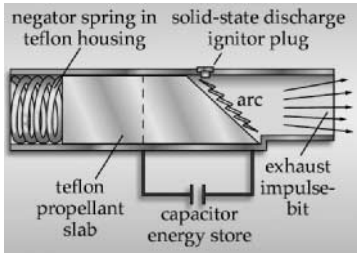


FIGURE 8.21 Pulsed-plasma thruster (PPT).

thrusters are the same as for ion thrusters, xenon is also the most widely used choice.

Unlike all other types of rockets we've looked at that operate continuously, PPTs operate in a noncontinuous, pulsed mode. Unlike ion and plasma thrusters, PPTs use a solid propellant, usually Teflon (PTFE). A high voltage arc pulses over the exposed surface of the propellant, vaporizing it and creating an instant plasma. The resulting induced magnetic field accelerates the plasma. Figure 8.21 shows a schematic for a simple PPT. A number of missions have used PPTs for spacecraft station keeping. Their advantage is precisely controlled, low thrust levels. Because they operate in a pulsed mode, they do not need continuous high power. Instead, much lower power can be stored up slowly in a capacitor for release in high-power bursts (the same technique used in a camera flash). This low-power, pulsed operating mode makes them suitable for many small satellite applications.

Compared to ion and plasma thrusters, PPTs are relatively low in energy-conversion efficiency (20%). However, they provide impressive I_{sp} (up to 1500 s) but with very low thrust (10^{-5} to 10^{-3} N). Their biggest potential advantage is in ease of integration. Because they do not require any additional propellant management, they can be built as simple, self-contained units. Table 8.10 summarizes key information about the electrodynamic rockets we have discussed.

System Selection and Testing

So far, we looked at the all the pieces that make up propulsion systems and many of the various rocket technologies options available. But many questions about propulsion system applications remain.

TABLE 8.10 Electrodynamic Rockets

Type	Propellant	Operating principle	Electrical efficiency	Thrust (N)	I_{sp} (s)
Ion (or electrostatic) thruster	xenon	Applied electric field is used to accelerate ionized propellant	90%	0.1–1.0	2000–10,000
Stationary plasma thruster (SPT)	xenon	Combined electric and magnetic fields produce a "Hall effect" that accelerates ions within a plasma	60%	0.1–1.0	~2000
Pulsed plasma thruster (PPT)	Teflon (PTFE)	An electric arc is discharged in a pulsed mode over a solid propellant, vaporizing it and creating a plasma. Interaction between the applied electric field and resulting magnetic field accelerates the plasma.	20%	10^{-5}–10^{-3}	~1500

Adapted from SPAD.

ROCKETS AND LAUNCH VEHICLES

- How do mission planners go about selecting the best technology from this large menu? How do researchers decide which is the best technology to pursue for future applications?
- How are new or improved systems tested and declared fit for flight?

As with most technology decisions, there is rarely one best answer for any given application. Sometimes, as in the case of our *FireSat* example, the severe constraints on volume, power and mass, coupled with the modest ΔV requirements leave only a few realistic options—cold-gas thrusters, or possibly a monopropellant system. Even when we narrow the field, the choice of the right propulsion system for a given mission depends on a number of factors that we must weigh together.

One way to trade off various rocket options is to select one with the lowest total cost. But here, cost represents much more than simply the engine's price tag. The *total cost* of a propulsion system includes at least eight other factors, in addition to the bottom-line price tag, that we must consider before making a final selection:

1. Mass performance—measured by I_{sp}
2. Volume performance—measured by I_{dsp}
3. Time performance—how fast it completes the needed ΔV, measured by total thrust
4. Power requirements—how much total power the EPS must deliver
5. Safety—how safe the system (including its propellant) is and how difficult it is to protect engineers working with the system
6. Logistics requirements—how difficult it will be to transport the system and propellant to the launch site and service it for flight
7. Integration cost—how difficult the system is to integrate and operate with other spacecraft subsystems and the mission operations concept
8. Technical risk—new, untried systems carry more risk to the mission than proven technology.

Different missions (and mission planners) naturally place a higher value on some of these factors than others. Other missions, such as a complex commercial mission, may place a high priority on reducing technical risk. For them, a new type of plasma rocket engine, even if it offers lower mass cost, may be too risky when they consider all other factors. When asking what the best option is for a given mission, "it depends" is usually the best answer!

Once a system is selected, engineers must conduct a rigorous testing and qualification process before it can be declared safe for use. New rocket development usually progresses from relatively crude engineering model testing under atmospheric conditions to more elaborate testing of flight models under high-altitude or vacuum conditions. Of course, for specialized systems such as electromagnetic thrusters (e.g., ion thrusters or SPTs), testing can only be done under vacuum conditions using highly accurate thrust stands to measure micronewtons (10^{-3} N) of thrust. During experimental testing, rocket scientist carefully measure mass flow rates, chamber pressures, temperatures, and other parameters and compare them to predicted values based on thermochemical and other models.

Because rockets typically involve high pressures, high temperatures, high voltages, and hazardous chemicals, safety issues are a primary concern. These concerns carry through from initial development of new rockets to servicing of proven sys-

tems in preparation for flight. In the case of launch vehicle propulsion, human lives may depend on safe, reliable operation. As discussed earlier, special loading procedures and equipment are used to ensure safe handling of hazardous propellants.

Ensuring system reliability involves a complex series of ground tests measuring performance over a wide range of conditions. These can range from relatively simple tests to ensure the system does not leak at flight pressure, to complicated tests that require widely varying O/F ratios and expansion conditions. In addition to performance, all the typical space environment testing done for other subsystems, such as thermal and vacuum testing must also be accomplished for the propulsion subsystem.

Exotic Propulsion Methods

Chemical rockets have given us access to space and taken spacecraft beyond the solar system. Electrostatic rockets offer a vast increase in mass efficiency making exciting, new missions possible. However, to really open space to colonization and allow humans to challenge the stars, we need new ideas. Exotic propulsion systems are those far out ideas still on the drawing boards. While there are many exotic variations to the rockets we have already discussed (such as using a high-energy density or meta-stable chemicals, nuclear fusion, or antimatter to create superheated products), here we will focus on even more unconventional types of propulsion—ones that produce thrust *without* ejecting mass:

- Solar sails
- Tethers

We will see how these far out concepts can be used to give us even greater access to the solar system. Then we will go beyond that to look at one of the unique challenges of interstellar flight.

Solar Sails. Just as a sail is used to harness the force of the wind to move a ship, a *very* large *solar sail* can be used to harness the force of solar pressure to prepell a spaceship without ejecting mass. Of course, the farther it goes from the sun, the less solar pressure it can collect, so a solar sail would work best inside Mars's orbit.

How large would a sail need to be? This force can be determined from

$$F = F_s/c\, A_s\, (1 + r) \cos I \tag{8.33}$$

where F_s = solar constant = 1358 W/m^2 at Earth's orbit around the sun
c = speed of light = 3×10^8 m/s
A_s = surface area (m^2)
r = surface reflectance (where $r = 1$ for a perfect reflector)
I = incidence angle to the sun (deg)

To produce just 5 N of thrust near Earth (about 1 lb), a 1 km^2 sail would be needed. To achieve escape velocity from a low-Earth orbit (assuming a total spacecraft mass of only 10 kg), this force would have to be applied for more than 17 years. Of course, a solar sail uses no propellant, so the thrust is basically "free." As long as no one is in a hurry, a solar sail offers a cheap way to get around. Solar sails have been proposed to maneuver mineral-rich asteroids closer to Earth to allow for orbital mining operations.

Tethers. Another imaginative means of propulsion that does not need propellant uses very long cables called *tethers*. Typically, these booms are only a few meters long. Using a small mass at the end of a very long tether, tens or even hundreds of kilometers long, produces the same stabilizing effect . But even more interesting effects become possible as well.

Picture a large spacecraft such as the Shuttle in a circular orbit. Now imagine a small payload deployed upward (away from Earth), from the Shuttle at the end of a very long tether as shown in Figure 8.22.

Typically, when we compute orbital velocities we assume we were dealing with point masses effected only by gravity. From an orbital mechanics standpoint, this point mass assumption is valid only at the center of mass of the Shuttle/payload system. If the payload mass is small with respect to Shuttle's mass, the system's center of mass will not move significantly when it deploys. Thus, the orbital velocity of the system will stay about the same. What does this mean for the payload? Secured by the tether, it is being pulled along in orbit at the Shuttle's orbital velocity. But the payload is above the Shuttle. Orbital velocity depends on distance from the center of the Earth. Therefore, because the payload is higher than the Shuttle, its proper circular, orbital velocity would be somewhat *lower* than the velocity it maintains due to the tether. Or, said another way, the tether forces it to travel *faster* than orbital mechanics would dictate for its altitude.

Now, what happens if we suddenly cut the tether? Orbital mechanics would take over and the payload would suddenly find itself at a velocity too fast for a circular orbit at that altitude. The situation would be as if its velocity were suddenly increased by firing a rocket. It would enter an elliptical orbit with a higher apogee one-half orbit later. Analysis indicates this new apogee altitude would be seven times the length of the tether higher than the original circular orbit. In other words, if the original altitude of the payload were 300 km and the tether length were 10 km, the new elliptical payload orbit would be 310 $\times$ 370 km, as illustrated in Figure 8.23.

If the payload were deployed downward instead of upward, the opposite would happen. Its orbit would shrink, so that half an orbit after the tether released, the

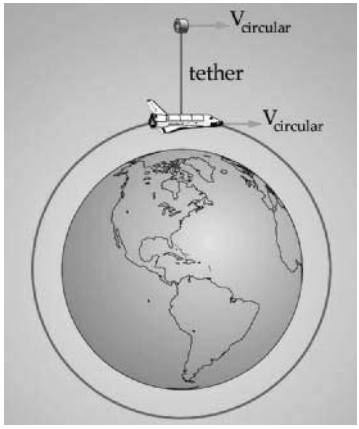


FIGURE 8.22 Space tether deployment.

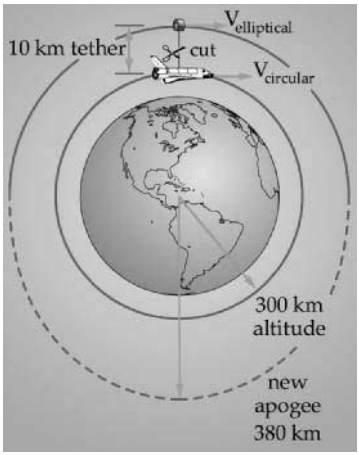


FIGURE 8.23 Tether orbit boost.

payload would reach perigee. This technique was used by the Small Expendable-tether Deployment System (SEDS) mission in 1993 to successfully deorbit a small payload (Humble et al. 1995).

Of course, tether propulsion is not completely "free." We still need to add the mass of the tether and its deployment motors and gears. And we need extra electrical power to operate the tether-deployment mechanisms. However, once we put these systems in place, we could conceivably use the tether system over and over again to boost or deorbit payloads.

Space Shuttle astronauts have performed a number of experiments to investigate the exciting possibilities of tethers. So far, these experiments have focused on the practical problems of deploying, controlling, and reeling in a small payload at the end of a long tether. Future applications for tethers are truly unlimited. A series of rotating tether stations could be used to sling-shot payloads, passing them from one to the other all the way from low-Earth orbit out to the Moon. Another exiting use of tethers is for power generation. A conducting tether passing through the Earth's magnetic field each orbit can generate large amounts of electrical power (Forward and Hoyt, 1999).

Interstellar Travel. The ultimate dream of space exploration is someday to travel to other star systems. Actually, the first human-built starships are already on their way out of the solar system. Launched in 1972 and 1973, NASA's Pioneer 10 and Pioneer 11 probes became the first spacecraft to leave our local planetary neighborhood and begin their long journey to the stars. Unfortunately, at their present velocities, they are not expected to pass near another stellar body in over 2 million and 4 million years, respectively!

Obviously, these travel times are far too long to be useful for scientists who want to be around to review the results from the mission. Hollywood's version of rocket science can take advantage of hyperspace and warp drive to allow round-trip times to nearby stars in the space of a single episode. Unfortunately, real-world rocket science is far from these amazing means of propulsion.

Assuming we could develop efficient on-board energy sources, such as fusion or antimatter, and rely on ion or other extremely efficient types of rockets to achieve very high specific impulse, there is still the limit imposed by the speed of light. If a rocket could thrust continuously for several years, even at a very low thrust level, it would eventually reach very high velocity.

One aspect of Albert Einstein's theory of relativity says that as an object's velocity approaches light speed, its perception of time begins to change relative to a fixed observer. This time adjustment leads to the so-called twin paradox. To visualize this concept, imagine a set of twins, a sister and brother. If the sister leaves her brother and sets off on a space mission that travels near the speed of light, when she returns, she'll find her brother much older than she is! In other words, while the mission seemed to last only a few years for her, tens or even hundreds of years would have passed for her brother.

We express this *time dilation* effect, sometimes called a tau (τ) factor, using the Lorentz transformation:

$$\tau = \frac{t_{\text{starship}}}{t_{\text{Earth}}} = \sqrt{1 - \frac{V^2}{c^2}} \tag{8.34}$$

where t_{starship} = time measured on a starship
t_{Earth} = time measured on Earth
V = starship velocity
c = speed of light = 300,000 km/s

The tau factor, τ, tells us the ratio of time aboard a speeding starship compared to Earth time. As the spacecraft's velocity approaches light speed, τ gets very small, meaning that time on the ship passes much more slowly than it does on Earth. While this may seem convenient for readers thinking about a weekend journey to the star Alpha Centauri (4.3 light years away), Einstein's theory also places a severe speed limit on would-be space travelers. As a spacecraft's velocity increases, its effective mass also increases. Thus, as the ship's velocity approaches light speed, it needs more thrust than it did at lower speeds to get the same velocity change. To attain light speed, it would need an infinite amount of thrust to accelerate it's infinite mass. For this reason alone, travel at or near the speed of light is well beyond current technology.

For years, scientists and engineers said travel beyond the speed of sound, the so-called sound barrier, was impossible. But in October 1947, Chuck Yeager proved them all wrong while piloting the Bell X-1. Today, jet planes routinely travel at speeds two and three times the speed of sound. Perhaps by the 23d century some future Chuck Yeager will break another speed barrier and take a spacecraft beyond the speed of light.

8.3 LAUNCH VEHICLES

Now that we have seen the types of rockets available, let's see how they are used to solve perhaps the most important problem of astronautics—getting into space. Launch vehicles come in many different shapes and sizes, from the mighty Space Shuttle to the tiny Pegasus. In this section, we will start by examining the common

elements of modern launch vehicles. Looking at launch vehicles as systems, we will review the various subsystems that work together to deliver a payload into orbit and focus on the unique requirements for the massive propulsion systems needed to do the job. Finally, we will look at staging to see why a launch vehicles are broken into pieces that are used and discarded on the way to orbit.

Launch Vehicle Systems

A launch vehicle needs most of the same subsystems to deliver a payload (the spacecraft) from the ground into orbit. The two biggest differences between a launch vehicle and a spacecraft are the total operation time (about 10 minutes versus 10^+ years) and the total ΔV needed (>10 km/s versus 0–1 km/s). Let us start by looking at the challenges of launch vehicle propulsion to see how we must adapt the technologies discussed earlier in this chapter to the challenging launch environment. Then we will briefly review the other subsystems needed to support these large rockets and safely deliver spacecraft (and people) into space.

Propulsion Subsystem. The launch vehicle propulsion subsystem presents several unique challenges that sets it apart from the same subsystem on a spacecraft. These include:

- Thrust-to-weight ratio—mmust be greater than 1 to get off the ground
- Throttling and thrust-vector control—may need to vary the amount and direction of thrust to decrease launch loads and allow for steering
- Nozzle design—nozzles face varying expansion conditions from the ground to space.

Let us go through each of these challenges in more detail.

Thrust-to-Weight Ratio. To get a rocket off the ground, the total thrust produced must be greater than the weight of the vehicle. We refer to the ratio of the thrust produced to the vehicle's weight as the *thrust-to-weight ratio*. Thus, a launch vehicle's propulsion system must produce a thrust-to-weight ratio greater than 1. For example, the thrust-to-weight ratio for the Atlas launch vehicle is about 1.2 and the Space Shuttle's about 1.6.

Even though chemical rockets are not as efficient as some options discussed in the last section, they offer very high thrust and, more importantly, very high thrust-to-weight ratios. For this important reason, only chemical rockets are currently used for launch vehicle applications.

Throttling and Thrust Vector Control. For virtually all spacecraft applications, rocket engines are either on or off. There is rarely a need to vary their thrust by *throttling* the engines. However, for launch-vehicle applications, throttling is often needed, greatly adding to the complexity (and cost!) of launch vehicle propulsion systems.

One reason for throttling has to do with the high launch forces that are generated on the vehicle as it flies through the atmosphere. Within the first minute or so of launch, the vehicle's velocity increases rapidly while it is still relatively low in altitude where the atmosphere is still fairly dense. The effect of passing through this dense atmosphere at high velocity produces *dynamic pressure* on a vehicle. Without careful attention to design and analysis, these launch loads could literally rip the vehicle apart. During design, some maximum value is assumed, based on

extensive analysis of expected launch conditions, that cannot be exceeded without risking structural failure. Prior to each launch, engineers carefully measure and analyze the winds and other atmospheric conditions over the launch site to ensure the vehicle won't exceed it's design tolerances. In many cases, this is based on an assumed thrust profile for the vehicle that decreases, or throttles down, during peak dynamic pressure. The Space Shuttle, for example, reduces the main engines' thrust from 104% to 65% and the solid rocket boosters are likewise tailored to reduce their thrust during this phase of flight to keep dynamic pressure below a predetermined, safe level.

Another reason for throttling is to keep total acceleration below a certain level. An astronaut strapped to the top of a launch vehicle feels the thrust of lift-off as an acceleration or *g-load* that pushes him or her back into the seat. From Newton's laws, we know the total acceleration depends on the force (thrust) and the total mass of the vehicle. If the engine thrust is constant, the acceleration will gradually increase as the vehicle gets lighter due to expended propellant. This means the acceleration tends to increase over time. To keep the overall g-load on the Space Shuttle under 3 g's, the main engines throttle back about six minutes into the launch to decrease thrust to match the propellant that is used.

Some vehicles also need throttling for landing. The decent stage engine in the Lunar Excursion Module (LEM) used during the Apollo missions allowed an astronaut to throttle the engine over a range of 10–100% so they could make a soft touch down on the lunar surface.

Finally, launch vehicle rockets often have the unique requirement to be able to vary their thrust direction to allow for steering. This *thrust vector control* (*TVC*) is usually managed by gimbaling the entire engine to point the thrust in the desired direction. The Space Shuttle, for example, can vary the thrust direction for each main engine by $\pm 10°$. Of course, the mechanical gears and hydraulic actuators needed to move massively thrusting rocket engines can be quite complicated. Earlier rockets used more simple methods of thrust vector control. The V-2 rocket, for example, used large, movable ablative vanes stuck into the exhaust to change direction. Other launch vehicles use separate steering rockets or direct injection of gases into the exhaust flow to change thrust direction.

Nozzle Design. In subsection 8.1 we discussed the importance of external pressure and nozzle expansion ratio to overall engine performance. We prefer not to have a rocket nozzle either overexpanded or underexpanded, but instead designed for ideal expansion. In comparison, spacecraft rocket engines always work within a vacuum and designers simply use the greatest expansion ratio possible for the best performance. For launch vehicle rocket engines, the choice of expansion ratio is not so simple.

During launch, the external pressure on the first stage engines goes from sea level (1 bar or 14.7 psi) to near zero bars (vacuum) in just a few minutes. Ideally, we would like the nozzle to increase its expansion ratio throughout the trajectory to change the exit pressure as atmospheric pressure decreases. Unfortunately, with current technology, the weight of the hardware to do this is far too great. Instead, we design the nozzle to achieve ideal expansion at some design altitude about two-thirds of the way from the altitude of engine ignition to the altitude of engine cut-off.

For example, if we design a rocket to go from sea level to 60,000 meters, a reasonable choice for the desired exit pressure would be the atmospheric pressure at about 40,000 meters altitude. As a result, our rocket would (by design) be over-expanded below 40,000 meters and under-expanded above 40,000 meters. As we

see in Figure 8.24, a nozzle designed in this way offers better overall performance than one designed to be ideally expanded only at sea level.

Guidance, Navigation, Control Subsystem. A launch vehicle must deal with the issue of attitude determination and control in a much more dynamic environment than that faced by a spacecraft. The guidance, navigation and control (GNC) subsystem keeps the launch vehicle aligned along the thrust vector to prevent dangerous side loads, keeps the thrust vector pointed according to the flight profile, and ensures the vehicle reaches the correct position and velocity for the desired orbit.

As with all control systems, the GNC subsystem has actuators and sensors. The primary launch vehicle actuators are the main engines, which make use of TVC and throttling to get the rocket where it needs to go. GNC sensors are typically include accelerometers and gyroscopes to measure acceleration and attitude changes. Even though the accuracy of these sensors drifts over time, they are usually sufficiently accurate for the few minutes needed to reach orbit. Future launch vehicles will use the Global Positioning System (GPS) for position, velocity, and attitude information.

Communications and Data Handling. Throughout launch, the vehicle must stay in contact with the Launch Control Center. There, flight controllers constantly monitor telemetry from the launch-vehicle subsystems to ensure they are functioning nominally. To do this, the vehicle needs a communication and data handling subsystem to process on-board data and deliver telemetry to the ground. Launch-vehicle data handling is very similar to a spacecraft's. Computers process sensor information and compute commands for actuators, as well as monitor other on-board processes. On expendable vehicles, these subsystems can be relatively simple because they need to work for only a few minutes during launch and will not be exposed to long periods of space radiation.

Communication equipment is also very similar in concept to that found on spacecraft. However, for safety reasons, an independent means of tracking the launch vehicle's location on the way to orbit is needed. In the Launch Control Center, Range Safety Officers monitor a launch vehicle's trajectory using separate tracking radar, ready to send a self-destruct command if it strays much beyond the planned flight path to endanger people or property.

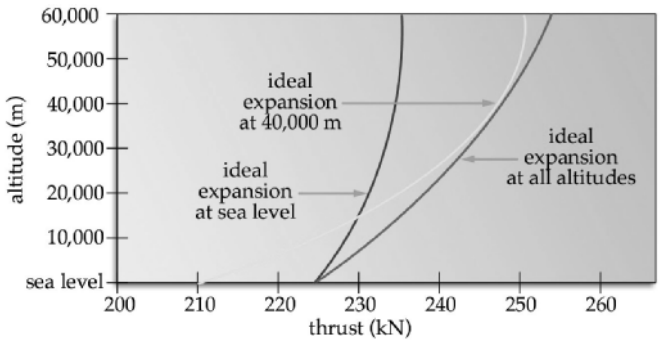


FIGURE 8.24 Thrust versus altitude for different nozzle designs.

Electrical Power. Launch vehicle electrical power requirements are typically quite modest compared to a spacecraft's. Launch vehicles need only enough power to run the communication and data handling subsystems, as well as sensors and actuators. Again, because of their limited lifetimes, expendable launch vehicles typically rely on relatively simple batteries for primary power during launch. The Space Shuttle uses fuel cells powered by hydrogen and oxygen.

Structure and Mechanisms. Finally, we must design the launch vehicle's structures and mechanisms to withstand severe loads and perform the numerous mechanical actuations and separations that must happen with split-second timing. A typical launch vehicle can have tens or even hundreds of thousands of individual nuts, bolts, panels, and load-bearing structures that hold the subsystems in place and take the loads and vibrations imposed by the engines' thrust and the atmosphere's dynamic pressure. Since the majority of a launch vehicle's volume is propellant tanks, these tend to dominate the overall structural design. Often the tanks become part of the primary load bearing structure. For the Atlas launch vehicle, used during the Mercury program, the thin-shelled tank literally inflated with a small positive pressure to create the necessary structural rigidity.

In addition to the problem of launch loads and vibrations, hundreds of individual mechanisms must separate stages and support other dynamic actions throughout the flight. These mechanisms tend to be larger than similar mechanisms on spacecraft. During staging, large sections of the vehicle's structure must literally break apart, usually by explosive bolts. Gimbaling the massive engines to change their thrust direction requires large hinges, hydraulic arms, and supporting structure.

Launch vehicle designers have the challenge of carefully integrating all of these structures and mechanisms with the engines, tanks and other subsystems to create a compact, streamlined vehicle. Sadly, for expendable vehicles, all the painstaking design and expensive construction and testing to build a reliable launch vehicle burns up or drops in the ocean within 10 minutes after launch.

Staging

Getting a payload into orbit is not easy. As we learned in subsection 8.2, the state of the art in chemical rockets (the only type currently available with a lift-to-weight ratio >1) can only deliver a maximum I_{sp} of about 450 s. Given the ΔV needed to get into orbit, and the hard realities of the rocket equation, that means most of a launch vehicle will be taken up by propellant. In fact, over 80% of a launch vehicle's lift-off mass is propellant. All of this propellant must be contained in large propellant tanks, that also add mass. Of course, the larger the mass of propellant, tanks, and other subsystems, the less mass is available for payload. One way of reducing the vehicle's mass is to somehow get rid of mass that is no longer needed. After all, why carry all that extra tank mass along when the rocket engine empties the tank steadily during launch anyway? Instead, why not split the propellant into smaller tanks and then drop them as they empty? Fighter planes flying long distances use this idea in the form of drop tanks. These tanks provide extra fuel for long flights and are dropped as soon as they are empty, thus lightening and streamlining the plane. This is the basic concept of staging.

Stages consist of propellant tanks, rocket engines, and other supporting subsystems that are discarded to lighten the launch vehicle on the way to orbit. As each

stage burns out, it is dropped off and the engines of the next stage ignite (hopefully) to continue the flight into space. As each stage drops off, the vehicle's mass decreases, meaning a smaller engine can keep the vehicle on track into orbit.

Table 8.11 gives an example of how staging can increase the amount of payload that can be delivered to orbit. For this simple example, notice the two-stage vehicle can deliver more than twice the payload to orbit as a similar-sized, single-staged vehicle with the same total propellant mass—even after adding 10% to the structure's overall mass to account for the extra engines and plumbing needed for staging. This added payload-to-orbit capability is why all launch vehicles currently rely on staging.

In Table 8.11, for both cases, the size of the payload delivered to orbit compared to the weight of the entire launch vehicle is pretty small—5% or less. About 80% of a typical vehicle is propellant. The other 15% or so is made up of structure, tanks, plumbing, and other subsystems. Obviously, we could get more payload into space if only our engines were more efficient. However, with engines operating at or near the state of the art, the only other option, as the examples show, is to shed empty stages on the way into orbit.

Now let's see how we use the rocket equation to analyze the total ΔV we get from a staged vehicle. We start with

$$\Delta V = I_{\text{sp}} g_o \ln \left(\frac{m_{\text{initial}}}{m_{\text{final}}} \right)$$

Recognize that for a staged vehicle, each stage has an initial and a final mass. Also, the I_{sp} may be different for the engine(s) in different stages. To get the total ΔV of the staged vehicle, we must add the ΔV for each stage. This gives us the following relationship for the ΔV of a staged vehicle with n stages.

$$\Delta V_{\text{total}} = \Delta V_{\text{stage 1}} + \Delta V_{\text{stage 2}} + \ldots + \Delta V_{\text{stage } n} \tag{8.35}$$

$$\Delta V_{\text{total}} = I_{\text{sp stage 1}}\, g_o \ln \left(\frac{m_{\text{initial stage 1}}}{m_{\text{final stage 1}}} \right) + I_{\text{sp stage 2}}\, g_o \ln \left(\frac{m_{\text{initial stage 2}}}{m_{\text{final stage 2}}} \right) + \cdots + I_{\text{sp stage } n} g_o \ln \left(\frac{m_{\text{initial stage } n}}{m_{\text{final stage } n}} \right) \tag{8.36}$$

where ΔV_{total} = total ΔV from all stages (m/s)
$I_{\text{sp stage } n}$ = specific impulse of stage n (s)
g_o = gravitational acceleration at sea level (9.81 m/s^2)
$m_{\text{initial stage } n}$ = initial mass of stage n (kg)
$m_{\text{final stage } n}$ = final mass of stage n (kg)

What is the initial and final mass of stage 1? The initial mass is easy; it is just the mass of the entire vehicle at lift-off. But what about the final mass of stage 1? Here we have to go to our definition of final mass when we developed the rocket equation. Final mass of any stage is the initial mass of that stage (including the mass of subsequent stages) less the propellant mass burned in that stage. So for stage 1,

$$m_{\text{final stage 1}} = m_{\text{initial vehicle}} - m_{\text{propellant stage 1}}$$

Similarly, we can develop a relationship for the initial and final mass of stage 2, stage 3, and so on.

ROCKETS AND LAUNCH VEHICLES

TABLE 8.11 Comparing a Single-Stage and Two-Stage Launch Vehicle

Launch vehicle	Parameters	Payload to orbit
Single stage	$\Delta V_{design} = 8000$ m/s $I_{sp} = 480$ s $m_{structure} = 250$ kg $m_{propellant} = 1500$ kg	$m_{payload} = 84$ kg
Two Stage	$\Delta V_{design} = 8000$ m/s **Stage 2** $I_{sp} = 480$ s $m_{structure} = 140$ kg $m_{propellant} = 750$ kg **Stage 1** $I_{sp} = 480$ s $m_{structure} = 140$ kg $m_{propellant} = 750$ kg	$m_{payload} = 175$ kg

$$m_{\text{initial stage 2}} = m_{\text{final stage 1}} - m_{\text{structure stage 1}}$$

$$m_{\text{final stage 2}} = m_{\text{initial stage 2}} - m_{\text{propellant stage 2}}$$

Overall, staging has several unique advantages over a single-stage vehicle. It

- Reduces the vehicle's total weight for a given payload and ΔV requirement
- Increases the total payload mass delivered to space for the same-sized vehicle
- Increases the total velocity achieved for the same-sized vehicle
- Decreases the engine efficiency (I_{sp}) required to deliver a same-sized payload to orbit

All of these staging advantages come with some drawbacks. These include

- Increased complexity because of the extra sets of engines and their plumbing
- Decreased reliability because we add extra sets of engines and the plumbing for the upper stages
- Increased total cost because more complex vehicles cost more to build and launch

Another interesting limitation of staging has to do with the law of diminishing returns. So far, you may be ready to conclude that if two stages are good, four stages must be twice as good. But this is not necessarily the case. Although a second stage significantly improves performance, each additional stage enhances it less. By the time we add a fourth or fifth stage, the increased complexity and reduced reliability offset the small performance gain. That is why most launch vehicles have only three or four stages.

Engineers working on single-stage to orbit (SSTO) concepts must overcome many technical challenges in propulsion and materials to make this idea feasible. However, the great promise of a completely reusable launch vehicle that can operate like an airplane offers the potential for operations far more cost effective than current expendable vehicles.

REFERENCES

Forward, R. L. and Hoyt, R. P. 1999. "Space Tethers," *Scientific American,* February, pp. 66–67.

Humble, R. J., Gary, H. N., and Larson, W. J. 1995. *Space Propulsion Analysis and Design,* McGraw-Hill, New York.

SECTION 9
AEROSPACE STRUCTURES

Section Editor: Glyn Davies

This section addresses the nature and role of aircraft structures. As appropriate for a handbook of this size, it avoids going into too much detail when the reader may have a broader interest in any of the topics in the other 17 sections. The aim is to help the reader understand the problems, the solutions, and the importance of getting the structure right. This latter point may seem obvious; i.e., the structure has to be safe for all likely loadings and for the operating life of the aircraft. However the safety should not be achieved in an overdesigned structure which is too heavy. A typical dry structure weight for a large civil aircraft can be 40% of the total all-up weight, when the payload may be only 20%. Every kilogram of structure overweight sacrifices 2 kg of payload. Or, to put it another way, if the structure is 10% too conservative (4% of all-up weight) the payload is down by 20%, and this can mean the difference between profit and loss for competitive routes.

The cost of the structure and its manufacture is clearly important, but so are the cost and duration of design and production. A revolution has taken place in the field of virtual everything (electronic mock-up, aerodynamics simulation, testing, systems, virtual manufacturing. . . . the list is endless). We will therefore discuss what can be simulated in structural performance and testing and how it is done with modern software tools. We will also demonstrate some very simple analysis just to show the physics of the structural behavior and the parameters which are important and those which are not. If the reader wishes to pursue structural analysis in any greater depth, there are many texts, ranging from the simple to the sophisticated (Megson 1997; Curtis 1997; Bruhn 1965; Kuhn 1956; Rivello 1969).

PART 1

AIRCRAFT LOADINGS

Glyn Davies

Limit loads are the maximum anticipated by experience, design, or wind tunnel tests. Under limit load the structure and the controls must not deform so much that the aircraft becomes unsafe or difficult to control.

Ultimate loads are limit loads $\times$ 1.5 for civil aircraft, and they must not cause structural failure. The factor 1.5 is a safety factor which experience has shown to be safe and reasonable. It represents the uncertainty in:

- Expected loads
- Structural calculations and tests
- Materials specification
- Dimensions of components

This factor of safety of 1.5 sounds generous, but if the errors in all four of the above items were all equal (an unreasonable extreme value case) then each one would have an uncertainty factor of $1.5^{1/4} = 1.106$. It would be an incautious engineer who believed estimates or specification were better than 10%!

Some of the variables, such as maximum velocity, acceleration, and maneuver rate, should be under the control of the pilot (or the computer), but gust loadings, for example, are not. The loadings and environment for which a civil aircraft has to be designed can be found in the Federal Aviation Administration's Federal Airworthiness Regulations (FAR25) in the United States. In Europe the joint airworthiness requirements (JAR25) are virtually identical. The military specification depends very much on the role for which the aircraft is designed, whether it be agile combat, long-range transport, stealthy bomber, and so on. In the end the airworthiness authorities (or the Departments of Defense) will negotiate with the manufacturer on any new aircraft which differs significantly from the conventional in configuration, flight regime, materials, etc.

The most severe loading cases are all actually dynamic due to, say, a pull-out maneuver or an upward gust which changes the effective incidence; therefore the lift is suddenly increased and the aircraft accelerates vertically. However, we find it convenient to represent the inertia loadings due to this acceleration in equilibrium with the enhanced aerodynamic loading, and then analyse the problem as one of static equilibrium, as shown for the wing in Figure 9.1 which has to support the fuselage and all wing masses, including fuel. Without going into quantitative details at this stage, this figure illustrates the structural advantages of mounting engines on the wing, since it is desirable to have the aerodynamic lift distribution opposed by a similar weight distribution. If both distributions were identical, the wing would not bend at all. This aspect is one attraction of the flying wing or the blended wing concept.

The extremes of the flight envelope, for which the structure has to be designed, are summarized in the n-V diagrams for maneuver and gust. The load factor n is the induced acceleration expressed as a multiple of g. The following definitions apply:

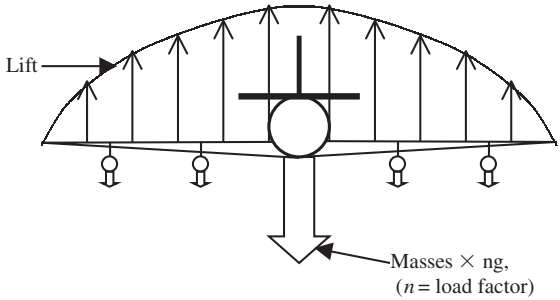


FIGURE 9.1 Aerodynamic and inertia forces during gust or maneuver.

V_c = design cruising speed
V_d = design dive speed sufficiently greater than V_c to provide for safe recovery after some possible emergency
V_A = maximum speed with flaps up

Figure 9.2 shows a typical n-V diagram. The stall line OA represents the maximum load factor n on the aircraft weight W which can be supported by the maximum wing loading, i.e.,

$$nW = \tfrac{1}{2}\rho V^2 S C_{L\max}$$

where C_L is the lift coefficient based on a wing area of S.

The maximum positive and negative values of n differ for transport, semiaerobatic, and aerobatic civil aircraft. Figure 9.3 shows a large civil aircraft. For military

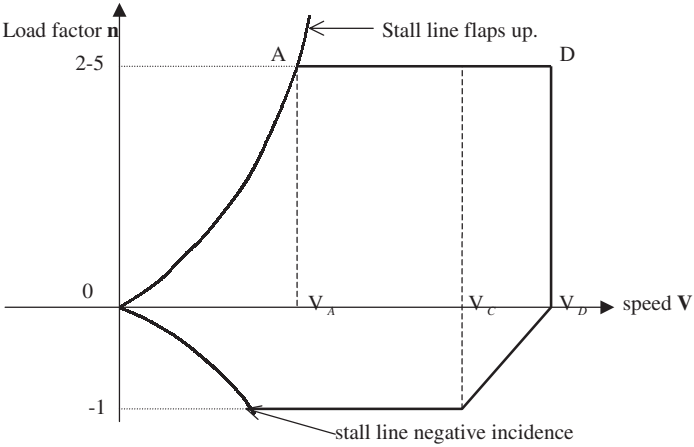


FIGURE 9.2 Maneuver envelope.

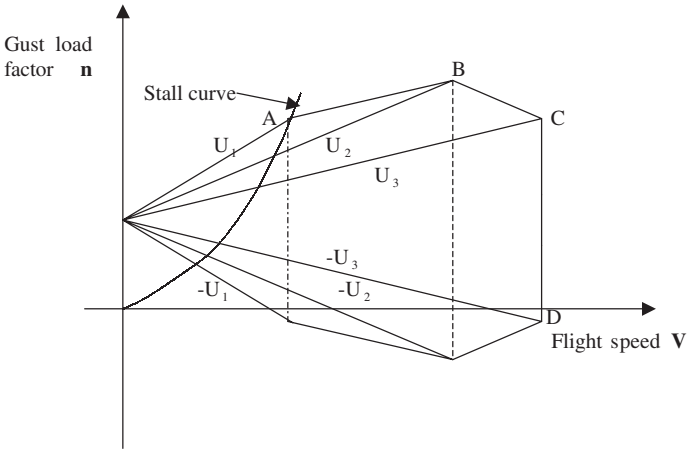


FIGURE 9.3 Gust envelope.

aircraft the design maneuver envelope depends entirely on the specification for the particular aircraft role.

A similar envelope can be constructed for the gust-induced lift and is shown in Figure 9.3. In the simplest concept a vertical sharp-edged gust velocity U causes a change in incident U/V and an instant increment in lift which causes an acceleration depending on the current weight. Thus, the increase in lift for an upward gust U is given by

$$\Delta L = \tfrac{1}{2}\rho V^2 \mathrm{S}\ \partial C_L/\partial\alpha \cdot U/V$$

giving an increase in load factor

$$\Delta n = \tfrac{1}{2}\rho V\ \partial \mathrm{C}_L/\partial\alpha \cdot U/w$$

where $w = W/S$, the wing loading. This simple value is for a sharp-edged gust. An alleviation factor may be introduced which depends on the finite dimensions of an aircraft, the wing loading w, $\partial C_L/\partial\alpha$, and the air density. The values of gust velocities may be selected to match the likelihood of occurrence at cruise or dive and also be reduced with altitude. The n-V diagram in the figure shows the intercepts for U_1, U_2, and U_3 where the airworthiness authorities allow values of

Rough air gust, U_3 = 20 m/s (66 ft/s)
High speed gust, U_2 = 15.25 m/s (50 ft/s)
Dive speed gust, U_1 = 7.5 m/s (25 ft/s)

All this is for a rigid aircraft. Nowadays the dynamic response of a flexible aircraft to a gust signature has to be evaluated to see if any overshoot transients occur, i.e., the resonant period of the aircraft may be close to the gust duration.

Although the above two envelopes can be used to design the overall wing and fuselage structures for symmetric wing loadings, some parts of the structure will be designed for different loading cases. For example, the landing-with-maximum all-up-weight may design the flaps and the undercarriage together with the struc-

tures to which they are attached. The fin (vertical stabilizer) will need to cope with the asymmetric engine thrust case in the event of one or two engines failing in a four-engine configuration. The tail-plane (horizontal stabilizer) may be designed by the loads necessary to achieve the design cases in Figure 9.2 or Figure 9.3. The fuselage critical design case may be, for the aft fuselage, the inertia and tail-plane loading during maneuver. But it is also likely to be a fatigue critical design case due to the pressure differential at altitude and occurring once per flight. There are also various crash or impact events for which civil or military aircraft need to be designed. For civil aircraft the heavy landing case is covered by stipulating the maximum forward and vertical decelerations (15 g and 5 g) and then designing the passenger seats not to break free, or the wing box for the fuel must not burst open in a swept wing with the crash-induced longitudinal hydrostatic pressure gradient. Other impact cases may be engine bursts, bird-strike, or low-velocity impact due to mishandling (this is particularly crucial for composite structures which are very vulnerable to impact loads, and the damage, delamination, or fiber breakage may be internal and invisible to visual inspection).

The uncertainty of predicting, analytically or by test, the behavior of the structure, particularly the fatigue life, is reflected in the emergence of a fail-safe structure, i.e., one that is so redundant that the failure of a component at limit load is not disastrous if the remaining structure can survive. This has now lead to the concept of *damage tolerance,* in which any cracks will not propagate to an unstable rate between inspections. The manufacturer is obliged to state the inspection period needed and the method of detection to suit customer requirements.

PART 2

PROPERTIES OF MATERIALS

Glyn Davies

To understand the behavior of aircraft structures, it is necessary to appreciate the properties of the basic materials from which they are made. We will look at the mechanical properties of metals and composites, but cost also is a factor in choosing the material and the fabrication process. Actually, the basic cost of materials is a small percentage of the whole (see Figure 9.4), so the cost should not be a main driver if significant weight savings can be made or if maintenance costs can be reduced with a material whose wear and tear or fatigue performance is superior. The situation in engines is quite different where exotic materials, able to withstand high temperatures, for example, are well worth the extra cost to the operator.

9.1 METALS

The most common aircraft materials are alloys of aluminum, titanium, and steel, all of which would be described as ductile, so some definitions are needed. Figure 9.5 shows a typical stress/strain curve. The initial portion is linear with a slope given by $\tan^{-1}E$, where E is Young's Modulus, or material stiffness. As the strain increases, eventually the dislocations propagate and irreversible plasticity occurs. If the material is unloaded it will recover in a linear fashion as shown where the residual strain of 0.002 has been selected. The stress from which this unloading occurred is known as the 0.2% proof stress. It is a convenient measure of the degree of early plasticity and is used in estimating proof load deformations. The local tangent modulus, E_t, decreases gradually with strain and ultimately a (tensile) failure occurs at the ultimate strain ε_u which is often used as a measure of the ductility.

Plasticity does lower the potential of a metal to resist high stresses, and most alloying techniques are aimed at delaying the dislocations. However, there are significant advantages as well. Clearly most metal-forming techniques need irreversible plasticity, whether the materials are being pressed, stretched, spun, or forged. Common mild steel is the most ductile of all materials and most consumer goods use it still, as does the civil engineering industry, which relies on cold-rolled sections. Another advantage of ductility is to render the structure less sensitive to local stress concentrations caused by discontinuities in geometry such as holes and fillets. If the stress concentration exceeds the proof stress, then local plasticity will defuse the stress field and leave behind an advantageous residual stress field. In the case of cyclic loading and fatigue, the working stress field may be elastic until a crack is formed, and then crack-tip plasticity can have a beneficial effect, as discussed later.

Table 9.1 compares the mechanical properties of the three alloys. The values for any alloy can be enhanced by heat treatment and work hardening, but this can make the material more brittle. The values quoted in the table are therefore average values for high-performance aerospace materials. We see that both the stiffness and the strengths increase as we go from aluminum to titanium to steel, but surprisingly the specific values (divided by density) do not vary that much. Titanium is the most expensive and is difficult to machine, but it does have superior toughness and

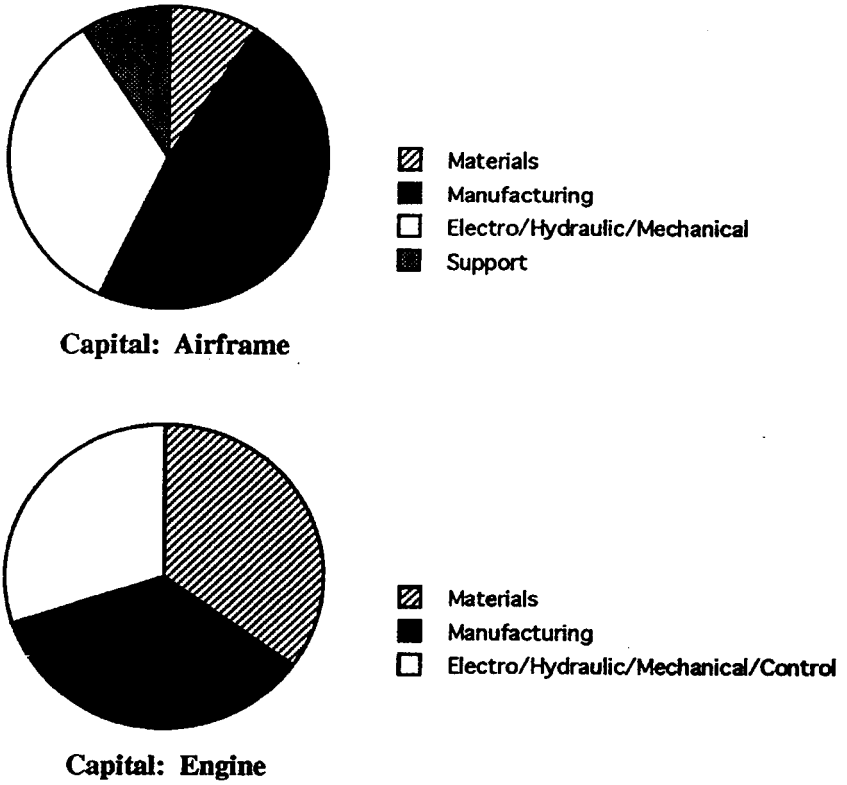


FIGURE 9.4 Relative capital cost of airframes and engines.

fatigue properties. The main reason for using it is that it retains its strength at much higher temperatures than aluminum. For example, aluminum alloys should not be used above 175°C, whereas titanium can be used up to 620°C. It is therefore to be found wherever a firewall is needed between a hot engine and the primary structure. As a matter of interest, titanium can be used for cheap manufacture of complex multilevel sheet components using the properties of superplastic forming which occur as the temperature approaches 1000°C. The metal flows like a fluid, and large strains can be accommodated without the sheet necking or creating voids. At the same time, diffusion bonding can occur between mated surfaces in which the microstructure is as good as the parent metal. Access doors in the Airbus wings are made this way, for instance. Titanium is favored for diffusion bonding since for aluminum alloys the surface oxidizes at the superplastic temperatures.

Steel has the superior strength and stiffness, but this can lead to very thin sheet at the design stresses, and buckling resistance will be poor. However, the smallest volume is a positive advantage for large forgings such as undercarriages where the available space may be limited—particularly for military aircraft.

Static strength is not the only requirement: fatigue and toughness can be important, particularly for aircraft where the loading history is cyclic due to gust maneuvers, pressurization, and landing and takeoff. A modern civil aircraft may

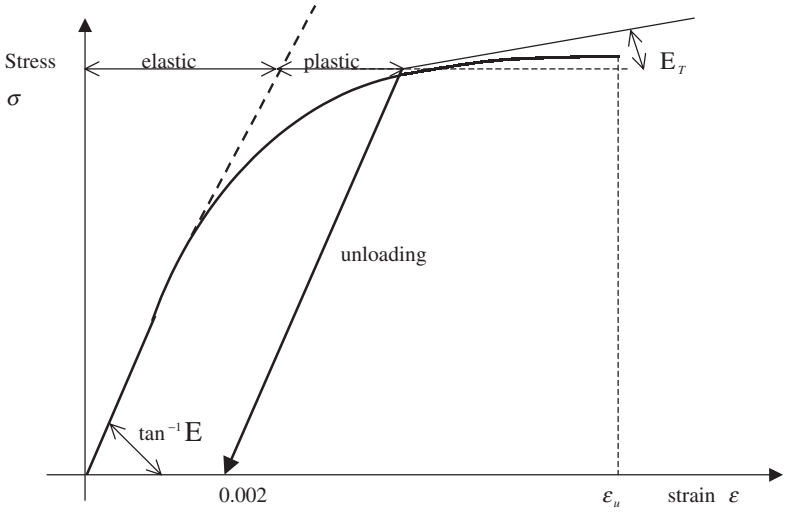


FIGURE 9.5 Stress–strain curve for ductile material.

TABLE 9.1 Comparison of Four Aerospace Metallic Materials

	E(GPa)	σ(UTS) (Mpa)	ε_U(%)	ρ (Mg/m^3)	E/ρ	σ/ρ
Aluminum alloy	71	482	3.0	2.80	25	172
Titanium alloy	110	1,000	2.5	4.43	25	226
Steel	207	1,723	7	7.78	26	221
Berylium alloy	304	207		2.0	151	104

have to be in operation for 25 years, covering 80–120,000 hours and 20,000 flights. For military aircraft however the figures would be about a tenth of this. In selecting materials with good fatigue performance the *S–N* curve is widely used. Figure 9.6 shows a typical family of these curves.

The *S–N* curve on its own does have limitations, primarily because the loading for typical aircraft spectrum does not have a constant stress amplitude, apart from pressure loading in the fuselage, and a way is therefore needed for estimating the cumulative damage for quite complex loading histories. A common way of estimating this is the Palmgren–Miner law (Miner 1954). If the number of cycles is n_i at a stress level where failure would occur at N_i cycles, the measure of cumulative damage is n_i/N_i, and for a variety of levels the Palmgren–Miner criterion for failure is

$$\sum (n_i/N_i) \leq 1$$

This so-called safe life philosophy works quite well but suffers if a particularly high tensile stress is applied, say early on in life. This could cause residual compressive stresses in an area of high stress concentration, and this would positively inhibit subsequent damage creation. The above cumulative damage law would be

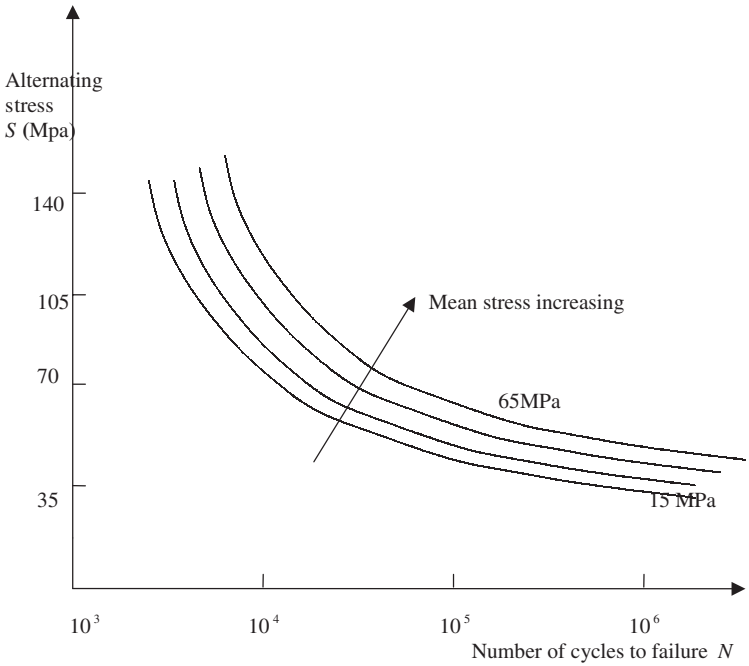


FIGURE 9.6 Typical *S-N* Curves for aluminum alloys.

pessimistic. The worst alternative likely, particularly for military aircraft or civil aircraft operating from rough terrains, would be damage in (say) an area of tensile stress concentration, and then a high compressive load being applied. This would leave a residual tensile stress concentration. Because of this uncertain empiricism, the tendency in all fields is now to adopt a damage-tolerance criterion. This comes from the original fail-safe ideas, where a structure was designed to be so redundant that a failure of one component would not be disastrous if the remaining load paths were adequate. The modern damage-tolerant view says that there may be a crack which propagates only in a stable fashion, under the given working stresses, until it reaches a critical length. The operator therefore has to have an inspection period which will reveal such a crack well before this length is reached. Experimental calibrations are usually necessary to convert a local stress field to a crack propagation law, but theoretical estimates are possible and indeed necessary for any designer. The first need is to define the conditions for propagation of a crack where the elastic stress concentration factor at a crack tip is theoretically infinite. This is most easily done by evaluating the elastic strain energy released when a crack increases by a small amount and equating this to the energy needed to create the new crack surface. The following model, greatly simplified, brings out the physics.

Imagine in Figure 9.7 a crack of length $2a$, in a plate of thickness t, subjected to an applied stress field σ. The stresses have to flow around the crack, leaving a zone where the stresses are now much reduced. If we make the radical assumption that this softened zone is circular and the stresses in it are now negligible, then the strain energy lost is given by

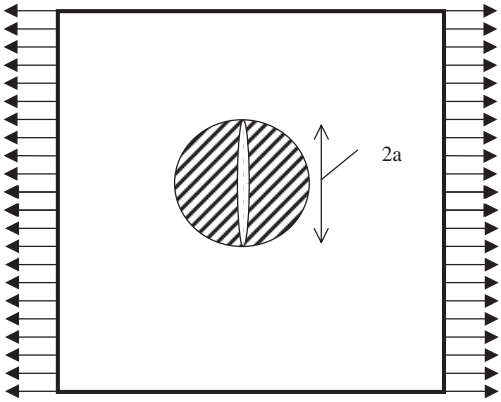


FIGURE 9.7 Crack propagation.

$$\iint \sigma d\varepsilon \cdot dV = (\sigma^2/2E) \cdot t \cdot \pi a^2$$

A perturbation of the crack by δa will release further energy, so, taking an increment of the above, we have

$$\sigma^2 t \pi a \cdot \delta a / E$$

The critical propagation threshold will be given when this is equal to the energy needed to create the new surface $t \cdot \delta a$, i.e., $G_c \cdot t \cdot \delta a$, where G_c is the critical energy release rate for the material and is a material property akin to surface tension. Equating the two gives the critical stress as

$$\sigma_c^2 = \beta \cdot EG_C / \pi a$$

so the critical stress varies like $a^{-1/2}$, which explains why there is a rapid decrease in strength as a crack grows. The stress field around the crack has been oversimplified in this demonstration. One would expect some unknown coefficient β to be present in the above formula. Catalogues are available of β or the equivalent stress intensity factor (explained next) for various cases such as cracks at the edges of holes or near a free edge or a stiffener—in fact, any situation where the stress field differs from that around a small crack in a large plate (Rooke and Cartwright 1976).

In the case of ductile metals, there is also a very efficient energy absorber in the form of local crack tip plasticity, which will leave behind a thin skin of residual stress as the crack propagates. This effect is included in the effective value of G_c found from coupon tests and is typically a factor of 10 greater than the energy needed simply to create a new surface.

An alternative explanation of the above uses the concept of stress intensity factor, which is a measure of the strength of the singular stress field at the crack tip. The stress intensity factor may be written as

$$K = \sigma \sqrt{\pi x a}$$

so we see that $K^2 = EG_C$. The concept of the strength of a stress field which is infinite is not as obvious as the energy release explanation, but the idea has been

around longer. Rooke and Cartwright (1976) is in fact a compendium of stress intensity factors. There are also a large number of software tools which will evaluate stress intensity factors for cracks of any size embedded in various stress concentrations.

All this discussion centers around a crack opening mode, but there are two other shear modes which have different critical values. Figure 9.8 shows all three possible modes. When we come to laminated composites, the possibility of cracks occurring as a delamination between laminae is a weakness of these materials, since a typical matrix will be brittle in nature. It is usually found that the critical energy release rate, or fracture toughness, almost doubles as the mode changes from I to II and then to III.

The energy release or stress intensity factor measure turns out to be useful in formulating a cyclic fatigue law for the damage tolerance approach which is able to cope with complex loading histories. The simple Paris et al. (1972) law says that the cyclic crack growth rate is given by

$$\partial a / \partial N = \mathrm{C}(\Delta K)^n$$

where ΔK is the amplitude of the stress intensity factor ($K_{\max} - K_{\min)}$ and C and n are constants to be found for the material. This law is not strictly true for the initiation stage, nor for the final stages when the growth becomes unstable, so several modifications have been made. The most comprehensive (Foreman 1967) is

$$\frac{da}{dn} = \frac{C \cdot K_{\max} \Delta K^n}{(K_{1C} - K_{\max})}$$

where K_{1C} is the critical value for the initiation growth under the static loading. A

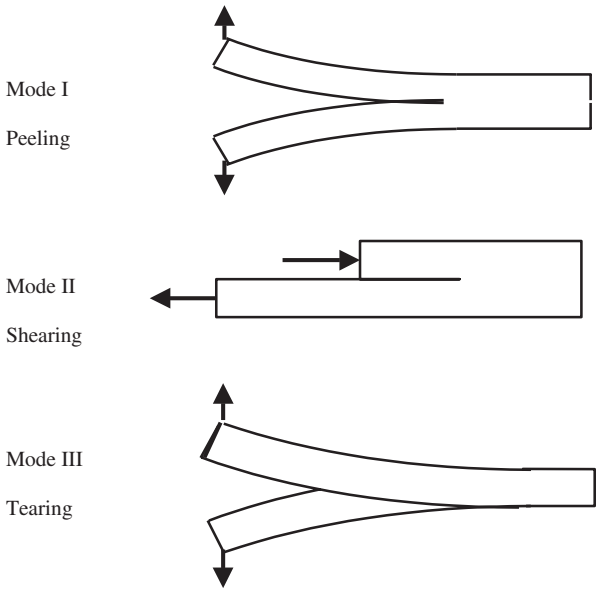


FIGURE 9.8 Three crack modes.

further mechanism for retarding the crack growth occurs if there is an overload causing an increase in crack tip plasticity. Thus, crack growth curves can be generated over a full range of stress and cycle numbers and hence an accurate inspection interval selected.

9.2 COMPOSITES

Composite aerospace structures have been around for more than 30 years. We shall mention only the most common type, where very fine fibers are embedded in a resin, forming unidirectional laminae (prepreg). The fibres are then laid up in a chosen stacking sequence to suit the designer. The most popular combination has been carbon fibers in an epoxy resin which is cured under a raised temperature and pressure (thermosets). Carbon is a brittle material which does not suffer dislocations but is sensitive to the most minor scratch or flaw. The fibers are only microns in diameter, and it has been found that a typical unflawed length is of order 15 to 30 diameters. These flaws in adjacent fibers will never coincide, so the resin is well able to provide a load path bypassing the fiber fracture. This is the basis for laminated composite structures.

Table 9.2 shows the values for strength and stiffness for pure fibres with no flaws and for unidirectional composites having fiber volume fractions of 60%. The previous metal figures are included in the table for comparison. Again, these values are average for aerospace materials. The strengths are higher than the metal figures and the stiffnesses comparable. Where composites score, of course, is in their low density. The specific stiffness and strength figures are much higher, with the exception of the glass stiffness. Of the three composites, the highest value of specific stiffness is for carbon composite and is better by a factor of 4 than that for any metal. It is stiffness which controls the buckling performance of structures, and more than 70% of the primary aircraft structure is designed by buckling. High-performance military aircraft need stiffness for their aeroelastic margins, and few could have been built without using carbon composites.

We have mentioned the poor transverse strength perpendicular to the fibers, and for most structures where the stress field is not everywhere unidirectional the designers will choose to assemble each ply into a laminate with a stacking sequence

TABLE 9.2 Mechanical Properties of Metallic and Polymer Composites

	E(GPa)	σ(UTS) (Mpa)	ε_U(%)	ρ(Mg/m^3)	E/ρ	σ/ρ
Aluminum	71	482	3.0	2.80	25	172
Titanium	110	1,000	2.5	4.43	25	226
Steel	207	1,723	7	7.78	26	221
E glass	70	2,200		2.54	27.6	866
Kevlar (aramid)	130	2,900		1.45	90	2,000
Carbon	380	2,700		1.86	204	1,450
E glass*	40	840		1.80	22.2	467
Kevlar*	82	1,500		1.39	59	1,080
Carbon*	200	1,500		1.55	129	970

*Denotes fiber volume fraction of 60%.

carefully chosen for the local stress field. Lay-ups vary from quasi-isotropic $(+45, -45,0,90°)_{ns}$ to cross-ply $(0, 90)_{ns}$, and angle ply $(+45, -45)_{ns}$. Subscript n denotes the number of sublaminates, and subscript s denotes symmetrical by using a mirror image repeated. Symmetrical lay-ups are almost universal since they do not distort when curing. Clever nonsymmetrical combinations have been used for aeroelastic tailoring where wing flexure can be accompanied by beneficial twisting, which would not happen in homogenous metals.

The weak transverse strength has been mentioned, and it is equally true for the through-thickness direction when the individual laminae can easily delaminate by failure of the resin matrix. This is the Achilles' heel of laminated composite structures, and currently several solutions to this problem are being evaluated. Textile technology such as weaving or stitching is possible for the dry fibers, which are then impregnated by infusion of the resin. This is also a cheap form of structure, but strengths will be down. Z-pinning is an option for laminated composites already cured. Here small pins of carbon, aramid, or metal are driven dynamically into the laminate from the outer surface. Some in-plane strength is sacrificed by the process, but the through-thickness strength and fracture toughness increase dramatically. The other alternative is to design a structure which has no through-thickness stresses, but this is virtually impossible near any 3D features such as joints and panel stiffeners. Also, low-velocity impact damage is a threat which can cause delamination and some 70% loss in the residual compressive strength.

Thermoplastic resins are an alternative which does not require a chemical change during manufacture, just application of heat to a flat laminated sheet. At the moment these appear to be too expensive, but they are much quicker to mass produce since a thermoset composite spends 1½–2 hours in an autoclave. Their properties are better, particularly the through-thickness fracture toughness where the thermoplastic matrix exhibits modest ductility. The Airbus 380, for example, uses thermoplastic in the leading edge, and these must withstand bird-strike without excessive losses in stiffness or strength.

Composite structures do have several other virtues. Their fatigue life is exceptional. Corrosion is not much of a problem, but beware moisture absorption in the matrix. It is also easier to design a stealthy aircraft with a small radar cross-section using composites than it is in metallic structures. The coefficient of expansion for carbon is virtually zero. The high-temperature performance (strength and creep) for thermoset materials is better than for aluminum alloys.

Finally, the structural properties of stiffness and strength of laminated composite structures are not formed until the manufacturing process is complete. They can be theoretically deduced from the unidirectional lamina properties, knowing the stacking sequence, but this is beyond the scope of this book. Several texts on composite structures are available (Hull 1981; Agarwal and Broutman 1990; Niu 1992).

9.3 SMART MATERIALS

Now that we have discussed composite structures, it is convenient to discuss the possibility of smart materials and structures, even if only to dispel some of the myths surrounding this high-profile name. There are two distinct forms of smart structures. First are the health and usage monitoring systems (HUMS), in which optic glass fibers are embedded in composite structures or stuck on metallic. These can measure strain, temperature, moisture, and fracture. The attraction of HUMS,

in which the signals are optical changes, is that the sensors are cheap and can be used extensively in arrays and the data can be processed readily to give the history and current state. The diameter of optic fibers is much larger than that of carbon, and some weakening of a composite is possible. Airlines have yet to be convinced. However, the same concept can be used to monitor manufacturing processes, and industry has warmed to this aspect.

The second form of smart structure is the use of active materials, such as piezoelectric and piezoceramic, which will strain under an applied voltage gradient. They can also produce an electrical potential gradient when strained. The strains are small, but design studies show it is feasible to have a control surface, such as a leading or trailing edge, sufficiently deformed to produce a respectable aerodynamic moment. Much larger strains can be produced in shape memory alloys, but these essentially involve a change of state which is not immediately reversible and thus would not be suitable for conventional control. By having a feedback control it is possible to use smart materials to suppress small strains, e.g., high-frequency (low-amplitude) noise in fuselages and cockpits. Flutter suppression is also possible since it is only necessary to inhibit the onset of flutter when the strains are small and simply prevent the amplitudes growing. Smart materials have already been used to dampen out and prevent growth of vibrations of large orbiting satellite telescopes where there is no aerodynamic damping and precision of the dish surface is crucial. Smart materials do certainly have a future, but the world is waiting for materials which can produce strains of the order of 0.1% at least. At the moment metals and ceramics will not do this, but there are some electroactive polymers which can be persuaded to undergo very large strains by an internally seeded liquid being forced to move under a voltage. These could be called "artificial muscles," and some applications are being developed for moving surfaces on satellites and extraterrestrial vehicles where conventional controls are at risk because the lubrication necessary ceases to function at extremely low temperatures.

PART 3
STRUCTURAL CONSIDERATIONS

Glyn Davies

9.4 STRUCTURAL ANALYSIS

As mentioned, this is not intended as an in-depth text on analysis. But it is easier to explain the choice of structural configurations if simple analytical models are to hand. It is also important to know where analysis is approximate and whether it is virtually exact. Therefore, to emphasize the simplicity of structural analysis we should note that there are really only four separate tasks necessary for all forms of structural analysis, whether the loadings be mechanical, thermal, aerodynamic, or environmental. Figure 9.9 illustrates this. The first stage recognizes that it is impossible to treat every structure as a 3D continuum. Some approximations have to be made—usually the nature of the displacements or strains in a 1D or 2D strain field. Most aircraft structures are thin-walled and therefore 2D. The next requirement is to satisfy equilibrium, i.e., the internal stresses should be in equilibrium with each other and with the applied loads. In dynamic cases we introduce inertial loadings for convenience and treat the problem as an exercise in equilibrium. In very rare cases this is sufficient information to solve for the stresses, i.e., the structure is *statically determinate*. This is almost unknown in aircraft for a very good reason, since if any small component fails a statically determinate structure fails, it becomes a mechanism. We therefore prefer a redundant or statically indeterminate structure, often referred to as *fail-safe.* If a structure is statically indeterminate, then extra information is needed. The word ''compatibility'' in the diagram is not self-explanatory but simply means using a geometrical argument to get strain in terms of displacements. For example, in a long, thin bar the strain is the end displacement divided by the length. This information is then linked to the stress in the equations of equilibrium by a stress–strain law, and the problem should then be solvable. The stress–strain laws are determined by tests. A typical one could be $\epsilon = \sigma/E + \alpha T$, i.e., some extensional strain is due to the tensile stress and some is due to the thermal expansion if there is a temperature rise T. All of the above strategies are independent and their equations are autonomous. A few simple examples illustrate these principles.

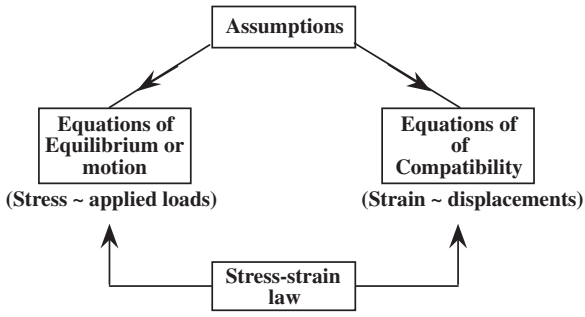


FIGURE 9.9 Basic tools of all structural analysis.

In Figure 9.10 a bar of length l and cross-sectional area A is heated through $T°$. Find the support reactions R.

1. *Assumption:* The stress and strain field is uniform along the entire length.
2. *Compatibility:* ϵ = (end displacements)/l. These are zero, so $\epsilon = 0$ everywhere.
3. *Stress–strain:* $\epsilon = \sigma/E + \alpha T = 0, \therefore \sigma = -E\alpha T$ (compressive)
4. *Equilibrium:* The end reactions are $R = A\sigma = AE\alpha T$

A more subtle example is shown in Figure 9.11.

Assumptions: All bars have a state of uniform stress and strain The displacements are small so that any component at right angles to a bar will not stretch it.

Equilibrium:

$$T_1 \cos\theta + T_2 + T_3 \cos\theta = W$$

$$T_1 \sin\theta - T_3 \sin\theta = H$$

There are not sufficient equations to solve (the structure is statically indeterminate), so we need to use:

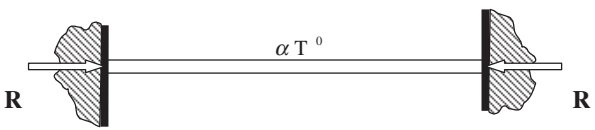


FIGURE 9.10 Heated bar.

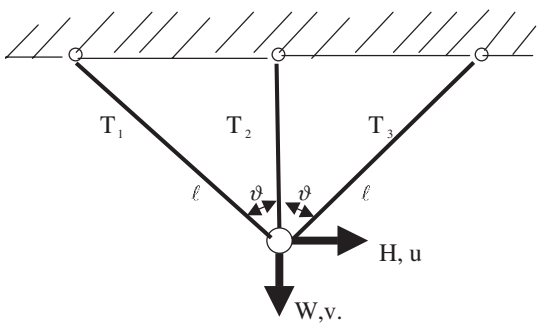


FIGURE 9.11 Three bars loaded by W and H.

Compatibility:

$$\epsilon_1 = (u \sin \theta + v \cos \theta)/l$$

$$\epsilon_2 = v/l \cos \theta$$

$$\epsilon_3 = (-u \sin \theta + v \cos \theta)/l$$

Stress–strain law:

$$\epsilon_1 = T_1/AE, \quad \epsilon_2 = T_2/AE, \quad \epsilon_3 = T_3/AE$$

This is all the information needed to solve the problem. The compatibility equations can be converted to bar forces using the above, and then substituting the bar forces into the equilibrium equations delivers two equations in terms of u and v.

We now use a different approach which paves the way to understanding the basis of the finite-element method. Known as the principle of virtual displacements, it uses the concept of *virtual work*. We define virtual work simply as the product of force and corresponding displacement—whatever the relationship between them is. (It may be nonlinear.) We then define virtual displacements $\underline{u}$ and $\underline{v}$ as arbitrary values of the real ones, for which the reason will become clear. It is still necessary to satisfy compatibility directly, but we now let virtual work take care of equilibrium by equating internal work to external work, thus

$$T_1\underline{\Delta}_1 + T_2\underline{\Delta}_2 + T_3\underline{\Delta}_3 = W\underline{v} + H\underline{u}$$

where $\underline{\Delta}$ are the bar extensions. We will relate them to $\underline{u}$ and $\underline{v}$ using compatibility, thus

$$T_1(\underline{u} \sin \theta + \underline{v} \cos \theta) + T_2\underline{v} + T_3(-\underline{u} \sin \theta + \underline{v} \cos \theta) = W\underline{v} + H\underline{u}$$

and rearranging

$$(T_1 \sin \theta - T_3 \sin \theta - H)\underline{u} + (T_1 \cos \theta + T_2 + T_3 \cos \theta - W)\underline{v} = 0$$

Now the reason for specifying virtual displacements becomes clear. They can be anything, i.e., all but one can be put to zero, for example. Thus, the individual coefficients of u and v must be separately zero. When we look at this we recognize that these are the equations of equilibrium. Clearly, had there been hundreds of unknown displacements this method would still deliver the correct number of equations for solving them. The reason for using the virtual work argument is that the separate equilibrium equations are replaced by a summation, and computers are very efficient at such bookkeeping. In this example the equations are all exact. In the finite-element method, where the computer takes over the summations, this is not so. A very basic explanation of the finite-element method now follows.

9.5 FINITE ELEMENT ANALYSIS

The trick is to imagine the structure metaphorically subdivided into small (but finite) elements which are so small that the *shape* of the displacement field over them can be assumed without fear of a large error. Figure 9.12 implies this. This then leaves

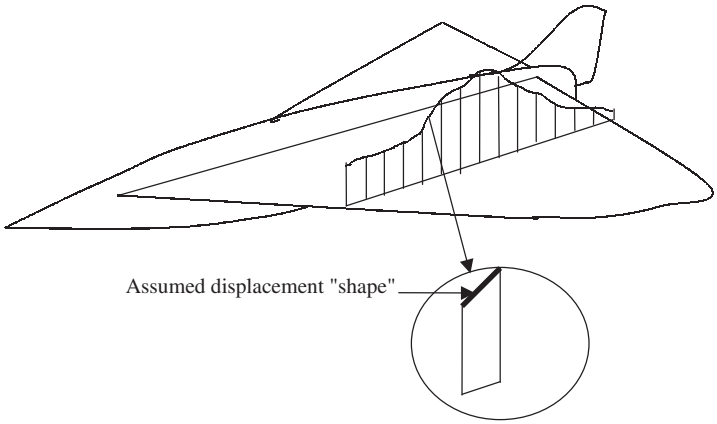


FIGURE 9.12 Finite element basis.

only the *magnitude* of the displacements to be solved, and these will be at selected element nodes, usually those points which decide the chosen boundaries of the element. For an assumed displacement shape function we can then derive exactly the strains. However, instead of the simple equations for a framework, such as $\varepsilon = \Delta/l$, we now have a continuous displacement field $u(x, y, z)$, $v(x, y, z)$, and $w(x, y, z)$. The compatibility equations for the various strain components can easily be found, such as

$$\varepsilon_x = \partial u/\partial x \text{ (a direct strain in the } x \text{ direction)}$$

and

$$\varepsilon_{xy} = \partial u/\partial y + \partial v/\partial x \quad \text{(a shear strain in the } x\text{–}y \text{ plane)}$$

The virtual work for an entire structure is still summed as

$$\text{Virtual work} = \sum (\textit{elements'VW})$$

where an individual element virtual work is simply an integral which can be evaluated since the strain field is deduced from the unknown nodal displacements and the associated shape functions, i.e.,

$$VW = \int \sigma^t \varepsilon \, dv \text{ over an element}$$

These finite elements may be one dimensional (bars and beams: see Part 4 for a fairly full treatment of beam elements), 2D plates and shells (triangular and quadrilateral), or 3D solids (tetrahedra or hexagonal bricks). The element sides or surfaces may also be curvilinear. The unknown displacements are selected to be at the corner points, mid-sides or mid-surfaces. There may be thousands of these unknown displacements, but the virtual work still delivers the right number of equations. These equilibrium equations will be approximate. For example, if the assumed displacement field is linear then the derived strains will be constant and so will the stresses via Hooke's law. If the stresses in each element are constant, then clearly for a complete structure there will be jumps in the stress as we move from one

element to another. This is no bad thing, and many codes use these jumps as a measure of how much refinement is still needed. The discontinuities are averaged or smoothed anyway. It can be shown that provided that the assumed element displacement field contains the rigid body movements (zero strain), the element will certainly be in equilibrium with its neighbors and the applied forces. The equilibrium conditions inside an element will be approximate but can be improved if necessary by using smaller elements or using displacement shape functions of a higher order such as quadratic or cubic.

The virtual work principle ensures that the number of equations of equilibrium matches the number of unknown displacements. If these are listed as a column matrix

$$\mathbf{u}' = [u_1 u_2 u_3 \cdots u_N]$$

then the virtual work delivers the equation

$$\mathbf{Ku} = \mathbf{F}$$

where **K** is the global stiffness matrix and **F** is the global column list of nodal applied forces. Both are readily summed. Not only static but also dynamic problems can be solved in this way as discussed in Part 9.4.

The finite-element method has revolutionized structural analysis. It has removed the need to simplify the geometry of very complex structures and enabled us to analyze approximately an exact structure, rather than the approach of yesteryear, which was to solve exactly an approximate structure. The hard work is now the creation of the finite element mesh and idealizations in the first place. Most commercial finite-element codes have a preprocessor which almost automates this part given a CAD model to work from (see Figure 9.58 for an example). The results obtained from the postprocessor can be maps or contours of displacement, stresses or strains, or damage/plasticity maps. The solution procedure for possibly hundreds of thousands of nodal displacements is routine and quick on modern PCs or workstations.

The possible trouble areas using a finite-element code can be in solving nonlinear problems which invariably arise when we try to model failure. Any nonlinear simulation involves an iteration (say during increasing the applied load in small steps) where a small change in the history is solved linearly, but using the sum of all previous displacements and the accumulated changes in the elements of the current stiffness matrix **K.** The nonlinearities which cause the stiffness matrix to change with displacements may be due to large accumulated displacements which cannot be ignored when satisfying equilibrium, or in the stress–strain law if plasticity occurs. The problems in iteration may occur when the residuals in the equilibrium equations fail to converge for algorithmic reasons rather than the physics. It would be nice to say that all finite element codes are robust and always converge, but it would not be true.

9.6 BEAMS

Long slender members can be analyzed making simple assumptions. Figure 9.13 shows one such member. First, any axial load T produces a stress T/A everywhere. However, there is also bending behavior due to the transverse loading, and this is

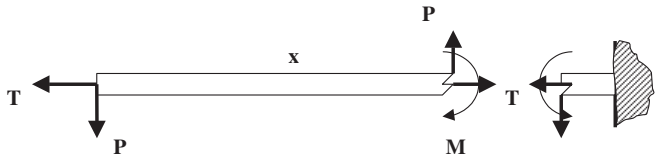


FIGURE 9.13 Cantilever.

best understood in terms of the internal local bending moment, M, illustrated in the figure as having a value $M(x) = P \cdot x$, from equilibrium.

The classical beam approximation, usually known as engineers' theory of bending, assumes that plane sections remain plane, which they would do by symmetry if the beam were subjected to a pure moment M and had a constant cross-section. Figure 9.14 illustrates the concept of plane sections remaining plane. Using purely geometrical assumptions, it can easily be shown that the bending strains in terms of the induced curvature $1/R$ are $\varepsilon = y/R$, where y is measured from the neutral axis—a point of zero strain. This equation is the beam equation of compatibility giving strains in terms of deformation. Then, using the stress–strain law to convert to stress, there are then two equations of equilibrium to be satisfied: axial and bending. The first shows that the neutral axis lies at the section centroid, i.e., $\int_A y dA = 0$. The moment equation delivers

$$M = EI/R$$

where $I = \int_A y^2 \, dA$, is a section property known as the second moment of area or sometimes incorrectly as the moment of inertia. It is often written as $I = Ak^2$, where k is the radius of gyration (cf. dynamics). Combining and rewriting these equations, we have

$$\sigma/y = M/I = E/R$$

These are probably the most used equations in structural mechanics. Provided a beam is slender, it is a very simple tool for initial design before getting down to the details. In addition to these direct stresses varying linearly with y we can examine the variation in shear stresses, τ, by using the equilibrium equation of a point looking in the axial direction z, i.e.,

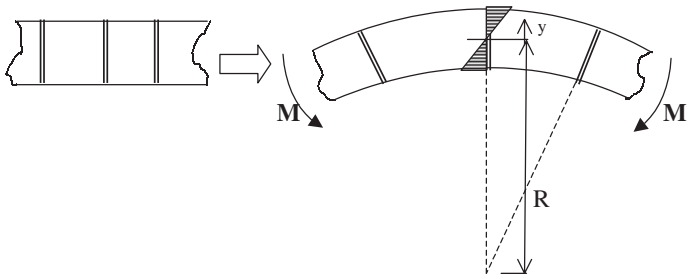


FIGURE 9.14 Simple beam bending.

$$\frac{\partial \sigma_z}{\partial z} + \frac{\partial \tau}{\partial s} = 0$$

Putting $\sigma_z = M \cdot y/I$, multiplying both sides by t, and integrating *wrt* s, we find

$$[\tau \cdot t]_0^s = -\int_0^s t \cdot \frac{F}{I} \cdot y \cdot ds$$

where again, using equilibrium, we have put $dM/dz = F$, the local shear force. This equation has been used to display the variation in shear stress around two sections in Figure 9.30, for example.

We can now see that a beam bending is not a very efficient structure. Consider the previous example, and suppose we apply an axial force $T = P$ as well as the transverse load P. The direct stress equals P/A. The bending stresses are given by

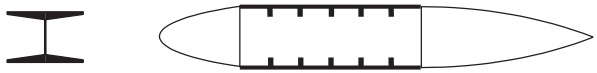
$$\sigma = M \cdot y/I = Ply/Ak^2$$

Now k is of the order of y_{max}, so the bending stresses are a factor l/k larger than those due to axial loading. It is desirable to maximize k, so the beam material should be as far from the neutral axis as possible, like those shown in Figure 9.15. The wing section is a classical example when the vertical dimension is small compared with the horizontal chord. Unfortunately, the aerodynamics demands this, of course. The inefficiency of beams in resisting bending is something that we just have to live with. Even for modest strains the wing deflections can be quite large. To vastly simplify the idea, suppose the wing of semi-span S shown in Figure 9.16 were bent into a circle of radius R, producing a tip deflection of δ. Simple geometry gives

$$R^2 = (R - \delta)^2 + S^2$$

If $\delta << R$, we then find $1/R = 2\delta/S^2$. The maximum bending strain, if the wing skin distance from the neutral axis is y_{max}, is now $\varepsilon = y/R = 2\delta y_{max}/S^2$.

Typical values for a large civil aircraft might be $S = 20m$, $y_{max} = 0.2m$, and putting the strain equal to the 0.2% we find $d = 0.5m$! This is clearly visible and can be seen by passengers in an aircraft at takeoff as the incidence is increased to generate the full wing loading. The flexibility of thin wings has important aeroelastic consequences since the deformations may affect the aerodynamic loading as discussed in Part 4. As the speed of military aircraft increased, wings became thinner and thinner to reduce the drag and because the aerodynamic lift coefficient did not have to be large at these much higher speeds. Their stiffness became the main design criterion, and this explains why carbon composites have become almost universal for such aircraft in the past 20 years. The flexibility of the beam has



Rolled steel joist Aircraft wing-box

FIGURE 9.15 Maximize I or k.

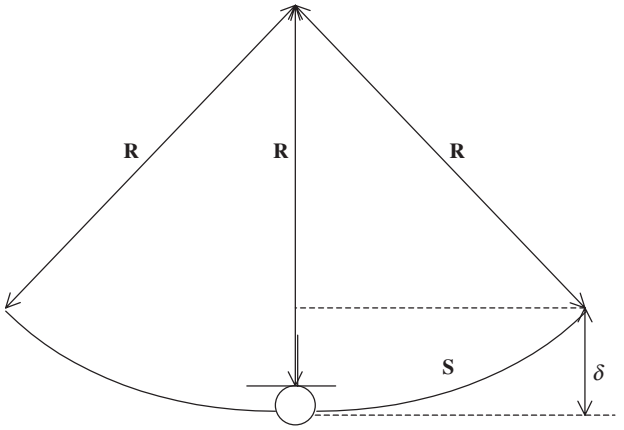


FIGURE 9.16 Simplified wing bending.

important consequences for the stability when loaded in compression, in which case they are often referred to as struts.

A pin-ended strut is shown in its buckled state in Figure 9.17. For equilibrium, to maintain this displaced position we see that $M = P \cdot v$, that is, the change in geometry $v(z)$ is no longer being ignored. The compatibility equation used earlier can be rewritten if we recognize that the curvature $1/R$ can be expressed in terms of $v(z)$ as $1/R = -\partial^2 v/\partial z^2$, so

$$M = EI/R = -EI\partial^2 v/\partial z^2$$

Equating the two expressions delivers the differential equation

$$\partial^2 v/\partial z^2 + \lambda^2 v = 0, \text{ where we have put } \lambda^2 = P/EI$$

The solution, satisfying the two boundary conditions $v(0) = v(l) = 0$, becomes

$$v = A \sin \lambda z, \quad \text{where } \lambda l = \pi$$

so the buckled state has an indeterminate amplitude (A) at the special values of $\lambda = \pi/l$ or

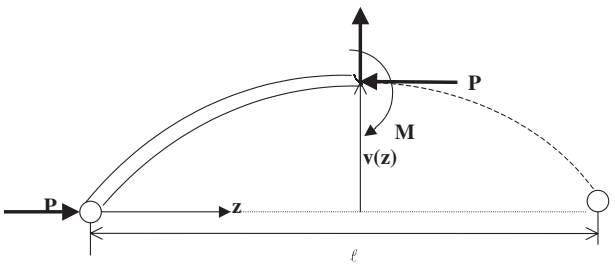


FIGURE 9.17 Strut buckling.

$$P_{\text{crit}} = \pi^2 EI/l^2$$

This is a simple model for a perfect strut, and if imperfections are taken into account a series of curves like Figure 9.18 can be drawn.

The perfect case has no displacement at all until the critical load P_{crit}, at which case the deflections are indeterminate. In real life if the strain has an allowable value, say ε_0, we see that the allowable loads decrease with imperfections.

The combined effect of imperfections and of the onset of material yielding can be expressed if we write the above buckling stress as

$$\sigma = P/A = \pi^2 EI/Al^2 = \pi^2 E(k/l)^2$$

The role of radius of gyration is seen to be important in raising the buckling stress. The parameter k/l is known as the slenderness ratio. However, the buckling stress σ in Figure 9.19 does not become infinite for small values of l/k, as the

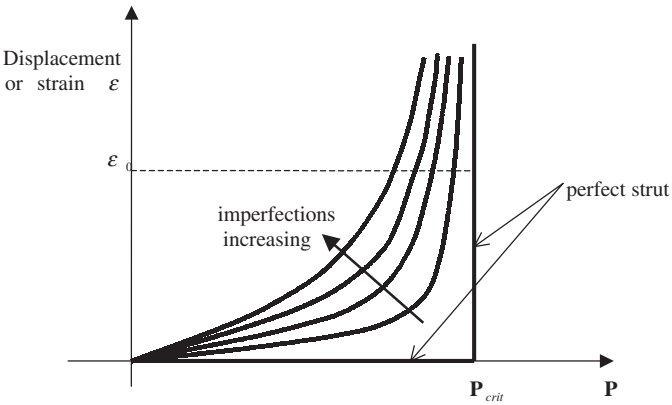


FIGURE 9.18 Effect of strut imperfections.

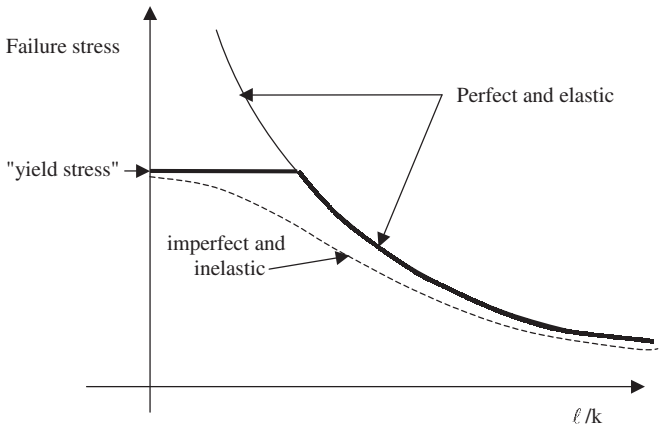


FIGURE 9.19 Strut failure as a function of slenderness ratio.

figure might imply. Eventually any ductile material will yield for high stresses, and this effect can be approximated by replacing E in the above equation by the tangent modulus E_T (Figure 9.5). A typical failure curve including ductility and imperfection is shown in Figure 9.19. A fuller treatment of strut behavior can be found in Timoshenko (1961) and Curtis (1997).

9.7 TUBES

Most of the behavior of beams under bending can be almost duplicated if a slender bar is twisted under torsion. They are usually called tubes in this case. The equivalent of engineers' beam theory can be applied immediately to a circular tube subjected to a torque T as shown in Figure 9.20. The compatibility equation, for which beam theory became the linear relation,

$$\varepsilon = y \cdot \frac{d}{dz}\left(\frac{dv}{dz}\right)$$

is now for the shear strains

$$\gamma = r \cdot \frac{d\theta}{dz}$$

where the analogy between the twist θ and the slope dv/dz is clear. Using the stress strain law $\tau = G\gamma$, (G is the material shear modulus) and integrating over the circle, we find

$$T = GJ\frac{d\theta}{dz}$$

where

$$J = \pi R^4/2$$

The expression GJ is known as the torsional stiffness, compared with EI, the flexural stiffness of beams. The value of J for a hollow tube of internal radius R_I and outside R_O is simply

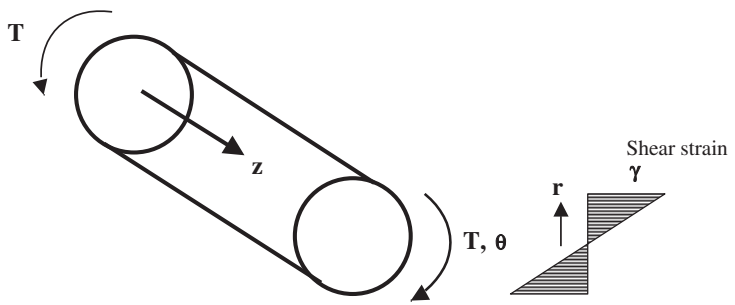


FIGURE 9.20 Circular tube twisted by torque T.

$$J = \frac{\pi}{2}(R_O^4 - R_I^4)$$

The most efficient tube is the thin one ($t << R$) where all the material is confined to a radius R, so putting $R_o = R$, and $R_I = R - t$, then J degenerates to $2\pi R^3 t$.

Unfortunately, these simple relationships are only true for circular tubes, the reason being that the above assumed shear stress contours do actually satisfy the boundary conditions that, at the surface, they should be parallel to that surface. For shapes other than circles the simple linear relationship between strain and radius is not true. Figure 9.21 shows the nature of shear stress contours for various shapes. The value of J is no longer a simple one except for a very thin section like the last in Figure 9.21, where it can be shown that J equals $\frac{1}{3}St^3$.

Returning to the thin circular tube where

$$J = 2\pi R^2 t \quad \text{and} \quad \tau = T/2\pi R^2 t$$

we can write this last term as $\tau \cdot t = T/2\ \pi\ R^2 = T/2A$, where $A = \pi R^2$ is the area enclosed by the tube. If a thin-walled closed tube of *any* shape has only shear stresses and no direct stress σ_z, then the equations of equilibrium show that the product $\tau \cdot t$ is constant and equal to $T/2A$ as above. It can also be shown that, whatever the shape of this closed thin-walled tube section,

$$\frac{d\theta}{dz} = \frac{1}{2AG}\oint \tau ds = \frac{1}{2AG} \cdot \frac{TS}{t2A}$$

where we see that J must be $J = 4A^2 t/S$, where S is the total perimeter length.

The big difference between closed tubes and open tubes can now be shown by considering what happens to a closed tube when it is cut as shown in Figure 9.22. Using the previous relationships and assuming a uniform thickness, we find the ratio

$$\frac{J_{\text{open}}}{J_{\text{closed}}} = \frac{\frac{1}{3}St^3}{4A^2 t/S}$$

$$= \frac{1}{12} \cdot \frac{S^2 t^2}{A^2}$$

Now the enclosed area is of order $(S/4)^2$ (for a square section), so

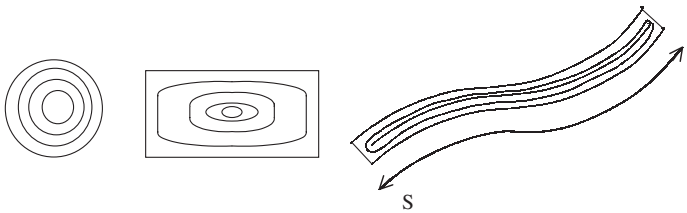
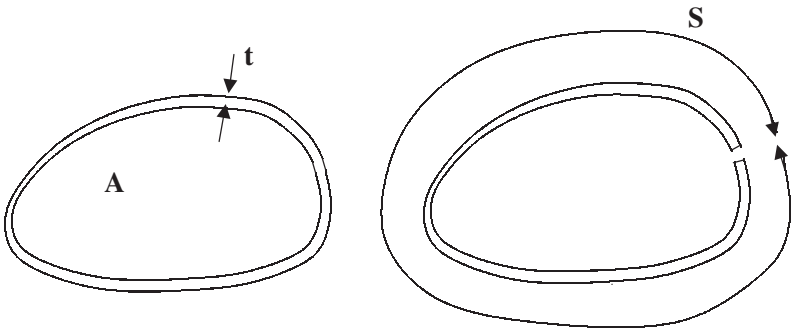


FIGURE 9.21 Shear stresses in various sections under torque.



(a) Closed tube (b) Open tube

FIGURE 9.22 Two different tubes.

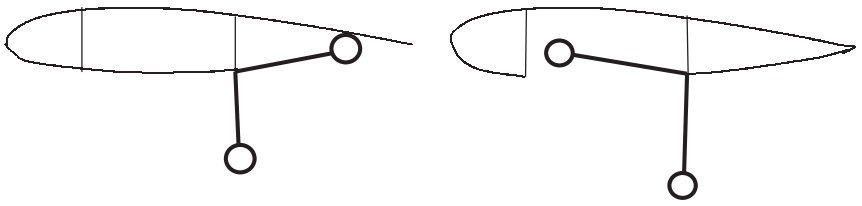
$$\frac{J_{\text{open}}}{J_{\text{closed}}} \approx \frac{\frac{1}{12} \cdot S^2 t^2}{(S/4)^3} \approx 20\ (t^2/S^2)$$

which is very small since $S >> t$.

If an aircraft structure needs a high torsional stiffness, then a closed tube is always selected. The main aircraft wing is a good example where low stiffness can give rise to aeroelastic problems (Part 4). Figure 9.23 shows a choice offered to designer of wing structures over 50 years ago when the first retractable undercarriages were used. (The wings did not carry fuel in those days.) Clearly (a) or something similar would be preferred.

9.8 PLATES AND SHELLS

Slender beams and tubes are common in small components and control systems, but the primary aircraft structure, like the sketch in Figure 9.23, consists of thin plates (flat) or shells (curved) which are held or attached to another structure on all sides. Thin-walled plates can be analyzed as 2D versions of beams, but it is



(a) (b)

FIGURE 9.23 Two options.

sufficient for our purposes simply to look at some effects. Like beams and struts, plates are not very efficient if loaded normal to their surface or allowed to buckle, but are efficient in resisting purely in-plane loading. However, it is rare to call upon thin-walled plates to resist large normal pressures. (The aerodynamic pressures responsible for lift do not cause significant bending stresses in the wing skin.) However, the buckling behavior of plates is sufficiently different from struts to be important.

Consider the rectangular plates thickness t and aspect ratio a/b in Figure 9.24. If the ends b are loaded in compression and the sides a free, then the plate buckles just like a strut. The buckling stress will vary like $(t/a)^2$ and rapidly decrease as a increases. However, the situation is quite different if the sides of the plate are held in some way like a simple support (free to rotate) or clamped. The buckling mode will look like the last figure, in which the half-wavelength is of order b. The buckling stresses can be shown to be of the form

$$\sigma_{\text{crit}} = KE(t/b)^2$$

where K is a constant depending on the nature of the edge supports and of the aspect ratio a/b. The latter effect is due to the need for the end boundary conditions to be satisfied. The number of half-wavelengths changes as a/b increases, but the buckling stress quickly becomes almost constant, as shown in Figure 9.25. For large values of a/b the values of K are found to be

Edge support	K
All clamped	6.32
All simply supported	3.62
One edge free	0.58

Similar coefficients can be found for buckling in shear or for combinations of compression, shear, and tension. Timoshenko (1961).

Since the size and configuration of most of an aircraft structure will be dictated by compression loading, it is important to note the consequences of this behavior. For example, consider the top surface of a wing skin which is called upon to resist

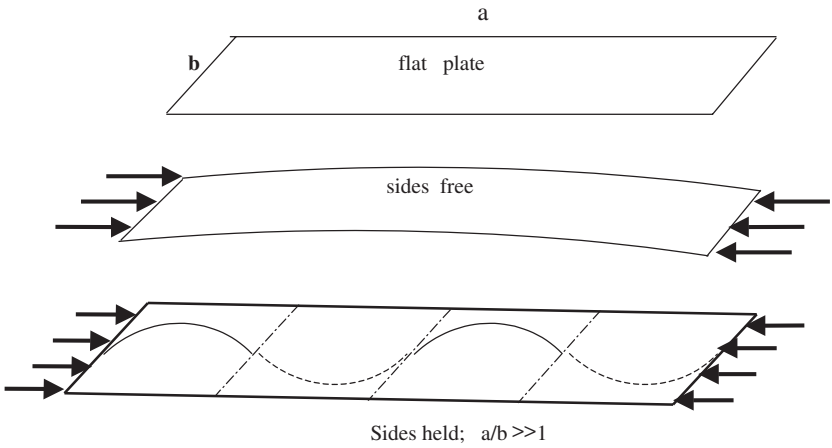


FIGURE 9.24 Plate buckling.

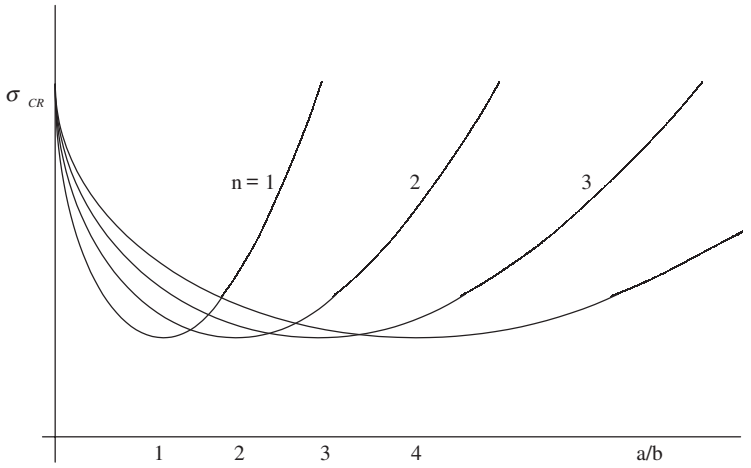


FIGURE 9.25 Buckling stress: change in number of half wavelengths.

high compression loads as the wing bends. If the skin were only supported at the two spars, then the buckling stress would behave like $(t/c)^2$ in Figure 9.26. A far higher stress, retaining the same skin thickness, can be achieved by adding stiffeners as shown in the figure, bringing the buckling strain up to order $(t/b)^2$. The stiffeners have to have sufficient rigidity to remain straight, and one criterion to satisfy is that the whole stiffened surface must not buckle globally as it would in (a). This global buckling stress depends on the distance along the wing between ribs and the effective thickness of the skin plus stiffeners ("effective" means the overall value of radius of gyration of the skin plus stiffener combination), but there is another criterion. The stiffeners themselves may have other local buckling modes in which the corners remain straight but the sides can buckle as if the edges were simply supported or free. Figure 9.27 shows some typical stiffener configurations. The top hat stiffener has the highest torsional rigidity and effectively supports the adjacent skin as a clamped edge. It is not popular, since internal corrosion will not be spotted. The Z section is a very popular choice, and the top flange can easily be bolted or bonded. The blade stiffener is the most inefficient, but it can be integrally machined from the solid rather than a separate stiffener as the figure implies. It is a lighter

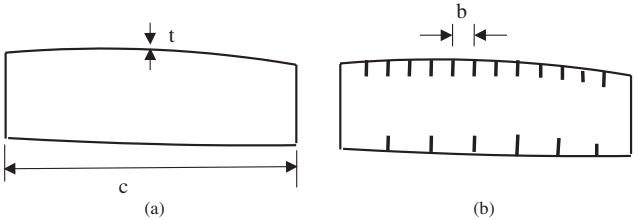


FIGURE 9.26 Wing box: prevention of skin buckling.

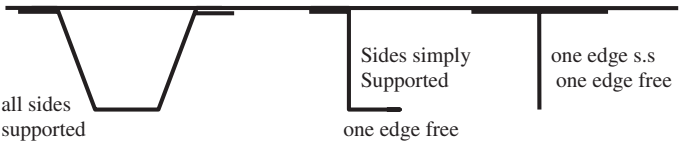


FIGURE 9.27 Stiffener types.

choice. The choice of the proportions of the Z stiffener could be made by choosing the flange and the web to have the same buckling stresses, using the coefficients quoted. Actually using the quoted buckling coefficients on the configurations shown is a slight oversimplification since the various elements cannot choose their own half-wavelength; they must couple with each other, and the actual buckling stresses will lie between the various separate values. Details on how to do these estimates and allow for inelastic effects can be found in Curtis (1997) and Timoshenko (1961).

If the aircraft is a high-speed military vehicle the wing will be thin and the compression loading high. To achieve a respectable buckling stress the stiffeners will have to be sufficiently deep, and if the wing is thin then the stiffeners on both surfaces will be so close that a multispar configuration is better, as shown in Figure 9.28. This configuration is also a much better fail-safe construction in that battle damage to a wing surface between spars is much less likely to propagate beyond adjacent spars, which can act as effective crack stoppers. The torsional stiffness of a multispar configuration is affected hardly at all by the additional internal spars. What counts is the enclosed area of the whole box section.

Another way in which plates buckle differently from struts lies in the behaviour after the initial buckling stress is exceeded, i.e, the *postbuckling* behavior. If the stiffened plate does not buckle globally and the edges remain straight, then the plate can resist further loads. The region midway between the (stiffened) edges will not carry much more load than the initial buckling stress, but the stress at the stiffened edges can be increased until the stiffener fails in a coupled mode or because the material approaches the yield stress. The postbuckling strength can exceed the initial buckling value by a factor of 1.5 to 2 for realistic configurations. In practice, it would not be a good idea to pass through the buckling stress under normal working loads since the postbuckled internal forces do place extra transverse loads on joints which may fail the rivets, bolts, or bonds. A reasonable criteria would be to buckle at proof and leave the postbuckling strength for the ultimate design load. Actually, before today's higher speeds demanded a very smooth wing surface, passenger aircraft used to have wing skins which buckled between spars and stiffeners, much to the concern of the observant passengers. Even today the lower surface of a fuselage skin can be seen to be buckled when standing on the ground and in com-

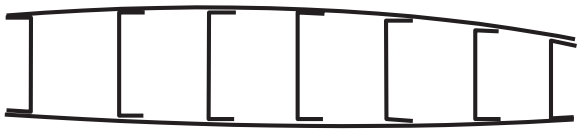


FIGURE 9.28 Multi-spar wing.

pression. In civil engineering, where aerodynamic needs are irrelevant, it is common to design beam shear webs to buckle at roughly 45° in shear. This is known as *diagonal tension*.

The buckling behavior of curved shells does not differ too much from that of flat plates if the curvature is small, as it is in wing and control surfaces. The change in the buckling coefficients can be accommodated (Curtis 1997). A circular section fuselage is stiffened just like a wing, and the skin between the stiffeners is just a slightly curved plate. However, the behavior of an *unstiffened* cylindrical shell is quite different. Simple classical buckling theory (Timoshenko 1961) shows that the buckling stress behaves like Et/R, where R is the shell radius. Unfortunately this value is not achieved in practice, and this becomes clear when we consider the postbuckling performance and the effect of imperfections. Figure 9.29 indicates the behavior roughly of struts, plates, and cylindrical shells. Struts have little postbuckling strength until the displacements become excessively large. Plates do have respectable performance, and the postbuckling slope may be about one-third of the initial slope. The cylindrical shell behavior, however, is quite different in that there is a complete and abrupt change in the mode at buckling, from an axisymmetrical wrinkle of small wavelength of order $\sqrt{Rt}$ to a diamond-shaped pattern where the diamond sizes approach the shell radius in size. The shell actually snaps through dynamically to the postbuckled shape if the load is constant and unable to reduce as the figure implies. The dotted lines indicate the effect of small imperfections, which is modest for flat plates but traumatic for the shell. A sphere under external pressure behaves similarly. Such shells are known as *imperfection sensitive*. Some stiffened plates are a little imperfection sensitive if all the individual local buckling stresses are designed to coincide.

Before passing on, it should be mentioned that curved cylinders and spheres do have one important thing going for them, and that is their *membrane stiffness*. We have seen how straight slender beams are very inefficient at resisting loads applied normal to their axis compared with loads applied axially. The span of a horizontal

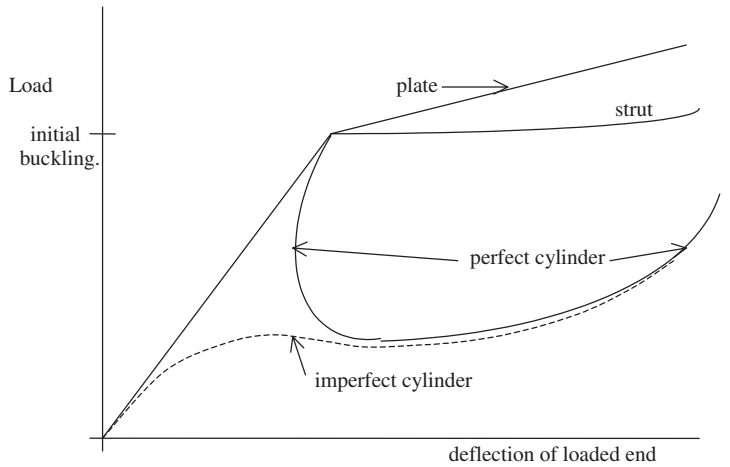


FIGURE 9.29 Post-buckling behavior.

beam loaded by its own weight will be much more limited than that of an arch since the latter has curvature which enables the internal axial forces to have a component in the direction of the local normal. It is possible to design an arch which has no bending stresses at all. Imagine suspending a chain (which has zero bending stiffness) and noting the shape into which it deforms (a catenary). If this shape is inverted into an arch, it will have a state of uniform compression with no bending. The same can be organized for thin shells, which can also have their shape chosen to avoid bending for a specified applied loading. This would be a pure membrane action. It is not always possible to achieve a pure membrane reaction for some shapes and loadings. For example, some balloons when blown up will manifest wrinkles, which are an indication that local buckling has occurred. The simplest case is that of internal pressure applied to a cylinder or sphere. In the case of a cylinder, under pressure p, equilibrium arguments show that the longitudinal stress is given by $pR/2t$ while the circumferential (hoop) stress is pR/t. In the case of the very efficient sphere, the stresses are pR/t everywhere. These membrane structures are so efficient that some very thin satellite launchers actually rely on an internal pressure to overcome buckling at launch.

A civil aircraft fuselage is the case where membrane action is exploited, although the presence of internal frames or stiffeners does modify this purity slightly. A cylinder has to be closed, of course, and ideally we would like to use a hemisphere at both ends (although the pure membrane action is not quite sustained at the intersection due to incompatible displacements). However, this is not very convenient for aerodynamic reasons at the front and a modern aircraft will compromise in designing complicated but smoothly varying surfaces. In the case of the rear closure it is common to use a portion of a sphere as a pressurized bulkhead and allow the rest of the tapered tail to be unpressurized. The aerodynamically attractive blended fuselage/wing configuration for a large civil airliner is a poor structure for resisting internal pressure but could be excellent at resisting lifting forces if the mass distribution is uniformly distributed like the lift distribution.

Having discussed the basic structural mechanisms for simple structural components, we can now turn to see how a real aircraft structure exploits them.

9.9 REAL STRUCTURES

Wings

We now look at some examples of modern aircraft structures and show how the simple basic principles we have outlined can be used to explain many of the features. The most important fact to remember is that any thin-walled structure should not be expected to resist a concentrated load applied normally to its local plane. We have seen that bending strains are best avoided and will be attributed to a bad design. The concentrated load may not actually be an applied load (such as that coming from an engine or undercarriage) but an internal force coming from one structure and needing a load path through to another. The easiest structure to look at is the main wing box of a civil aircraft since it is similar in all aircraft, being of high aspect ratio and therefore beam and tube-like in behavior. A wing has to resist primary loads such as:

- The maximum aerodynamic lift during a maneuver or gust, and the resisting inertial forces of the wing and fuel plus the fuselage to which it is attached
- The inertia loads from one or two engines on each wing during the above cases
- The undercarriage loads during a heavy landing

There are also secondary loads from:

- All controls, such as leading edge slats and trailing edge flaps and ailerons
- Internal fuel pressures adding to the inertia loading

There are other threats to a wing, briefly mentioned here such as:

- Flutter and divergence (see Part 4)
- Fatigue, which we will cover when discussing stress concentrations
- Lightning strikes. These are best alleviated by ensuring a conducting path from any high points and then safely around the wing box fuel tank. This is difficult for composites.
- Impact threats such as bird-strike, runway debris, and engine compressor or turbine failure. Vulnerable areas such as leading edges need to absorb or deflect impact energy.

Basic wing designs are very similar. It used to be popular to use three-spar configurations to demonstrate that the structure was truly fail-safe and could lose one spar and still have a respectable torque box. Some very large civil aircraft do have a third spar in the center regions where the chord of the wing box becomes unacceptably large (see Figure 9.40). However, this philosophy was succeeded by the damage tolerance strategy, whereby crack growth does not lead to failure before the next inspection. Thus the lower (tensile) wing surface can consist of up to five separate integrally stiffened planks bolted together along their spanwise edges. However it is possible to convince the authorities that integrally machined blade stiffeners can be designed to act as efficient crack-stoppers without needing a bolted joint. Such blades are also machined into the lower (tensile) half of front and rear spar webs. For very large civil aircraft in the future we are likely to see quite thick skins having blades welded to them by friction stir-welding.

We will now consider how the wing reacts to applied forces, particularly vertical (shear) forces and especially the problem of concentrated forces. Consider a vertical shear force applied symmetrically to an I-beam or a rectangular box. The equilibrium equations in the section on beams may be used to evaluate the variation in shear stress around the section in terms of the shear force F. Figure 9.30 shows roughly to scale the nature of these resisting shears. In this case these two simple sections have only vertical and horizontal components and the applied shear force is resisted entirely by the vertical webs. The variation in shear stress in these webs does not differ much from the average stress (given by the shear force divided by the total web area). If a pure torque is applied then the open I-section is a very poor resistor, of course, but the closed section will resist with a shear stress given by the Batho theory, $\tau \cdot t = T/2A$.

Thus, if a wing box is subjected to an inertia load, say from a cantilevered engine, or an undercarriage load, then the applied load will be a sum of shear and torsion. Figure 9.31 shows a typical civil aircraft podded engine on the end of a pylon. The pylon has to be a sturdy box beam itself and is attached to the wing

AEROSPACE STRUCTURES

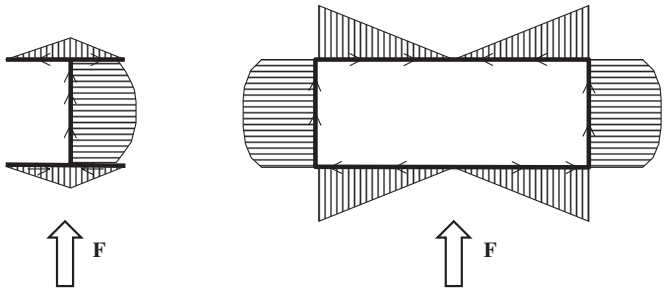


FIGURE 9.30 Variations in shear stresses.

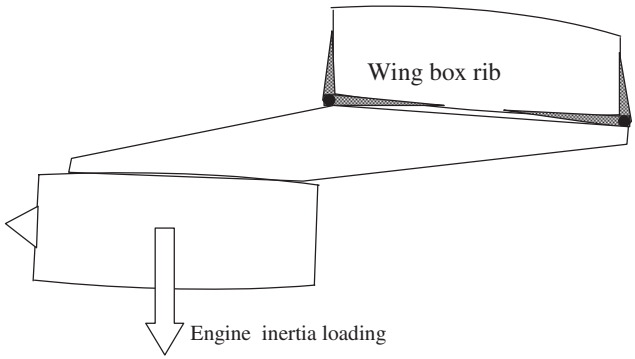


FIGURE 9.31 Cantilevered podded engine.

box at discrete supports shown here near the front and rear spars. The reason for this is to allow the attachments to have a breakable shear pin or mechanical fuse so that in a crash landing the engine will break free and not tear open the wing box full of fuel. The loading and the resisting shear stresses around the wing box can therefore be summarized as in Figure 9.32.

The shear stresses sketched in this figure are those obtained by applying simple beam theory and Batho torsion theory. For these distributions to be realized there

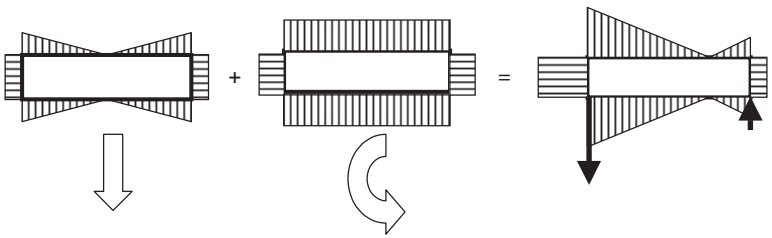


FIGURE 9.32 Combined shear and torsion loading.

must be a substantially stiff rib to take the loadings to the skin and spars. Even a stiff rib may not be sufficient if the applied loads are concentrated as described. There would therefore have to be additional diffusing members to take the concentrated loads and to diffuse into the spar and rib flanges, as indicated crudely by the shaded portions in Figure 9.31. Ideally, such diffusing members should have an exponentially decreasing area, allowing the transfer to be a shear stress which is substantially constant. Figure 9.39 shows several such members necessary to take concentrated loads into the rear spar on an Airbus wing. It would also be desirable to hang flap tracks at the rear of such a stiff rib, but in any case all sources of concentrated forces coming from the undercarriage, slats, flaps, and ailerons will need ribs at the attachments to diffuse the loads around the wing box skin. All these sources will probably dictate many of the wing box rib positions, but there will still be a need to have many secondary ribs to keep the global skin buckling stress high without needing excessively deep stiffeners. Figure 9.40 shows the large number in an Airbus wing.

There is another reason for secondary ribs. Imagine the wing bending due to the distributed aerodynamic lift. The upper surface will go into compression. However, due to the induced curvature this compression will have a (downward) normal component rather like the membrane action discussed earlier. At the lower surface, which will be in tension, there will be an upward component so the combined effect is to squash the box, and closely spaced ribs will be needed to resist this. These crushing loads due to the bending curvature are known as Brazier loading in the United Kingdom. Actually, on the upper wing surface at least this loading and the aerodynamic lifting pressure will tend to compensate. Because the wing box is also a fuel tank, the ribs will have lightening holes or cutouts to permit free fuel flow. The ribs and spars will carry shears as shown, and to improve their buckling performance they will have vertical stiffeners in the same way that the wing covers have longitudinal stiffness.

The complete wing, subjected to aerodynamic lift and structure/fuel inertia loading, will eventually take all the loads via the spars and skin covers to the fuselage. The shear loads will need to be taken to the fuselage and then diffused into the fuselage skin and eventually reacted by the fuselage inertia and tail-plane loadings. The wing covers develop the main stresses due to bending, and in the symmetrical loading case these loads will pass across the center section and equilibrate those from the other wing. In the case of a swept wing there will also be a large torque coming from both wings, which will not cancel but will add and be reacted by the fuselage. This effect can be visualized in detail if we consider the root rib and the wing skin loading in (say) the upper compression surface, as shown in Figure 9.33. The components at right angles to the root rib will equilibrate those from the other wing, as in the unswept case. The forward/aft components will find their way into the rib flange and similarly in the opposite sense on the lower wing surface. Thus, the root rib for a swept wing will take a large distributed torque, which needs to be reacted by two or more fuselage frames as shown in Figure 9.34. It is therefore expedient to consider the nature and problems of forces in the fuselage structure.

Fuselages

The problem of the forces coming from the wing is a good starting point. We have already seen how the root rib will need to be substantial for swept wings. Even for unswept wings the total lift on each wing (a maximum in maneuver or gust load-

AEROSPACE STRUCTURES

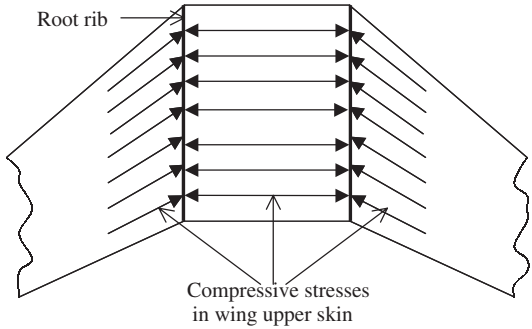


FIGURE 9.33 Root rib function.

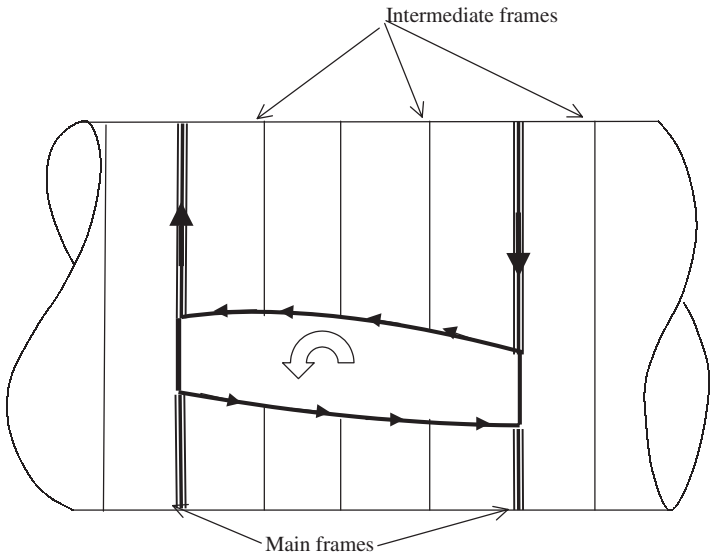


FIGURE 9.34 Wing torque/fuselage transfer.

ings) will be taken at the root ribs into the fuselage via the frames and hopefully be resisted by a distributed shear system in the fuselage skins as rapidly as possible. Thereafter it will be reacted by the fuselage/payload mass together with the tailplane loads appropriate to the gust or maneuver case. The total wing force $2P$ should be reacted by a sinusoidal variation of shear accorded to simple beam theory, and it is the function of the fuselage main frames to achieve this as best they can. If we consider the system of forces acting on the frame, from both the wing and the fuselage shears, as in Figure 9.35, it is clear that these frames must take considerable bending effects. These mainframes therefore must be strong and stiff. They would be much more efficient structures if they did not have to leave the fuselage interior free to carry the payload. This may not be a requirement in military

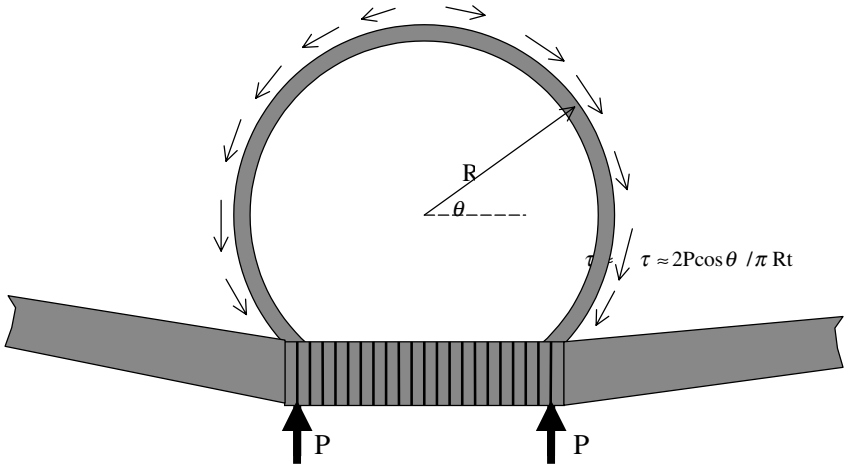


FIGURE 9.35 Fuselage main frame loading from wing.

aircraft, where the only interception is the engine ducting, and here the frames become like thin-walled bulkheads, with reinforced cut-outs for the engine ducts. The heavy mainframes at the wing pick-up points have to merge with the front and rear spars, and this combination can become complicated. It is common for the lower region to use a construction machined from forgings, but to use simpler rolled sections like channels for the upper half of the mainframes. The carry-through wing center section is not expected to be pressurized like the fuselage. The fuselage is of course an efficient membrane structure, but at the wing box upper surface the wing structure has to take the full pressure loading (half an atmosphere in most civil aircraft). The wing upper surface has very little curvature, so it has to take the bending action, often assisted by corrugations or sandwich panels.

Similar heavy frames will be needed if the undercarriage loads go straight to the fuselage rather than via the wing. There will also be a need for strong frames or bulkheads where the tail-plane (horizontal stabilizer) and the fin (vertical stabilizer) need to react loads into the fuselage. It is doubtful whether any of these stiff frames will be stiff enough to induce the simple beam theory shears shown in Figure 9.35. This may be a good starting point in estimating their desirable performance, but in the end a finite-element analysis will probably be necessary to allow the finite bending stiffness of these mainframes to act their role accurately. The use of finite-element codes such as NASTRAN and CATIA/ELFINI is now routine in the aerospace industry.

As in the wing box structure, in addition to these special frames, there will be a large number of secondary lightweight frames at regular intervals—in the case of a civil aircraft, at twice the window pitch. These *standard* frames will act like wing ribs in resisting global buckling between them. They also act as a resistor to the Brazier effect, but fuselage bending deformations are much less than the wing. A simple criterion for the necessary frame stiffness (EI) sufficient to inhibit this sort of buckling can be expressed as (Niu 1988) $EI = MD^2\,1600\,L$, where M is the local bending moment, D is the fuselage diameter, and L is the frame pitch. The other role of the standard fuselage frames is to act as effective crack stoppers if

the pressure cycle causes fatigue cracks which need to be found, at regular inspections, without becoming unstable.

The circular cross-section is an efficient membrane in resisting internal pressure, but other variations have been used to optimize passenger/cargo payload or to allow more than one passenger deck. The essential membrane action can still be exploited provided a stiff member reacts the kink loads across the section. Figure 9.36 show the possibilities (not to scale).

Some large transport/freighter aircraft need to have low-access freight doors and will therefore use shorter undercarriages retracting into the fuselage as shown in Figure 9.37. This effectively removes the lower portion of the fuselage and to compensate and maintain the flexural stiffness, a longitudinal center-section or keel beam is added. The diffusion of the keel beam loads back into the fuselage at the end of this undercarriage bay is an interesting exercise in design, aided of course

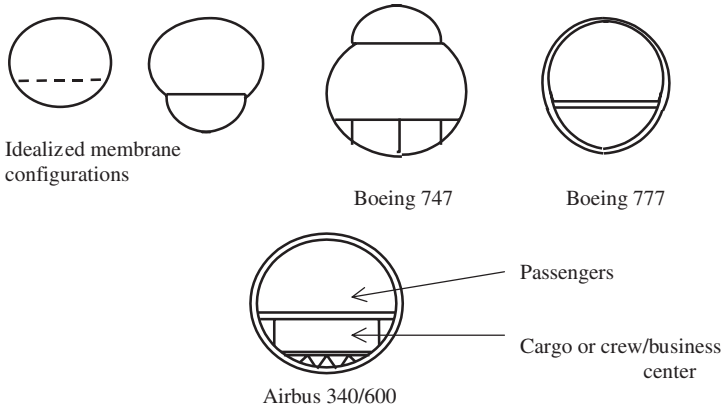


FIGURE 9.36 Various fuselage configurations.

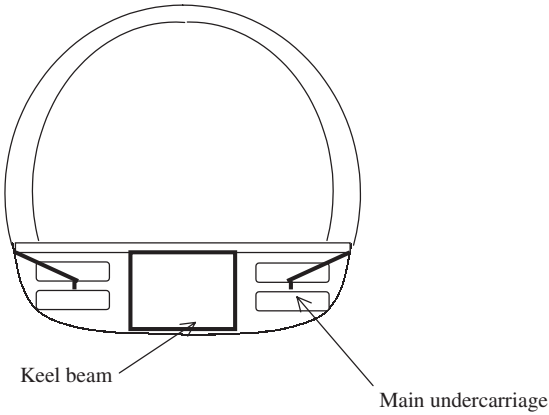


FIGURE 9.37 Low loader transporter.

by finite-element analysis. Keel beams occur frequently in military aircraft whose fuselage is interrupted in many places by weapons and engine ducts/exhausts. For stealthy aircraft all bombs and weapons now have to be stored internally so the lower fuselage may be predominantly of the keel beam type with the outer skin a series of non-load carrying doors and removable panels.

The connection of the tail-plane and the fin to the fuselage is a little easier than the main wings since this part of a civil aircraft will generally be aft the rear pressure bulkhead. All the tail-plane and fin spars can therefore run smoothly into complete bulkheads without the need to have open frames.

9.10 STRESS CONCENTRATIONS

So far we have mostly looked at structures in an overall (beam or tube) sense, in which the stress fields are very simple. We then looked at the effects of concentrated loads which will cause departures from these simple smooth fields unless special structures and local diffusing members are used. There is, however, another cause of local stress concentration—the presence of cut-outs, holes, fillets, in fact any local curvature to a boundary. These local stress concentrations can be confined to a small zone, but if the stress intensity factor (Part 2) is large enough and the fracture toughness small then a crack may propagate either statically or in fatigue. There are literally hundreds of small stress raises in a typical aircraft, and not all of them will be successfully designed out or accounted for. This is one reason why a full-scale fatigue test will probably always be mandatory.

Probably the most infamous case of fatal stress concentration was that of the Comet failures in the 1950s. This aircraft was the first commercial jetliner to fly above 10,000 meters, at a higher pressure differential than any other. After several cases of midair explosive failures the cause was diagnosed as unstable crack growth from the windows. These windows were oval in shape with the major axis horizontal—very pleasing aesthetically but structurally disastrous, as we now show. A hole in a uniformly stressed skin will cause the stress fields to flow around it. This analogy is quite a good one since idealized (potential) flow around a 2D obstacle can be readily imagined, as in Figure 9.38. Point A is a stagnation point and point B is the point of maximum suction, in terms of $\frac{1}{2}\rho V^2$ the pressures being -1 and $+3$. These are also the stress concentrations due to a circular hole in a uniform stress field.

As it happens, the stress concentration around an elliptical hole, with semi-axis a and b (b at right angles to the applied stress) can be solved analytically as $1 + 2b/a$. For the circle ($b = a$) the value is 3. For a crack-like ellipse aligned with the stresses ($b = 0$) there is no stress concentration, but for one aligned at right angles ($a = 0$) the stress concentration becomes infinite. If we look at the special case of a pressurized fuselage with a hoop stress twice that of the longitudinal stress, the stress concentration at the horizontal axis points is $0.5 + 2a/b$ and that on the vertical axis is $-0.5 + b/a$.. Clearly the stress concentration depends on the shape, which is a cheap commodity. For example if we choose an ellipse of ratio $b/a = 2$, the stress concentration everywhere around the boundary would be 1.5. However, it is possible to reduce such stress concentrations by reinforcing the edge of the hole with a compact reinforcing member, which, in the case of a pressure cabin window, it is necessary to have anyway to carry the edge of the window and seal. It is actually possible to design a neutral hole with no stress concentrations at

AEROSPACE STRUCTURES

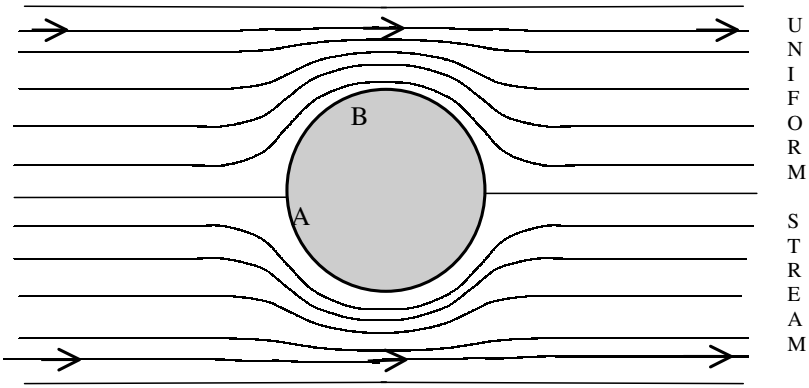


FIGURE 9.38 Idealized potential flow around a cylinder.

all, by having an axis ratio of $b/a = \sqrt{2}$, but the reinforcement has to be a complicated and costly shape. However even with a uniform cross-section the stress concentration factor is very close to 1 and the reinforcement about three times the weight of the hole removed. After the Comet experience several European aircraft manufacturers used this shape, with extraordinary large windows. The Fokker F27 is still flying. Most aircraft nowadays settle for a rather smaller window having a square shape with rounded corners. The much larger doors will also be rectangular with rounded corners and will be reinforced with a frame and significant amounts of an additional patch concentrated near the corners. The reinforcement needed, whether a reinforcing plate or additional frame stiffening or both, will need a finite-element analysis to evaluate any stress concentrations.

Another good example is the large number of access panels needed on the lower surface of a modern high-speed civil-aircraft wing structure, in order to be able to inspect (or repair). These are oval (aligned correctly with the span-wise tension field) with suitable edge reinforcements, rather than using a load-carrying panel which would rely on close-fitting removable bolts. The stress field for high-aspect ratio wings is predominantly uniaxial in this region, but if the reader cares to examine the access panels on the underside of the Concorde's triangular wings the varying direction of the stress field becomes immediately apparent.

One of the victims of stress concentrations is almost any composite structure, since laminated composites have very little strength in the through-thickness direction.

9.11 COMPOSITE STRUCTURES

Laminated composite structures do seem to have everything going for them when we judge the specific strength and stiffness properties summarized in Table 9.2. As mentioned, for military aircraft, where performance has been the main driver, particularly aeroelastic efficiency, there has been little choice other than to use carbon/epoxy composites. Now that affordability is probably the main driver for military as well as civil aircraft, we might expect composites to become less attractive. This

is not now so since composite structures have made great strides in reducing cost. The unitization concept, which involves reducing dramatically the separate number of parts, is now being exploited for composite structures. The degree of automation in manufacturing has advanced to the use of automated tow or tape-laying machines whose head can move with 5 degrees of freedom so that double curvature and any choice of stacking sequence and orientation is now possible. It also looks as if techniques borrowed from textile technology will drastically reduce manufacturing costs with a small loss of performance. These techniques use dry or tacky fiber preform, often a woven fabric, and then infuse the resin under pressure and heat. Resin transfer molding (RTM)) introduces the resin at selected points, and resin infusion methods deploy film or tiles which can be sandwiched between layers of the preform and then infused as the temperature increases. One of the reasons for high costs is the necessary use of an autoclave for thermosetting resins. We expect to see increasing use of room temperature cure, probably using vacuum-assisted RTM. So if we look at the use of carbon fiber composites in civil aircraft (where the acquisition cost certainly has to be justified), they are used, for example, in the Airbus A340-600 (long-range version) for the fin, the fuel-carrying tail-plane, rudder and elevator, ailerons, wing-fuselage fairings, floors, engine cowlings, nose and main undercarriage doors, and so on, but no composites yet for the primary wing and fuselage structure. And yet the fatigue performance of carbon fiber composites is exceptional. All this is not excessive conservatism. In fact, the wings of the single-aisle long-range Boeing and Airbus aircraft will almost certainly be composites, and cured in one piece to avoid heavy bolted joints. Herein lies the weakness of laminated composites. They are vulnerable to stresses transverse to the fiber direction where failure values in tension or shear may be only 4% of the in-plane strength. As stress concentrations give rise to local stress gradients, and if there is a gradient in stress along the direction of the fibers, the equations of equilibrium (pg 9.21) show that there will be an interlaminar shear, and if these also have a gradient there will be a through-thickness compression or tension (peeling) stress. Industry is well aware of this and has useful design rules to alleviate the problem. For example, when laminates are tapered a minimum taper ratio of 1:20 is aimed for. However, this does not solve the problem of complex parts such as any joint, whether bonded or bolted. It has to be admitted that highly 3D bolted joints are designed very conservatively and are probably too heavy. However, the ability of structural designers to recognize, analyze, and design for these stress concentrations will undoubtedly improve.

Finally, a novel hybrid composite material called Glare is being introduced for newer aircraft like the Airbus 380. This material is a sandwich of unidirectional glass fiber and very thin sheets of aluminum alloy, produced in any stacking sequence that the customer needs. The main reason for this hybrid was to exploit the fiber bridging capability of high-strain glass fiber if the aluminum alloy developed a crack. This it does admirably, although the basic cost of Glare is high. However, when it is manufactured then all the very thin aluminum sheets can be laid down in a staggered fashion so that their edges are not on top of each other. This means that very much larger sheets can be made than that of the basic aluminium alloy. The very large fuselage of the Airbus 380 can then be constructed using only one or two sheets of Glare fore and aft of the wing support frames. This saves more manufacturing costs than the basic increase in the cost of the material. It should also be welcomed by airline operators since its fatigue performance is without equal in homogenous metallic structures.

Further details of some composite manufacturing processes are given in Part 5.

AEROSPACE STRUCTURES

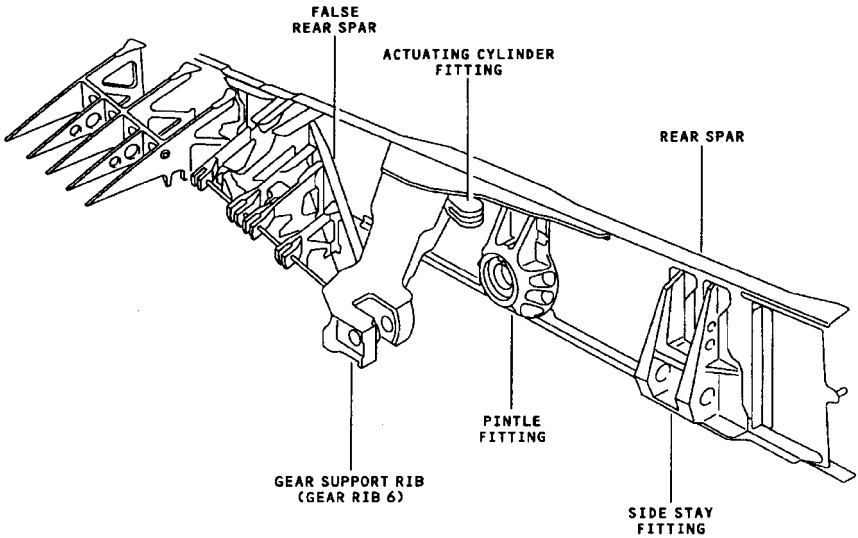


FIGURE 9.39 Attachment points of landing gear to rear spar on Airbus 340. (Courtesy of Airbus UK Ltd.)

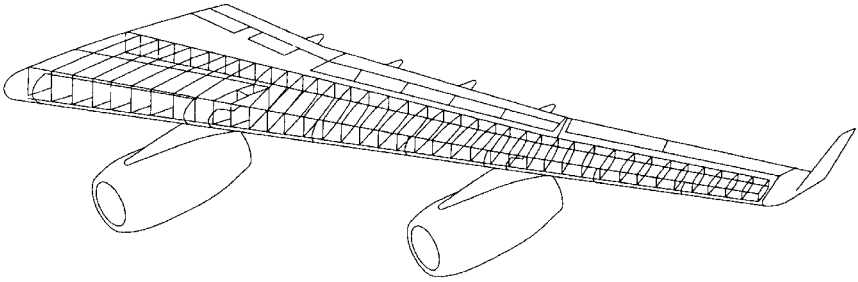


FIGURE 9.40 Layout of ribs in Airbus 340. (Courtesy of Airbus UK Ltd.)

9.12 STRUCTURAL TESTS

The emphasis in this long Part 3 on structural considerations has been on understanding the nature of structures, how they react to loads, and the factors governing their efficiency. This has been done without any great excursion into structural analysis, although the finite-element method has been described as the basic tool for analyzing structures of great complexity. Analysis is part of the total design process, and in a modern aerospace design organization the analytical (commercial) software design tools must be embedded in an easily accessible form into the integrated product development system. This is a networked managed suite consisting of everything from the original digital assembly (electronic mock-up) through the aerodynamics (CFD), structural analysis, systems engineering and control, virtual

manufacturing, process modeling, etc. and finishing at in-service maintenance. Moreover, all stages are linked and updated concurrently so that the evolution from concept to final production can be achieved without any stage waiting unduly for the latest update upstream or downstream.

It would be nice if the structural analysis were such an accurate simulation of reality that no structural testing would be required; that is, that all tests could be virtual. Industry is certainly moving in this direction since the potential saving, in time-to-market and acquisition cost, by simulating all of the product development stages is 50–60%.

However, the airworthiness authorities are unlikely ever to accept virtual alternatives to some tests of the complete airframe. The reason, of course, is the likelihood of errors in the analytical model, the material properties, and the physics of the failure mechanisms when predicting the ultimate load or fatigue life. Although modern finite-element codes and their tools for creating a model and interpreting the answers, have achieved great sophistication, human creation is not infallible. One of the errors in the finite-element modeling stage is an omission of a local feature, which gives rise to a stress concentration. As discussed earlier, the stress concentrator may be very local but failure propagation from it may be unstable in fracture or buckling. This fact alone will make authorities insist on a full-scale fatigue test. For composite structures, which are very vulnerable to stress concentrations, even for static loading, the authorities will expect much better models than currently available.

Full-scale tests will attempt to reproduce the worst loading case, but this is difficult since different parts of the structure may be critically stressed under different loading conditions. Thus a safe envelope is aimed at, based on the vertices of Figures 9.2 and 9.3. The distributed aerodynamic loading is usually applied through a "whiffle tree," crudely shown in Figure 9.41, and the inertia loading simulated by a series of jacks at selected points plus a few ground reactions.

The loading will be applied in increments while the structure is heavily instrumented with many strain gauges and displacement transducers. Accelerometers may also be used in case the sequence of events during the very rapid final collapse is important. These may also trigger high-speed cameras. It is common for some minor failure to occur before even the proof load, usually because a component was thought not to be structurally important or load carrying and was not included in the modeling. It is relatively rare for complete structural failure to occur before the ultimate design load. The lot of structural engineers is tough. If failure occurs at 95% of UDL they are in trouble. If failure occurs at 105% they are still in trouble for designing an overweight airframe.

The point of using as much instrumentation as possible for tests is not only to verify the UDL but also to check the strain and displacement fields to verify that

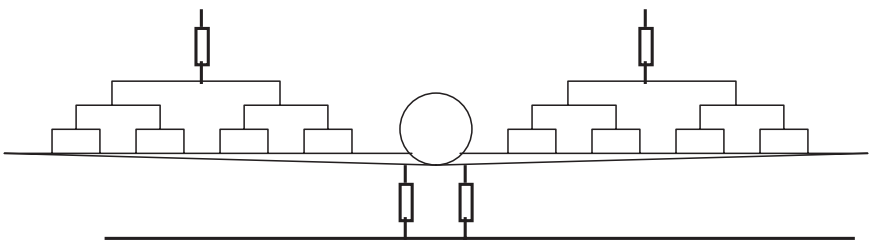


FIGURE 9.41 Loading frames for a "whole aircraft" structural test.

the analysis was sound. This is important since civil aircraft will undoubtedly be stretched in future Marks to increase payload or range. The certification authorities will accept analysis to requalify a new variant, the philosophy being that any problematic stress concentrations will be the same.

The fatigue test will be a separate exercise on another airframe. The mechanical loading will be cycled in a complicated block diagram sequence representing a typical history for that aircraft and its customers. In addition to mechanical loading in a civil aircraft, the fuselage will need to be pressure cycled. This will be done in a surrounding water tank since the explosive energy released from compressed air is unacceptable.

In addition to full-scale aircraft tests, there will usually be a series of component tests for items which are not covered by the envelope, such as landing gear, flap tracks, and the like, or perhaps representative sections such as a stiffened compression panel. However, the latter type is being replaced by analysis since it has traditionally been used where analysis was in doubt. After all, the full airframe test covers the certification requirement, but that is not the time for the manufacturer to discover its sums were wrong.

Finally, there may be a whole series of tests of components or portions of structure where the loadings are not simple aerodynamic or mechanical. This is likely to happen for military vehicles whose environment is very hostile. Examples are ballistic or blast loading; highly energetic acoustic loadings near an engine exhaust; or shock resonance loading inside an opened weapons bay in a supersonic aircraft. Another similar example is flutter, where the aerodynamic forcing terms are very difficult to simulate. This is covered in Part 4.

PART 4
STRUCTURAL DYNAMICS

Jonathan Cooper

9.13 INTRODUCTION

Aerospace vehicles are subject to a wide variety of dynamic loads both in flight (e.g., maneuvers, gusts, buffet, acoustic excitation, etc.) and on the ground (e.g., landing and taxiing). These loads cause the structures to vibrate due to their inherent flexibility. It is essential that these vibrations be constrained within certain design limits so that structural failure does not occur (either due to excessive loads or fatigue), the desired maneuverability and performance are achieved, and passenger comfort and integrity of equipment and systems are maintained.

Great emphasis is placed upon obtaining accurate mathematical models of aerospace vehicles to ensure that the predicted behaviour in flight is accurate. These models are validated by comparison with vibration and control system tests on the ground, and also by wind tunnel model testing. Flight flutter testing is undertaken to demonstrate that the aircraft is flutter-free throughout its desired flight envelope. This approach not only enables the dynamic responses to stay within desirable limits but also ensures that costly redesigns do not have to be made once prototype vehicles have been built.

This Part provides a brief overview of techniques for modeling continuous structures as multiple-degree of freedom (MDOF) systems and predicting the vibration response to some predefined input. A number of aeroelastic phenomena are also described, as well as consideration of the effect of noise on aerospace vehicles and some specific vibration problems encountered in helicopters. Subsection 9.15 on vibration provides the background on single-degree of freedom (DOF) systems relevant to this item.

9.14 MULTIPLE DEGREE OF FREEDOM VIBRATION

The key to understanding MDOF vibration is the concept of a mode of vibration, whereby the vibration response of a structure to some input is obtained as the summation of the response of each mode (we shall consider only linear systems). Each mode is characterized by a natural frequency, damping ratio, and mode shape, and its response depends upon the amplitude, frequency content, and position of the input. In this section, it will be shown how the MDOF equations of motion are derived and the response of a system to a forced input is modeled using frequency response functions (FRFs)

9.15 MDOF FREE VIBRATION—LUMPED MASS MODELS

Consider the bending behavior of the simple uniform wing in Figure 9.42 that we wish to model as, say, three equal sections. The displacement at the center of each

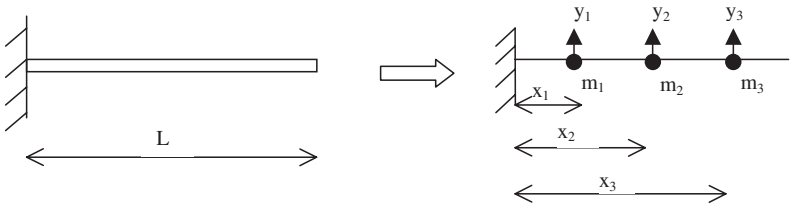


FIGURE 9.42 Modeling of simple cantilever wing using lumped masses.

section is defined as y_i with the mass distributed as a number of lumped masses positioned at each center.

Setting up Equations of Motion—Flexibility Matrix Approach

The flexibility influence coefficients δ_{ij} are defined as the displacement at position i due to a unit load at position j (units m/N) as shown in Figure 9.43.

The values of δ_{ij} can be determined from simple beam bending theory or one of the many energy methods (e.g., unit load method). The influence coefficients of a uniform beam are found as

$$\delta_{ij} = \int_0^{y_i} \frac{(y_i - y)(y_j - y)}{EI}\, dy = \frac{y_i^2}{6\,EI}(3y_j - y_i) \qquad y_i \le y_i$$

Becauses the system is linear, Maxwell's reciprocity theorem holds and therefore $\delta_{ji} = \delta_{ij}$. Having found the flexibility influence coefficients, the total deflection y_i at station i due to forces F_j at all stations j is

$$y_i = \delta_{i1}\mathrm{F}_1 + \delta_{i2}\mathrm{F}_2 + \delta_{i3}\mathrm{F}_3$$

In matrix form, this expression leads to

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} \delta_{11} & \delta_{12} & \delta_{13} \\ \delta_{21} & \delta_{22} & \delta_{23} \\ \delta_{31} & \delta_{32} & \delta_{33} \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} \qquad \text{or } \mathbf{y} = \Delta\mathbf{F} \Rightarrow \mathbf{F} = \mathbf{K}\mathbf{y}$$

where Δ is the flexibility matrix and its inverse is $\mathbf{K}$, the stiffness matrix. When the system is allowed to vibrate, the beam introduces forces to the idealized lumped masses, causing them to accelerate and introduce reactive inertial forces on the beam such that

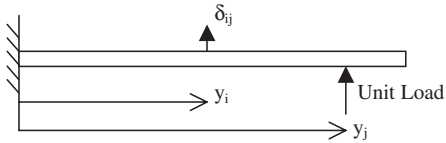


FIGURE 9.43 Definition of flexibility influence coefficient δ_{ij}.

$$F_j = -m_j \ddot{y}_j$$

Combining the above two equations gives the second order differential equation

$$\begin{bmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{bmatrix} \begin{bmatrix} \ddot{y}_1 \\ \ddot{y}_2 \\ \ddot{y}_3 \end{bmatrix} + \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \mathbf{M}\ddot{\mathbf{y}} + \mathbf{K}\mathbf{y} = \mathbf{0}$$

Increasing the number of masses will improve the model. The method can be extended to include rotary inertia and torsional effects. External forces can be included as long as they are applied at the masses and in the direction of the degree of freedom. For example, if a force P is applied to mass 3 acting vertically upwards, then the equations of motion become

$$\begin{bmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{bmatrix} \begin{bmatrix} \ddot{y}_1 \\ \ddot{y}_2 \\ \ddot{y}_3 \end{bmatrix} + \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ P \end{bmatrix} \Rightarrow \mathbf{M}\ddot{\mathbf{y}} + \mathbf{K}\mathbf{y} = \mathbf{P}$$

9.16 SETTING UP EQUATIONS OF MOTION—FINITE-ELEMENT APPROACH

The finite element method has been discussed in Part 3 in principle. It is convenient to illustrate one element type here—a beam element (see Figure 9.44).

To ensure kinematic continuity from one element to another we choose as the nodal unknowns the displacements d_1 and d_3 and the end rotations d_2 and d_4. A cubic polynomial shape function is defined uniquely in terms of these four nodal displacements. The strains follow from the curvature $1/R = d^2y/dx^2$, and integrating the internal and external virtual work we derive 4 equilibrium equations for an element as

$$\mathbf{p} = \mathbf{k}\mathbf{d}$$

where the stiffness matrix $\mathbf{k}$ emerges as

$$\mathbf{k} = \frac{EI}{L^3} \begin{bmatrix} 12 & 6L & -12 & 6L \\ 6L & 4L^2 & -6L & 2L^2 \\ -12 & -6L & 12 & -6L \\ 6L & 2L^2 & -6L & 4L^2 \end{bmatrix}$$

This expression assumes the x-axis is along the beam but can be amended to allow for an element inclined at any angle.

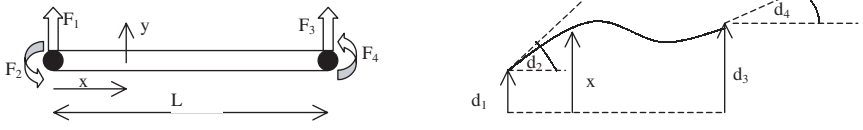


FIGURE 9.44 Definition of degrees of freedom for two-node beam finite element.

The corresponding mass matrix is found by considering the beam element to be in motion. The simplest (lumped) representation is to consider the mass of the element of be distributed equally at each end. Alternatively, a more consistent mass matrix can be found by considering the mass to be distributed uniformly along the beam, whence the virtual work integral delivers the inertia body forces. These mass matrices are found as

$$\mathbf{m}_{\text{lump}} = \frac{\rho AL}{24}\begin{bmatrix} 12 & 0 & 0 & 0 \\ 0 & L^2 & 0 & 0 \\ 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & L^2 \end{bmatrix}$$

and

$$\mathbf{m}_{\text{uniform}} = \frac{\rho AL}{420}\begin{bmatrix} 156 & 22L & 54 & -13L \\ 22L & 4L^2 & 13L & -3L^2 \\ 54 & 13L & 156 & -22L \\ -13L & -3L^2 & -22L & 4L^2 \end{bmatrix}$$

The beam element equations of motion then become

$$\mathbf{m}\ddot{\mathbf{d}} + \mathbf{kd} = \mathbf{p}$$

9.17 GLOBAL STIFFNESS AND MASS MATRICES—ELEMENT ASSEMBLY

In order to model the overall behavior of a structure, global mass and stiffness matrices must be found by using the local mass and stiffness matrices of all of the individual elements along with their position on the structure. This process is best illustrated using the uniform wing example with two equal beam elements with local mass and stiffness matrices m^1, m^2, k^1, and k^2 as shown in Figure 9.45.
The corresponding nodal deflections between the element and global systems are mapped in order to ensure compatibility at the nodes such that

$$\{D_1\ D_2\ D_3\ D_4\} = \{d_1\ d_2\ d_3\ d_4\}_{\text{element 1}} \text{ and } \{D_3\ D_4\ D_5\ D_6\} = \{d_1\ d_2\ d_3\ d_4\}_{\text{element 2}}$$

Thus, the global system equations of motion can be formulated by placing the elements of the local mass and stiffness matrices onto the relevant parts of the global matrices **M** and **K** such that

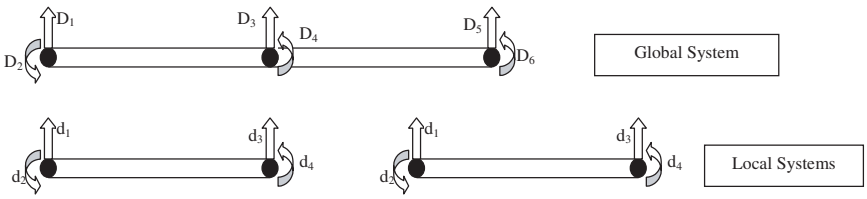


FIGURE 9.45 Definition of degrees of freedom for local and global systems.

$$\mathbf{M}\ddot{\mathbf{D}} + \mathbf{K}\mathbf{D} = \begin{bmatrix} m^1 & & \\ & m^2 & \end{bmatrix}\begin{bmatrix} \ddot{D}_1 \\ \ddot{D}_2 \\ \ddot{D}_3 \\ \ddot{D}_4 \\ \ddot{D}_5 \\ \ddot{D}_6 \end{bmatrix} + \begin{bmatrix} k^1 & \\ & k^2 \end{bmatrix}\begin{bmatrix} D_1 \\ D_2 \\ D_3 \\ D_4 \\ D_5 \\ D_6 \end{bmatrix} = \mathbf{P}(t)$$

or

$$\begin{bmatrix} m_{11}^1 & m_{12}^1 & m_{13}^1 & m_{14}^1 & 0 & 0 \\ m_{21}^1 & m_{22}^1 & m_{23}^1 & m_{24}^1 & 0 & 0 \\ m_{31}^1 & m_{32}^1 & m_{33}^1 + m_{11}^2 & m_{34}^1 + m_{12}^2 & m_{13}^2 & m_{14}^2 \\ m_{41}^1 & m_{42}^1 & m_{43}^1 + m_{21}^2 & m_{44}^1 + m_{22}^2 & m_{23}^2 & m_{24}^2 \\ 0 & 0 & m_{31}^2 & m_{32}^2 & m_{33}^2 & m_{34}^2 \\ 0 & 0 & m_{41}^2 & m_{42}^2 & m_{43}^2 & m_{44}^2 \end{bmatrix}\begin{bmatrix} \ddot{D}_1 \\ \ddot{D}_2 \\ D_3 \\ \ddot{D}_4 \\ \ddot{D}_5 \\ \ddot{D}_6 \end{bmatrix}$$

$$+ \begin{bmatrix} k_{11}^1 & k_{12}^1 & k_{13}^1 & k_{14}^1 & 0 & 0 \\ k_{21}^1 & k_{22}^1 & k_{23}^1 & k_{24}^1 & 0 & 0 \\ k_{31}^1 & k_{32}^1 & k_{33}^1 + k_{11}^2 & k_{34}^1 + k_{12}^2 & k_{13}^2 & k_{14}^2 \\ k_{41}^1 & k_{42}^1 & k_{43}^1 + k_{22}^2 & k_{44}^1 + k_{22}^2 & k_{23}^2 & k_{24}^2 \\ 0 & 0 & k_{31}^2 & k_{32}^2 & k_{33}^2 & k_{34}^2 \\ 0 & 0 & k_{41}^2 & k_{42}^2 & k_{43}^2 & k_{44}^2 \end{bmatrix}\begin{bmatrix} D_1 \\ D_2 \\ D_3 \\ D_4 \\ D_5 \\ D_6 \end{bmatrix} = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \\ P_6 \end{bmatrix}$$

These equations need to be partitioned so that the boundary conditions (in this case $D_1 = D_2 = 0$) can be accounted for, thus

$$\mathbf{M}\ddot{\mathbf{D}} + \mathbf{K}\mathbf{D} = \left[\begin{array}{c|c} M_{aa} & M_{ab} \\ \hline M_{ba} & M_{bb} \end{array}\right]\begin{bmatrix} \ddot{D}_a \\ \ddot{D}_b \end{bmatrix} + \left[\begin{array}{c|c} K_{aa} & K_{ab} \\ \hline K_{ba} & K_{bb} \end{array}\right]\begin{bmatrix} D_a \\ D_b \end{bmatrix} = \begin{bmatrix} P_a \\ P_b \end{bmatrix}$$

where $D_a = (D_1 D_2)^T$, $D_b = (D_3 D_4 D_5 D_6)^T$, P_a are the reactions at DOF 1 and 2, P_b are the forces vector at the nodes. The dynamic equations then become

$$\mathbf{M}_{bb}\ddot{\mathbf{D}}_b + \mathbf{K}_{bb}\mathbf{D}_b + \mathbf{P}_b$$

This process can be readily extended to any finite-element model of a complicated structure.

9.18 MDOF VIBRATION—SETTING UP EQUATIONS OF MOTION—RAYLEIGH-RITZ METHOD

An alternative approach uses assumed mode shapes over the entire structure and clearly is most powerful when the structure has a fairly simple geometry. The

Rayleigh–Ritz method characterizes the motion of the structure as the summation of a series of assumed shapes multiplied by unknown constants. The equations of motion can then be defined in terms of the unknown constants through the use of either energy or work principles. Each shape must satisfy the geometric boundary conditions and can either be polynomial or trigonometric functions. The more shapes that are used, the more accurate the solution becomes.

Consider the uniform cantilever wing as above (mass per unit length m and flexural rigidity EI) and approximate its bending deflection to be made up of two shapes (quadratic and cubic) such that

$$y = a_1x^2 + a_2x^3$$

The unknown a_i terms are the generalized coordinates that need to be determined. This shape satisfies the boundary conditions $\mathbf{y} = \mathbf{y}' = 0$ at $x = 0$.

The work/energy method can be expressed in terms of Lagrange's equations, which are:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right) + \frac{\partial D}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i} + \frac{\partial U}{\partial q_i} = Q_i$$

where q_i is the ith generalized coordinate, T is the kinetic energy, D is the dissipation function, U is the potential energy, and Q_i is the ith generalized force $= \partial(\delta W)/\partial(\delta q_i)$, with δW being the work done by the forces moving through incremental displacements. The kinetic energy is found as

$$T = \frac{1}{2}\int_0^L m\dot{y}^2\,dx = \frac{1}{2}\int_0^L m(\dot{a}_1x^2 + \dot{a}_2x^3)^2\,dx$$

and the potential energy due to bending deformation (strain energy) is

$$U = \frac{1}{2}\int_0^L \sigma\varepsilon\,dx = \frac{1}{2}\int_0^L EI\left(\frac{\partial^2 y}{\partial x^2}\right)^2 dx = \frac{1}{2}\int_0^L \mathrm{EI}(2a_1 + 6a_2x)^2\,dx$$

Applying Lagrange's equations for both generalized coordinates gives

$$m\begin{bmatrix} 1/5 & L/6 \\ L/6 & L^2/7 \end{bmatrix}\begin{bmatrix} \ddot{a}_1 \\ \ddot{a}_2 \end{bmatrix} + \frac{EI}{L^4}\begin{bmatrix} 4 & 6L \\ 6L & 12L^2 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

which is again in the general form

$$\mathbf{M}\ddot{\mathbf{y}} + \mathbf{K}\mathbf{y} = \mathbf{0}$$

More accurate solutions are obtained when more shapes are used. It is possible to include changes in the cross-sectional geometry through defining I as a function of x and also other types of stress/strain (e.g., torsion, shear, etc.). External forces are included through the use of generalized force terms in Lagrange's equations. For example, if a load P is applied at the wingtip in the vertical direction, then the generalized force term comes from the work

$$\delta W = P(\delta a_1x^2 + \delta a_2x^3)_{x=L} = P(\delta a_1L^2 + \delta a_2L^3)$$

and the equations of motion become

$$m\begin{bmatrix} 1/5 & L/6 \\ L/6 & L^2/7 \end{bmatrix}\begin{bmatrix} \ddot{a}_1 \\ \ddot{a}_2 \end{bmatrix} + \frac{EI}{L^4}\begin{bmatrix} 4 & 6L \\ 6L & 12L^2 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = P\begin{bmatrix} L^2 \\ L^3 \end{bmatrix} \Rightarrow \mathbf{M\ddot{a} + Ka = P}$$

Undamped Free Vibration

The equations of motion having been determined, the vibration properties of the system can be calculated and the response to any given input determined. The natural frequencies, mode, shapes and damping ratios are obtained by considering the free vibration response (system given an initial displacement, or velocity, and then allowed to vibrate). At first we shall only consider the case with no damping.

For the M DOF system equations, obtained using one of the above methods, $\mathbf{M\ddot{y} + Ky = 0}$, we can assume a solution of the form $\mathbf{y} = \mathbf{y_0} \sin \omega t$ and hence obtain the expressions

$$(-\omega^2\mathbf{M} + \mathbf{K})\mathbf{y}_0 = \mathbf{0} \text{ or } (\mathbf{M}^{-1}\mathbf{K} - \mathbf{I}\,\omega^2)\mathbf{y}_0 = \mathbf{0}$$

These equations are in the form of an *eigenvalue* problem. The eigenvalues of matrix $\mathbf{M^{-1}K}$ are the square of the natural frequencies. The corresponding *eigenvectors* are the mode shapes that determine the relative position of all parts of the structure for each mode during each vibration cycle, but not the relative amplitude of the modes. If the deflection of a particular mode is known at one position on the structure, then the position of all the other points can be determined. For the Rayleigh–Ritz approach, the resulting eigenvectors give the values of the a_i coefficients for each mode, and the corresponding mode shapes are then found using the defined deflection shapes.

Consider the system equations derived using the flexibility matrix method for a uniform wing split into three equal sections ($m_1 = m_2 = m_3 = m_0L/3$ and $x_1 = L/6$, $x_2 = L/2$, $x_3 = 5L/6$). The flexibility and mass matrices are found as

$$\Delta = \frac{L^3}{648\,EI}\begin{bmatrix} 1 & 4 & 7 \\ 4 & 27 & 54 \\ 7 & 54 & 125 \end{bmatrix} \quad \mathbf{M} = \frac{m_0L}{3}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

corresponding to displacements (y_1, y_2, y_3). The natural frequencies and mode shapes are shown in Table 9.3 and Figure 9.46.

For the undamped case the mode shapes are *real*. This means that the shapes are either exactly in (0°) or out (180°) of phase, resulting in simultaneous max/min or zero deflection for all parts of the structure in that mode. Node lines (points of zero deflection) remain the same throughout the vibration of the mode shape. The values for the first mode are close to the exact values, whereas the estimates for the second and third modes are not good. They would be improved by including more masses in the model.

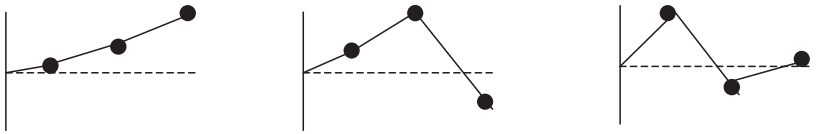
Damped Free Vibration—Proportional Damping

The most common approach to modeling damping is to add what is known as proportional damping. The equations of motion become $\mathbf{M\ddot{y} + C\dot{y} + Ky = 0}$, where the damping matrix $\mathbf{C} = \alpha\mathbf{M} + \beta\mathbf{K}$, with α and β being constants.

To solve the equations of motion, it is most convenient to rewrite the equations of motion in first order form in terms of the vector $[y\ \dot{y}]^t$ such that

TABLE 9.3 Natural Frequencies, Damping Ratios, and Mode Shapes for Three-Mass Beam

	No damping			Proportional damping			Nonproportional damping		
	Mode 1	Mode 2	Mode 3	Mode 1	Mode 2	Mode 3	Mode 1	Mode 2	Mode 3
Natural freq/ $\sqrt{\frac{EI}{m_0 L^4}}$	3.61	24.17	77.69	3.61	24.17	77.69	3.61	24.17	77.69
Damping ratio				0.018	0.121	0.389	0.0405	0.2094	0.7335
Mode shape	$\begin{bmatrix} 0.0591 \\ 0.4433 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.3273 \\ 1 \\ -0.4626 \end{bmatrix}$	$\begin{bmatrix} 1 \\ -0.2943 \\ 0.0713 \end{bmatrix}$	$\begin{bmatrix} 0.0591 \\ 0.4433 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.3273 \\ 1 \\ -0.4626 \end{bmatrix}$	$\begin{bmatrix} 1 \\ -0.2943 \\ 0.0713 \end{bmatrix}$	$\begin{bmatrix} 0.059 - 0.004i \\ 0.44 - 0.006i \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.30–0.08i \\ 1 \\ -0.44 + 0.04i \end{bmatrix}$	$\begin{bmatrix} 1 \\ -0.12 + 0.15i \\ 0.027 - 0.04i \end{bmatrix}$



Mode 1. Mode 2 Mode 3

FIGURE 9.46 Mode shapes for undamped beam modeled with three lumped masses.

$$\begin{bmatrix} -\mathbf{K} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} \end{bmatrix}\begin{bmatrix} \dot{\mathbf{y}} \\ \ddot{\mathbf{y}} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{K} \\ \mathbf{K} & \mathbf{C} \end{bmatrix}\begin{bmatrix} \mathbf{y} \\ \dot{\mathbf{y}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \Rightarrow \begin{bmatrix} \dot{\mathbf{y}} \\ \ddot{\mathbf{y}} \end{bmatrix} - \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} \end{bmatrix}\begin{bmatrix} \mathbf{y} \\ \dot{\mathbf{y}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \Rightarrow \dot{\mathbf{z}} - \mathbf{Q}\mathbf{z} = \mathbf{0}$$

The eigenvalues of matrix $\mathbf{Q}$ appear as complex conjugate pairs for an oscillating system and can be shown to take the form $\lambda = -\zeta\omega \pm j\omega\sqrt{1-\zeta^2}$, thus the natural frequency and damping can be found. The corresponding eigenvectors also occur in complex conjugate pairs in the form $[\boldsymbol{\Psi}/\boldsymbol{\lambda}\boldsymbol{\Psi}]$, where $\boldsymbol{\Psi}$ are the mode shapes.

Consider the same system as above, but now with a damping matrix of $\mathbf{C} = \mathbf{K}/100$. Estimates for the natural frequencies, damping ratios, and mode shapes are found in Table 9.3. Note that although there is now some damping acting upon the system, and for a free decay each mode vibrates at its damped natural frequency, the mode shapes are exactly the same as if there were no damping.

Free Vibration—Nonproportional Damping

Consider now adding the same damping matrix $\mathbf{C} = \mathbf{K}/100$ but also with the top left hand term in $\mathbf{C}$ being multiplied by 2. This now breaks up the proportional damping used in the previous case. The estimates of the natural frequency, damping ratio and mode shapes are also shown in Table 9.3. It can be seen that the mode shapes now contain some complexity. Although the points making up the mode shape maintain their relative amplitude and phasing, the complexity means that there is now nonsimultaneous maximum/minimum excursion and the position of the nodal points varies throughout each vibration cycle.

Generalized Coordinates

For systems with a large number of degrees of freedom, such as an aircraft, it is necessary to make a transformation into so-called generalized (or principal or modal) coordinates q. This operation is achieved by first determining the matrix of mode shapes (the eigenvectors) and the modal matrix $\boldsymbol{\Psi}$, and then performing the transformation $\boldsymbol{\Psi}\mathbf{q} = \mathbf{y}$. The differential equations of motion are then multiplied by $\boldsymbol{\Psi}^T$ such that

$$\mathbf{M}\ddot{\mathbf{y}} + \mathbf{C}\dot{\mathbf{y}} + \mathbf{K}\mathbf{y} = \mathbf{P} \Rightarrow \boldsymbol{\Psi}^{\mathbf{T}}\mathbf{M}\boldsymbol{\Psi}\ddot{\mathbf{q}} + \boldsymbol{\Psi}^{\mathbf{T}}\mathbf{C}\boldsymbol{\Psi}\dot{\mathbf{q}} + \boldsymbol{\Psi}^{\mathbf{T}}\mathbf{K}\boldsymbol{\Psi}\boldsymbol{\Psi}$$
$$= \mathbf{M}_q\ddot{\mathbf{q}} + \mathbf{C}_q\dot{\mathbf{q}} + \mathbf{K}_q\mathbf{q} = \boldsymbol{\Psi}^{T}\mathbf{P}$$

and the transformed mass and stiffness matrices, $\mathbf{M}_q$ and $\mathbf{K}_q$, become diagonal due to the orthogonal properties of the eigenvectors. If proportional damping holds, then the damping matrix also becomes diagonal. Consequently, each mode becomes uncoupled and can be considered as a separate single-degree of freedom system.

If the damping is nonproportional, then the modal transformation does not diagonalize the damping matrix and the reduction to separate single degree of freedom systems is not possible. In this case, the equations must be considered as above in first order form $\mathbf{A}\dot{\mathbf{z}} + \mathbf{B}\mathbf{z} = \mathbf{0}$.

The modal matrix is then made up of 2n eigenvectors such that

$$\mathbf{R} = [\mathbf{z}_1, \mathbf{z}_1^*, \ldots, \mathbf{z}_n\mathbf{z}_n^*]$$

where * denotes the complex conjugate. The transformation of generalized coordinates is achieved in a similar manner by letting $\mathbf{z} = \mathbf{R}\mathbf{q}$ and premultiplying by $\mathbf{R}^{\mathrm{T}}$ thus $\mathbf{R}^{\mathbf{T}}\mathbf{A}\mathbf{R}\dot{\mathbf{q}} + \mathbf{R}^{\mathbf{T}}\mathbf{B}\mathbf{R}\mathbf{q} = \mathbf{0} \Rightarrow \mathbf{A}_q\dot{\mathbf{q}} + \mathbf{B}_q\mathbf{q} = \mathbf{0}$

The equations in first order form have now been uncoupled and there is no requirement for the damping to be proportional. However, note that the $\mathbf{A}_q$, $\mathbf{B}_q$, and $\mathbf{q}$ terms are now complex.

MDOF Forced Vibration and Response

The response of a MDOF system is often required to some general or harmonic excitation, for example response of an aircraft to a gust sequence or landing loads or a helicopter to harmonic rotor aerodynamic loading. Let us consider the response to steady state harmonic excitation of circular frequency ω.

The equations of motion for a MDOF forced system are

$$\mathbf{M}\ddot{\mathbf{y}} + \mathbf{C}\dot{\mathbf{y}} + \mathbf{K}\mathbf{y} = \mathbf{f}$$

where $\mathbf{y}$ is the response vector and $\mathbf{f}$ is the excitation force vector. Now, if a harmonic excitation of $\mathbf{f} = \mathbf{F}(\omega)e^{i\omega t}$ is applied, then it can be assumed that the response is $\mathbf{y} = \mathbf{Y}(\omega)e^{i\omega t}$ where $\mathbf{Y}(\omega)$ is some complex vector. Substituting and cancelling $e^{i\omega t}$ give

$$[\boldsymbol{K} - \omega^2\boldsymbol{M} + i\omega\boldsymbol{C}]\,\boldsymbol{Y}(\omega) = \boldsymbol{F}(\omega) \Rightarrow \boldsymbol{Y}(\omega) = [\boldsymbol{K} - \omega^2\boldsymbol{M} + i\omega\boldsymbol{C}]^{-1}\boldsymbol{F}(\omega) = \boldsymbol{H}(\omega)\boldsymbol{F}(\omega)$$

where $\mathbf{H}$ is the frequency response function, which relates the response at a measurement point to the input at some position at all frequencies ω. The FRF having been determined, the response of a linear system can be determined by

1. Fourier transforming (see Part 1) the time history of input $\mathbf{f}$ to give $\mathbf{F}(\omega)$
2. Determining the frequency content of the response $\mathbf{Y}(\omega)$ by multiplying $\mathbf{F}(\omega)$ by the FRF
3. Inverse Fourier transforming $\mathbf{Y}(\omega)$ to give the time response history $\mathbf{y}$.

This process can be expanded to consider the multiple input–multiple output (MIMO) case.

Assuming that proportional damping holds, the equations of motion can be written in an uncoupled single-DOF form

$$m_j\ddot{q}_j + c_j\dot{q}_j + k_jq_j = f_j = \Psi_j^T f$$

where m_j, c_j, and k_j are the respective jth elements of the modal mass, damping, and stiffness matrices. For a harmonic excitation as above, the response of each mode at a particular frequency is given by

$$(k_j - \omega^2 m_j + i\omega c_j)q_j = f_j \qquad j = 1, 2, \ldots, n$$

and thus

$$Y = \left[\sum_{j=1}^{n} \frac{Y_jY_j^T}{k_j - \omega^2 m_j + i\omega c_j}\right] \qquad F = \left[\sum_{j=1}^{n} \frac{Y_jY_j^T}{m_j(\omega_j^2 - \omega^2 + 2i\omega\zeta_j\omega_j)}\right] F$$

which shows other forms of FRF. Typical FRFs are shown in Figures 9.47 and 9.48 for the three stations on the previous uniform wing example related to an input at station 1.

It is also possible to determine the response of a system in the time domain using some form of numerical integration method, such as Runge–Kutta or Newmark–Beta. Such methods, although often slow, are essential for the prediction of the response of nonlinear systems.

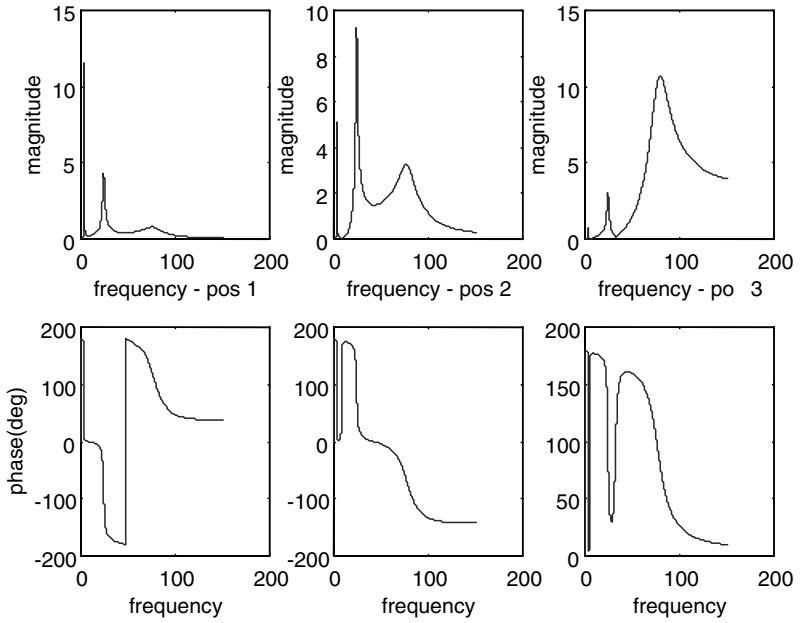


FIGURE 9.47 FRFs for the three beam stations relative to input from station 1—Bode plot.

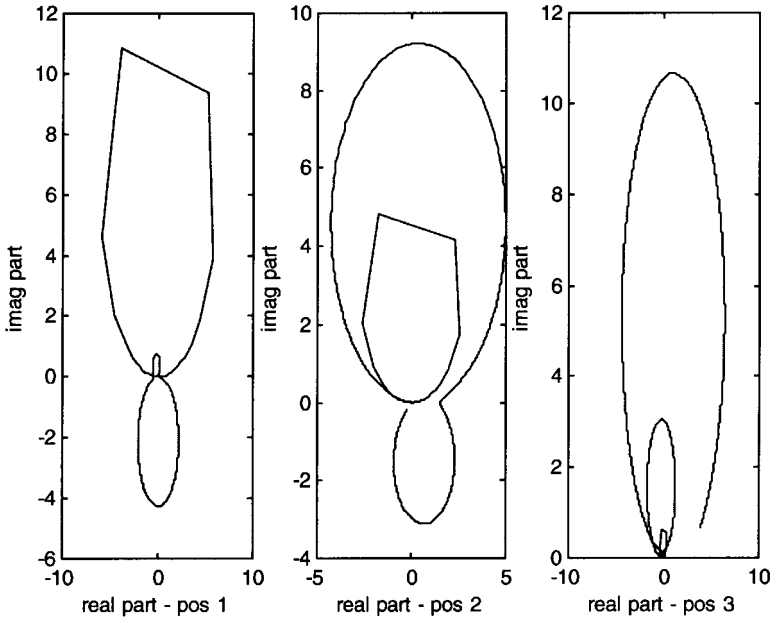


FIGURE 9.48 FRFs for the three beam stations relative to input from station 1—Nyquist plot.

Use of Frequency Response Functions for Modal Testing

A great deal of information about the modal characteristics of a system can be deduced from the FRFs. The science of modal testing has been developed to estimate vibration characteristics from measured test data, and the most widely used model is the FRF. FRFs are typically determined using measured force and acceleration data. Figure 9.47 shows a typical frequency response function in Bode (amplitude and phase versus frequency) form and Figure 9.48 shows the Nyquist (real versus imaginary parts for increasing frequencies) format. For modes that are not close (i.e., influenced by the presence of other modes), much information can be deduced without needing advanced curve-fitting techniques. Modes can be distinguished by a peak in the Bode amplitude plot (corresponding to a sudden change in phase passing through $\pm 90°$ for displacement or acceleration response) and a circular arc on the Nyquist plot. The peaks of the FRF become narrower the lighter the damping of a particular mode, as seen in the first mode. It is possible to determine the damping ratios of each mode on the Bode plot through the use of the half power points. The highest frequency mode does not appear totally circular in the Nyquist plots due to an inadequate frequency resolution.

The mode shapes can be determined through the use of FRFs measured at a number of different measurement stations. At each resonance frequency, the amplitude and phase of the mode are determined. The amplitudes for the various measurement points on the structure are plotted upwards if the phase is 90° and downwards if the phase is $-90°$.

9.19 AEROELASTICITY

Aeroelasticity is the study of the interaction of aerodynamic forces on elastic bodies. It is most famously characterized by Collar's Aeroelastic Triangle (Figure 9.49), which shows the interdependence of aerodynamic, elastic, and inertial forces. For example, if an aerodynamic load is applied to, say, a wing, this will cause the wing to deflect. However, this deflection will alter the manner in which the aerodynamic forces act on the wing, and so on. Most aeroelastic effects are not desirable and in some cases can lead to structural failure. Consequently, the study of aeroelasticity is very important for the design of aerospace structures as well as other structures such as racing cars, bridges, chimneys, power cables, etc.

The various aeroelastic phenomena can be categorized as to whether they are static or dynamic, involve an attached or separated air flow, or behave in a linear or nonlinear fashion. We are particularly interested in the nature of each phenomenon and the deflections and loads that occur. For aerospace structural design it is important to determine the critical air speed at which aeroelastic phenomena occur and the behavior just before any instability occurs.

The aeroelastic equations of motion for an n-DOF system, such as an aircraft, in terms of coordinate system $\mathbf{y}$ can be written as a second order differential equation such that

$$\mathbf{A}\ddot{\mathbf{y}} + (\rho V\mathbf{B} + \mathbf{D})\dot{\mathbf{y}} + (\rho V^2\,\mathbf{C} + \mathbf{E})\mathbf{y} = \mathbf{0}$$

where $n \times n$ matrices $\mathbf{A}$, $\mathbf{D}$, and $\mathbf{E}$ are, respectively, the inertial, structural damping, and stiffness matrices as before (note the change in notation). Compared to the equations described earlier in this section, extra terms are included to represent the aerodynamic damping ($\rho V\mathbf{B}$) and aerodynamic stiffness ($\rho V^2\mathbf{C}$), which depend upon the density ρ (and hence the altitude) and airspeed V. The damping term reflects the effective change in incidence due to a vertical motion of velocity y' in an airflow of velocity V. The stiffness term reflects the change in incidence due to structural rotations. Thus, the modal parameters of an aeroelastic system vary with the flight condition. Note that the aerodynamic terms also depend upon the frequency of vibration, due to the so-called unsteady aerodynamic behavior, and these effects must be accounted for in a complete dynamic aeroelastic analysis.

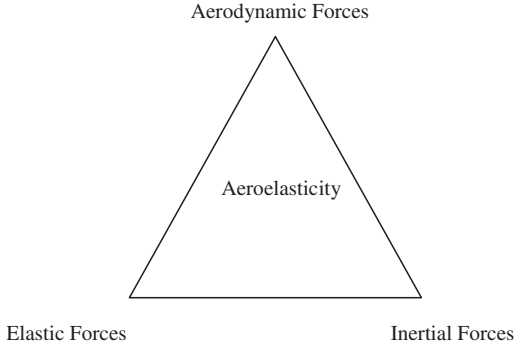


FIGURE 9.49 Collar's aeroelastic triangle.

Divergence

Divergence is a static aeroelastic phenomenon that results in structural failure. When an aerodynamic load or moment is applied, say to a wing, then there is a resultant deflection, and once equilibrium has been achieved the aerodynamic forces and moments are balanced by the structural restoring forces and moments. Classically, divergence occurs when the aerodynamic moment overcomes the structural restoring force and the structure fails.

For example, consider the simple rigid aerofoil in Figure 9.50 with initial angle of incidence α, eccentricity ec between the lift acting on the aerodynamic center (at the ¼ chord) and flexural axis, chord c, 2D lift curve slope a_1, unit span and torsional stiffness k_θ. The moment due to the aerodynamic lift (considered to be proportional to the dynamic pressure q) is balanced by the spring restoring moment, thus,

$$\text{Lift} \times \text{eccentricity} = qc^2a_1(\alpha + \theta)e = k_\theta\theta \Rightarrow \theta = \frac{qc^2a_1e\alpha}{k_\theta - qc^2a_1e}$$

As the airspeed (and hence dynamic pressure) increases, the angle of twist θ increases. At the critical divergence speed, $k_\theta = qc^2a_1e$ and structural failure occurs.

Control Effectiveness / Reversal

The flexibility of wings means that the application of the control surfaces can have a different effect, depending upon the flight speed. With zero control surface angle, the lift occurs at the aerodynamic centre at the ¼ chord, causing a nose-up pitching moment. However, if a control angle of β is applied, the resultant extra lift occurs somewhere around the ⅔ chord point, causing a nose-down moment to be applied. If the aerofoil were rigid, no torsional deflection would occur; however, in practice the aerofoil will rotate downwards. This rotation reduces the lift obtained through application of the control angle. Figure 9.51 shows how the effectiveness (here defined as the ratio between the lift for flexible and rigid wings) varies with airspeed. The effectiveness drops with increasing speed until it reaches the reversal speed, when it becomes zero. At this speed, application of the control surface has no effect. Beyond this speed, the effectiveness becomes negative and application of the control will have the opposite effect to that intended. Although not disastrous, this phenomenon is undesirable, as control of the aircraft is poor around the reversal speed.

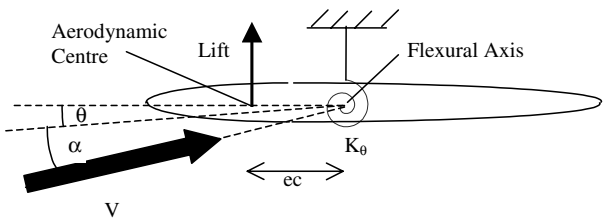


FIGURE 9.50 Divergence example.

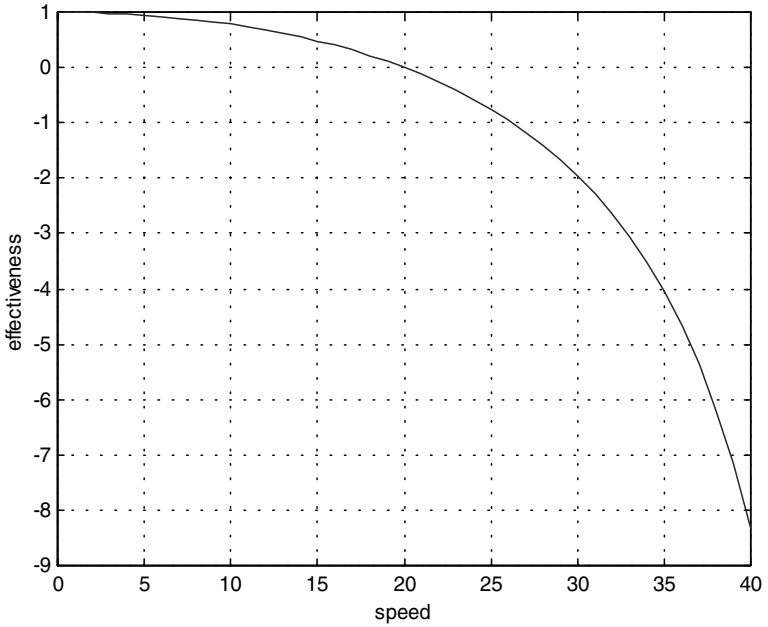


FIGURE 9.51 Typical control effectiveness versus speed.

Flutter

Flutter is a violent unstable oscillation that results in structural failure. It is the most important aeroelastic phenomenon, and considerable effort is spent in the design and prototype testing stages of aircraft to ensure that it cannot occur. Flutter classically occurs when two modes (wing bending and torsion) interact at a certain flight condition and effectively extract energy from the airflow. Figure 9.52 shows how the frequency and damping ratio of a binary aeroelastic system change with speed. The frequencies move closer together but do not necessarily join together. One of the damping ratios gets very large, whereas the other eventually becomes negative. The point where one of the damping ratios becomes zero is the critical flutter speed. Beyond this speed the damping ratio becomes negative and any small disturbance will cause an unstable vibration with disastrous consequences. Although the flutter analysis of an aircraft contains many more modes, the critical flutter mechanism is nearly always binary in nature.

Nonlinear Aeroelastic Effects

If the system contains either structural, aerodynamic, or control nonlinearities, it is possible that the flutter oscillations become limited to some constant amplitude. Such an effect is known as a limit cycle oscillation (LCO), see Figure 9.53. Traditional linear analysis techniques are unable to predict LCO, which, although undesirable, is not immediately catastrophic. Other nonlinear aeroelastic effects with

AEROSPACE STRUCTURES

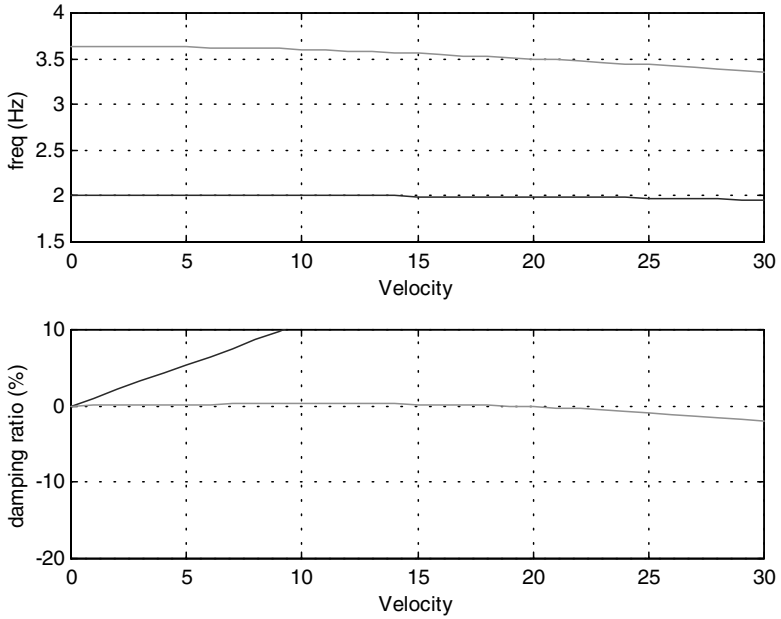


FIGURE 9.52 Frequency and damping trends for binary flutter system.

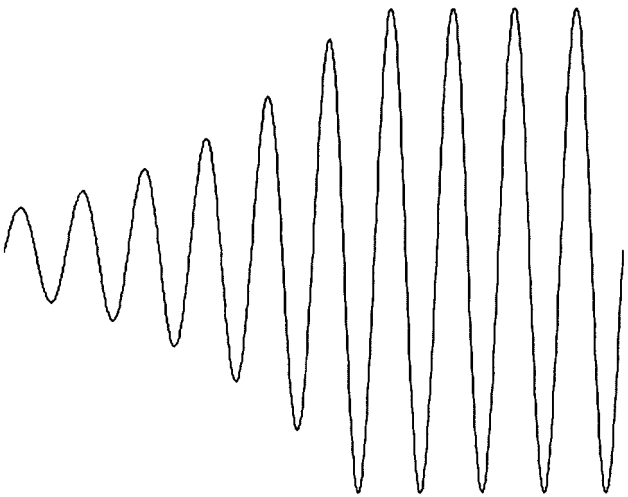


FIGURE 9.53 Typical limit cycle oscillation.

limited amplitude oscillations include the transonic phenomenon of control surface buzz, where a vibration is caused by the movement of a shock over the control surface. Stall flutter occurs when an aerofoil reaches an angle of attack such that the flow separates and lift is lost. This results in the angle of attack reducing and the flow reattaching and cause the lift to increase the angle of attack, and so on.

Buffet / Buffeting

A further important aeroelastic phenomenon is buffeting, where turbulent separated flows (buffet) from one part of a structure impinge on another part, causing a vibration known as buffeting. It rarely produces an instantaneous catastrophic failure, but the loads can be severe, resulting in reduced fatigue lives. This effect is currently a severe problem for twin-finned military aircraft.

Vortex Shedding

Vortex shedding is very important effect for the design of chimneys and buildings. Figure 9.54 shows how under certain Reynolds numbers the flow around a cylinder results in a Von Karman vortex street, whereby two streams of alternating vortices form downstream of the body. These vortices give rise to a sinusoidal force perpendicular to the flow. The frequency of the shedding ω is related to the air speed V by the Strouhal number

$$S = \frac{\omega D}{V}$$

where D is the diameter of the cylinder.

Should the frequency of the force correspond to one of the natural frequencies of the structure, then large deflections can result causing fatigue problems. Solutions to this include the use of helical shrouds on the upper one-third of a vertical chimney to break up the vortex formation.

Negative Damping—Galloping

Galloping is a phenomenon associated with power transmission cables where in strong winds vertical vibrations of 10 m in a span of 150 m have been observed. A purely circular cross-section cannot gallop, but certain cross-sections are prone

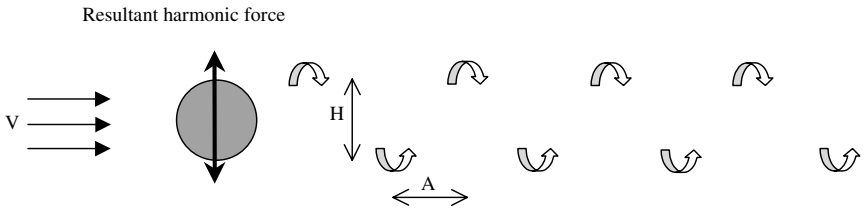


FIGURE 9.54 Vortex shedding.

to this phenomenon either through the formation of ice on the cable or certain configurations of cable winding (e.g., the Severn suspension bridge). These cross-sections lead to negative aerodynamic damping, resulting in increasing oscillations.

9.20 NOISE

Mention of aircraft noise usually brings to mind environmental problems and the effect that aircraft have on communities surrounding airports. However, aircraft designers also have to consider means of reducing the noise inside aircraft for passenger comfort and also the effects of high-level noise on the aircraft structure itself and structural fatigue problems that may result. Noise is produced when the air pressure varies rapidly, and consequently engines are a large source of noise due to extreme changes in pressure and temperature. Recent developments in high-bypass engines have lead to significant reductions in noise levels. The noise inside the aircraft can be reduced through the use of passive and active damping systems. In military aircraft acoustic fatigue can be a problem near engine exhausts.

The airframe is also a considerable source of noise and is particularly important in the approach phase to landing. Sources of the airframe that generate noise are those areas where the airflow direction is changed and where there are discontinuities in the surface profile. They include control surfaces, intakes, high-lift devices, landing gear, and airbrakes.

9.21 HELICOPTERS

Helicopters are particularly susceptible to vibration, primarily caused by the response of the helicopter airframe to the rotor hub forces and moments. Other less important sources are the engine and transmission, and also the effect of rotor downwash on the fuselage. The rotor-induced vibration is characterized by a harmonic excitation at multiples of N/rev, where N is the number of blades. The forces from all the blades cancel out at the hub except for those harmonics at multiples of N/rev which are transmitted to the fuselage. These N/rev harmonics dominate the vibration caused by real rotors.

When a helicopter is in hover, the aerodynamic forces acting on the rotor are exactly the same regardless of position, and any vibration is of a relatively low level. However, once the helicopter begins to fly forwards, the symmetry of the rotor system is broken. The advancing blades have the same flight and rotational velocity directions, whereas for a retreating blade the flight and rotational velocities are in opposite directions. The resulting vibration is partly due to:

- The pitch angle of the blades being varied to allow for the asymmetric lift distribution
- Blade vortex interactions as the rotor blades move through the vortex wake produced by the preceding blades, and complex transonic shock behavior acting at the outer part of the advancing blade

Vibration levels can be reduced through design of the fuselage to avoid harmonics of the rotor speed. Other approaches have included the use of passive and

active vibration suppression systems and of active trailing edge flaps on the rotor blades. Noise generated by helicopters is primarily through impulsive noise acting on the main rotor. There are two main types: blade vortex interaction in descent or level flight at low–medium speeds and high-speed impulsive noise that occurs due to transonic aerodynamic effects acting on the advancing blade. It is possible to reduce these effects somewhat through the advanced design of rotor blades through accurate prediction of the vortex/structural interactions. The other key source of helicopter noise is the turboshaft, which can be a significant problem on takeoff.

9.22 AIRCRAFT AIRWORTHINESS CERTIFICATION

Aircraft manufacturers must demonstrate to the airworthiness authorities that all new aircraft are safe to fly. The airworthiness regulations cover the full range of operational aspects and possible types of failure (stress, fatigue, etc.). Here we shall be concerned with items relating to structural dynamics.

Once the dynamic mathematical model of the aircraft has been determined, it is possible to predict the response to the many dynamic loads that may be encountered, e.g., maneuvers, takeoff/landing, store release, etc. The aeroelastic behavior, e.g., flutter boundaries and response to gusts, may also be predicted, following the addition of an aerodynamic model. It is mandatory for certification purposes that testing must be performed in order to validate the models, and there are two major vibration tests that must be undertaken: ground vibration testing and flight flutter testing.

Ground Vibration Testing

Ground vibration testing (GVT) is performed to measure the aircraft's modal characteristics—natural frequencies, damping ratios, and mode shapes. These results are then used to validate the dynamic model (usually determined using finite elements) and can be used as the basis for adjusting (updating) the model.

The aircraft must be freely supported so that the natural frequencies of the support do not overlap with those of the aircraft. Bungees, inflatable airbags, and semideflated tires are all methods that are commonly used as supports. For very large aircraft, such as the A380, the fundamental vibration frequencies are so low that it will not be possible to support the structure at a lower frequency. It is likely, in this case, that the mathematical model will have to include the support mechanism.

The aircraft is excited using electromechanical shakers that can be controlled to give the required input of prescribed frequency and amplitude. It is usual to use at least four shakers, although for large aircraft it is more likely that up to four may be used. The response to the excitation is measured using accelerometers, with typically upwards of 500–1000 being used for the test of a large civil aircraft. The whole data-acquisition process is computer controlled, the exact procedure used depending upon which type of analysis method is to be employed.

The traditional (phase separation) approach is to excite the structure with broadband random, or stepped sine, signals and to calculate the frequency response functions (FRFs) from the measured data. The FRFs are then curve-fitted using system

identification methods to estimate the natural frequencies, damping ratios, and mode shapes. This approach is likely to produce complex mode shapes which, while not being erroneous, can lead to problems when comparing with the finite-element model (invariably based upon proportional damping and given real modes). Consequently, the aerospace industry has traditionally employed the force appropriation (phase resonance) approach, whereby the structure is excited at each individual frequency and the amplitude and phase of each shaker are adjusted until only the normal mode at that frequency is excited. It is straightforward to compare the resultant normal modes with the finite-element model.

Flight Flutter Testing

Once an aerodynamic model is added to the validated structural model, it is possible to predict the frequency and damping behaviour against speed or Mach number (e.g., Figure 9.52). Flight flutter testing is performed to demonstrate that the aircraft is flutter free throughout the design flight envelope. Figure 9.55 illustrates the typical flight envelope clearance procedure that is used. There are three steps to the process:

1. The aircraft is flown at some constant flight condition and excited using one of a number of approaches: aerodynamic vanes, control surfaces, explosive devices, eccentric masses, or simply atmospheric turbulence. The response to this excitation is measured using accelerometers in the same way as the GVT, except that far fewer are used.
2. The measured excitation and response data are curve-fitted using system identification methods to determine frequencies and damping ratios.

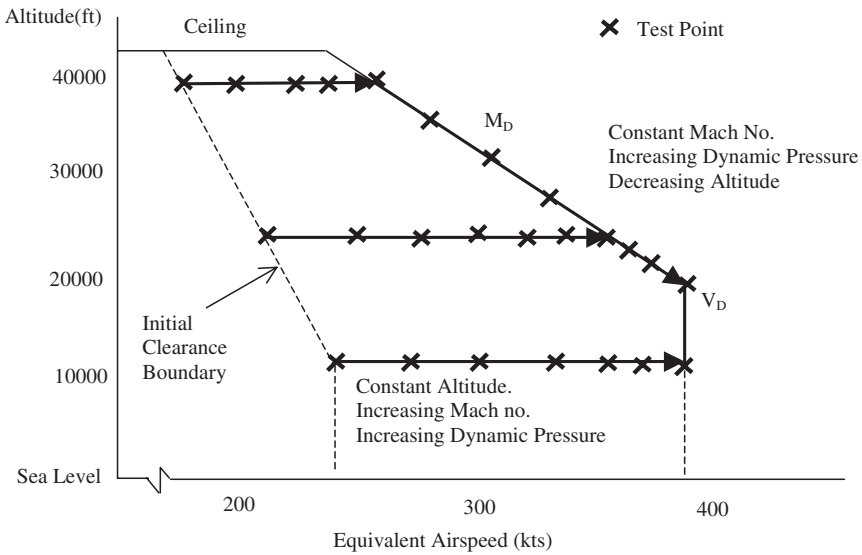


FIGURE 9.55 Typical flight envelope clearance.

3. The decision is made to move to the next flight test point, traditionally based upon damping ratio versus speed trends, although it is also possible to predict the flutter speed using the frequency and damping ratios obtained during the tests.

This procedure is repeated until the entire envelope is cleared. Extrapolation of results is used to establish a safety margin, typically of 20%.

9.23 AEROELASTIC DESIGN

Wing flutter usually occurs due to the interaction of two modes, one bending and the other torsion. However, this simplistic behavior becomes much more complicated with the addition of engines, stores, and control surfaces. Ideally, the designer should aim to place the inertial and flexural axes on the aerodynamic centre (¼ chord), in which case it is impossible for flutter to occur. However, whereas it is not too difficult to place the flexural axis close to the aerodynamic center, it is not as easy to do so with an inertial axis which lies aft of the aerodynamic center.

In order to keep the flutter speed as high as possible, the following general rules can be followed:

- The distance between the flexural axis and the aerodynamic center must be kept as small as possible.
- The ''wind-off'' natural frequencies should be kept well separated.
- An increase in the torsional stiffness will increase the flutter speed.
- Control surface flutter can be avoided by mass-balancing the control surfaces.

Should the aeroelastic prediction show that the flutter speed is less than the desired design speed, some alteration of the structure is required. Traditionally this has been achieved by adding mass to move the inertial axis forwards, though this process is by no means obvious due to the interaction between all of the modes. Care must be taken to ensure that while curing a problem between two modes, a different instability is not caused instead.

PART 5

SPACECRAFT STRUCTURES

Constantinos Stavrinidis

The objectives of this section are to provide a synthetic overview of processes and techniques involved in the design, manufacturing, and verification of spacecraft structures and to show the interaction of spacecraft structures with other spacecraft parts or subsystems and equipment.

9.24 ROLE OF SPAECRAFT STRUCTURES AND VARIOUS INTERFACES

Structures form the skeleton of all spacecraft and support key spacecraft components in desirable locations where various constraints need to be adequately considered, for example thermal control, fields of view for antenna and sensors, and lengths and weights of cables. The stowed configuration must fit within the launch vehicle's payload envelope, yet the design must provide access for installing and maintaining components.

Structures need adequately to protect spacecraft components from dynamic environments during ground operations, launch deployment, and mission operations. They need to cover successfully the deployment of antennas and sensors and provide adequate structural stability during their operation. Structures must of course be sufficiently light to allow useful payload mass.

The structural stiffness needs to cover adequately launcher frequency constraints in order not to interfere with the dynamics of the launch vehicle during flight. Similarly, the structural frequency characteristics of the spacecraft in its deployed configuration must not interfere with its own control system.

The materials used must survive ground, launch, and on-orbit environments, e.g. time-varying applied forces, pressure, humidity, radiation, contamination, thermal cycling, and atomic particles, without rupturing, collapsing, excessively distorting, or contaminating critical components. The structure materials need to support thermal control, and in some cases electrical conductivity.

The manufacturing, handling, and storage processes must be selected/conceived so as not to introduce life-limiting factors.

Structure categories (see Figure 9.56):

- The primary structure is the backbone, or major load path, between the components of the spacecraft and the launch vehicle.
- Secondary structures might include support beams, booms, trusses, and solar panels.
- Tertiary structures refer to the smallest structures such as boxes that house electronics and brackets that support electrical cables.

Typical characteristics of these categories are:

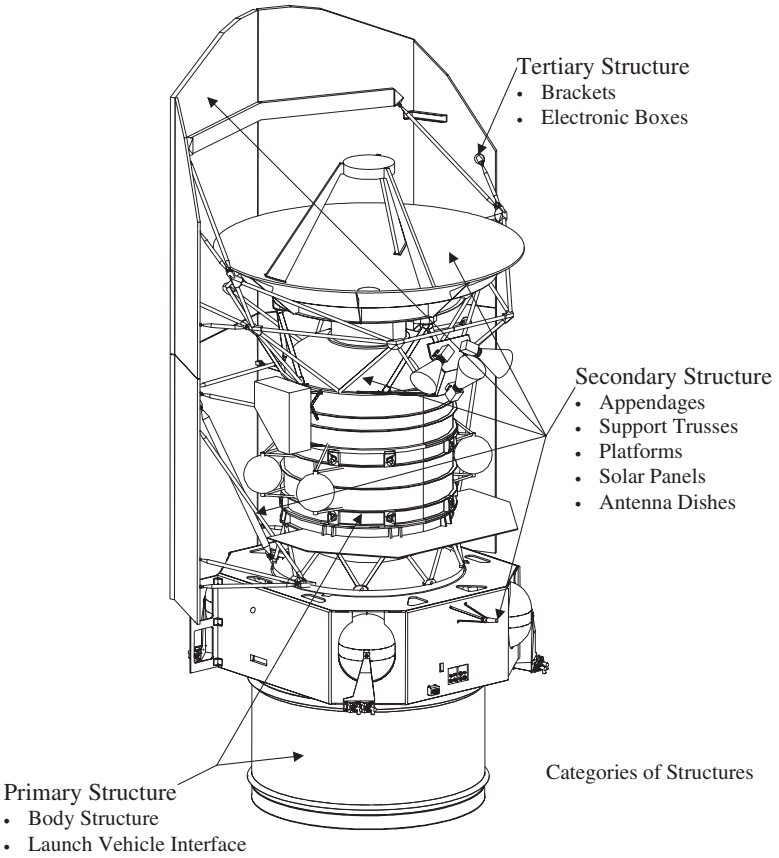


FIGURE 9.56 Categories of structures. HERSCHEL infra-red telescope. (Courtesy of the European Space Agency.)

- The primary structure is usually designed for stiffness, or natural frequency, and to survive steady-state accelerations and transient loading during launch. A major characteristic of the primary structure is to interface adequately with the launcher.
- The design of secondary structures is influenced significantly by steady-state accelerations and transient loading during launch. However, loads generated by mission operations, on-orbit thermal cycling, and acoustic pressure during launch is often a more severe environment for deployable appendages.
- Stiffness and structural stability over a lifetime including thermal cycling are other driving requirements for primary, secondary, and tertiary structures.

Interfaces

The structural subsystem interfaces on the one hand with the launcher (launcher/payload adapter) and on the other hand with the payload and the other spacecraft (S/C) subsystems.

Being the physical support medium for all hardware, the structural subsystem is usually designed before all other S/C elements, and therefore interdependence of the structure and other S/C elements necessitates close interaction between structures engineers and other engineers of the spacecraft team.

9.25 MECHANICAL REQUIREMENTS

Mechanical requirements are defined, and typical requirement sources are now identified, for spacecraft structures.

Strength is the amount of load, applied once to a structure that must not cause rupturing, collapsing, or deforming enough to jeopardize the mission. This basic structural requirement for all space programs applies to all life-cycle events. Strength inadequacy can lead to spectacular failures.

Structural life is the number of loading cycles a structure can withstand before its materials fatigue (rupture from cyclic loading), or the duration of sustained loads a structure can withstand before its materials creep (deform excessively from sustained loading) or crack from stress corrosion. Again, this applies to all life-cycle events.

Structural response attenuation is the magnitude of vibration in response to external loads which must be contained so as to avoid damaging critical components or the launch vehicle.

Natural frequency is the frequency at which a structure will vibrate, when excited by a transient load and then left undisturbed, depends on mass properties and stiffness, as outlined in Part 4. Each structure has an infinite number of natural frequencies, corresponding to different mode shapes of vibration. Natural frequencies must usually be above a given value or outside a certain range. The lowest fundamental frequency of a structure is of particular concern and corresponds to its first mode of operation. In the stowed configuration the natural frequencies of the combined spacecraft/launch vehicle system must not interfere with the control system of the launch vehicle or causing excessive loads. Likewise, for on-orbit configuration, vibrations must not affect the control system of the spacecraft.

Stiffness is often prescribed for substructures to achieve the required natural frequency for a larger assembly, or alternatively, to provide the necessary structural stability for a sensor or antenna.

Damping is the dissipation of energy during vibration; it is a structural characteristic that limits the magnitude and duration of response to input forces. It is designed to control loads and ensure that any vibration decays before influencing the control system of the spacecraft.

Mass properties include mass, center of mass, and mass moments and products of inertia. These are imposed by the launch vehicle and allocated to all substructures to achieve the required natural frequencies of a larger assembly.

Dynamic envelope is the physical space the spacecraft or substructure must stay within while deflecting under loads to avoid contact between the spacecraft and the payload fairing of the launch vehicle or between parts of the spacecraft.

Structural stability is the ability to maintain location or orientation within a certain range. Typical concerns are thermoelastic distortions, material yielding, and shifting of mechanical joints. The intention is to ensure that critical instruments, such as antennas and sensors, will find their targets, since inadequacy in this area results in performance degradation.

Mechanical interface includes features, such as flatness and locations of bolt holes, that define how structures and components attach and may also include stiffness. It is derived from designs of mating structures to ensure fit and avoid excessive deformations and loads.

9.26 SPACE MISSION ENVIRONMENT AND MECHANICAL LOADS

Environments on earth, during launch, and in space are the design drivers for spacecraft structures. When applicable, on-orbit performance of spacecraft structures such as pointing accuracy and structural stability are further important design drivers. Activities related to space mission environments and mechanical loads are identified:

- Structures must not only survive the environments to which the spacecraft is subjected but also protect the spacecraft nonstructural components and allow them to function.
- The selected materials, nonstructural as well as structural, must not degrade before and during the mission.
- Ground testing needs to envelop mission environments with margin. As a result, test environments need to be defined at an early stage and structures need to be designed adequately to cover them.
- Mechanical loads can be static or dynamic. Mechanical loads can be external (e.g., engine thrust, sound pressure, and gusts of wind during launch) or self-contained (e.g., mass loading of a vibrating satellite during environmental testing or in space after the force that caused the excitation is removed).

The design of a spacecraft and its subsystems must cater to all the loads which will be experienced. Typical loading events are:

- Testing (tailored to cover adequately flight, in-orbit, and ground transportation environment)
- Ground handling and transportation
- Liftoff
- Stage separations
- Stage ignition
- Stage or main engine cut-off
- Maximum aerodynamic pressure
- Spin-up and deployments
- Attitude control system (ACS) firings
- Reentry
- Emergency landing (for shuttle space transportation system [STS])

The above events induce accelerations, shock, and vibration to the structure. Careful attention is needed to cover these effects adequately. In general, the limiting factors in the structural design are set by the dynamic effects rather than by steady-

state accelerations. The primary source of loads occurs during the launch phase. All types of launchers apply different levels according to their design; for example:

- Solid rocket engines have combustion chambers which run the length of the fuel column, and they burn continuously and cannot be throttled.
- Liquid fuel rockets have a combustion chamber fed by separate fuel tanks; they can be (but rarely are) throttled to reduce thrust loads at critical aerodynamic phases of the flight and to give a more controlled trajectory. The tanks of liquid fuel act to some extent as damper and attenuate vibrations from the engines.

The launcher agency will issue standard guidelines for the design and qualification of spacecraft. These apply to the mounting interfaces at its base and are concerned with both steady and dynamic forces. They will be stated as flight limit loads, i.e., levels which one would not expect to be exceeded in 99% of launches. The launcher agency will, however, require the spacecraft designer to demonstrate (by a combination of test and analysis) that the design can withstand these levels with significant margins, differing from those of aircraft systems, and which typically are:

- Flight limit loads: 1.0
- Flight acceptance: 1.1
- Design qualification: 1.25–1.4

Therefore, test levels for flight acceptance are set for structures and equipment which have previously (in prototype form) passed tests at the design qualification level. The implication here is that if a one-model program is followed (protoflight approach, i.e., the prototype is actually flown), then the model must clear the higher qualification test. Separate factors for the materials may also be specified by the launch agency. Typically these are 1.1 at yield and 1.25 at ultimate stress for metals, and up to 2.0 for composites. Therefore, a structure clearing acceptance test levels plus material safety factors will have a healthy margin over actual flight loads.

- Margin of safety: Terms used in strength analysis are identified below.
- Load factor, n: This is the same as for aircraft (Part 2), i.e., a dimensionless multiple of g's (gravitational acceleration) that represents inertia force. Often used to define limit load.
- Limit load (or design limit load): The maximum acceleration, force, or moment expected during the mission or for a given event, at a statistical probability defined by the selected design criteria (usually between 99% and 99.87%).
- Limit stress: the predicted stress level corresponding to limit load.

The allowables and factors of safety are somewhat different from those discussed for aircraft.

- Yield failure: Permanent deformation
- Ultimate failure: Rupture or collapse
- Allowable load (or stress): The minimum strength (load or stress) of a material or a structure at a statistical probability defined by the selected design criteria (usually 99% probability at 95% confidence level)

- Allowable yield load (or stress): The highest load (or stress) that, based on statistical probability, will not cause yield failure.
- Allowable ultimate load (or stress): The highest load (or stress) that, based on statistical probability, will not cause ultimate failure.
- Yield factor of safety, FS_y: A factor applied to the limit load or stress for the purpose of decreasing the chance of detrimental deformation; usually between 1.0 and 2.0 for flight structures, depending on the option, whether personnel safety is at risk, and how sensitive the mission is to small deformations.
- Ultimate factor of safety, FS_u: A factor applied to the limit load or stress to decrease the chance of ultimate failure; usually between 1.25 and 3.0 for flight structures, depending on the test option and whether people are at risk.
- Design yield load: Limit load multiplied by the yield factor of safety; this value must be no greater than the allowable ultimate load.
- Design ultimate load: Limit load multiplied by the ultimate factor of safety; this value must be no greater than the allowable ultimate load.
- Design yield stress: Predicted stress caused by design yield load; this value must not exceed the allowable yield stress.
- Design ultimate stress: Predicted stress caused by the design ultimate load; this value must not exceed the allowable ultimate stress.
- Yield margin of safety, MS_y:

$$MS_y = \frac{\text{Allowable yield load (or stress)}}{\text{Design yield load (or stress)}} - 1$$

- Ultimate margin of safety, MS_u:

$$MS_u = \frac{\text{Allowable ultimate load (or stress)}}{\text{Design ultimate load (or stress)}} - 1$$

9.27 PROJECT OVERVIEW: SUCCESSIVE DESIGNS AND ITERATIVE VERIFICATION OF STRUCTURAL REQUIREMENTS

The assessment of load distribution within a spacecraft structure is largely an iterative process starting with generalized launcher predictions. These are used as a basis for the design of the initial spacecraft concept and therefore provide subsystem target specifications according to their location.

The first step in structural design is to convert the mission requirements into a spacecraft concept and specify the underlying parameters with as much detail as can be expected at the concept stage. This may be quite mission specific, contrasting the thermal dissipation of a communication satellite with the precision requirements of large antennas and telescopes.

In general, the following requirements need to be covered from early stage:

- Overall configuration meeting mission objectives
- Accommodation for the payload and spacecraft systems

- Ability to withstand launch loads
- Stiffness
- Provision of environmental protection
- Alignment
- Thermal and electrical paths
- Accessibility

At this stage the configuration effort will dominate as the distribution of masses becomes established. The concept of a load path from the interface with the launch vehicle (from which all accelerations are imparted) through the spacecraft structure to the mounting points of individual systems or units is followed. At each mounting point individual sets of interface requirements (alignment, thermal, field of view, screening, connections, and accessibility) will be generated which therefore set some of the constraints to be applied to the design of the structure. Subsequently, the problem of providing a structure to meet the specification might be well outlined. However, the need for high mass efficiency and reliability presents challenges. The selection of materials is often dominated by stiffness and structural dynamics rather than the stress levels. Testing as a means of validation is an inherent part of the design, by full test or by limited testing supported with analysis and modeling. The final design stage requires a coupled loads analysis which combines the characteristics of the full assembly of the spacecraft with those of the launch rocket.

In the case of large structures such as a space station, reconfigurations well beyond the initial concept will occur on orbit during its 25-year lifetime. Addition of modules and long beams will change the very low-frequency dynamic characteristics during attitude maneuvers and orbit boosting, presenting an additional design criterion. The design methodology is summarized in Figure 9.57.

9.28 ANALYTICAL EVALUATIONS

With the developement of modern fast-access large-memory computers it is now possible, with a readily available commercial finite-element code, to model in detail satellite mechanical systems and to examine their behavior under various static or time dependent load conditions. The finite-element method was outlined in Part 3. Commercial packages are available (NASTRAN, ASKA, SAMCEF, ANSYS, SYSTUS) which enable a spacecraft structure to be evaluated under static, dynamic, and thermal loads. The natural frequencies and modes of vibration are determined as a first step in dynamic evaluations to assess the characteristics of the structure. These are employed to evaluate the compatibility of the satellite with launcher requirements and for modal response analysis of the structure excited by external loads to determine the satellite response at different locations. The finite-element model of the satellite system is employed in coupled loads dynamic analysis (CLDA) with the launcher to examine the launch assembly configuration and determine representative test excitation (tailoring of test profile and level). Figure 9.58 shows a model of a highly three-dimensional satellite.

In such analytical evaluations it is particularly important to assess qualitatively and quantitatively the analytical predictions. A number of qualitative checks are presented here which assist practicing engineers to identify mathematical errors.

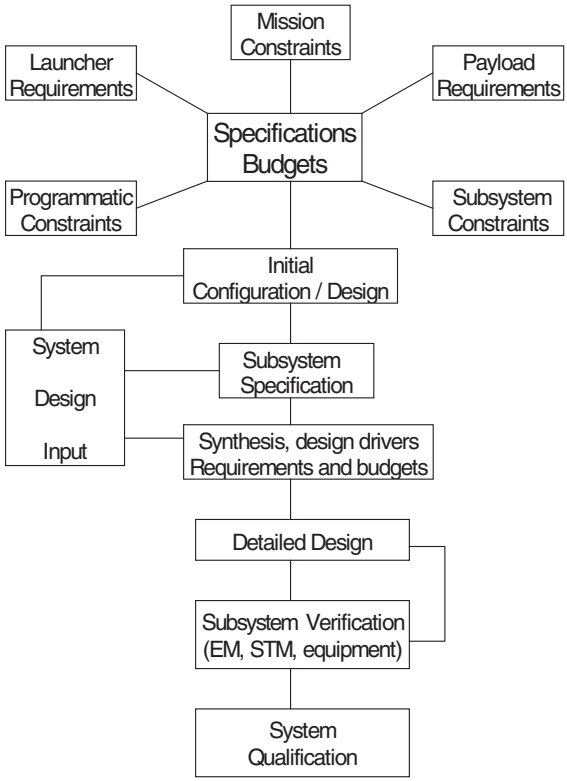


FIGURE 9.57 Design methodology.

(Quantitative checks are covered in the following subsection). The underlying physics of the problem should be respected in analytical modeling, and these assist in qualitative checks of the finite-element mathematical model. For example:

- No stress/loads should be generated due to rigid body translations and rotations of a structure.
- The mass matrix of the structure (often simplified for dynamic analysis) should represent the total mass and inertia properties of the structure.
- No stress/loads should be generated in a homogeneous structure free to expand when subject to a uniform temperature field.

9.29 TEST VERIFICATION, QUALIFICATION, AND FLIGHT ACCEPTANCE

The objective of a structural test is to engender confidence in the analytical predictions which support satellite development and ultimately to support the qualifi-

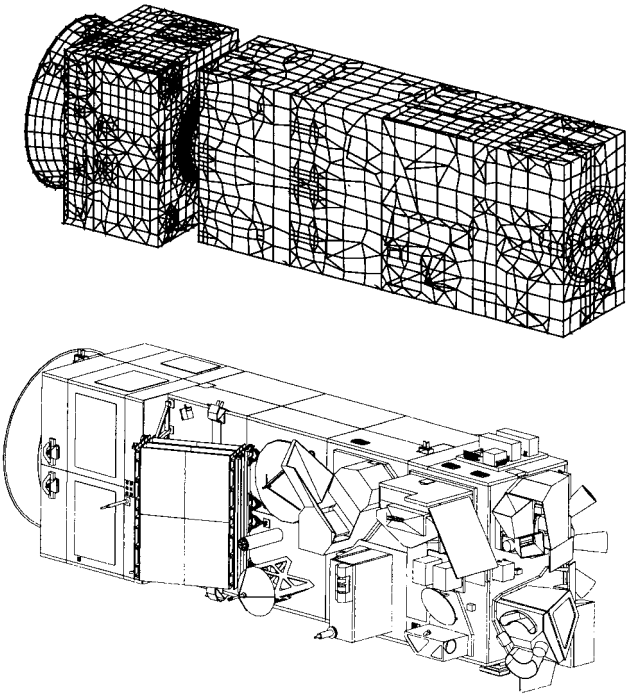


FIGURE 9.58 The ENVISAT satellite and FE model. (Courtesy of the European Space Agency.)

cation and flight acceptance of the satellite system. The types and purpose of the different tests are now presented.

The *static test* is achieved by subjecting the structure to discrete loads or by centrifuge test to simulate inertia loading. The static test provides insight on the validity of the stiffness matrix but is used mainly to qualify the strength adequacy of the primary structure and critical structural interfaces, e.g., satellite/launcher and satellite/payload interfaces.

The *modal survey test* is achieved by exciting the structure with small exciters to determine the natural frequencies and modes of vibration of the satellite. Since the applied excitation forces at resonance are compensated only by damping, high responses can be achieved with small excitation forces. Modal survey tests identify the satellite natural frequencies, modes of vibration, and damping to determine the dynamic compatibility of the satellite with the launcher and support the verification of the finite element model which is used in launch vehicle (LV)/spacecraft (S/C) coupled loads dynamic analysis (CLDA) in the loads cycle assessment and to tailor the vibration test level.

The *shaker vibration* (*sine*) *test* supports the verification of the mathematical model used in forced frequency response predictions and is particularly useful to determine the amplification of the excitation input from the L/V–S/C interface to various elements of the satellite (amplification factor Q = output/input). The main purpose of the shaker vibration sine test is to qualify the adequacy of the secondary

structures when subjected to dynamic environment, and also to verify the adequacy of the satellite system by performing functional tests after satellite system qualification and flight acceptance shaker tests.

The *shaker vibration random test* supports the verification of satellite units subjected to random dynamic environment which might be experienced during flight. The latter usually results from acoustic excitation of structural interfaces.

The *shock test* supports the verification and qualification of the satellite structure and instruments subjected to shock environment due to pyro and latching loads, e.g., release of the L/V–S/C interface clampband, and release of booms, solar panels, antennas, etc.

9.30 SATELLITE QUALIFICATION AND FLIGHT ACCEPTANCE

Classical Approach

In the classical approach, structural, in some cases engineering, and flight models are tested prior to satellite flight. The purpose of these models and the tests which are performed are summarized below.

Structural model (representative of satellite flight structure, units might be mechanically representative or with electronics represented as mass dummies)	• Static/centrifuge test at qualification load level to check and qualify the structural strength in particular primary structure and critical interfaces. • Modal survey test identify vibration natural frequencies, natural modes, and damping to support verification of mathematical model which is used in loads cycle and coupled loads dynamic analysis (CLDA) of the satellite with the launcher. • Shaker sine vibration test at qualification test input level to qualify satellite secondary amplification from S/C–L/V interface to various parts of the spacecraft.
Flight model	• Modal survey test is in general not performed with the flight model unless significant design changes have been introduced or there are buitl standard differences with the structural model • Shaker sine vibration test at flight level

	• Acoustic test at flight level • Shock test to determine compatibility of the satellite system with shock environment in particular payload units.
Engineering model	• The engineering model is usually not structurally representative and is employed for electrical purposes. In a number of cases the structural model may be employed as engineering model after the mechanical qualification tests.

Protoflight Approach

The major difference between the protoflight approach and the classical approach is that the flight hardware is subjected to qualification loads. The employment of the hardware model and the test which are performed are summarized below.

Structural model (this is refurbished after qualification test as the flight model—units might be mechanically representative or with electronics represented as mass dummies)	• Static/centrifuge test at qualification load level • Modal survey test • Shaker sine vibration test at qualification test input level • Acoustic test at qualification level • Shock test
Flight model (structural model refurbished to flight model standard after qualification test)	• Shaker sine vibration test at flight level • Acoustic test at flight level (qualification level acceptance duration) • Shock test

9.31 MATERIALS AND PROCESSES

The selection of appropriate materials for spacecraft structures applications requires a knowledge of the way each material property can best be used and where various limitations must be adequately recognized. Selection criteria encompass the specific strength, specific stiffness, stress corrosion resistance, thermal parameters, sublimation and erosion, and ease of manufacture and modification.

Ground and launch environments have to be considered when selecting materials for handling, moisture absorption/desorption, shocks and quasistatic and dynamic loads. The space environment in particular influences the selection of materials. Among the most important factors are:

- Atomic oxygen affects mostly polymeric materials (important exception: Ag).
- Charged particles require conductive materials to prevent local electrical charging (ITO conductive coating).
- Thermal radiation and uneven temperatures lead to thermal distortions (particularly multilayer composite materials).
- Vacuum facilitates outgassing (important selection criterion).
- Micrometeoroids and debris

Analysis of the possible impact of an environment aspect and verification of suitability/acceptability of selected materials is performed using a structured and documented approach. In addition, costs, availability of data, previous experience, and aptitude to be easily repaired can affect the choice of materials

Two classes of materials are suitable for spacecraft structures: metal alloys and composites. Metals are homogeneous and isotropic. Composites are inhomogeneous and orthotropic. The basic properties are summarized in Tables 9.1 and 9.2. The most expensive metallic alloy is that of beryllium, which is justified for satellite structures in view of its outstanding specific stiffness,

Typical applications of metals and composites:

- Aluminum: Skins, shells, truss members, face sheets, and core of sandwich structures
- Titanium: Attachments fittings, fasteners, pressure vessels
- Beryllium: Stiffness-critical appendages, support structures for optical equipment
- Steel: Fastening hardware, ball bearings

Composites are used in stiffness-driven applications, such as as for reflectors or in the case of Kevlar, where transparency to electromagnetic waves is required (see Figure 9.59)

Jointing is important and relies basically on three approaches: mechanically fastened joints, bonded joints, and welding. In general, mechanical joints lead to stress concentrations, which require special attention for load introduction, in particular

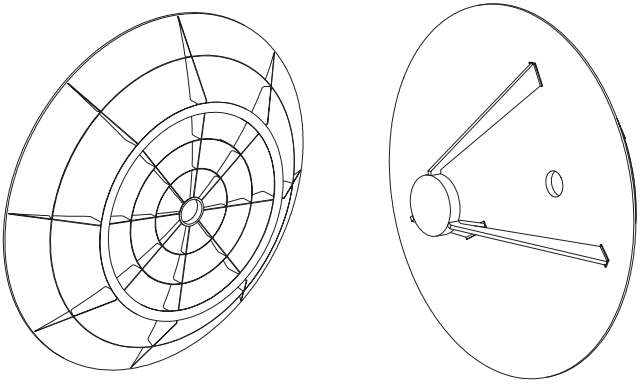


FIGURE 9.59 Composites reflector.

for composite materials which are susceptible to such concentrations. Bonding is the most elegant approach as it guarantees a better continuity of the load transfer. However, it requires proper shaping of the parts to be assembled to have a progressive load transfer so as to minimize through-thickness stresses. In composites attention shall be given to avoid or limit the peeling forces.

The tolerance between parts and cleanliness of the surfaces is also very important. Welding is reserved to metals and provides excellent tightness. Fusion associated with the welding process introduces weak zones in the joint, particularly when the heat-affected zone is not small, as in the tungsten inert gas process. High-performance welded joints are required for pressure vessels and can be obtained by electron beam welding or laser welding.

Two applications are leading the search for improved materials: high stability of scientific instruments of antennas and high-temperature materials for reentry vehicles. For high-stability, new materials like carbon fiber-reinforced cyanate ester resins are considered due to their low moisture absorption/desorption. For high temperatures, carbon-carbon and carbon silicon carbide offer good mechanical strength under high thermal fluxes leading to temperatures well above 1,000°C. However, oxidation of these materials under severe heat fluxes requires the development of adequate coatings.

9.32 MANUFACTURING OF SPACECRAFT STRUCTURES

Manufacturing of metallic or composites parts is fundamentally different. Because the processing of composites starts from raw materials and not semifinished products as for metals, it is possible to produce fairly complex parts as single piece (unitization), saving on manufacturing and assembly costs (see Figure 9.60).

Metals

Machining is the process of removing material with cutting or grinding tools. Many machining operations are automated to some extent, although for low-volume production as spacecraft manufacturing, manual processes are often cost effective.

Forming is one of the most economical methods of fabrication. The most limiting aspect of designing formed parts is the bend radius, which must be quite large to limit the amount of plastic strain in the material. Super-plastic forming and diffusion bonding, at temperatures up to 1,000°C, can produce complex components but only for titanium alloys since others are prone to surface oxidization, which inhibits the diffusion bonding. Spin forming has been successfully applied for manufacturing of aluminium pressure vessels.

Forging, in which structural shapes are produced by pressure, is well suited for massive parts such as load introduction elements. Where large sheets have to be manufactured (as for the skin of main thrust cylinders), chemical milling is frequently used to reduce the thickness of the sheet where possible. This process is more reliable than machining when processing very thin elements, but adequate tolerances have to account for various inaccuracies as thickness variations of the original piece of material are achieved by masking.

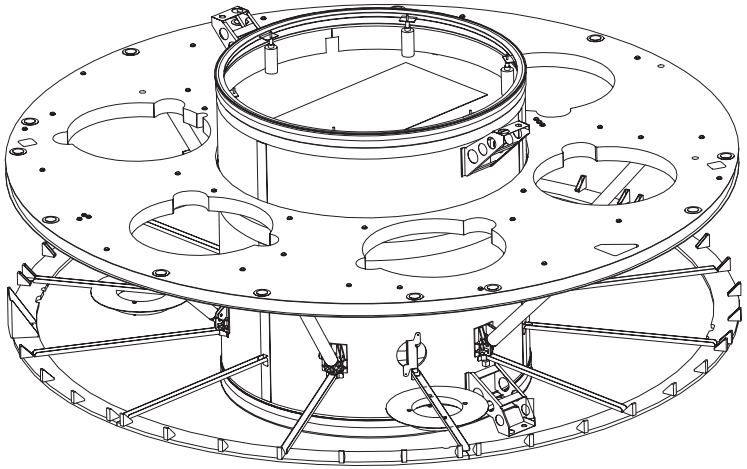


FIGURE 9.60 CLUSTER spacecraft structural layout. (Courtesy of the European Space Agency.)

Casting can be used to produce parts of complex shapes. However, the quality of a casting is difficult to control because as the material solidifies, gas bubbles can form, resulting in porosity. Material strength and ductility are not as high as with most other processes.

Composites

Impregnated tapes are the most widely used precursor for composite manufacturing. Fibers can be woven to form fabrics. The key parameters when specifying the material are the types of fibers and resin, resin content, tack, and drape.

The fiber will control the major mechanical properties of the part, such as strength and stiffness. The resin which binds the fibers together will determine the maximum temperature under which the part can be safely used. It will also influence moisture absorption and desorption, with possible effects on geometrical distortions under thermal and moisture cycling.

Tack is a measure of how much prepreg sticks to itself and to other layers. Prepregs that have too much tack can be difficult to handle because a misplaced layer is difficult to reposition without disrupting the resin or fiber direction. Pregregs with little tack are difficult to keep in place as more plies are applied. Lack of tack often indicates the prepreg resin has cured beyond an acceptable limit, which means the composite part will not cure properly.

Drape is the ability of the pregreg to form around contours and complex shapes. Drape is influenced by the fiber material and the diameter of the filaments as well as the cross-section of the tow. It is also conditioned by the weaving pattern. For flat panels, a low tack and drape are acceptable, but for complex shapes and cavities, tack and drape are important to keep the composite in place during laying-up and preparation for curing.

Manual lay-up is a costly technique which is well adapted to production of single parts or very small series. When the number of identical parts to be produced grows, filament winding, resin transfer molding, and braiding are cost-effective alternatives.

Filament winding consists of wrapping bands of continuous fiber or strands or rovings over a mandrel in a single machine-controlled operation. A number of layers of the same or different patterns are placed on the mandrel. The fibers may be impregnated with the resin before winding (wet winding), preimpregnated (dry winding), or post-impregnated. The first two winding sequences are analogous to wet or dry lay-up in the reinforced plastic fabrication methods. The process is completed by curing the resin binder and removing the mandrel. Curing is normally conducted at elevated temperature without pressure. Finishing operations such as machining or grinding are usually not necessary.

Resin transfer molding is closed-mold low pressure process. The fiber reinforcement is placed into a tool cavity, which is then closed. The dry reinforcement and the resin are combined within the mold to form the composite part. This process allows the fabrication of composites ranging in complexity from simple, low-performance small parts to complex elements of large size.

In braiding operation, a mandrel is fed through the center of a braiding machine at a uniform rate and the fibers or yarns from the carriers are braided around the mandrel at a controlled angle. The machine operates like a maypole, the carriers working in pairs to accomplish the over-and-under braiding sequence. Parameters in the braiding operation include strand tension, mandrel feed rate, braider rotational speed, number of strands, width, and the perimeter being braided. Interlaced fibers result in stronger joints. Applications include lightweight ducts for aerospace applications.

REFERENCES

Agarwal, B. D. and Broutman, L. J. 1990. *Analysis and Performance of Fiber Composites,* 2d ed., John Wiley, New York.

Bruhn, E. F. 1965. *Analysis and Design of Flight Vehicle Structures,* Tri-State Offset, Cincinnati.

Curtis, H. D. 1997. *Fundamentals of Aircraft Structural Analysis,* Times-Mirror HE Group, Los Angeles.

Dowell, E. H. 1995. *A Modern Course in Aeroelasticity,* Kluwer Academic, Dordrecht.

Foreman, R. G., Kearney, V. E., and Engle, R. M. 1967. Numerical Analysis of Crack Propagation in Cyclically Loaded Structures,'' *ASME Transactions, Journal of Basic Engineering* vol. 89(D), pp. 459–464.

Hull, D. 1981. *An Introduction to Composite Materials,* Cambridge University Press, Cambridge.

Kuhn, P. 1956. *Stresses in Aircraft Shell Structures,* McGraw-Hill, New York.

Megson, T. H. 1997. *Aircraft Structures for Engineering Students,* 3d ed. Edward Arnold, London.

Miner, M. A. 1954. ''Cumulative Damage in Fatigue,'' *ASE Transactions, Journal of Applied Mechanics,* vol. 67, p. A159.

Niu, M. C. Y. 1988. *Airframe Structural Design,* CONMILIT, Hong Kong.

Niu, M. C. Y. 1992. *Composite Airframe Structures,* CONMILIT, Hong Kong.

Paris, P. C., Bucci, R. J., Wessel, E. T., Clark, W. G., and Mager, T. R. 1972. "Extensive Study of Low Fatigue Crack Growth Rates in A533 and A508 Steels," *ASTM STP* 513, pp. 141–176.

Rivello, R. M. 1969. *Theory and Analysis of Aircraft Structures,* McGraw-Hill, New York.

Rooke, D. P. and Cartwright, D. J. 1976. *Compendium of Stress Intensity Factors,* HMSO, London.

Timoshenko, S. and Gere, J. M. 1961. *Theory of Elastic Stability,* McGraw-Hill, New York.

SECTION 10
AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

Section Editors: Andrew J. Niven and Trevor M. Young

10.1 INTRODUCTION

This section covers the three areas of aerodynamics, airplane performance, and stability and control. Each area is accompanied by a dedicated list of references and nomenclature.

The development of aerodynamic theory started in the early 19th century. There is thus a vast amount of literature available; many of the key publications are included in the aerodynamics reference list. Although early aerodynamic theory often requires certain restrictions to be imposed on the flow field before an analytical solution can be obtained, the end result provides extremely useful information about how the wing geometry governs the airloads. Since the arrival of the computer, many of these analytical methods have been replaced by numerical models, which, although they have a wider applicability, require specialized computer programs to obtain the solution. It is probably fair to say that, at present, these numerical models have now been surpassed by computational fluid dynamics software. The general aim of the aerodynamics sections is to provide various equations which can be used to estimate the airloads produced by an arbitrary wing geometry in either subsonic or supersonic flow. Thus, with only a few exceptions, only analytical aerodynamic models have been described. An attempt has been made to state the applications of each aerodynamic model if appropriate. One big advantage of many analytical models is that solutions can be obtained, for a wide range of airfoils and wings, using spreadsheet software.

Subsections 10.12 to 10.22 give a detailed description of airplane performance characteristics, such as its speed, rate of climb, range, endurance, maneuverability, and takeoff and landing distances, and the aerodynamic parameters that influence these measures. There are essentially three types of airplane performance analysis routinely conducted in the industry. The first is the prediction of the performance of a new airplane at the design stage; the second is the reduction of flight-test data from a prototype test airplane; and the third is the determination of the expected

performance of an aircraft for planning operations, taking into account actual weather conditions. Although different calculations have to be carried out in these cases, the basis of analysis is the same. For performance prediction at the conceptual and preliminary design stages of a new airplane it is usual to make a number of assumptions as to the behavior of the airplane and its environment, enabling problems to be solved analytically. In later design stages where more accurate performance predictions are required, and in the reduction of flight test data, it is likely that numerical methods will be used. Mathematical models are used as a framework for this analysis. These models are not absolute laws but approximations to physical measured quantities, and are only valid within specified limits.

Subsections 10.23 to 10.27 cover the subject of aircraft stability and control. The aim is to give the reader sufficient information to construct simple aircraft simulations (linear and nonlinear) and the ability to analyze the stability characteristics of fixed-wing aircraft. In Section 10.23 the nonlinear equations of motion are introduced and explained, and a basic longitudinal simulation is developed. In Subsection 10.24 the nonlinear equations are linearized and expressed in state space form both in their compact representation and in their nondimensional form. Expressions for the stability derivatives are presented in Subsection 10.25 which allow the derivatives to be estimated from simple aircraft parameters. The methods used to analyze aircraft stability are presented in Subsection 10.26 along with a series of approximations allowing the aircraft stability modes to be estimated from stability derivatives. Finally, analytical methods for calculating aircraft response to control inputs and atmospheric disturbances are developed in Subsection 10.27.

PART 1

AERODYNAMICS

Andrew J. Niven

NOTATION

a	local speed of sound, m/s
A	area, m^2
A'	axial force per unit span, N/m
A_R	aspect ratio of wing
b	wing span, m
c	airfoil chord, m
$c_{l,i}$	ideal lift coefficient
c_g	mean geometric chord, m
c_p	specific heat at constant pressure, J/K/kg
$c_{l\alpha}$	airfoil lift curve slope, per radian or per degree
$\overline{c}$	mean aerodynamic chord, m
C_L	wing lift coefficient
D'	drag per unit span, N/m
f_x, f_y, f_z	components of body force, N/kg
$\dot{G}$	heat generated per unit volume, J/m^3
h	maximum z-distance between camber line and chord line, m
$\hat{h}$	nondimensional maximum camber, h/c
i	specific internal energy, J/kg
k	fluid thermal conductivity, W/m/K
l	distance along specified path, m
L'	lift per unit span, N/m
m	mass, kg
M'	aerodynamic moment per unit span, N
M_∞	free-stream Mach number
N'	normal force per unit span, N/m
p	pressure, N/m^2
$\vec{q}$	local or instantaneous velocity vector, m/s
$\vec{q}$	turbulence velocity vector, m/s
$\vec{Q}$	free-stream or mean velocity, m/s
R	specific gas constant, J/K/kg
R'	resultant aerodynamic force per unit span, N/m
s	specific entropy, J/K/kg
S	wing area, m^2
t	maximum airfoil thickness, m

$\hat{t}$	nondimensional maximum thickness, t/c
T	air temperature, K
u, v, w	components of local or instantaneous velocity vector, m/s
$\tilde{u}$, $\tilde{v}$, $\tilde{w}$	local disturbance velocities, m/s
u', v', we'	components of turbulence velocity vector, m/s
U_∞, V_∞, W_∞	components of free stream or mean velocity, m/s
x, y, z	Cartesian coordinate system
z	a complex number, x + iy
$\hat{x}$	nondimensional x-coordinate, x/c
$\hat{y}$	nondimensional y-coordinate, 2y/b
$\hat{z}$	nondimensional z-coordinate, z/c
$\hat{z}_t$	nondimensional thickness, z_t/c

Greeks

α	angle of attack, degrees or radians
α_i	ideal angle of attack, degrees or radians
α_{l0}	airfoil zero lift angle of attack, degrees or radians
α_{L0}	wing zero lift angle of attack, degrees or radians
β_0	defined as $\sqrt{1 - M_\infty^2}$
β_1	defined as $\sqrt{M_\infty^2 - 1}$
γ	vortex sheet strength, m/s
γ	isentropic index
Γ	circulation, m^2/s
δ	flow turning angle, degrees or radians
ε	apex half angle of delta wing, degrees or radians
ε_t	wing tip rake angle, degrees or radians
ε_a	aerodynamic twist, degrees or radians
ε_g	geometric twist, degrees or radians
$\dot{\varepsilon}$	fluid strain rate, m/s
H	doublet strength, m^3/s
θ	an arbitrary angle, degrees or radians
λ	wing taper ratio
Λ	sweep angle, degrees
μ	viscosity, N/s
ξ	source sheet strength, m^2/s
ρ	density, kg/m^3
σ	oblique shock wave angle, degrees or radians
σ_M	Mach wave angle, degrees or radians

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

τ viscous shear stress, N/m^2
τ constant used in wing lift curve slope calculation, radians
ψ stream function, m^2/s
ϕ velocity potential function, m^2/s
$\tilde{\phi}$ disturbance velocity potential function, m^2/s
$\vec{\omega}$ vorticity vector, $(\omega_x, \omega_y, \omega_z)$

Subscripts

f skin friction value
inc incompressible flow value
l lower surface
le value at leading edge
n value normal to leading edge of swept wing
pl value appropriate to a flat plate supersonic airfoil
r wing root
t wing tip
u upper surface
w value at wall
$c/4$ value at the quarter chord location
$3c/4$ value at the three-quarter chord location
∞ free-stream value

10.2 AIRFOIL GEOMETRIC AND AERODYNAMIC DEFINITIONS

Airfoil Geometry

Figure 10.1 illustrates the terminology and the geometric parameters used to systematically define an airfoil (sometimes referred to as a wing section). In general, most airfoil profiles are generated by combining a mean line (or camber line) and a thickness distribution. The upper and lower surface coordinates are related to the camber line and thickness in the following manner:

$$\hat{x}_{u,l} = \hat{x}_c \mp \hat{z}_t \sin\theta \text{ and } \hat{z}_{u,l} = \hat{z}_c \pm \hat{z}_t \cos\theta.$$

where the upper sign indicates the upper surface. To increase the accuracy of the leading edge geometry, a circle is centered on a line defined by the tangent of the camber line at $\hat{x}_c = 0.005$ and the radius specified such that the circumference passes through the leading edge.

The camber line is defined by a function $\hat{z}_c = f_c(\hat{x})$ which has a maximum value, $\hat{h}$, at a particular chord position, $\hat{x}_{c,m}$. This value is normally referred to as the camber of the airfoil rather than the maximum camber.

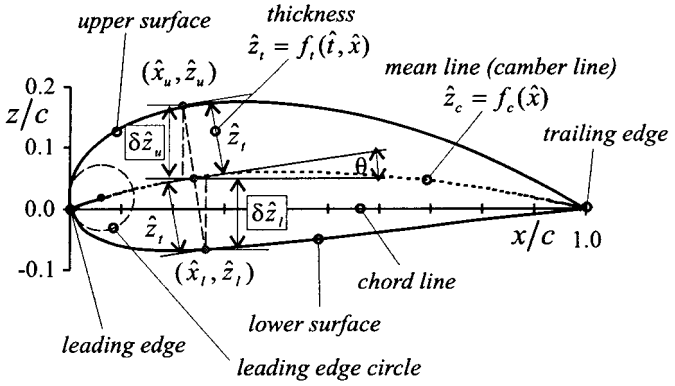


FIGURE 10.1 Airfoil geometrical definitions.

The thickness distribution is function of chord and a specified maximum thickness, i.e., $\hat{z}_t = f_t(\hat{t},\hat{x})$. The maximum thickness, $\hat{t}$ (simply known as the thickness), occurs at a particular chord position, $\hat{x}_{t,m}$. The thickness function actually defines a symmetrical airfoil in its own right which is often referred to as the basic thickness form.

The NACA Series of Airfoils

In 1929 the National Advisory Committee for Aeronautics (NACA) embarked upon the systematic development of various families of airfoils using a combination of theoretical methods and wind tunnel testing. Each airfoil was assigned a number which represented specific geometrical and aerodynamic properties of the airfoil. The equations for the camber line, $\hat{z}_c = f_c(\hat{x})$, and the basic thickness form, $\hat{z}_t = f_t(\hat{t},\hat{x})$, for most of the NACA airfoils are given in Abbott and von Doenhoff (1959). Once these equations are known, the airfoils may be easily plotted via spreadsheet software.

The first family of airfoils to be developed was designated a four-digit number, e.g., NACA 2412. The meaning of this number is described in Table 10.1 and the airfoil profile is shown in Figure 10.2(b). The basic thickness form is shown in Figure 10.2(a). The effect of varying the camber and thickness is illustrated in Figure 10.2(c), which displays the NACA 6408 airfoil.

TABLE 10.1 The NACA Four-Digit Numbering System

NACA Four-digit	Meaning	Equation
2	Maximum camber	$100(h/c)$
4	Distance along chord from leading edge to point of maximum camber	$10(x_{c,m}/c)$
12	Maximum thickness at $\hat{x}_{t,m} = 0.3$	$100(t/c)$

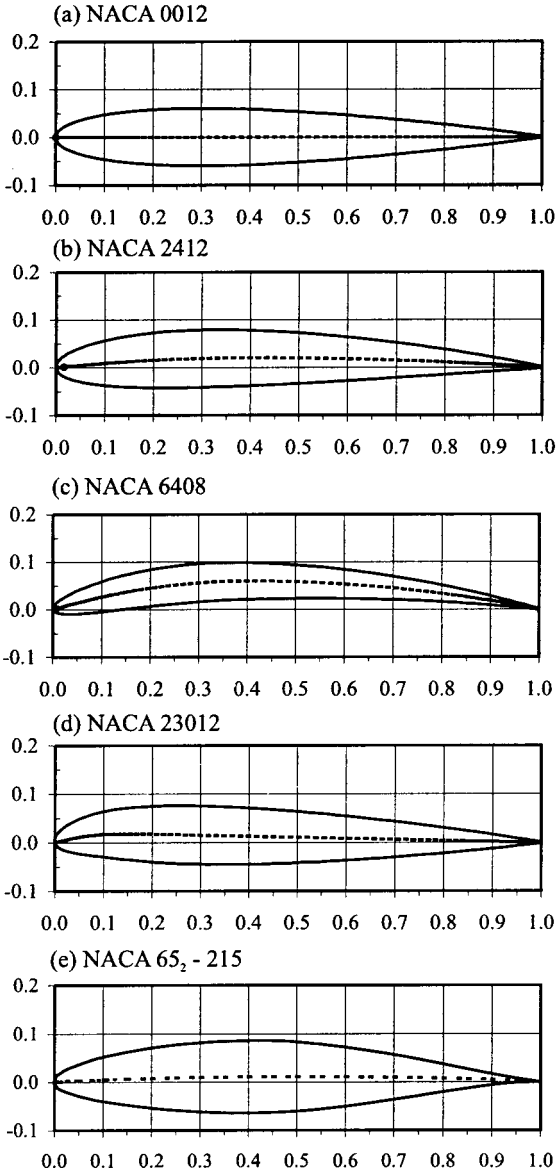


FIGURE 10.2 Examples of NACA airfoil profiles.

The NACA five-digit family of airfoils was designed to have point of maximum camber within the first quarter chord. The example profile shown in Figure 10.2(d) is the NACA 23012, and the meaning of the number is given in Table 10.2.

The NACA 6-series was designed, using the methods described in Section 10.7, to possess low drag at the design lift coefficient. Figure 10.2(e) displays the NACA 65_2-215 profile, and Table 10.3 explains the numbering scheme. The mean lines and basic thickness forms for the NACA six series are given in Abbott and von Doenhoff (1959).

There are many more types of NACA airfoils: modified four- and five-digit families, the 1, 2, 3, 4, 5, and 7 series airfoils and modified 6 series airfoils. Further details on the theoretical development of the NACA airfoils, along with extensive wind tunnel data, are given by Abbott and von Doenhoff (1959); this reference is the definitive work on airfoil design and development and should always be consulted when choosing an airfoil for a particular application.

Airfoil Aerodynamic Forces and Moments

When an airfoil moves through the air, each surface element is subjected to a normal pressure stress, p, and a tangential viscous shear stress, τ. Although the magnitude of these stresses will vary greatly around the airfoil contour, Figure 10.3 shows they may be integrated around the entire surface such that a resultant aerodynamic force and moment is obtained. Since an airfoil is a two-dimensional shape, surface areas are based on unit wing span which gives rise to forces and moments per unit length (denoted by a prime). The magnitude of the moment will depend on the

TABLE 10.2 The NACA Five-Digit Series Numbering System

NACA five-digit	Meaning	Equation
2	Strictly related to the ideal lift coefficient Also approximately equal to the maximum camber.	$20(c_{l,i}/3)$ $\approx 100(h/c)$
30	Distance along chord from leading edge to point of maximum camber	$200(x_{c,m}/c)$
12	Maximum thickness at $\hat{x}_{t,m} = 0.3$	$100(t/c)$

TABLE 10.3 The NACA Six-Series Numbering System

NACA six-series	Meaning	Equation
6	Series designation	
5	Nondimensional distance along chord line from leading edge to minimum pressure for the basic symmetrical section at zero lift	$10(\hat{x}_{mp})$
2	Defines the width of the low drag region whose limits are given by $(c_{l,i} - \Delta c_l)$ and $(c_{l,i} + \Delta c_l)$ on a drag polar	$10(\Delta c_l)$
2	Defines the ideal lift coefficient	$10(c_{l,i})$
15	Maximum thickness at $\hat{x}_{t,m} = 0.4$	$100(t/c)$

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

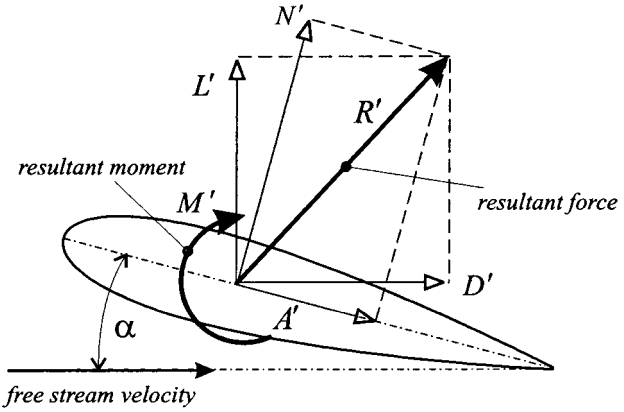


FIGURE 10.3 Forces and moments on an airfoil.

chosen reference point about which the elemental surface moments are calculated. The angle α is known as the angle of attack and is measured between the chord line and free-stream velocity; it is defined as positive when the chord line is rotated anticlockwise onto the free-stream velocity vector (often referred to as *nose up*). The moment is taken as positive in the direction which increases the angle of attack As indicated in Figure 10.3 the resultant aerodynamic force can be split into various components: normal force, axial (or chord) force, lift, and drag.

When discussing the aerodynamic characteristic of an airfoil it is common practice to deal with the following nondimensional groups.

Pressure coefficient: $$C_p = \frac{p - p_\infty}{\frac{1}{2}\rho_\infty Q_\infty^2}$$

Section lift coefficient: $$c_f = \frac{\tau}{\frac{1}{2}\rho_\infty Q_\infty^2}$$

Skin friction coefficient: $$c_l = \frac{L'}{\frac{1}{2}\rho_\infty Q_\infty^2 c}$$

Section drag coefficient: $$c_d = \frac{D'}{\frac{1}{2}\rho_\infty Q_\infty^2 c}$$

Section pitching moment coefficient: $$c_m = \frac{M'}{\frac{1}{2}\rho_\infty Q_\infty^2 c^2}$$

Section normal force coefficient: $$c_n = \frac{N'}{\frac{1}{2}\rho_\infty Q_\infty^2 c}$$

Section axial force coefficient: $$c_a = \frac{A'}{\frac{1}{2}\rho_\infty Q_\infty^2 c}$$

The above integrated force coefficients are related to the local surface pressure and skin friction coefficients in the following manner:

$$c_n = \int_0^1 (C_{p,l} - C_{p,u})d\hat{x} + \int_0^1 \left(c_{f,u} \frac{d\hat{z}_u}{d\hat{x}} + c_{f,l} \frac{d\hat{z}_l}{d\hat{x}} \right) d\hat{x}$$

$$c_n = \int_0^1 \left(C_{p,u} \frac{d\hat{z}_u}{d\hat{x}} - C_{p,l} \frac{d\hat{z}_l}{d\hat{x}} \right) d\hat{x} + \int_0^1 (c_{f,u} - d_{f,t})d\hat{x}$$

$$c_{m,le} = \int_0^1 (C_{p,u} - C_{p,l})\hat{x}d\hat{x} - \int_0^1 \left(c_{f,u} \frac{d\hat{z}_u}{d\hat{x}} + c_{f,l} \frac{d\hat{z}_l}{d\hat{x}} \right) \hat{x}d\hat{x}$$
$$+ \int_0^1 \left(C_{p,u} \frac{d\hat{z}_u}{d\hat{x}} + c_{f,u} \right) \hat{z}_u d\hat{x} + \int_0^1 \left(-C_{p,l} \frac{d\hat{z}_l}{d\hat{x}} + c_{f,l} \right) \hat{z}_l d\hat{x}$$

The lift and drag coefficients can be obtained from

$$c_l = c_n \cos \alpha - c_a \sin \alpha$$

$$c_d = c_n \sin \alpha + c_a \cos \alpha$$

The Center of Pressure

When integrating the surface pressures and shear stresses, a particular moment reference point can be chosen about which there will be no net moment. This point is known as the center of pressure. If the resultant force is split into the normal and axial components, the distance from the leading edge to the center of pressure is given by

$$\hat{x}_{cp} = \frac{x_{cp}}{c} = \frac{M'_{le}}{N'c}$$

For small angles of attack, N' can be replaced by L' if required.

The Aerodynamic Center

The aerodynamic center is defined as the moment reference point about which the pitching moment does not significantly change with lift. Essentially, the pitching moment remains at the zero lift value. The aerodynamic center for a thin airfoil is theoretically located at the quarter chord position, $\hat{x}_{ac} = 1/4$, for subsonic flow (Subsection 10.7) and the half chord position for supersonic flow (Subsection 10.10), $\hat{x}_{ac} = 1/2$.

The Reynolds Number and Mach Number

The aerodynamic characteristics of an airfoil are governed by its shape and two nondimensional groups known as the Reynolds number and the Mach number, respectively defined as

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

$$\mathrm{Re}_\infty = \frac{\rho_\infty Q_\infty c}{\mu_\infty}$$

$$M_\infty = \frac{Q_\infty}{a_\infty}.$$

It is shown in Subsection 10.4 that if two airfoils are geometrically similar and have equal Reynolds and Mach numbers they will produce identical aerodynamic

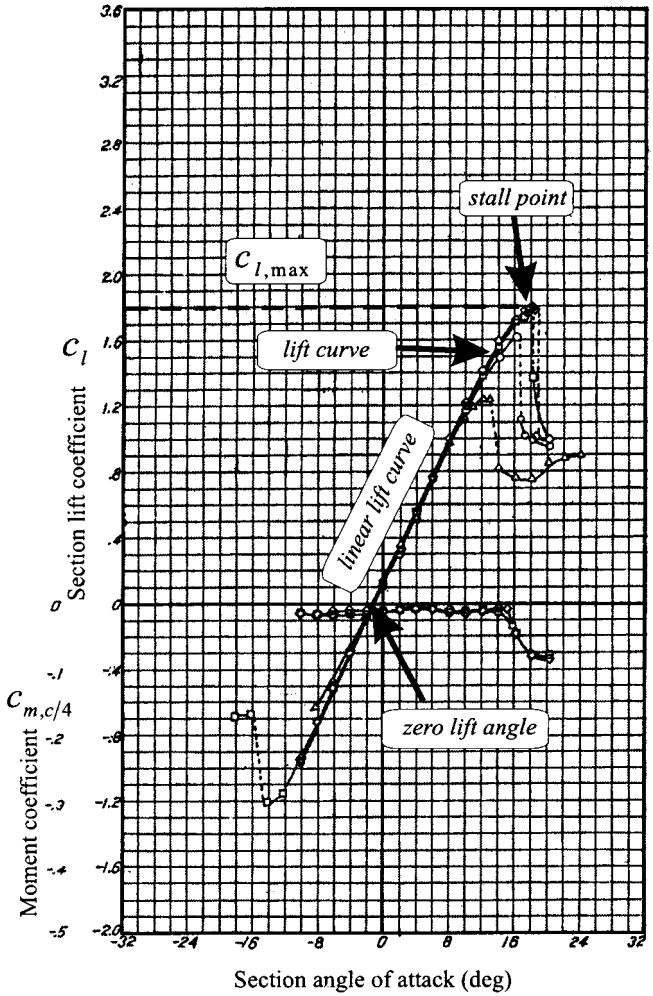


FIGURE 10.4 Lift and pitching moment characteristics for a NACA 23012 (from Abbott et al. 1954, courtesy of NASA).

force and moment coefficients. When this condition is satisfied, the flow fields are said to be dynamically similar. The influences the Reynolds number and Mach number have on the flow field are respectively discussed in Subsections 10.6 and 10.9.

Typical Wind Tunnel Aerodynamic Data

Figures 10.4 and 10.5 display typical wind tunnel data obtained for the NACA 23012 airfoil. Figure 10.4 illustrates the terminology used when describing aerodynamic load characteristics.

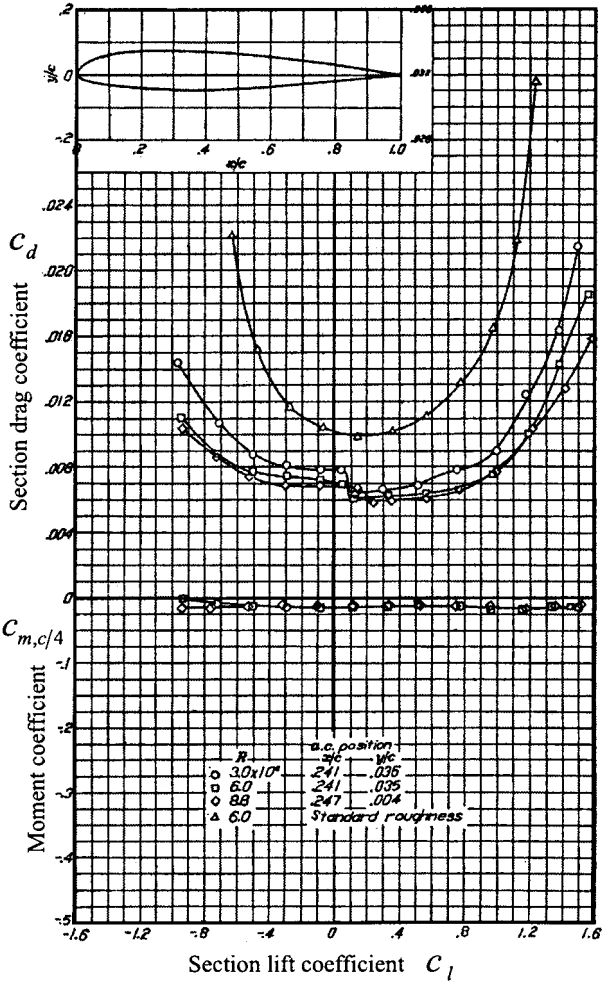


FIGURE 10.5 Drag polar for NACA 23012 (from Abbott et al. 1945 courtesy of NASA).

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

As indicated in Figure 10.4, every airfoil will have one particular angle of attack at which zero lift will be produced; this is known as the zero lift angle, α_{l0}. In general, symmetrical airfoils will have a zero lift angle of zero and airfoils with a net positive camber will have a small negative value. The lift curve slope is defined as the slope of the linear portion of the lift curve which passes through the zero lift angle. Within this linear region the lift coefficient can be computed from

$$c_l = \frac{dc_l}{d\alpha}(\alpha - \alpha_{l0}) = c_{l\alpha}(\alpha - \alpha_{l0})$$

where $c_{l\alpha}$ is known as the lift curve slope.

When the drag coefficient is plotted against the lift coefficient, as in Figure 10.5, the resulting curve is known as a drag polar. The marked decrease in lift coefficient and accompanying increase in drag coefficient are referred to as airfoil stall, and this, as described in Subsections 10.6 and 10.7, is due to various boundary layer phenomena. Figure 10.6 displays the generally accepted terminology used to describe an airfoil pressure distribution. This figure illustrates the variation in upper surface pressure distribution with angle of attack for a modified NACA 23012 airfoil (Niven 1988); this particular airfoil stalls due to trailing edge separation which reduces the net upper surface suction force and thus the lift.

10.3 WING GEOMETRIC AND AERODYNAMIC DEFINITIONS

Wing Geometry

Figure 10.7 illustrates the main terminology and geometric parameters used to define a wing. The following parameters are also used to help define the wing:

Taper ratio: $\lambda = c_t/c_r$

Nondimensional spanwise position: $\hat{y} = 2y/b$

Wing area: $S = 2\int_0^{b/2} c\,dy = b\int_0^1 c\,d\hat{y}$

Mean geometric chord: $c_g = S/b$

Aspect ratio: $A_g = b/c_g = b^2/S$

Definitions of Wing Twist

It will be shown in Subsection 10.7 that it is aerodynamically beneficial to twist a tapered wing. Figure 10.8 shows that the geometric twist at an arbitrary spanwise position (which lies between the root and tip), $\varepsilon_g(\hat{y})$ is defined as the angle between the root chord and the chord of the wing section at the location of interest. Twist is defined as positive when the section is rotated nose-up (relative to the root chord), and is referred to as wash-in. If the section is rotated nose-down the twist is negative and is called wash-out.

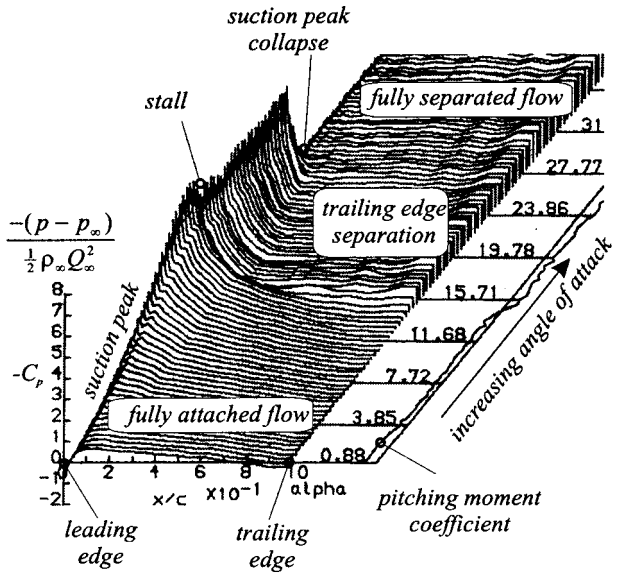


FIGURE 10.6 Typical upper surface pressure distribution (from Niven 1988).

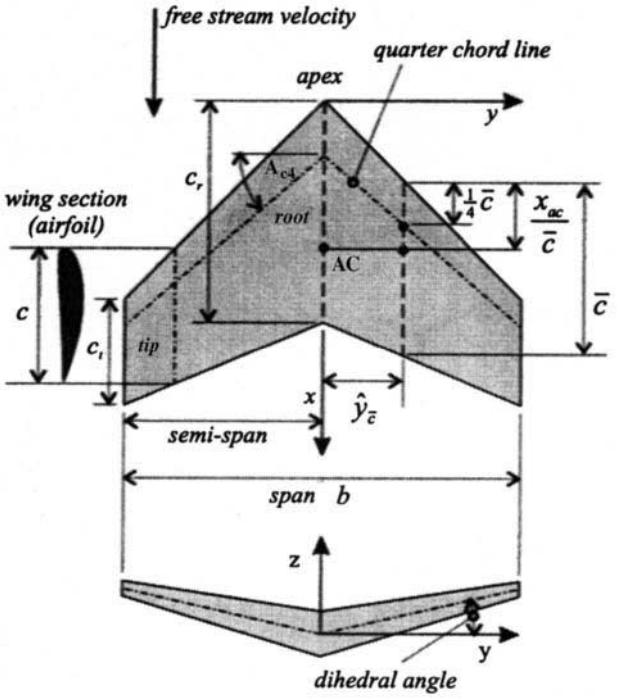


FIGURE 10.7 Wing geometric definitions.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

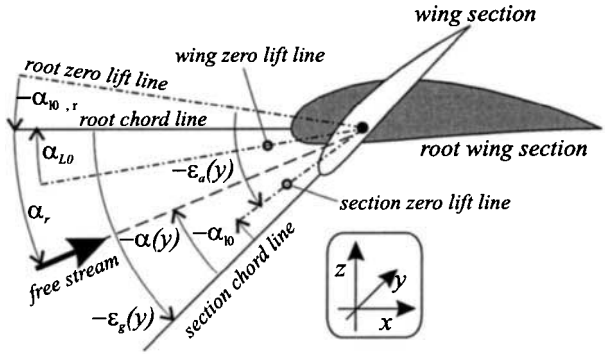


FIGURE 10.8 Wing twist geometric definitions.

The aerodynamic twist, $\varepsilon_a(\hat{y})$, is defined as the angle between the section zero lift line and the root zero lift line. The definition allows for the airfoil profile to change along the span and is thus defined as

$$\varepsilon_a(\hat{y}) = \varepsilon_g(\hat{y}) + \alpha_{l0,r} - \alpha_{l0}(\hat{y})$$

The aerodynamic twist at the wing tip, $\varepsilon_{a,t}$, is often simply referred to as the wing twist and denoted by ε_t. A common spanwise distribution of aerodynamic twist would be a linear variation given by

$$\varepsilon_a(\hat{y}) = \hat{y}(\varepsilon_{a,t})$$

Wing Aerodynamic Forces and Moments

When discussing the aerodynamic characteristic of a wing, the following nondimensional groups are used:

Wing lift coefficient: $$C_L = \frac{L}{\frac{1}{2}\rho_\infty Q_\infty^2 S}$$

Wing drag coefficient: $$C_D = \frac{D}{\frac{1}{2}\rho_\infty Q_\infty^2 S}$$

Figure 10.8 shows that the wing angle of attack is defined as the angle between the root chord line and the free stream velocity. The section angle of attack is related to the wing angle of attack by

$$\alpha(\hat{y}) = \varepsilon_g(\hat{y}) + \alpha_r$$

Every wing will have a particular angle of attack at which no net lift is produced; this is known as the wing zero lift angle, α_{L0}, and can be obtained using the methods described in Subsection 10.8. In a manner analogous to an airfoil, the wing lift coefficient can be plotted against the wing angle of attack. Over the linear portion of the lift curve the wing lift coefficient is given by

$$C_L = \frac{dC_L}{d\alpha}(\alpha_r - \alpha_{L0}) = C_{L\alpha}(\alpha_r - \alpha_{L0})$$

where $C_{L\alpha}$ is known as the wing lift curve slope and may be computed using the methods described in Section 10.8.

The pitching moment coefficient of the wing is defined as

$$c_m = \frac{M}{\frac{1}{2}\rho_\infty Q_\infty^2 S\bar{c}}$$

where the mean aerodynamic chord of a wing is given by

$$\bar{c} = \frac{2}{S}\int_0^{b/2} c^2\, dy = \frac{b}{S}\int_0^1 c^2\, d\hat{y}$$

and is located at a spanwise position equal to

$$\hat{y}_{\bar{c}} = \frac{b^2}{2S}\int_0^1 c\hat{y}d\hat{y}$$

The mean aerodynamic chord represents the chord of a untwisted, unswept, rectangular wing which produces the same lift and pitching moment as the actual wing. For a straight tapered wing it is given by

$$\frac{\bar{c}}{c_r} = \frac{2}{3}\left(\frac{1 + \lambda + \lambda^2}{1 + \lambda^2}\right)$$

The Wing Aerodynamic Center

The wing aerodynamic center is located on the root chord and is defined as the moment reference point about which the wing pitching moment does not significantly change with wing lift. Figure 10.7 shows that the approximate position of the wing aerodynamic center can be found (Schlichting and Truckenbrodt 1979) by first locating the aerodynamic center on the mean aerodynamic chord, $x_{ac}/\bar{c}$, and then projecting this point onto the root chord. When the flow is subsonic $x_{ac}/\bar{c} = \frac{1}{4}$, while for supersonic flow $x_{ac}/\bar{c} = 1/2$.

10.4 FUNDAMENTALS OF VECTOR FLUID DYNAMICS

Control Volume Analysis

The laws of mass, momentum, and energy conservation are applied to a fluid using well-defined regions of the flow field known as control volumes. When the chosen volume is very small it is referred to as a differential control volume or a fluid element. The geometry of the fluid element can be defined using either Cartesian (as used here), cylindrical, or spherical coordinate systems. The conservation laws can be applied either to a moving or a stationary fluid element.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

Generalized Motion of a Fluid

Consider a point in a steady flow where the fluid velocity in the x-direction is u_1. A small distance away from this point the velocity, u_2, can be related to u_1 via the truncated Taylor series expansion below:

$$u_2 = u_1 + \left(\frac{\partial u}{\partial x}\right) dx + \left(\frac{\partial u}{\partial y}\right) dy + \left(\frac{\partial u}{\partial z}\right) dz$$

The fluid velocity can now be written in the form

$$u_2 = u_1 + \left[0 - \frac{1}{2}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) dy + \frac{1}{2}\left(\frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}\right) dz\right] \left[\left(\frac{\partial u}{\partial x}\right) dx + \frac{1}{2}\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right) dy + \frac{1}{2}\left(\frac{\partial w}{\partial x} + \frac{\partial u}{\partial z}\right) dz\right]$$

Similar equations can be written for the v-velocity and the w-velocity to give the following tensor equation (i.e., a matrix of vector quantities):

$$\begin{bmatrix} u_2 \\ v_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} u_1 \\ v_1 \\ w_1 \end{bmatrix} + \begin{bmatrix} 0 & -\Omega_z & \Omega_y \\ \Omega_z & 0 & -\Omega_x \\ -\Omega_y & \Omega_x & 0 \end{bmatrix} \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} + \begin{bmatrix} \dot{\varepsilon}_{xx} & \dot{\varepsilon}_{xy} & \dot{\varepsilon}_{xz} \\ \dot{\varepsilon}_{yx} & \dot{\varepsilon}_{yy} & \dot{\varepsilon}_{yz} \\ \dot{\varepsilon}_{zx} & \dot{\varepsilon}_{zy} & \dot{\varepsilon}_{zz} \end{bmatrix} \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix}$$

When written in this form, the motion of the fluid can be considered to be composed of a translation, a rotation (the Ω matrix) and a deformation (the $\dot{\varepsilon}$ matrix). The mathematical form of the rotation and deformation matrices is described below.

The Curl of the Velocity Vector

By considering the motion of a small rectangular fluid element it can be shown (Eskinazi 1967) that the local angular velocity of the fluid is related to the velocity field in the following manner:

$$(\Omega_x, \Omega_y, \Omega_z) = \vec{i}\,\frac{1}{2}\left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}\right) + \vec{j}\,\frac{1}{2}\left(\frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}\right) + \vec{k}\,\frac{1}{2}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)$$

$$= \frac{1}{2}(\nabla \times \vec{q}) = \frac{1}{2}\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ u & v & w \end{vmatrix} \tag{10.1}$$

where

$$\nabla = \vec{i}\,\frac{\partial}{\partial x} + \vec{j}\,\frac{\partial}{\partial y} + \vec{k}\,\frac{\partial}{\partial z}$$

is referred to as a vector differential operator and is commonly called del. The vector product $(\nabla \times \vec{q})$ is known as the curl of the velocity vector. In fluid dynamics twice the angular velocity is known as the vorticity vector and is denoted

$\vec{\omega} = (\omega_x, \omega_y, \omega_z)$. The right-hand screw rule is used to indicate the positive sense of both the angular velocity and the vorticity vectors. An irrotational flow is one in which $\nabla \times \vec{q} = 0$ throughout the flow field.

The Divergence of the Velocity Vector

Utilizing the fluid element again, it can be shown (White 1991) that the deformation matrix consists of normal and shear strain rates which consist of various groups of the local velocity gradients. This can be compactly written in matrix form as

$$\begin{bmatrix} \dot{\varepsilon}_{xx} & \dot{\varepsilon}_{xy} & \dot{\varepsilon}_{xz} \\ \dot{\varepsilon}_{yx} & \dot{\varepsilon}_{yy} & \dot{\varepsilon}_{yz} \\ \dot{\varepsilon}_{zx} & \dot{\varepsilon}_{zy} & \dot{\varepsilon}_{zz} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial u}{\partial x} & \dfrac{1}{2}\left(\dfrac{\partial v}{\partial x} + \dfrac{\partial u}{\partial y}\right) & \dfrac{1}{2}\left(\dfrac{\partial w}{\partial x} + \dfrac{\partial u}{\partial z}\right) \\ \dfrac{1}{2}\left(\dfrac{\partial u}{\partial y} + \dfrac{\partial v}{\partial x}\right) & \dfrac{\partial v}{\partial y} & \dfrac{1}{2}\left(\dfrac{\partial w}{\partial y} + \dfrac{\partial v}{\partial z}\right) \\ \dfrac{1}{2}\left(\dfrac{\partial u}{\partial z} + \dfrac{\partial w}{\partial x}\right) & \dfrac{1}{2}\left(\dfrac{\partial v}{\partial z} + \dfrac{\partial w}{\partial y}\right) & \dfrac{\partial w}{\partial z} \end{bmatrix} \tag{10.2}$$

The volumetric strain rate is given by

$$\dot{\varepsilon}_V = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = \nabla \cdot \vec{q} \tag{10.3}$$

where $\nabla \cdot \vec{q}$ is known as the divergence of the velocity vector.

The Gradient of a Scalar Field

At any point within a scalar field, there will be a single value and direction for the maximum spatial rate of change of the scalar. The gradient of a scalar field is a vector field given, in Cartesian coordinates, as

$$\nabla\phi = \vec{i}\,\frac{\partial \phi}{\partial x} + \vec{j}\,\frac{\partial \phi}{\partial y} + \vec{k}\,\frac{\partial \phi}{\partial z}$$

As described in Subsection 10.5, an important application of the gradient of a scalar field is the velocity potential function. Another important parameter, used in many aerodynamic theories, is the unit vector normal to an arbitrary surface in the flow field. This surface may be defined by

$$F(x_s, y_s, z_s) = z_s - f_x(x_s, y_s) = 0$$

The unit normal surface vector can be obtained at any location using the expression

$$\vec{n} = \frac{\nabla F}{|\nabla F|} = \frac{1}{|\nabla F|}\left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z}\right)$$

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

Circulation and Integral Theorems

Circulation is defined as the line integral, around a closed curve, of the component of the local velocity tangential to the path of integration. The right-hand screw rule is used to indicate the positive direction along the integration path. If l and $\vec{l}$ respectively denote the distance around the path and the unit vector along the path, the circulation is given by

$$\Gamma = \oint_c (\vec{q} \cdot \vec{l})dl \tag{10.4}$$

Figure 10.9 illustrates a small differential area placed in a flow which lies in the y–z plane. The circulation around $abcd$ is given by

$$\delta\Gamma = v\delta y + w\delta z - \left(v + \frac{\partial v}{\partial z}\delta z\right)\delta y - \left(w + \frac{\partial w}{\partial y}\delta y\right)\delta z$$
$$= \left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}\right)\delta y\delta z$$

This result states that circulation is the product of vorticity and area and may be generalized in the form

$$\frac{d\Gamma}{dA} = (\nabla \times \vec{q}) \cdot \vec{n}$$

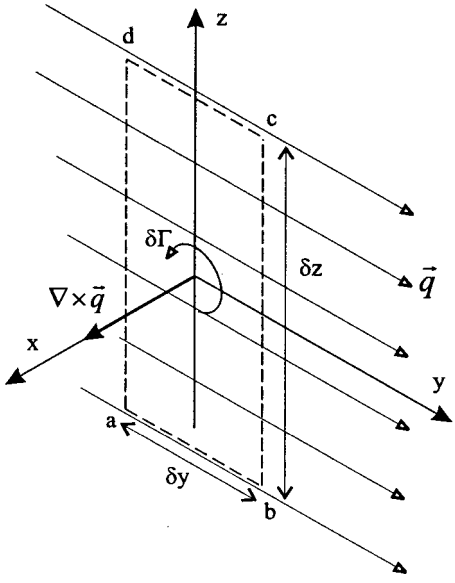


FIGURE 10.9 Relationship between circulation and vorticity.

where $\vec{n}$ is the unit normal of area dA. Stokes's theorem states that the net vorticity over an arbitrary three-dimensional surface can be equated to the circulation around any closed curve which bounds the surface, i.e.,

$$\Gamma = \oint_C \vec{q} \cdot \vec{l}\, dl = \iint_{CS} (\nabla \times \vec{q}) \cdot \vec{n}\, dA \tag{10.5}$$

There are two other useful relationships, known as the gradient theorem and Gauss's theorem, which are respectively given by

$$\oint_{CS} p dA = \iiint_{CV} (\nabla p) dV \quad \text{and} \quad \oint_{CS} (\vec{q} \cdot \vec{n}) dA = \iiint_{CV} (\nabla \cdot \vec{q}) dV$$

Definitions of Inviscid and Incompressible Flow Fields

Figure 10.10 displays the surface stresses experienced by a fluid element. These stresses are due to the pressure of the surrounding fluid and the action of viscosity. Viscous stresses arise when the fluid element is distorted over a period of time and, as described in Subsection 10.6, are a direct result of the microscopic behavior of the fluid molecules. When these forces are assumed to be zero, the fluid is said to be inviscid and the fluid element only experiences pressure forces. Furthermore, an inviscid fluid will not exhibit the effects of mass diffusion or thermal conduction. An incompressible flow arises when the density of the fluid is assumed to be constant throughout.

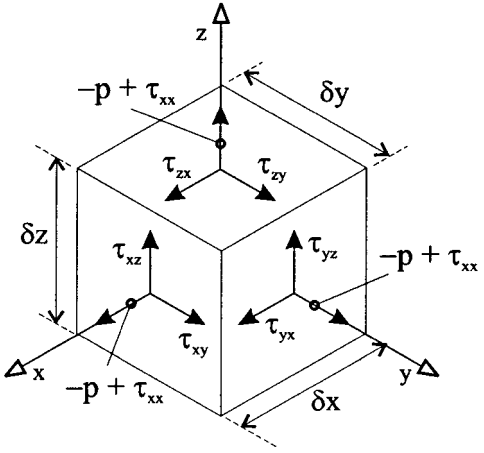


FIGURE 10.10 Pressure and viscous stresses on a fluid element.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

The Conservation of Mass

Applying the conservation of mass principle to a stationary fluid element results in the equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{q}) = 0$$

This equation is often referred to as the continuity equation, which for an incompressible flow reduces to $\nabla \cdot \vec{q} = 0$. This agrees with the earlier definition of volumetric strain rate, which must be zero if the flow is incompressible.

The Navier–Stokes Equations

The Navier–Stokes equations are the result of applying the conservation of linear momentum to a differential control volume along each direction of the chosen coordinate system.

In addition to the surface forces which act on a fluid element (see Figure 10.10), there are body forces which act directly on the mass of the fluid element, e.g., gravity, centrifugal, and electromagnetic. For a stationary fluid element, the sum of these forces is set to the rate of change of momentum flow through the element, which is represented by the left-hand side of the equations below.

$$\frac{\partial(\rho u)}{\partial t} + \nabla \cdot (\rho u \vec{q}) = -\frac{\partial p}{\partial x} + \frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} + \frac{\tau_{xz}}{\partial z} + \rho f_x$$

$$\frac{\partial(\rho v)}{\partial t} + \nabla \cdot (\rho v \vec{q}) = -\frac{\partial p}{\partial y} + \frac{\partial \tau_{yx}}{\partial x} + \frac{\partial \tau_{yy}}{\partial y} + \frac{\partial \tau_{yz}}{\partial z} + \rho f_y$$

$$\frac{\partial(\rho w)}{\partial t} + \nabla \cdot (\rho w \vec{q}) = -\frac{\partial p}{\partial z} + \frac{\partial \tau_{zx}}{\partial x} + \frac{\partial \tau_{zy}}{\partial y} + \frac{\partial \tau_{zz}}{\partial z} + \rho f_z$$

The viscous stresses are related to the fluid strain rates using Stokes's deformation laws, which consist of the following three postulations:

1. The stress is a linear function of the strain rate (originally proposed by Newton in 1686, hence the name Newtonian fluid).
2. The fluid displays no preferential direction of deformation, i.e., the fluid is isotropic.
3. When the strain rates reduce to zero (stationary fluid) the normal stresses must become equal and opposite to the local fluid static pressure (the hydrostatic condition, where the static pressure is a function of temperature and depth from a reference point).

These three postulations can be used to develop the following relationships between viscous stresses and strain rates (Prandtl and Tietjens 1934; White 1991):

$$\begin{bmatrix} \tau_{xx} & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \tau_{yy} & \tau_{xz} \\ \tau_{zx} & \tau_{yz} & \tau_{zz} \end{bmatrix} \tag{10.6}$$

$$= \begin{bmatrix} 2\mu \dfrac{\partial u}{\partial x} - \tfrac{2}{3}\mu\nabla \cdot \vec{q} & \mu\left(\dfrac{\partial u}{\partial y} + \dfrac{\partial v}{\partial x}\right) & \mu\left(\dfrac{\partial u}{\partial z} + \dfrac{\partial w}{\partial x}\right) \\ \mu\left(\dfrac{\partial v}{\partial x} + \dfrac{\partial u}{\partial y}\right) & 2\mu \dfrac{\partial v}{\partial y} - \tfrac{2}{3}\mu\nabla \cdot \vec{q} & \mu\left(\dfrac{\partial v}{\partial z} + \dfrac{\partial w}{\partial y}\right) \\ \mu\left(\dfrac{\partial w}{\partial x} + \dfrac{\partial u}{\partial z}\right) & \mu\left(\dfrac{\partial w}{\partial y} + \dfrac{\partial v}{\partial z}\right) & 2\mu \dfrac{\partial w}{\partial z} - \tfrac{2}{3}\mu\nabla \cdot \vec{q} \end{bmatrix}$$

The Conservation of Energy

Applying the conservation of energy principle to a stationary fluid element results in the following equation, which involves the internal energy of the fluid;

$$\frac{\partial(\rho i)}{\partial t} + \nabla \cdot (\rho i \vec{q}) = -p\nabla \cdot \vec{q} + \nabla \cdot (k\nabla T) + \Phi + \rho(\dot{G} + \vec{f} \cdot \vec{q})$$

The term Φ is known as the viscous dissipation term and is given by

$$\Phi = \mu \begin{bmatrix} 2\left(\dfrac{\partial u}{\partial x}\right)^2 + 2\left(\dfrac{\partial v}{\partial y}\right)^2 + 2\left(\dfrac{\partial w}{\partial z}\right)^2 \\ +\left(\dfrac{\partial u}{\partial y} + \dfrac{\partial v}{\partial x}\right)^2 + \left(\dfrac{\partial u}{\partial z} + \dfrac{\partial w}{\partial x}\right)^2 + \left(\dfrac{\partial v}{\partial z} + \dfrac{\partial w}{\partial y}\right)^2 \\ -\dfrac{2}{3}\left(\dfrac{\partial u}{\partial x} + \dfrac{\partial v}{\partial y} + \dfrac{\partial w}{\partial z}\right)^2 \end{bmatrix}$$

The Euler and Bernoulli Equations

The Euler equations can be obtained from the Navier–Stokes equations by using the assumption that the fluid is incompressible and inviscid, which gives

$$\frac{\partial u}{\partial t} + \vec{q} \cdot \nabla u = \frac{1}{\rho}\frac{\partial p}{\partial x} + f_x$$

$$\frac{\partial v}{\partial t} + \vec{q} \cdot \nabla v = -\frac{1}{\rho}\frac{\partial p}{\partial y} + f_y$$

$$\frac{\partial w}{\partial t} + \vec{q} \cdot \nabla w = -\frac{1}{\rho}\frac{\partial p}{\partial z} + f_z$$

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

Let h denote the positive vertical direction opposite to that of gravity. Considering only gravitational force on the fluid element, the body forces terms can now be written as

$$f_x = -\frac{\partial(gh)}{\partial x}, \quad f_y = -\frac{\partial(gh)}{\partial y} \quad \text{and} \quad f_z = -\frac{\partial(gh)}{\partial z}$$

Now assume that the flow is irrotational (see Subsection 10.5), which gives

$$u = \frac{\partial \phi}{\partial x}, \quad v = \frac{\partial \phi}{\partial y} \quad \text{and} \quad w = \frac{\partial \phi}{\partial z}$$

Using this condition along with the three xyz Euler equations and then integrating each resulting equation with respect to the appropriate coordinate direction leads to the conclusion that

$$\frac{\partial \phi}{\partial t} + \frac{p}{\rho} + \tfrac{1}{2}q^2 + gh = F(t) \tag{10.7}$$

This is known as Bernoulli's equation and in this form is applicable to a fluid which is unsteady, inviscid, irrotational, and incompressible. For steady flow, $(\partial\phi/\partial t) = 0$ and the value of F is constant throughout the entire flow field. It can be shown (White 1991) that Bernoulli's equation is also valid for irrotational flow, but the value of F is only constant along a given streamline and varies from streamline to streamline.

The Reynolds Number and Mach Number

The nondimensionalized version of the x-direction Navier–Stokes equation for two-dimensional steady flow is given by

$$\frac{\partial(\hat{\rho}\hat{u}\hat{u})}{\partial \hat{x}} + \frac{\partial(\hat{\rho}\hat{u}\hat{v})}{\partial \hat{y}} = -\frac{1}{\gamma M_\infty^2}\frac{\partial \hat{p}}{\partial \hat{x}} + \frac{1}{\mathrm{Re}_\infty}\frac{\partial}{\partial \hat{y}}\left[\hat{\mu}\left(\frac{\partial \hat{v}}{\partial \hat{x}} + \frac{\partial \hat{u}}{\partial \hat{y}}\right)\right]$$

where

$$\bar{x} = x/c, \quad \hat{y} = y/c, \quad \hat{u} = u/Q_\infty, \quad \hat{v} = v/Q_\infty, \quad \hat{p} = p/p_\infty$$

$$\hat{\rho} = \rho/\rho_\infty, \quad \hat{\mu} = \mu/\mu_\infty$$

and

$$M_\infty = Q_\infty/a_\infty \quad \text{and} \quad \mathrm{Re}_\infty = (\rho_\infty Q_\infty c)/\mu_\infty$$

If the flow over two geometrically similar bodies has the same Mach number and Reynolds numbers the solution of the nondimensional Navier–Stokes equations will be numerically identical. When this occurs, the two flows are said to be dynamically similar.

10.5 FUNDAMENTALS OF POTENTIAL FLOW

The Potential Function

Consider an inviscid irrotational flow defined by $\vec{\omega} = \nabla \times \vec{q} = 0$ (see Subsection 10.4 for expansion of vector product). Using the rules of partial differentiation, this condition is also satisfied if

$$\vec{\omega} = \nabla \times (\nabla\phi) = 0$$

Thus, for an irrotational flow, the velocity vector field is the gradient of a scalar field (see Subsection 10.4), which is known as the velocity potential, ϕ, i.e.,

$$\vec{q} = (u,v,w) = \nabla\phi = \left(\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}, \frac{\partial\phi}{\partial z}\right)$$

The velocity potential can also be defined as $\vec{q} = -\nabla\phi$ without any effect on the irrotational flow condition. If the flow is both irrotational and incompressible, we have

$$\nabla \cdot \vec{q} = \nabla \cdot (\nabla\phi) = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} = \nabla^2\phi = 0$$

This is a second order linear partial differential equation for the velocity potential and is known as Laplace's equation. A flow which is both incompressible and irrotational is known as a potential flow. Laplace's equation may be solved analytically or numerically subject to the boundary conditions discussed later on in this section. Many analytical solutions of the Laplace equation can be associated with simple flow field patterns; the most frequently used solutions are described below.

The Disturbance Velocity Potential

A concept used frequently in aerodynamic theories is one which considers the flow field velocity potential function to consist of two parts: one due to the free stream flow and one, known as the disturbance potential, which accounts for the presence of any body within the flow. Mathematically this is written as $\phi = \phi_\infty + \tilde{\phi}$, where $\phi_\infty = U_\infty x + V_\infty y + W_\infty z$ and $\tilde{\phi}$ is the disturbance velocity potential. The local velocity field is given by $u = U_\infty + \tilde{u}$, $v = V_\infty + \tilde{v}$, and $w = W_\infty + \tilde{w}$, where $\tilde{u} = \partial\tilde{\phi}/\partial x$, $\tilde{v} = \partial\tilde{\phi}/\partial y$, and $\tilde{w} = \partial\tilde{\phi}/\partial z$. The solution to the flow field is often found using the small disturbance approximation as described later in this section.

The Stream Function

A stream line is defined as a line whose spatial variation in gradient, at any given instant in time, corresponds to the variation in flow direction of the local velocity vector. For two-dimensional flow, a streamline is mathematically defined (using Cartesian coordinates) by the relationship

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

$$\frac{dy}{dx} = \frac{v}{u}$$

The conservation of mass principle (Subsection 10.4) for two-dimensional incompressible flow is given by

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0$$

This equation can also be satisfied, using the rules of partial differentiation, by another function, called the stream function, and defined by

$$u = \frac{\partial \psi}{\partial y} \quad \text{and} \quad v = -\frac{\partial \psi}{\partial x}.$$

When the flow field is irrotational, the stream function satisfies Laplace's equation, i.e.,

$$\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = \nabla^2 \psi = 0$$

It should be noted that whereas both streamlines and the potential function exist for all irrotational three-dimensional flow fields, the stream function can only be defined in three dimensions when the flow is axisymmetric (Vallentine 1959).

Two-Dimensional Solutions of the Laplace Equation

Table 10.4 gives the potential and stream functions in polar coordinates for a number of elementary types of flow which form the basis of most aerodynamic mathematical models. In the uniform flow case, α is the angle between the free stream flow and the horizontal. For the vortex, Γ is taken positive according to the right-hand screw rule, which is anticlockwise when the vortex lies in the x–y plane and $\psi = 0$ when $r = r_0$. Figure 10.11 illustrates the stream line patterns associated with these flows.

TABLE 10.4 Elementary Flow Patterns

Flow	Potential function	Stream function
Uniform flow	$\phi = Q_\infty r \cos(\theta - \alpha)$	$\psi = Q_\infty r \sin(\theta - \alpha)$
Source	$\phi = \frac{\Lambda}{2\pi} \ln r$	$\psi = \frac{\Lambda}{2\pi} \theta$
Vortex	$\phi = \frac{\Gamma}{2\pi} \theta$	$\psi = -\frac{\Gamma}{2\pi} \ln \frac{r}{r_0}$
Doublet	$\phi = \frac{H}{2\pi} \frac{\cos \theta}{r}$	$\psi = -\frac{H}{2\pi} \frac{\sin \theta}{r}$

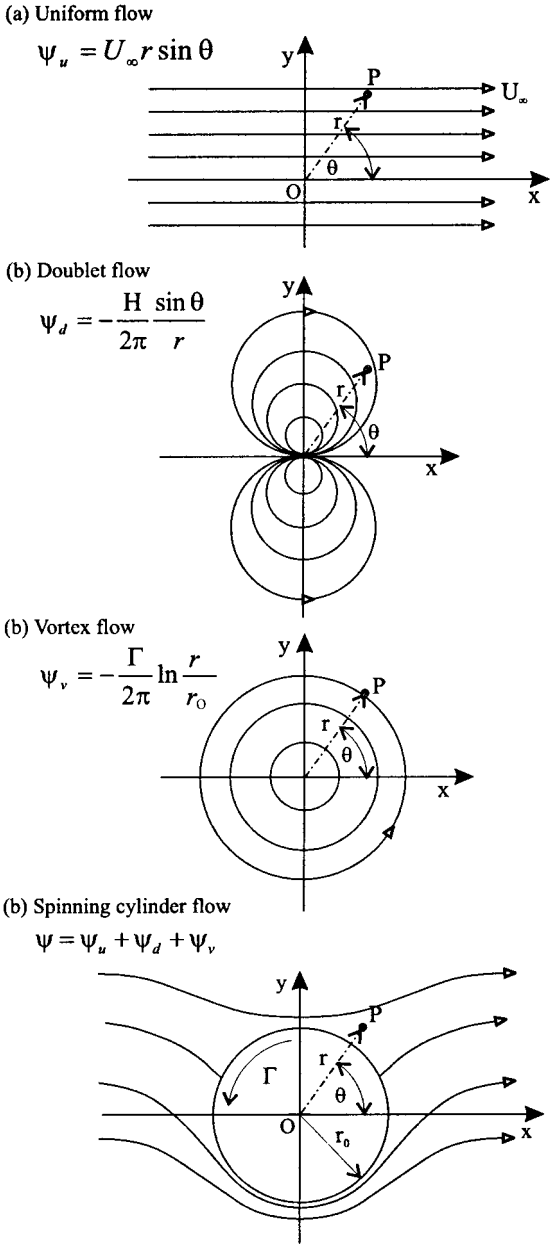


FIGURE 10.11 The superposition principle.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

The components of the local velocity vector, $\vec{q}$, in polar coordinates are related to the potential and stream functions as follows:

$$q_r = \frac{\partial \phi}{\partial r} = \frac{1}{r}\frac{\partial \psi}{\partial \theta} \quad \text{and} \quad q_\theta = \frac{1}{r}\frac{\partial \phi}{\partial \theta} = -\frac{\partial \psi}{\partial r}$$

With respect to the vortex flow field, the circulation around any closed curve which encloses the center has the value Γ. However, if the closed curve does not surround the vortex center the circulation will be zero. Thus, the entire flow field is irrotational and the vortex center is said to be a singularity.

The Principle of Superposition

Solutions of the Laplace equation, which are relevant to aerodynamic work, are normally achieved by adding together elementary potential flow solutions which is known as superposition. For example, Figure 10.11 shows that the lifting flow over a cylinder can be solved by adding together the flow fields due to uniform flow, a doublet and a vortex (the circulation can be set to any value). The final stream function for the spinning cylinder would be given by

$$\psi = U_\infty \left(r - \frac{r_0^2}{r}\right) \sin\,\theta - \frac{\Gamma}{2\pi} \ln \frac{r}{r_0} \tag{10.8}$$

where $r_0^2 = \dfrac{H}{2\pi U_\infty}$

The Kutta–Joukowski Theorem of Lift

Using the stream function, given by Equation (10.8), the tangential component of the velocity vector, at any point in the flow field, is given by

$$q_\theta = -\frac{\partial \psi}{\partial r} = -U_\infty \left(1 + \frac{r_0^2}{r^2}\right) \sin\,\theta + \frac{\Gamma}{2\pi r}$$

The velocity on the cylinder surface is given when $r = r_0$, i.e.,

$$q_0 = -2U_\infty \sin\,\theta + \frac{\Gamma}{2\pi r_0}$$

The surface pressure coefficient becomes

$$C_p = 1 - \left(\frac{q_\theta}{U_\infty}\right)^2 = 1 - 4\sin^2\theta + \frac{4\Gamma}{2\pi U_\infty r_0}\sin\,\theta - \left(\frac{\Gamma}{2\pi U_\infty r_0}\right)^2$$

The lift coefficient for the cylinder can be obtained from

$$c_l = -\tfrac{1}{2}\int_0^{2\pi} C_p \sin\,\theta d\theta = \frac{-\Gamma}{U_\infty r_0}$$

which gives the lift per unit span as

$$L' = \tfrac{1}{2}\rho_\infty Q_\infty^2 2r_0 c_l = -\rho_\infty U_\infty \Gamma$$

This result can be shown to apply to the flow around any body which has some value of circulation associated with it and is known as the Kutta–Joukowski theorem. The pressure drag on the cylinder is given by

$$c_d = -\tfrac{1}{2}\int_0^{2\pi} C_p \cos\theta d\theta = 0$$

The fact that inviscid flow theory always predicts zero drag which is never the case in reality is known as d'Alembert's paradox. As discussed in Subsections 10.6 and 10.7, this difference can be explained using the viscous phenomena of boundary layer formation and separation.

Three-Dimensional Vortex Flows

A vortex line is defined as a line whose spatial variation in gradient, at any given instant in time, corresponds to the spatial variation in direction of the local vorticity vector.

When a number of vortex lines bunch together and pass through a common differential cross-sectional area, dA, a vortex filament is said to have formed. The strength of this filament is given by the circulation around the perimeter of dA and is related to the local vorticity via Stokes's theorem (Subsection 10.4), i.e., $d\Gamma = \vec{\omega} \cdot \hat{n} dA$. A vortex tube is defined as a bundle of vortex filaments.

The behavior of a vortex tube is governed by Helmholtz's laws of vorticity, which state that:

1. The strength of a vortex filament is constant along its length.
2. A vortex filament must either form a closed path, extend to infinity, or terminate on a solid boundary. This is a consequence of the first law, which effectively states that the product of vorticity and the filament cross-sectional area must remain constant.
3. A vortex filament always consists of the same fluid elements.
4. The strength of a vortex filament remains constant as it moves throughout the flow field.

The last two laws are consequences of the inviscid flow assumption, which does not allow any diffusion of flow properties. For more information on vortex flows and proofs of Helmholtz's laws, see Lugt (1995) and Eskinazi (1967).

The velocity field induced by a vortex filament is given by the Biot–Savart law. Figure 10.12 shows a straight-line vortex filament lying along the z-axis. The velocity induced by a small element, δl, has the magnitude

$$\delta q_\theta = \frac{\Gamma}{4\pi}\frac{\sin\theta}{r^2}\delta l$$

and points in the direction $\delta\hat{l} \times \vec{r}$ where $\delta\hat{l}$ is the unit vector in the direction l. Integrating between the points A and B results in

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

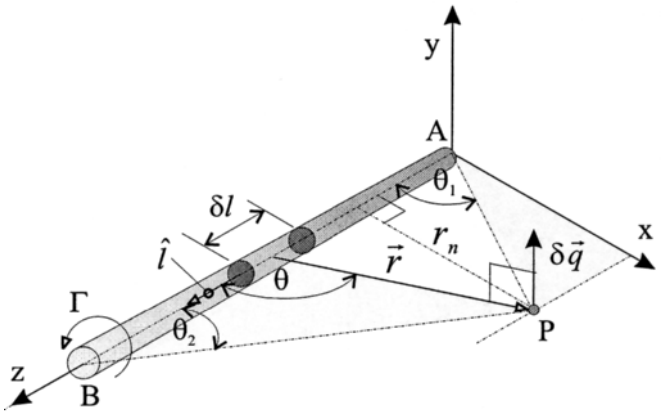


FIGURE 10.12 A straight-line vortex filament.

$$q_\theta = \frac{\Gamma}{4\pi r_n}(\cos\theta_1 + \cos\theta_2)$$

As will be discussed in Subsection 10.8, the concept of a semi-infinite vortex filament will be utilized to model the flow field around a finite wing. Referring again to Figure 10.12, a semi-infinite vortex is obtained $\theta_1 = \pi/2$ and $\theta_2 \to 0$, which causes q_θ to lie in the x–y plane and have the magnitude

$$q_\theta = \frac{\Gamma}{4\pi r_n}$$

Proofs of the Biot–Savart law can be found in Eskinazi (1967), Karamcheti (1980), and Katz and Plotkin (1991).

An important concept in aerodynamic theory is the vortex sheet, shown in Figure 10.13, which is defined as a large number of vortex filaments whose axes lie parallel to each other within a mutual plane. The circulation (see Subsection 10.4) around the path *abcd* is given by

$$\delta\Gamma = \left(\frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}\right)\delta z\delta x = \delta u\delta x - \delta v\delta z$$

For a thin vortex sheet this can be approximated by

$$\delta\Gamma = \delta u\delta x = (u_u - u_l)\delta x = \gamma\delta x$$

where γ is known as the vortex sheet strength. Referring to Figure 10.13, the velocity potential and induced velocity at point P given by a small segment of the vortex sheet lying at the origin are respectively given by

$$\delta\phi = \frac{\gamma\delta x}{2\pi}\theta \quad \text{and} \quad \delta q_\theta = \frac{\gamma\delta x}{2\pi r}$$

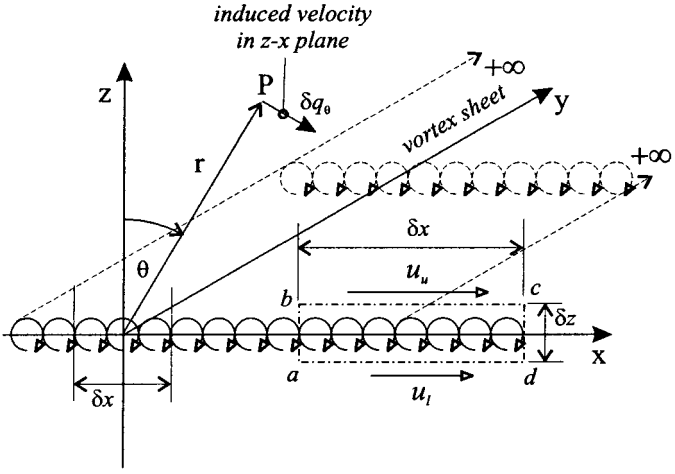


FIGURE 10.13 A vortex sheet lying in the x–z plane.

Conformal Transformation

In a similar fashion to the function $y = f(x)$, a function of a complex number $w = f(z)$ can be formulated. When this occurs, z is known as a complex variable and w can be written in the following form:

$$w = f(z) = f(x + iy) = \phi(x,y) + i\psi(x,y)$$

Just as z defines a point on the z-plane with x as the abscissa and y as the ordinate, w defines a point on the w-plane which has ϕ as the abscissa and ψ as the ordinate. The phrase function of a complex variable is conventionally restricted to a type of function known as analytic (or holomorphic or regular). A function $w = f(z)$ is classified as analytic when the following two conditions are satisfied:

1. For each value of z there is only one finite value of w.

2. dw/dz is single-valued and neither zero or infinite.

Although there can be exceptions to these conditions, known as singularities, these points can often be mathematically excluded from the transformation process. Condition (1) is normally always satisfied, and it can be shown that condition (2) is satisfied when

$$\frac{\partial \phi}{\partial x} = \frac{\partial \psi}{\partial y} \quad \text{and} \quad \frac{\partial \phi}{\partial y} = -\frac{\partial \psi}{\partial x}$$

These are known as the Cauchy–Riemann equations, which, after differentiation, result in

$$\nabla^2 \phi = 0 \quad \text{and} \quad \nabla^2 \psi = 0$$

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

This result indicates that ϕ can be regarded as representing the velocity potential function and ψ as the stream function. Thus, the pattern on the w-plane represents uniform inviscid incompressible irrotational flow and the function $w = f(z)$ is often referred to as the complex potential. Since both ϕ and ψ satisfy Laplace's equation and are related by the Cauchy–Rieman equations, they are known as conjugate harmonic functions. Under the inverse function $z = f^{-1}(w)$, Figure 10.14 shows that the pattern on the z-plane can be associated with a particular case of nonuniform potential flow; in this example, the function $w = z^2$ is used and the flow on the z-plane can be taken to represent the internal flow around a right-angle corner.

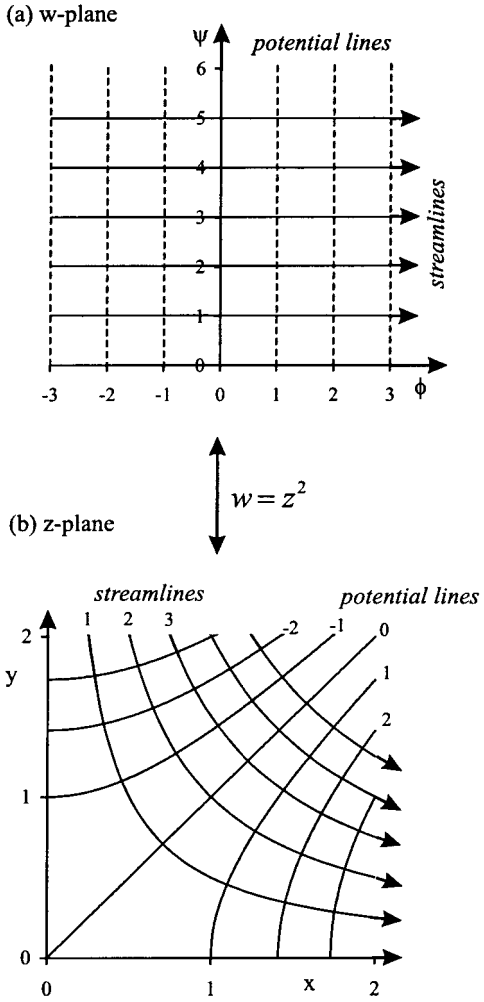


FIGURE 10.14 An example of conformal transformation.

The derivative dw/dz can be regarded as a complex operator where a small line δz on the z-plane is mapped onto a corresponding line δw on the w-plane, i.e., $\delta w = (dw/dz)\delta z$. Since dw/dz is itself a complex number, this mapping consists of a rotation and a change of scale. Furthermore, it can be shown that the angle of intersection between any two lines is preserved on both the z and w planes. When this characteristic occurs the transformation is known as conformal. Alternatively, dw/dz can be thought of as a complex velocity since

$$dw = \left(\frac{\partial w}{\partial x}\right) dx + \left(\frac{\partial ws}{\partial y}\right) dy = \left(\frac{\partial \phi}{\partial x} + i\,\frac{\partial \psi}{\partial x}\right) dx + \left(\frac{\partial \phi}{\partial y} + i\,\frac{\partial \psi}{\partial y}\right) dy$$

$$= (u - iv)dz$$

Thus,

$$\frac{dw}{dz} = V^{-i\theta} \quad \text{where } V = \sqrt{u^2 + v^2} \quad \text{and } \theta = \tan^{-1}(v/u)$$

A typical solution procedure for a flow field using conformal transformation could be as follows:

1. Find $w = f(z)$, which represents the flow field of interest, by adding together standard transformation functions (see Table 10.5).
2. Split w into ϕ and ψ components.
3. Use both lines of constant ϕ and ψ to draw flow pattern on z-plane.
4. Differentiate ϕ or ψ with respect to either x or y to obtain local velocity components u and v, respectively, velocity magnitude Q_∞, and flow angle θ at any point of interest.

Flow Field Boundary Conditions

Any potential flow field can be obtained by solving the Laplace equation either analytically or numerically using the following boundary conditions:

TABLE 10.5 Elementary Complex Potential Functions

Flow in z-plane	Transformation
Uniform flow	$w = Az$
Source, centered at z_0	$w = m \ln(z - z_0)$
Vortex, centered at z_0 (positive Γ anticlockwise)	$w = -\dfrac{i\Gamma}{2\pi} \ln(z - z_0)$
Doublet, centered as z_0	$w = \dfrac{\mu}{z - z_0}$
Cylinder, radius r_0^2 with circulation	$w = Q_\infty \left(z + \dfrac{r_0^2}{z}\right) w = -\dfrac{i\Gamma}{2\pi} \ln \dfrac{z}{r_0}$

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

1. There is zero normal fluid velocity relative to the body surface, which, for steady flow, is given by $\vec{q} \cdot \hat{n} = (\nabla\phi) \cdot \hat{n} = 0$. This criterion is sometimes referred to as the flow tangency condition or the Neumann problem.
2. The velocity components away from the body should equal those of the free stream.

In many aerodynamic theories the first boundary condition is applied using a technique known as the small disturbance approximation. This approximation can be described by first considering the flow around a finite wing, shown in Figure 10.15, whose surface is defined as

$$F(x_s,y_s,z_s) = z_s - f_x(x_s,y_x) = 0$$

The potential flow solution will be given by $\nabla^2\phi = 0$, where $\phi = \phi_\infty + \tilde{\phi}$, $\phi_\infty = U_\infty x + W_\infty z$, and $\tilde{\phi}$ is the disturbance velocity potential ($V_\infty = 0$). The zero normal velocity boundary condition must be satisfied on the entire wing surface which, for steady flow, is given by

$$\nabla\phi \cdot \hat{n}_s = \frac{\partial f_s}{\partial x}\left\{U_\infty + \left(\frac{\partial\tilde{\phi}}{\partial x}\right)_s\right\} + \frac{\partial f_s}{\partial y}\left(\frac{\partial\tilde{\phi}}{\partial y}\right)_s - \left(\frac{\partial\tilde{\phi}}{\partial z}\right)_x - W_\infty = 0$$

where, as shown in Subsection 10.4, $\hat{n}_s = \nabla F/|\nabla F|$. The small disturbance approximation is given by

$$\frac{|\tilde{u}|}{Q_\infty} << 1 \quad \frac{|\tilde{v}|}{Q_\infty} << 1 \quad \frac{|\tilde{w}|}{Q_\infty} << 1$$

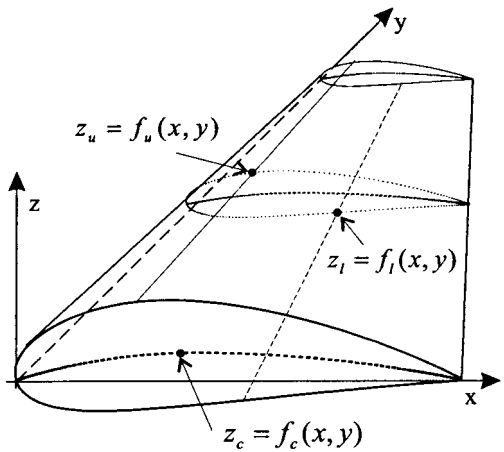


FIGURE 10.15 Airfoil surface definitions.

and

$$\left|\frac{\partial f_s}{\partial x}\right| << 1 \quad \left|\frac{\partial f_s}{\partial y}\right| << 1 \quad \left|\frac{W_\infty}{Q_\infty}\right| \approx \alpha << 1 \quad (\alpha \text{ in radians})$$

Using this approximation results in the following simplified form of the flow tangency condition:

$$\left(\frac{\partial \tilde{\phi}}{\partial z}\right)_s = Q_\infty \left(\frac{\partial f_s}{\partial x} - \alpha\right)$$

The left-hand side of this equation can be expanded using the Taylor series expansion. For example, application to the upper wing surface results in

$$\left(\frac{\partial \tilde{\phi}}{\partial z}\right)_u = \frac{\partial \tilde{\phi}}{\partial z}(x,y,0) + z_u \frac{\partial^2 \tilde{\phi}}{\partial z^2}(x,y,0) + O(z_u^2)$$

Linearizing the flow field means all derivatives higher than first order are ignored and the boundary condition becomes

$$\frac{\partial \tilde{\phi}}{\partial z}(x,y,0) = Q_\infty \left(\frac{\partial f_s}{\partial x} - \alpha\right)$$

Essentially, linearization has reduced the problem to one which has to find the disturbance velocity potential on the wing planform in the x–y plane rather than over the entire wing surface. As will be discussed in the following sections, many aerodynamic models utilize that fact that an airfoil or wing introduces a disturbance velocity in the z-direction on the x–y plane. This is known as downwash, which is given by

$$w(x,y,0) = \frac{\partial \tilde{\phi}}{\partial z}(x,y,0)$$

It should be noted that the small disturbance approximation is not valid near stagnation points or the leading edge.

10.6 ELEMENTARY BOUNDARY LAYER FLOW

Kinetic Theory

When considering the macroscopic properties of a real fluid, we have to examine the behavior of the individual molecules that constitute the fluid. The subject which deals with the microscopic behavior of the fluid is known as kinetic theory.

The molecules which constitute any fluid are in a constant state of random motion. For a stationary fluid the molecular velocity is random in both magnitude and direction. When the fluid is moving in a particular direction the instantaneous velocity of each molecule is the vector sum of the fluid velocity and the instantaneous molecular velocity. Because the fluid velocity is superposed over the molecular velocity, it is often referred to as the ordered velocity, while the molecular

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

velocity is known as the random velocity. This terminology also applies to other fluid properties such as its momentum and energy.

The No-Slip Condition

The surface of an airfoil is made up of molecules which leave spaces between each other of sufficient size to allow the air molecules to penetrate into them. In the 18th century Maxwell suggested that diffuse reflection (i.e., scattering in all directions) would result from the penetration of fluid molecules into the pores of the airfoil surface, where they would strike several times before escaping back into the air flow. Essentially, the air molecules are reflected from the airfoil surface irrespective of their initial direction of impact. This can cause the ordered velocity component of an individual molecule to change direction. When the average ordered velocity is taken over all the molecules, lying just above the surface, its value is found to be zero relative to the surface (i.e., after impacting the airfoil surface as many molecules drift upstream as drift downstream). This behavior is known as the no-slip condition since it causes the air in direct contact with the airfoil surface to acquire the velocity of the surface.

Viscosity and Boundary Layer Formation

When the random movement of air molecules transport ordered momentum from one place to another within a moving mass of air, it is referred to as the action of viscosity. At some distance, y_e, away from the airfoil surface the air is free to move with the ordered velocity, u_e, unaware that there is a solid surface below. By considering the exchange of ordered momentum across an imaginary plane lying parallel to, and just above, the airfoil surface, it can be shown that the action of viscosity diffuses the no-slip condition out into the airflow such that the local air velocity varies from zero at the wall to u_e. This region is known as the boundary layer and thus u_e is known as the boundary layer edge velocity.

The concept of the boundary layer was first introduced by Prandtl in 1904 and is considered as the region in which the effects of viscosity are concentrated. Numerous experimental studies of boundary layer flows have revealed two distinct types of flow behavior, known as laminar and turbulent. Some of the features of these boundary layers are discussed below.

The Laminar Boundary Layer

Figure 10.16 shows that over the airfoil's leading edge the air particles move downstream in smooth and regular trajectories without appreciable mixing between different layers of air. This type of flow is known as a laminar boundary layer. The nondimensional group, which heavily influences the development of any boundary layer, is the Reynolds number, based on the surface distance from the origin of the boundary layer to the point in question. When working with boundary layer flows it is convenient to use a body fitted (or curvilinear) coordinate system in which the x-direction is taken to represent the distance traveled along the surface and the y-direction is taken normal to the surface. Figure 10.17 illustrates the variation in

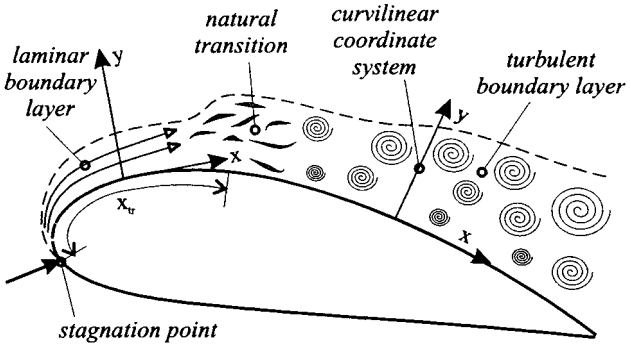


FIGURE 10.16 Typical boundary layer phenomena.

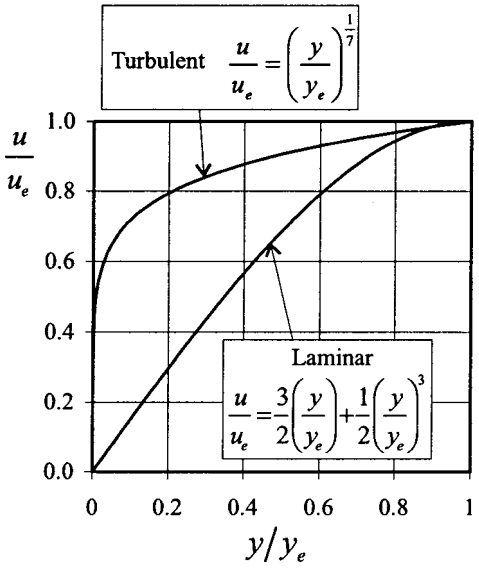


FIGURE 10.17 Velocity profiles for laminar and turbulent flow.

local velocity within a laminar boundary layer; this variation is referred as a velocity profile.

Skin Friction Drag

As described in Subsection 10.4, velocity gradients in a viscous fluid are always accompanied by viscous stresses. Figure 10.17 shows that there is a large velocity gradient at the airfoil surface, which induces a large viscous shear stress to act on

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

the surface in the direction of fluid motion. When these shear stresses are integrated over the entire airfoil surface a drag force, known as the skin friction drag, is obtained.

Transition from Laminar to Turbulent Flow

The phenomena which transform the smooth laminar flow into a chaotic flow, known as a turbulent flow, are collectively known as transition. From the point of view of mathematical modeling, transition comprises of two main processes: the stability of laminar flow to small perturbations, and the amplification of these disturbances such that the transition becomes inevitable. Numerous experimental investigations of pipe flows, boundary layers, and jets have established that transition includes the following phenomena, in order of appearance:

1. Amplification of small disturbances
2. Development of isolated large-scale vortical structures
3. Formation of pockets of small-scale vortical structures known as turbulent spots
4. Growth and coalescence of turbulent spots into a fully developed turbulent flow

The exact details of transition are further complicated by factors such as free stream turbulence, pressure gradient, wall roughness, heat transfer, and Mach number. For further information on transition see Schlichting and Truckenbrodt (1979) and White (1991).

Experimental investigations of the flow over flat plates have indicated that the maximum length of travel of a laminar boundary layer, without undergoing transition to turbulent flow, is given by

$$\mathrm{Re}_x = \frac{\rho Q_\infty x}{\mu} < 500{,}000$$

In practice, however, this value is reduced when the effects of high free stream turbulence, surface roughness, and adverse pressure gradients are considered.

The Turbulent Boundary Layer

Figure 10.16 shows that the flow within a turbulent boundary layer consists of vortex structures known as eddies. In 1885, Reynolds suggested that the instantaneous value of any fluid property within a turbulent flow could be separated into a mean value (which stays constant over some specified time period) and a turbulence value which fluctuates about the mean value. Thus the instantaneous velocity in the x-direction is written in the form $u = U + u'$.

Figure 10.17 illustrates a typical mean velocity profile found in a turbulent boundary layer. The random eddy motion transports high-momentum air, at the outer regions of the boundary layer, towards the airfoil surface. This behavior causes the velocities close to the airfoil surface to be larger than those found in a laminar boundary layer.

The transportation of ordered momentum by eddy motion is a similar action to that of molecular viscosity, although it can be up to three orders of magnitude greater. Because of this similarity, the action of the eddies can be represented by a

variable known as the eddy viscosity. Unfortunately, unlike the molecular viscosity, eddy viscosity is a variable quantity which depends on the flow field itself (plus boundary conditions) rather than a constant value fluid property. The subject which deals with relating the eddy viscosity to the mean velocity gradients (e.g., $\partial U/\partial y$) is known as turbulence modeling. For an excellent introduction to the subject Wildox (1994) should be consulted.

Figure 10.18 details the generally accepted terminology used to describe the various important regions found within a turbulent boundary layer. The inner region covers between 10 and 20% of the overall thickness and the total shear stress (molecular plus turbulent) is almost constant and equal to the wall value (White 1991). The inner region is further broken down into the following three sublayers:

1. The linear sublayer, where molecular viscosity dominates (the turbulent motion is restricted by the presence of the wall)
2. The buffer layer, where molecular and eddy viscosity are of similar magnitude
3. The log-law, where turbulent stresses dominate

Within the linear sublayer the viscous shear remains very close to the wall value (White 1991), which gives

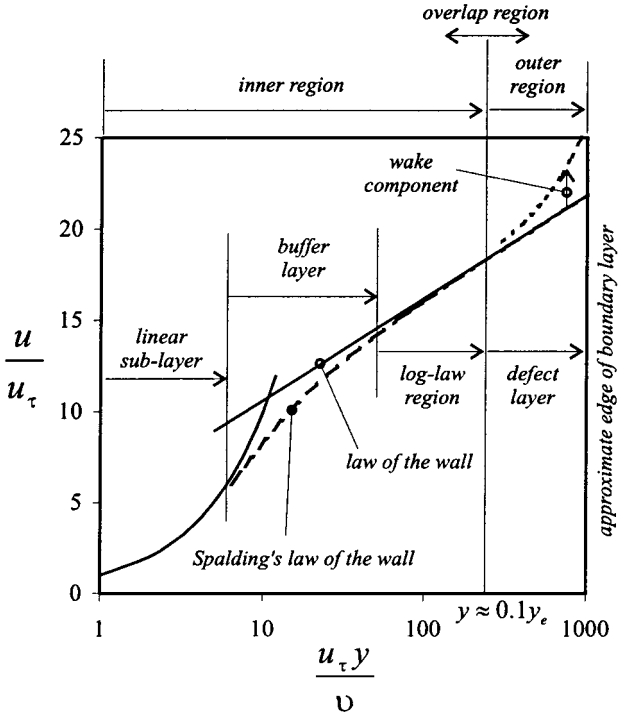


FIGURE 10.18 Turbulent boundary layer definitions.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

$$\mu\left(\frac{dU}{dy}\right) = \tau_w \Rightarrow U = \frac{\tau_w}{\mu} y$$

In nondimensional terms this can be written as

$$\frac{U}{u_\tau} = \frac{\rho u_\tau y}{\mu} \quad \text{or} \quad u^+ = y^+ \quad \text{where} \quad u_\tau = \sqrt{\frac{\tau_w}{\rho}}$$

where u_τ is known as the friction velocity.

As indicated in Figure 10.18, the entire velocity profile is plotted using these nondimensional groups. Moving out of the linear sublayer, Prandtl suggested the following velocity variation, known as the log-law or the law of the wall:

$$u^+ = \frac{1}{\kappa} \ln y^+ + B$$

where experimental measurements have indicated that $\kappa = 0.41$ (known as von Karman's constant) and $B = 5.0$.

Experimental data have indicated a smooth variation of u^+ with y^+ between the linear sublayer and the law of the wall, which is known as the buffer layer. Spalding's law of the wall covers both the buffer and the log-law layers and has the form

$$y^+ = u^+ + e^{-\kappa\beta}(e^{\kappa\beta^+} - 1 - \kappa u^+ - \tfrac{1}{2}(\kappa u^+)^2 - \tfrac{1}{6}(\kappa u^+)^3)$$

The outer region is often referred to as the defect layer or the law of the wake and is the region where turbulent stresses begin to decrease. The velocity variation is given by the law of the wake

$$\frac{u_e - u}{u_\tau} = -\frac{1}{\kappa} \ln\left(\frac{y}{y_e}\right) + A$$

where the value of A depends on the magnitude and direction of the pressure gradient applied to the boundary layer.

Boundary Layer Separation

Figure 10.19 illustrates the flow through a Venturi. When the flow is incompressible, the conservation of mass and energy state that the velocity must decrease and the pressure increase as the flow moves from the throat to the outlet. The increase in pressure is known as an adverse pressure gradient. The opposite case occurs between the inlet and throat and a favorable pressure gradient forms in this region. When a boundary layer is subjected to an adverse pressure gradient the slower moving fluid elements can be forced to change direction and move upstream. This is referred to as reversed flow, and it causes the boundary layer to separate from the surface and form a free shear layer. A turbulent boundary layer is more resistant to separation than a laminar layer due to the more uniform velocity profile. Both laminar and turbulent separation have a great influence on the lift produced by an airfoil, as discussed in Subsection 10.7.

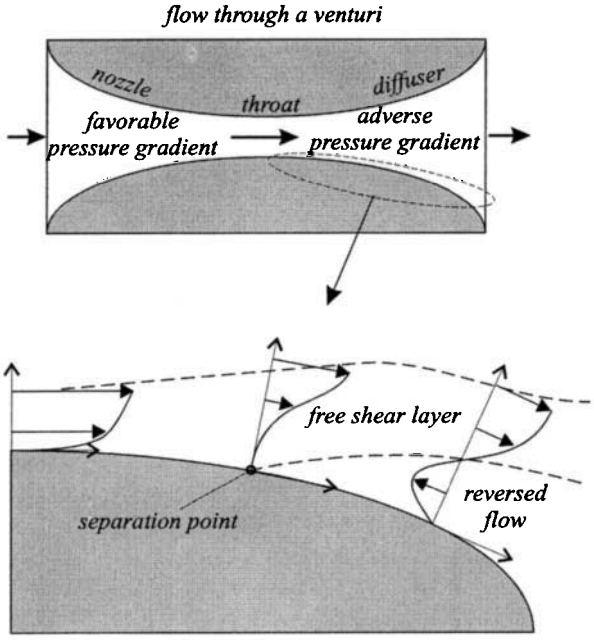


FIGURE 10.19 Boundary layer separation.

The Laminar Separation Bubble

Figure 10.20 shows a type of transition, known as free shear layer transition, which is commonly found over the leading edge of many airfoils. Under certain conditions the laminar boundary layer will separate in the leading edge region. The resulting laminar flow free shear layer quickly undergoes transition which expands in a wedge-like shape. If the now turbulent wedge touches the airfoil surface it will reattach as a turbulent boundary layer and a laminar separation bubble will form. Ward (1963) gives an excellent review of work done on laminar separation bubbles and their effect on the stall characteristics of airfoils (see also Section 10.7).

The Boundary Layer Equations

The boundary layer equations are essentially the conservation equations of mass, momentum, and energy (Section 10.4), which have been simplified using an order of magnitude analysis. Prandtl's fundamental boundary layer assumption is that the layer is very thin in comparison to the characteristic length of the body over which it flows. Using this assumption, it can be shown that the x- and y-direction Navier–Stokes equations can be respectively reduced to the following forms:

$$\frac{\partial(\rho uu)}{\partial x} + \frac{\partial(\rho uv)}{\partial y} = -\frac{\partial p}{\partial x} + \mu \frac{\partial^2 u}{\partial y^2} \quad \text{and} \quad \frac{\partial p}{\partial y} = 0$$

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

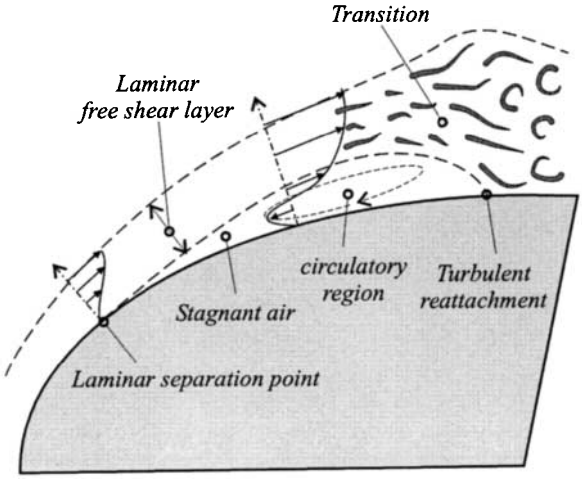


FIGURE 10.20 Flow phenomena associated with a laminar separation bubble.

When the flow is turbulent, steady, and incompressible the instantaneous variables (i.e, $u = U + u'$, $v = V + v'$ and $p = P + p'$) are substituted into the above equations and then time averaged. This gives the Reynolds averaged boundary layer equations in the form

$$\frac{\partial(\rho UU)}{\partial x} + \frac{\partial(\rho UV)}{\partial y} = -\frac{\partial P}{\partial x} + \mu \frac{\partial^2 U}{\partial y^2} + \frac{\partial(\tau'_{xy}/\rho)}{\partial y}$$

and

$$\frac{\partial P}{\partial y} = 0$$

where $\tau'_{xy} = -\rho\overline{u'v'}$ is known as a Reynolds stress. The overbar stands for the time averaged value and is given by

$$\overline{u'v'} = \lim_{T\to\infty} \frac{1}{T} \int_0^T u'v'dt$$

In terms of the eddy viscosity the x-direction turbulent boundary layer equation becomes

$$\frac{\partial(\rho UU)}{\partial x} + \frac{\partial(\rho UV)}{\partial y} = -\frac{\partial P}{\partial x} + (\mu + \mu_t) \frac{\partial^2 U}{\partial y^2}$$

where

$$\tau'_{xy} = -\rho\overline{u'v'} = \mu_t \left(\frac{\partial U}{\partial y}\right)$$

The subject area of boundary layer flows has been well researched and documented. The definitive work which compiles a great deal of this work is Schlichting (1991).

10.7 INCOMPRESSIBLE FLOW OVER AIRFOILS

Lift Generation in Subsonic Flow

Numerous experimental investigations of the flow around an airfoil have indicated four fundamental facts:

1. The presence of the airfoil produces large amounts of streamline curvature as the air is forced to travel around the airfoil profile.
2. When referenced to the free-stream static pressure, p_∞, Figure 10.21 shows that regions of positive $(p - p_\infty)$ and negative $(p - p_\infty)$ are found to act on the airfoil surface. For positive angles of attack, regions of negative $(p - p_\infty)$ are found to exist over the upper surface of the airfoil. Lift is generated from the net resultant force exerted by the distribution of static pressure around the airfoil surface.

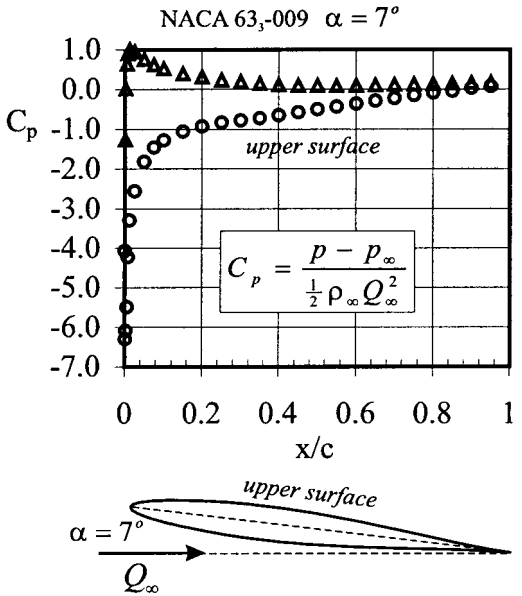


FIGURE 10.21 Typical airfoil pressure distribution (data from McCullough and Gault 1951, courtesy of NASA).

3. These regions of high and low relative pressure are respectively accompanied by pockets of negative and positive $(q - Q_\infty)$ where q is the local velocity and Q_∞ is the free-stream value. Conservation of energy states that these local velocity values can be related to the local static pressure via Bernoulli's equation.
4. Both the upper and lower surface air streams flow smoothly off the trailing edge. This phenomenon is referred to as the Kutta condition. This condition is due to the inability of a viscous fluid to negotiate tight corners due to the presence of extremely high viscous shear stresses. All inviscid mathematical models of lift require the use of the Kutta condition to introduce the correct amount of circulation into the flow field.

Consider an air particle traveling alone a curved streamline. Newton's second law states that the net force acting normal to the streamline must be proportional to the centripetal acceleration of a fluid particle. Assuming an inviscid fluid, without any body forces, the centripetal acceleration is balanced by a radial pressure gradient in the following manner (often referred to as the Euler-n equation);

$$\left(\frac{\partial p}{\partial r}\right)_R = \frac{\rho q^2}{R}$$

Japikse and Baines (1994) put forward a compelling argument which utilizes the Euler-n equation to explain the origin of subsonic lift. This argument is now elaborated upon below. Figure 10.22 displays the steady flow streamline pattern as viewed by an observer traveling with the airfoil. The Euler-n equation implies that the airfoil induces a static pressure distribution by forcing the air to follow a curved path. The airflow farther away from the airfoil has less streamline curvature and, thus, the static pressure relaxes back to the free stream value. The local velocity distribution, around the airfoil profile, is created as the air accelerates and decel-

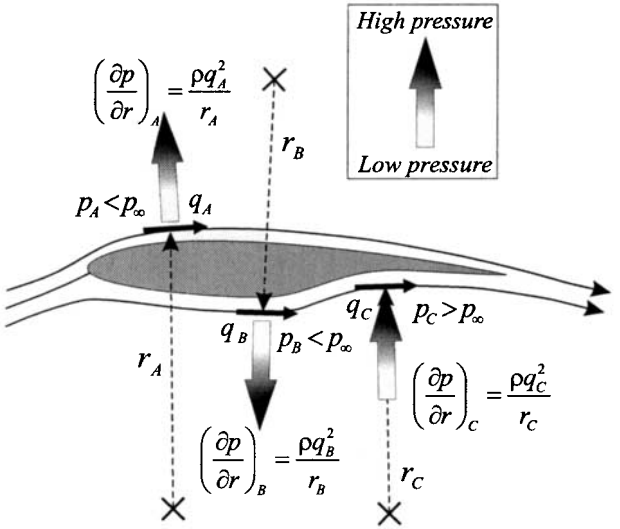


FIGURE 10.22 Subsonic flow around an airfoil.

erates from regions of differing static pressure. The Kutta condition forces the flow to adopt one particular pattern of streamline curvature and thus controls the lift produced by the airfoil. If the boundary layer separates (Subsection 10.6) from the airfoil contour, the Kutta condition is removed from the trailing edge and the amount of lift is reduced. When the stream function is utilized to visualize the incompressible flow around an airfoil, the regions of high velocity are indicated by stream tubes which contract such that the conservation of mass is obeyed.

An Overview of Mathematical Models of Lift

Nearly all mathematical models of lift use techniques which predict the velocity field around the airfoil first and then use Bernoulli's equation to obtain the pressure field. To obtain the correct velocity and pressure fields, the Kutta condition is applied in a form appropriate to the mathematical model (examples are given later). The Kutta–Joukowski lift theorem (see Subsection 10.5) is then used, in conjunction with the circulation, which accompanies the velocity field, to obtain the lift produced.

Figure 10.23 illustrates the essence of the circulation theory of lift, which can be summarized into the following steps:

1. Restrict the analysis to inviscid, incompressible, irrotational flow over an arbitrary airfoil in a free stream flow.
2. Replace the airfoil surface with a vortex sheet of unknown variable strength. It can be argued that in a viscous flow the boundary layer (Subsection 10.6) pro-

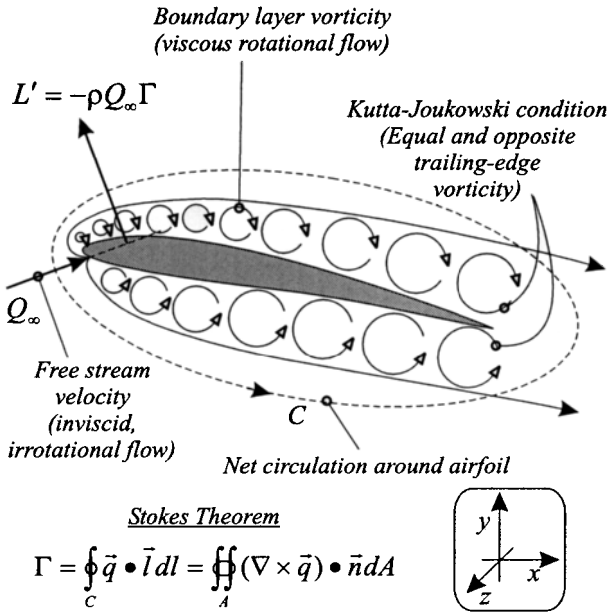


FIGURE 10.23 The circulation theory of incompressible lift.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

duces the vorticity, and this can be related to the total circulation around the airfoil via Stokes theorem (Subsection 10.4).

3. Calculate the variation of vortex sheet strength subject to the following boundary condition: (i) when the vortex sheet velocity field is added to the free-stream velocity there is no normal component of velocity at every point on the airfoil surface; and (ii) that the suction and pressure surface vortex sheet strengths at the trailing edge are equal (the Kutta condition).
4. The total circulation is the net value due to the entire vortex sheet and the resulting lift per unit span is calculated using the Kuttal–Joukowski theorem.

Small-Disturbance Airfoil Theory

Using the principle of superposition (see Subsection 10.5), the velocity distribution around an airfoil can be decomposed into the following three separate and independent components:

1. The distribution due to the basic airfoil thickness form at zero lift
2. The distribution due to the camber line at the ideal angle of attack (explained below)
3. The distribution due to angle of attack

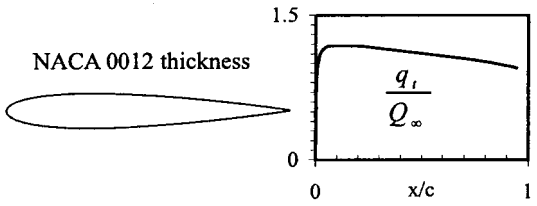
When the first two components are added together this is known as the basic velocity distribution. It is only dependent on the geometric properties of the airfoil. The third component is known as the additional velocity distribution and is strongly dependent on angle of attack and weakly dependent on airfoil thickness. Referring to Figure 10.24, the following points should be noted:

1. Velocity component one is denoted q_t/Q_∞ and can be found using either conformal transformation or singularity methods.
2. Velocity component two is denoted $\Delta q_c/Q_\infty$ and is obtained using thin airfoil theory. The ideal angle of attack, α_i, is defined as the angle which places a stagnation point exactly on the foremost position of the camber line. Theodorsen (1931) referred to this as ''the angle of best streamlining.'' The lift coefficient which corresponds to the ideal angle of attack is called the design lift coefficient and is denoted $c_{l,i}$.
3. Velocity component three is denoted $\Delta q_\alpha/Q_\infty$ and is obtained using conformal transformation methods and includes the influence of airfoil thickness on the angle of attack term.

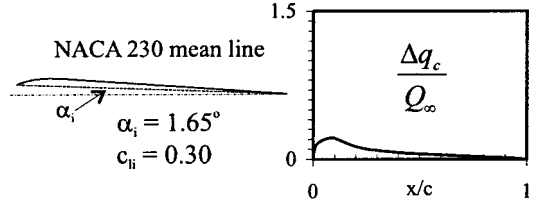
The final velocity distribution is given by

$$\frac{q}{Q_\infty} = \frac{q_t}{Q_\infty} \pm \frac{\Delta q_c}{Q_\infty} \pm \frac{\Delta q_\alpha}{Q_\infty}$$

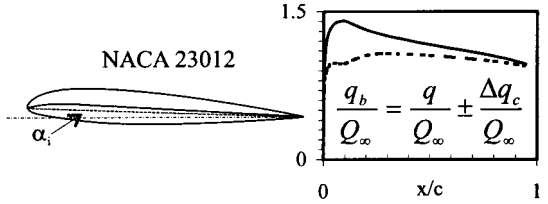
where the plus stands for the upper surface and the negative the lower surface. Since each velocity distribution is essentially a solution to the Laplace equation, the values of $\Delta q_\alpha/q_\infty$ and $\Delta q_c/Q_\infty$ scale linearly with angle of attack and airfoil geometry respectively. For example, if the camber line ordinates are multiplied by a constant factor, the velocity distribution, ideal angle of attack and design lift



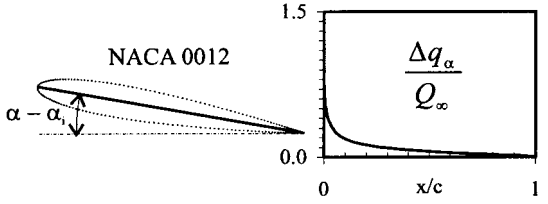
(a) Symmetrical airfoil at zero angle of attack



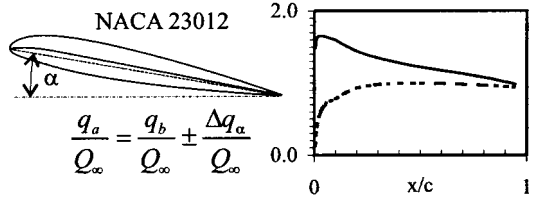
(b) Camber line at ideal angle of attack



(c) The basic velocity distribution



(d) The additional velocity distribution



(e) The final velocity distribution

FIGURE 10.24 The velocity superposition principle used by NACA.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

coefficient all change by the same factor. Thus, Abbott and von Doenhoff (1959) tabulate the following values for NACA airfoils:

$$\frac{q_t}{Q_\infty}, \quad \frac{\Delta q_\alpha / Q_\infty}{(c_{.} - c_{l,t})} \quad \text{and} \quad \frac{\Delta q_c / Q_\infty}{c_l}$$

As an example, consider the NACA 23012 airfoil, which has the following design parameters:

$$c_{l,i} = 0.3, \quad \alpha_i = 1.65^\theta \quad \text{and} \quad c_{mi,/4} = -0.014$$

Using the tabulated data, given in Abbott and von Doenhoff (1959), the following velocity components were obtained at the point $x/c = 0.5$ and for a lift coefficient, c_l, equal to 0.5:

$$\frac{q_t}{Q_\infty} = 1.108, \quad \frac{\Delta q_\alpha}{Q_\infty} = 0.149(0.5 - 0.3), \quad \text{and} \quad \frac{\Delta q_c}{Q_\infty} = 0.054(0.5)$$

The quarter-chord pitching moment and the zero-lift angle are respectively given by

$$c_{m,c/4} = -0.014(c_l) \quad \text{and} \quad \alpha_{l0} = \alpha_i - \frac{57.3}{2\pi} c_{l,i}$$

Several more examples are given for other NACA airfoils in Abbott et al. (1945).

Conformal Transformation

(*Applications—inviscid, incompressible, irrotational flow around a family of airfoils of arbitrary thickness and camber known as Joukowski airfoils.*) Although the following analysis is applicable to a particular family of airfoils, the results obtained clearly demonstrate the following important characteristics which are common to most airfoil profiles: the influence of profile thickness on the lift curve slope is indicated along with the effect of camber on the zero-lift angle and the zero-lift pitching moment. Also, conformal transformation has played an important role in the development of the NACA family of airfoils (Theodorsen 1931).

As discussed in Subsection 10.5, the flow around a lifting cylinder can be obtained using conformal transformation techniques. The Joukowski transformation treats the lifting cylinder flow as an intermediate mapping, which is then subjected to a further transformation to obtain the flow around a cambered finite thickness airfoil. The transformation which maps a circle on the z_n-plane, of radius r_0 and center z_0, onto an airfoil on the z-plane is given by

$$z = z_n + \frac{b^2}{z_n}$$

where b is approximately equal to a quarter of the final airfoil chord and the values of z_0 and r_0 (in relation to b) control the final airfoil profile on the z-plane. Theodorsen (1931) and Theodorsen and Garrick (1932) recognized that the inverse transformation $z_n = \frac{1}{2}(z \pm \sqrt{z^2 - 4b^2})$ could be applied to an arbitrary airfoil to produce a near circle in the z_n-plane. The flow about this near circle was then related to the flow about a true circle and hence the velocity distribution around the airfoil surface was obtained.

Figure 10.25 illustrates the transformations involved in forming the lifting flow around a cambered Joukowski airfoil at an arbitrary angle of attack. The following three transformations are used:

$$w = -Q_\infty \left(z_1 + \frac{r_0^2}{z_1} \right) - \frac{i\Gamma}{2\pi} \ln \frac{z_1}{r_0}$$

$$z_n = z_1 e^{-i\alpha} + m e^{i\delta}$$

$$z = z_n + \frac{b^2}{z_n}$$

A Joukowski airfoil profile, of specified thickness and camber, may be constructed using the following equations:

$$\hat{y}_{u,l} = \hat{y}_c \pm \Delta\hat{y}$$

where $$\hat{y}_c = \hat{h}(1 - 4\hat{x}^2),$$

and $$\Delta\hat{y} = \frac{2i}{3\sqrt{3}} (1 - 2\hat{x}) \sqrt{1 - 4\hat{x}^2}$$

The maximum thickness occurs at the quarter-chord point, while the maximum camber occurs at the mid-chord position. It should be noted that $\Delta\hat{y}$ is an additional distance added to the local camber line y-coordinate in a direction perpendicular to the chord line rather than the camber line as in the generation of the NACA airfoils (see Subsection 10.2). Also, an x–y coordinate system has been used here rather than the x–z coordinate system normally used for airfoil definition and analysis, to avoid confusion with the complex number z.

Referring to Figure 10.25, the complex velocity in the z-plane is given by

$$u - iv = \frac{dw}{dz} = \frac{dw}{dz_1} \frac{dz_1}{dz_n} \frac{dz_n}{dz} = \left\{ -Q_\infty \left(1 - \frac{r_0^2}{z_1^2} \right) - \frac{i\Gamma}{2\pi z_1} \right\} \{ e^{-i\alpha} \} \left\{ \left(1 - \frac{b^2}{z_n^2} \right)^{-1} \right\}$$

The circulation required to place the rear stagnation point on the airfoil trailing edge (i.e., the Kutta condition) is given when $(dw/dz) = 0$ and $z_1 = r_0 e^{-i(\alpha+\beta)}$, which results in

$$\Gamma_k = 4\pi r_0 Q_\infty \sin(\alpha + 2\hat{h}) \quad \text{(positive anticlockwise)}$$

Geometrical considerations result in

$$r_0^2 = c^2 \left(\frac{1}{4} + \frac{\hat{t}}{3\sqrt{3}} \right)^2 + \hat{h}^2$$

Using the Kutta–Joukowski lift theorem, the airfoil lift coefficient is given by

$$c_l = 2\pi \left(1 + \frac{4}{3\sqrt{3}} \hat{t} \right) \sin(\alpha + 2\hat{h}) \approx c_{l\alpha}(\alpha - \alpha_{l0}) \tag{10.9}$$

where $c_{l\alpha} \approx 2\pi$ and $\alpha_{l0} = -2\hat{h}$.

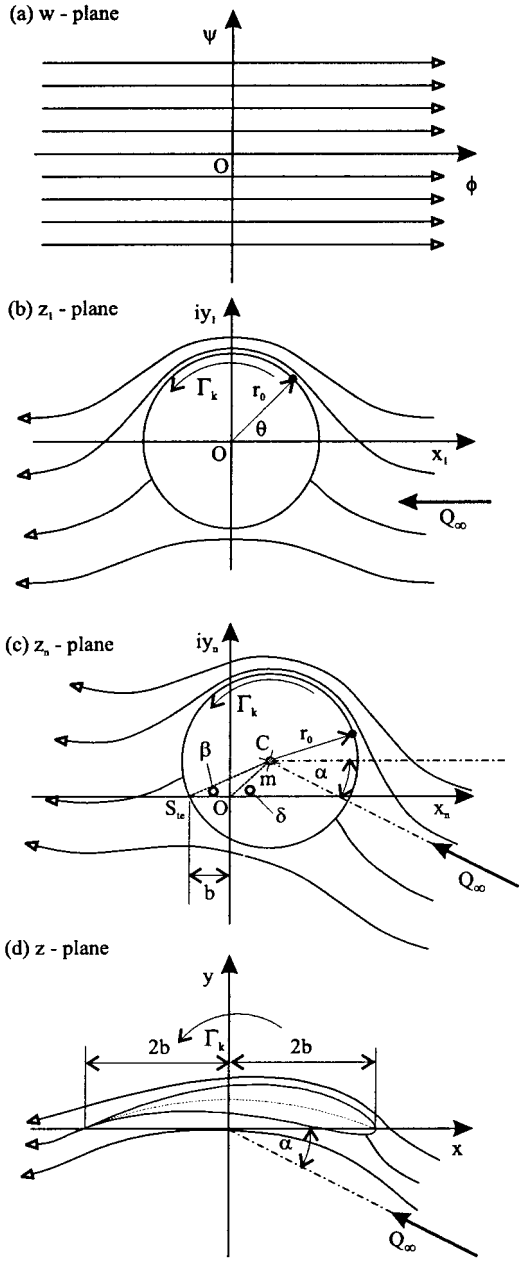


FIGURE 10.25 The Kutta–Joukowski transformation.

The pitching moment around the leading edge has the form

$$c_{m,le} = -\pi(\hat{h}) - \frac{c_l}{4} = c_{m0} - \frac{c_l}{4}$$

The pitching moment around the quarter-chord position is given by

$$c_{m,c/4} = c_{m,le} + \frac{c_l}{4} = -\pi\hat{h} \tag{10.10}$$

which means that, by definition, the quarter chord locates the aerodynamic center. The variation of the center of pressure with lift coefficient can be obtained from

$$\frac{x_{cp}}{c} \approx -\frac{c_{m,le}}{c_l} = \frac{\pi\hat{h}}{c_l} + \frac{1}{4}$$

and is plotted in Figure 10.26 for a 4% cambered Joukowski airfoil.

If required, the local velocity (and hence the local pressure coefficient) at any point in the z-plane can be calculated using the complex velocity. Figure 10.27 illustrates the pressure distribution around a Joukowski airfoil at the angle of zero lift. Although this distribution produces zero life a negative pitching moment is still induced. This explains why the center of pressure, shown in Figure 10.26 tends to infinity as the lift tends to zero.

Singularity Methods (Teardrop Theory)

(*Applications—inviscid, incompressible, irrotational nonlifting flow around a symmetrical airfoil.*) Figure 10.28 shows that the effect of airfoil thickness can be modeled using a distribution of source strength along the x-axis. Applying continuity to the control volume $ABCD$ and linearizing the result gives the source strength per unit length as

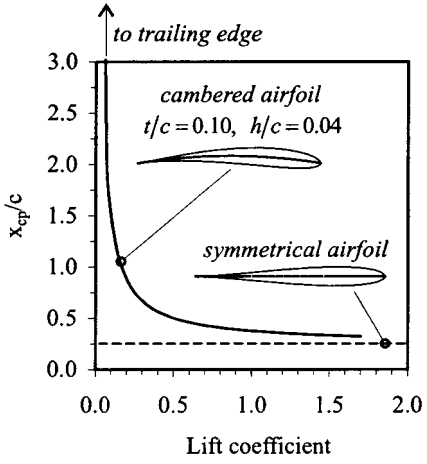


FIGURE 10.26 Variation of center of pressure with lift.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

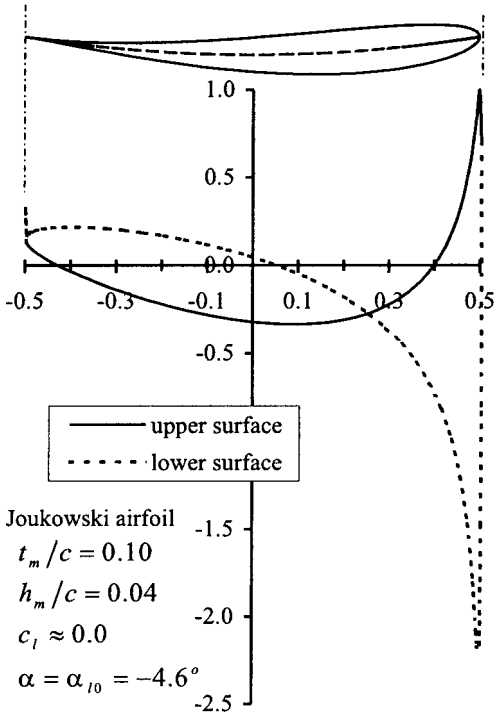


FIGURE 10.27 Zero-lift c_p distribution.

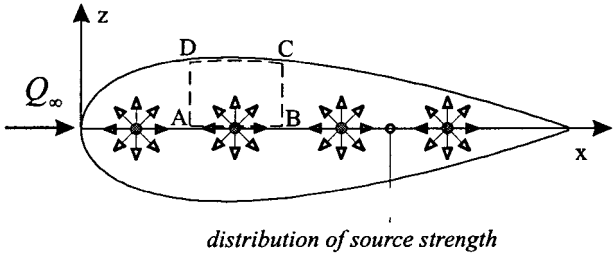


FIGURE 10.28 Thickness modeled using source distribution.

$$\xi(\hat{x}) = 2Q_\infty \frac{dz_u(\hat{x})}{d\hat{x}} \quad \text{where} \quad \hat{x} = x/c.$$

For thin airfoils, the disturbance velocity components on the airfoil surface are approximately equal to those on the x-axis, which gives

$$\tilde{u}(\hat{x}) = \frac{1}{2\pi}\int_0^1 \frac{\xi(\hat{x}_o)d\hat{x}_o}{(\hat{x} - \hat{x}_o)} \quad \text{and} \quad \tilde{w}(\hat{x}) = \pm \tfrac{1}{2}\xi(\hat{x})$$

The local surface velocity can now be obtained from

$$\frac{q(\hat{x})}{Q_\infty} \approx 1 + \frac{\tilde{u}(\hat{x})}{Q_\infty} = 1 + \frac{1}{\pi}\int_0^1 \frac{\xi(\hat{x}_0)d\hat{x}_0}{(\hat{x} - \hat{x}_0)}$$

This equation can be solved by first expressing the airfoil upper surface profile by a Fourier series

$$z_u(\theta) = \tfrac{1}{2}\sum_{n=1}^{N} b_n \sin n\theta \quad \text{where} \quad \hat{x} = \tfrac{1}{2}(1 + \cos\theta)$$

which results in

$$\frac{q(\hat{x})}{Q_\infty} = \frac{1}{r_i(\theta)}\left[1 + \sum_{n=1}^{N}\frac{nb_n \sin n\theta}{\sin\theta}\right]$$

where $r_i(\theta)$ is known as the Riegels factor and is given by

$$r_i(\theta) = \sqrt{1 + \frac{1}{\sin^2\theta}\left(\frac{dz_u(\theta)}{d\theta}\right)^2}$$

This procedure can be applied to any symmetric airfoil. For example, take a symmetric Joukowski airfoil whose surface contour is defined by

$$z_u(\theta) = (\tfrac{1}{2})\varepsilon \sin\theta(1 - \cos\theta)$$

where

$$\varepsilon = \frac{4\hat{t}}{3\sqrt{3}}$$

which gives the local surface velocity at any point as

$$\frac{q(\hat{x})}{Q_\infty} = \frac{1 + \varepsilon(1 - 2\cos\theta)}{\sqrt{1 + \varepsilon^2\left(\dfrac{\cos\theta - \cos 2\theta}{\sin\theta}\right)^2}}$$

Figure 10.29 displays the surface velocity distribution for various thickness ratios (in steps of 0.05). As discussed in Subsection 10.10, the maximum surface velocity is highly important when considering the behavior of an airfoil at high subsonic velocities.

Thin Airfoil Theory

(*Applications—inviscid, incompressible, irrotational flow around airfoils of thickness less than 12% chord and camber less than 2% chord.*) Figure 10.30 shows that to make the camber line of a thin airfoil a streamline of the flow, the upwash, at all points along the camber line, due to the angle of attack has to be equal to the downwash induced by the entire vortex sheet. This criterion gives the fundamental equation of thin airfoil theory:

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

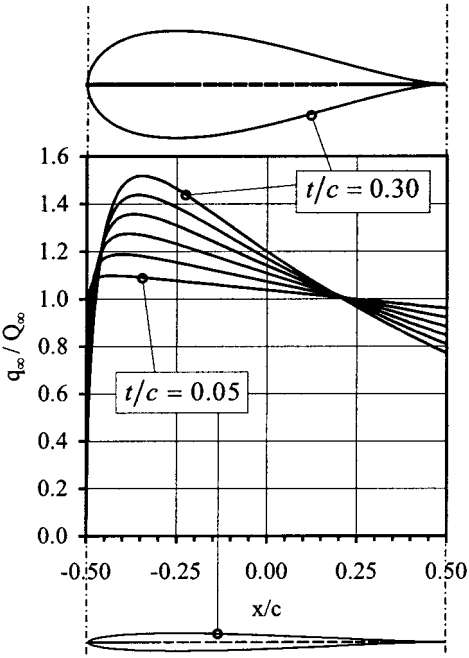


FIGURE 10.29 Effect of thickness on velocity distribution (at zero lift).

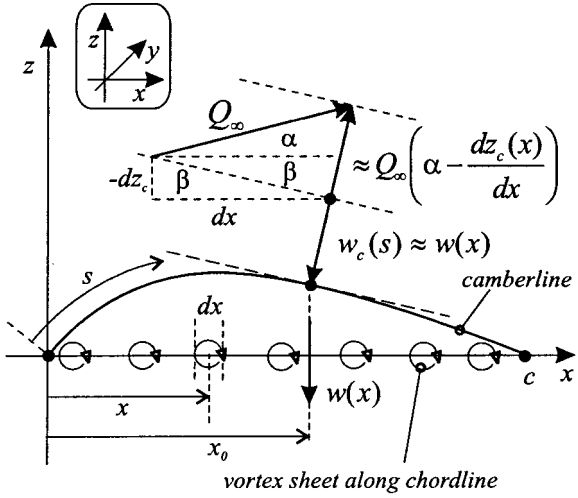


FIGURE 10.30 Thin airfoil mathematical model.

$$\int_0^c \frac{\gamma(x)dx}{2\pi(x_0 - x)} = Q_\infty \left(\alpha - \frac{dz_c(x_0)}{dx}\right) \tag{10.11}$$

To solve equation (10.11), the coordinate transformation $x = (\frac{1}{2})c(1 - \cos\theta)$ is first used to give the following form of the fundamental equation:

$$\int_0^c \frac{\gamma(\theta)\sin\theta d\theta}{2\pi(\cos\theta - \cos\theta_0)} = Q_\infty \left(\alpha - \frac{dz_c(\theta_0)}{dx}\right)$$

The solution to this equation can be expressed in the form

$$\gamma(\theta) = 2Q_\infty \left(A_0 \frac{1 + \cos\theta}{\sin\theta}\right) + 2Q_\infty \sum_{n=1}^{\infty} A_n \sin n\theta \tag{10.12}$$

The fundamental form of this vorticity distribution can be obtained using conformal transformation methods applied to flat plate and circular arc airfoils (Houghton and Brock 1970). It should be noted that equation (10.12) satisfies the Kutta condition of smooth trailing-edge flow since the vortex sheet strength tends to zero when $\theta = 180°$. Using this vorticity distribution reduces equation (10.11) into the form

$$\frac{dz_c(\theta_0)}{dx} = (\alpha - A_0) + \sum_{n=1}^{\infty} A_n \cos n\theta_0$$

Since this is a Fourier series, the coefficients are now given by

$$A_0 = \alpha - \frac{1}{\pi}\int_0^\pi \frac{dz_c(\theta_0)}{dx} d\theta_0 = \alpha - \alpha_i$$

$$A_n = \frac{2}{\pi}\int_0^\pi \frac{dz_c(\theta_0)}{dx} \cos n\theta_0 d\theta_0$$

For any given airfoil camber line, the A-coefficients can be obtained using either analytical or numerical integration. The value α_i is known as the ideal angle of attack. When the free-stream velocity is set at the ideal angle of attack, $A_0 = 0$ and the infinite leading edge vorticity (and hence infinite local velocity), inherent in equation (10.12), is avoided.

The total circulation around the airfoil is given by

$$\Gamma = \int_0^c \gamma(x)dx = \int_0^\pi \gamma(\theta)\tfrac{1}{2}c\sin\theta d\theta = \pi c Q_\infty(A_0 + \tfrac{1}{2}A_1)$$

Using the Kutta–Joukowski theorem gives the lift coefficient, and the lift curve slope respectively as (α in radians):

$$c_l = 2\pi(A_0 + \tfrac{1}{2}A_1) = c_{l\alpha}(\alpha - \alpha_{l0}) \tag{10.13}$$

where

$$c_{l\alpha} = 2\pi \quad \text{and} \quad \alpha_{l0} = -\frac{1}{\pi}\int_0^{\pi} \frac{dz_c(\theta_0)}{dx}(\cos\,\theta_0 - 1)d\,\theta_0$$

The lift coefficient corresponding to the ideal angle of attack is known as the ideal or design lift coefficient and is given by

$$c_{l,i} = 2\pi(\tfrac{1}{2}A_1) = 2\pi(\alpha_i - \alpha_{l0}) \tag{10.14}$$

Pitching moment characteristics can be obtained from

$$c_{m,le} = -\tfrac{1}{2}\pi(A_0 + A_1 - \tfrac{1}{2}A_2)$$

$$c_{m,c/4} = \tfrac{1}{4}\pi(A_2 - A_1)$$

Since A_1 and A_2 are independent of α, the quarter-chord point locates the aerodynamic center (see Subsection 10.2) for all thin airfoils. The center of pressure is given by

$$x_{cp} \approx -\frac{c_{m,le}c}{c_l} = \frac{c}{4}\left[1 + \frac{\pi}{c_l}(A_1 - A_2)\right]$$

All thin symmetric airfoils thus have the center of pressure (Subsection 10.2) at the quarter chord point since $A_1 = A_2 = 0$ for this type of airfoil.

Information regarding the pressure difference between the upper and lower surface of a thin airfoil can be obtained by considering Bernoulli's equation with a velocity perturbation along an arbitrary stream line.

$$p_\infty + \tfrac{1}{2}\rho_\infty Q_\infty^2 = p + \tfrac{1}{2}\rho_\infty(Q_\infty + \tilde{u})^2 \Rightarrow p \approx p_\infty - \rho_\infty Q_\infty \tilde{u}$$

The pressure difference across the thin airfoil is thus given by

$$\Delta p(\theta) = (p_l - p_u) = \rho_\infty Q_\infty(\tilde{u}_u - \tilde{u}_l) = \rho_\infty Q_\infty \gamma(\theta)$$

Figure 10.31 illustrates the pressure difference across a symmetrical thin airfoil, which can be written in the form

$$\frac{\Delta p}{\tfrac{1}{2}\rho_\infty Q_\infty^2 \alpha} = 4\left(\frac{1 + \cos\,\theta}{\sin\,\theta}\right) = \sqrt{\frac{1 - x/c}{x/c}} \tag{10.15}$$

where α is in radians.

The Lumped Vortex Model and the Rear Aerodynamic Center

Figure 10.32 shows that, from a far field point of view, the continuous vortex sheet, used in thin airfoil theory, can be replaced by a single vortex placed at the one quarter-chord position. This single bound vortex requires only one control point where the flow tangency boundary condition is satisfied. Denoting this point by *kc* gives

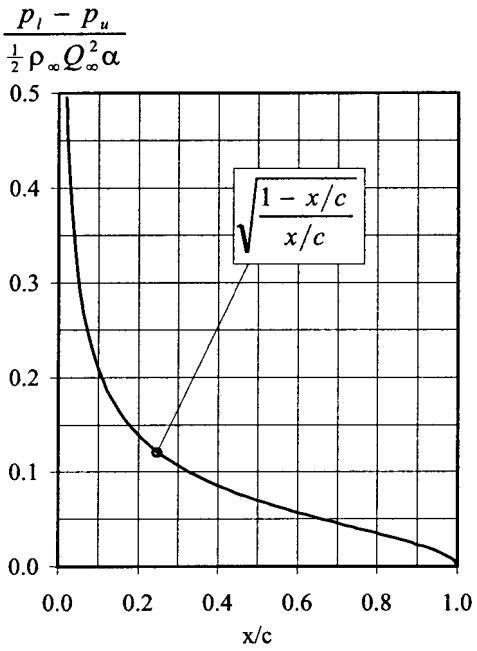


FIGURE 10.31 Pressure difference across a symmetrical airfoil.

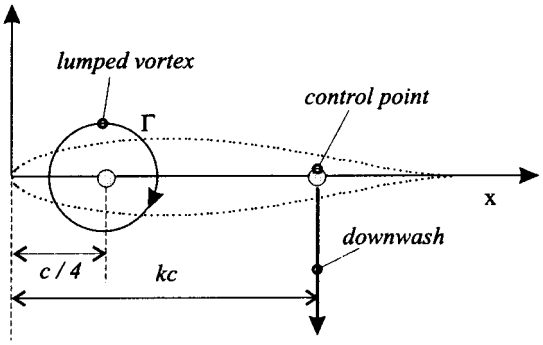


FIGURE 10.32 The lumped vortex model.

$$\frac{-\Gamma}{2\pi(kc - \frac{1}{4}c)} + Q_\infty \alpha = 0$$

The value kc can be calculated for any cambered airfoil by using the thin airfoil result for the total circulation. For simplicity, a symmetrical airfoil is used, which gives

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

$$\Gamma = \pi c Q_\infty \alpha \quad \text{and hence} \quad k = \tfrac{3}{4}$$

Katz and Plotkin (1991) demonstrate that the lumped vortex model is an excellent method for estimating the effect on the lift due to close proximity of other airfoils (e.g., biplanes) or solid surfaces (e.g., airfoils in ground effect).

For an airfoil of zero thickness, thin airfoil theory shows that the slope of the camber line at the three quarter chord point is equal to the zero-lift angle (in radians). For example, the camber line of a circular arc airfoil can be represented by the equation $z_c = h \sin^2 \theta$, which results in the following:

$$\alpha_{l0} = -2\hat{h} = (dz_c/dx)_{3c/4}$$

When a similar calculation is carried out on the 12% thick NACA 23012, the following results are obtained:

$$\alpha_{l0} = -1.09^\circ \quad \text{and} \quad (dz_c/dx) = -1.26^\circ$$

Thus, to a first approximation, for any thin airfoil

$$\alpha_{l0} \approx \left(\frac{dz_c}{dx}\right)_{3c/4}$$

Because of the two important properties described above, the three quarter chord point is often referred to as the rear aerodynamic center. As discussed in Subsection 10.8, the rear aerodynamic center is utilized in certain finite wing models to aid in the prediction of spanwise lift distributions over swept wings.

Vortex Panel Methods

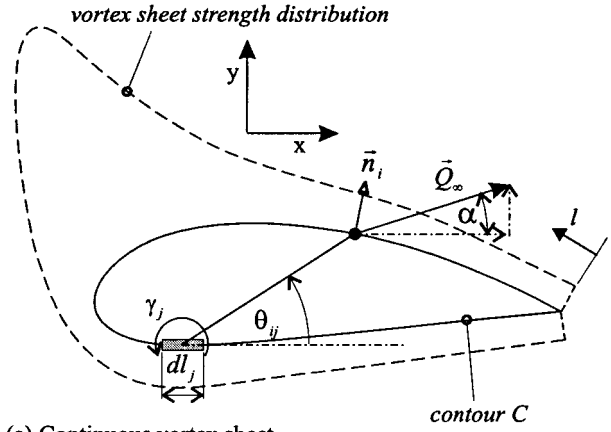
(*Applications—inviscid, incompressible, irrotational flow around airfoils of arbitrary thickness and camber.*) Figure 10.23 illustrates that the boundary layer can be thought of as a vorticity layer wrapped around the airfoil contour. Figure 10.33(a) shows that the airfoil surface can be replaced by a vortex sheet (see also Subsection 10.5) whose variation in strength can be related to the free stream velocity by using the airfoil surface flow tangency boundary condition in the form

$$\vec{Q}_\infty \cdot \vec{n}_i + \frac{1}{2\pi} \oint_c \gamma_j \frac{\partial \theta_{ij}}{\partial n_i} dl_j = 0 \tag{10.16}$$

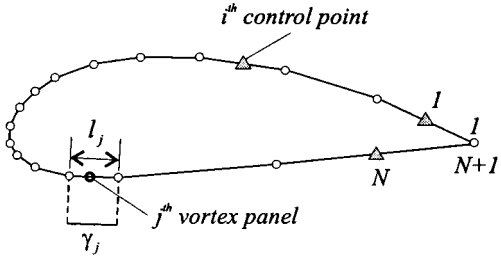
Figure 10.33(b) illustrates that this integral equation may be numerically solved by decomposing the airfoil surface into a number of vortex panels over which the vorticity is piecewise constant. This is known as a first order vortex panel method, and equation (10.16) takes the form

$$\vec{Q}_\infty \cdot \vec{n}_i + \sum_{j=1}^{N} \frac{\gamma_j}{2\pi} \int_{l_j} \frac{\partial \theta_{ij}}{\partial n_i} dl_j = 0 \tag{10.17}$$

Applying equation (10.17) to $N - 1$ panels produces $N - 1$ simultaneous equations for N unknown values of γ_j. The Nth equation is obtained by applying the Kutta–Joukowski in the form $\gamma_1 = \gamma_{N+1}$. The lift per unit span is then obtained using



(a) Continuous vortex sheet



(b) Vortex panels

FIGURE 10.33 The vortex panel method.

$$L' = \rho_\infty Q_\infty \sum_{j=1}^{N} \gamma_j l_j$$

For further information on vortex panel methods and other numerical models of aerodynamics Katz and Plotkin (1991) should be consulted.

Low-Speed Airfoil Stalling Characteristics

Airfoil stall is defined as the flow conditions which accompany the first peak in lift coefficient. Figure 10.34 illustrates the dominant boundary layer phenomena associated with the various types of subsonic airfoil stall. Figure 10.35 shows the variation in lift coefficient with angle of attack which accompany each stall type. Additional information regarding the various boundary layer phenomena can be found in Subsection 10.6. A short description of each stall is given below.

1. *Thin airfoil stall:* this stall is sometimes referred to as long bubble stall because it involves the formation of a laminar separation bubble which grows in length towards the trailing edge with increasing values of angle of attack. The small discontinuity in the lift curve, indicated in Figure 10.35, occurs when the bubble

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

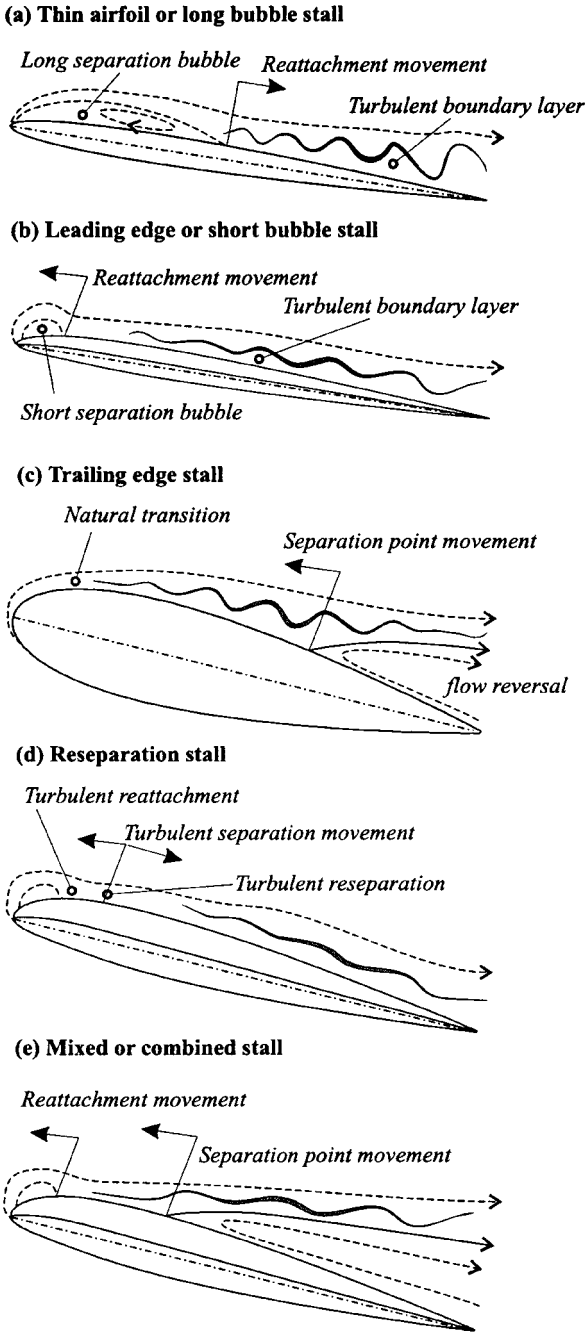


FIGURE 10.34 Types of airfoil stall.

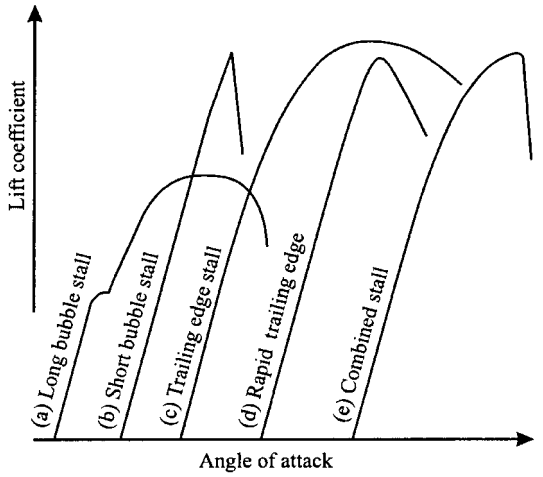


FIGURE 10.35 Types of subcritical airfoil stall.

reaches a size large enough to initiate a considerable reduction in the leading edge suction peak (see Figure 10.36(a)). As the bubble grows in length, the lift starts to gradually decrease. Maximum lift is low relative to other stall types and occurs when the bubble reattachment point reaches the trailing edge. Any further attempts to increase the angle of attack results in bubble thickening, followed by bursting, with a rapid loss of the remaining lift.

2. *Leading edge stall:* this stall is also known as short bubble stall because it involves the formation of a laminar separation bubble which decreases in length as the lift is increased. Eventually turbulent reattachment fails to take place and the bubble bursts resulting in a catastrophic loss of lift. Since the bubble, prior to bursting, is small, it has little effect on the leading edge suction peak (see Figure 10.36(b)), and thus the maximum lift is larger than that associated with a long bubble stall.
3. *Trailing edge stall:* this stall involves turbulent boundary layer separation which starts at the trailing edge and progresses towards the leading edge as the angle of attack is increased. The rate of forward movement of the separation point is dependent on the airfoil, Reynolds number, and angle of attack. This can occasionally result in a lift curve which would normally be attributed to a leading edge stall. Figure 10.36(c) shows that trailing edge separation induces a constant-pressure region to form over the rear of the airfoil, which causes a reduction in the leading edge suction peak.
4. *Reseparation stall:* this type of stall involves the sudden separation of the turbulent boundary layer just downstream of the reattachment point of a short laminar separation bubble. This behavior causes the bubble to burst. Evans and Mort (1959) have provided evidence which suggests that reseparation and failure to reattach after transition are different phenomena.
5. *Combined stall:* sometimes referred to as a mixed stall this type of stall can be thought of as a race between bubble bursting and trailing edge separation for the determination of maximum lift. For example, a noticeable amount of trailing

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

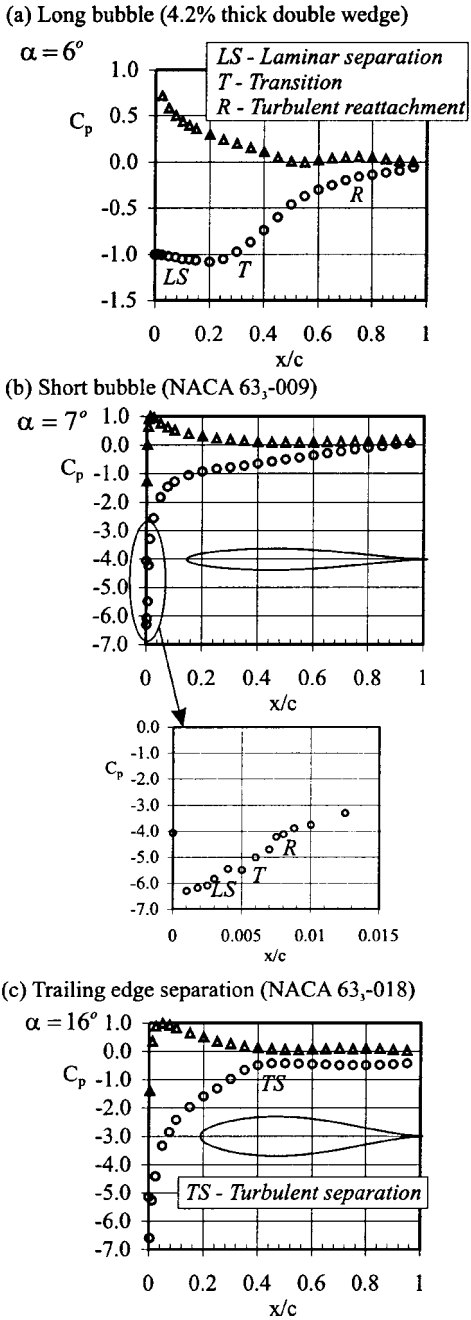


FIGURE 10.36 Typical pressure distributions associated with various boundary layer phenomena (data from McCullough and Gault 1951, courtesy of NASA).

edge separation may form prior to bubble bursting, resulting in a rounding of the lift curve preceding an abrupt loss of lift (Figure 10.35).

Gault (1957) studied the stall characteristics of 150 airfoils (obtained by numerous investigators) over a range of Reynolds number. He found a very useful correlation could be made between the type of stall and the upper surface ordinate at the 1.25% chord. Figure 10.37 displays this correlation, and Table 10.6 gives the relevant ordinate for most of the NACA series airfoils. The correlation is only strictly valid for airfoils with aerodynamically smooth surfaces, no high-lift, devices and tested in low-turbulence free stream flows.

Types of Incompressible Flow Drag over Airfoils

When the flow is two-dimensional and incompressible there are two types of drag:

1. *Pressure drag:* this rearward facing force, sometimes referred to as form drag, arises from boundary layer thickness and separation effects which do not allow the trailing edge pressures to recover fully to those found in the leading edge region.
2. *Skin friction drag:* as described in Subsection 10.6, skin friction drag arises from viscous shear stresses which act on the airfoil surface.

A boundary layer cannot exist if the entire flow field is deemed to be inviscid, and thus both the pressure and skin friction drag terms will be zero.

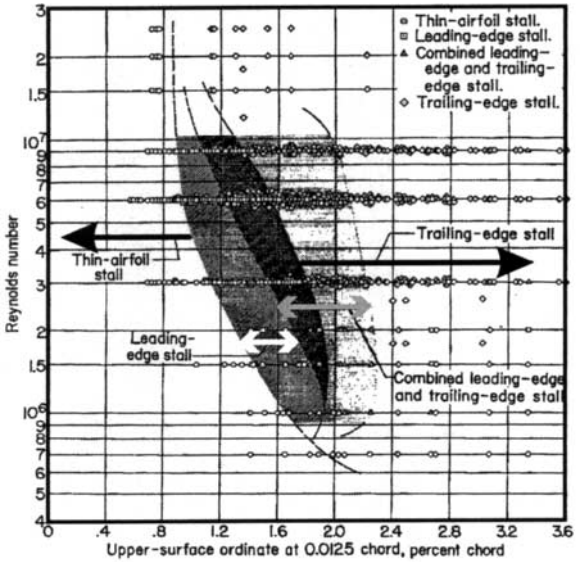


FIGURE 10.37 Gault's low-speed stall correlation (Gault 1957, courtesy of NASA).

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

TABLE 10.6 Upper Surface 1.25% Ordinate for various NACA Series Airfoils

NACA Series	Thickness ratio (% chord)									
	0	6	8	9	10	12	15	8	21	24
00xx	0.00	0.95	1.26	1.42	1.58	1.89	2.37	2.84	3.31	3.95
24xx	0.12	1.11	1.44	1.62	1.78	2.15	2.71	3.28	3.87	4.44
44xx	0.25	1.25		1.81		2.44	3.07	3.76	4.45	5.2
230xx	0.36	1.42		2.02		2.67	3.34	4.09	4.87	5.65
16-0xx	0.00	0.65		0.97		1.29	1.61	1.94	2.26	
63-0xx	0.00	0.77		1.15	1.27	1.52	1.88	2.22	2.53	
63-2xx	0.11	0.89		1.29	1.41	1.69	2.08	2.46	2.78	
63-4xx	0.21	1.02		1.44		1.85	2.30	2.76	3.10	
64-0xx	0.00	0.75	1.01	1.13	1.25	1.49	1.84	2.18	2.52	
64-2xx	0.11	0.88	1.14	1.26	1.40	1.65	2.03	2.49	2.77	
64-4xx	0.21	1.01		1.41		1.83	2.25	2.70	3.06	
65-0xx	0.00	0.72	0.95	1.06	1.17	1.39	1.70	2.01	2.30	
65-2xx	0.11	0.84		1.19	1.30	1.54	1.88	2.21	2.54	
65-4xx	0.21	0.97		1.32	1.47	1.70	2.07	2.43	2.78	
66-0xx	0.00	0.69	0.92	1.03	1.14	1.36	1.67	1.95	2.24	
66-2xx	0.11	0.81		1.16	1.28	1.51	1.85	2.15	2.46	
66-4xx	0.21	0.93		1.30		1.67	2.03	2.36	2.68	
63A0xx	0.00	0.75	1.00		1.25	1.49	1.84			
64A0xx	0.00	0.74	0.98		1.23	1.46	1.81			
65A0xx	0.00	0.72	0.95		1.18	1.41	1.75			

Data taken from Gault 1957 curtesy of NASA.

10.8 INCOMPRESSIBLE FLOW OVER FINITE WINGS

Prandtl's Lifting Line Model

(*Applications—inviscid, incompressible, irrotational, flow over unswept, tapered and twisted wings with aspect ratios greater than 3.*) From 1912 to 1918, Ludwig Prandtl and his colleagues in Germany developed an incompressible theory of finite wing aerodynamics which could be split into two parts; the study of two-dimensional flow around a wing section (an airfoil); and the modification of each span-wise airfoil flow to account for the three-dimensional flow which occurs over a finite wing. The strength of Prandtl's model lies in the fact that the airfoil characteristics can be obtained either from theory (see Subsection 10.7) or from wind tunnel testing.

Figure 10.38 shows that since the tip of a finite wing cannot sustain a differential pressure between the upper and lower surfaces, the lift, and hence the circulation, must reduce to zero. As the bound vortex reduces in strength, Helmholtz's laws (Subsection 10.5) state that the difference between the old and new circulation must be shed downstream as a trailed vortex filament. Thus, Prandtl's lifting line model replaces the wing with a bound vortex, located at the one-quarter chord position, and a wake consisting of an infinite number of trailed vortex filaments.

Using the Biot–Savart Law (see Subsection 10.5), Figure 10.38 illustrates that the downwash (velocity component in the z-direction), due to the entire trailed wake, at an arbitrary point along the bound vortex is given by

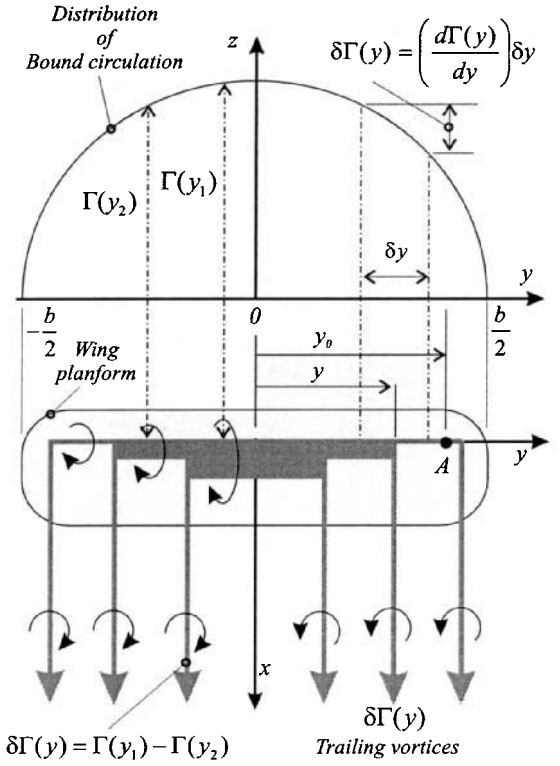


FIGURE 10.38 Prandtl's lifting model.

$$w(y_0) = -\frac{1}{4\pi}\int_{-b/2}^{b/2} \frac{\{d\Gamma(y)/dy\}dy}{y_0 - y} \tag{10.18}$$

Figure 10.39 shows that the downwash reduces the wing section angle of attack and cants the local lift vector rearward, which gives rise to a drag component known as the induced drag.

Using Figure 10.38 to define the relationship between the effective and geometric angles of attack and combining equation (10.18) with the Kutta–Joukowski lift theorem and the definition of the section lift coefficient results in the formulation of Prandtl's simple lifting line equation:

$$\alpha(y_0) = \frac{2\Gamma(y_0)}{Q_\infty c(y_0) c_{l\alpha}(y_0)} + \alpha_{l0}(y_0) + \frac{1}{4\pi Q_\infty}\int_{-b/2}^{b/2} \frac{(d\Gamma(y)/dy)dy}{y_0 - y} \tag{10.19}$$

To solve this equation, we first use the coordinate transformation $y = -\frac{1}{2}b\cos\theta$ along with the following general circulation distribution:

$$\Gamma(\theta) = \tfrac{1}{2}Q_\infty c_r c_{l\alpha,r}\sum_{n=1}^{N} A_n \sin n\theta \tag{10.20}$$

This procedure results in the following form of the Prandtl's lifting line equation:

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

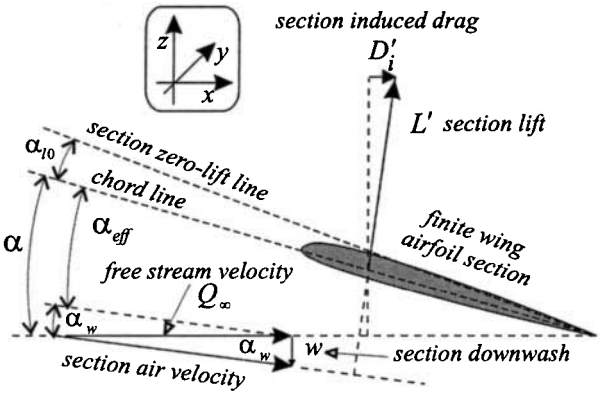


FIGURE 10.39 Effect of finite wing trailed wake.

$$\alpha(\theta) = \frac{c_r c_{l\alpha,r}}{c(\theta) c_{l\alpha}(\theta)} \sum_{n=1}^{N} A_n \sin n\theta + \alpha_{l0}(\theta) + \frac{c_r c_{l\alpha,r}}{4b} \sum_{n=1}^{N} \frac{n A_n \sin n\theta}{\sin \theta} \tag{10.21}$$

To numerically solve for the lift distribution over an arbitrary wing planform, equation (10.21) is applied at a chosen number of spanwise locations, i.e., at different values of θ. This will result in N equations in N unknown coefficients A_1, A_2, . . . , A_N. All coefficients will be involved when the lift distribution is asymmetric, while only the odd numbered coefficients will be required when the distribution is symmetrical. For symmetrically loaded rectangular wings it is normally only necessary to retain only the first three or four coefficients. Once the A-coefficients have been obtained the following quantities can be calculated:

1. Section induced angle of attack:

$$\alpha_w(\theta) = -\frac{w(\theta)}{Q_\infty} = \frac{c_r c_{l\alpha,r}}{4b} \sum_{n=1}^{N} \frac{a A_n \sin n\theta}{\sin \theta} \tag{10.22}$$

2. Section lift coefficient:

$$c_l(\theta) = \frac{L'(\theta)}{\frac{1}{2}\rho_\infty Q_\infty^2 c(\theta)} = \frac{c_r c_{l\alpha,r}}{c(\theta)} \sum_{n=1}^{N} A_n \sin n\theta \tag{10.23}$$

3. Lift per unit span:

$$L'(\theta) = \tfrac{1}{2}\rho_\infty Q_\infty^2 c(\theta) c_l(\theta) = \rho_\infty Q_\infty \Gamma(\theta) \tag{10.24}$$

4. Wing lift coefficient:

$$C_L = \frac{L}{\frac{1}{2}\rho_\infty Q_\infty^2 S} = \frac{2}{Q_\infty S} \int_{-b/2}^{b/2} \Gamma(y) dy = \frac{c_r c_{l\alpha,r}}{4b} \frac{b^2}{S} \pi A_1 = \frac{c_r c_{l\alpha,r}}{4b} A_R \pi A_1 \tag{10.25}$$

5. Induced drag coefficient:

$$C_{D,ti} = \frac{D_i}{\frac{1}{2}\rho_\infty Q_\infty^2 S} = \frac{2}{Q_\infty S}\int_{-b/2}^{b/2} \Gamma(y)\alpha_i(y)dy$$
$$= \frac{C_L^2}{\pi A_R}\left(1 + \sum_{n=2}^{N} \frac{aA_n^2}{A_1^2}\right) = \frac{C_L^2}{\pi A_R}(1 + \delta) \tag{10.26}$$

Figures 10.40 and 10.41 present spanwise lift distributions obtained for various tapered and twisted wings. These data were obtained by solving Prandtl's lifting

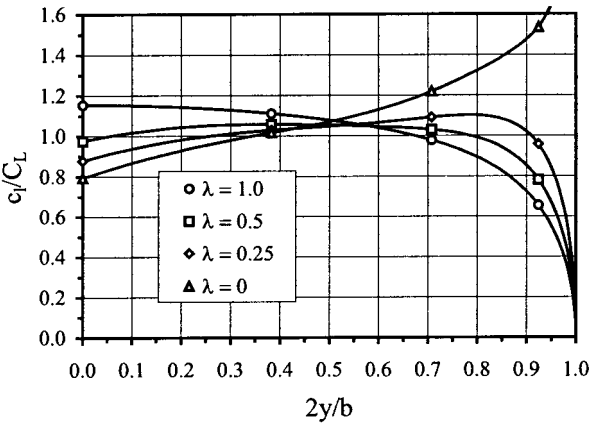


FIGURE 10.40 Effect of taper on spanwise lift distribution.

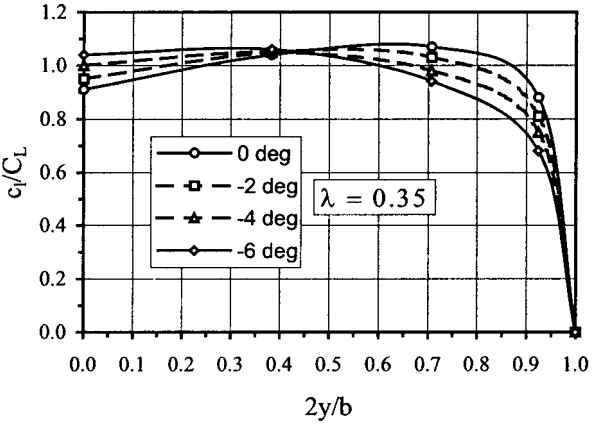


FIGURE 10.41 Effect of twist on the spanwise lift distribution of a tapered wing.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

line equation using the four spanwise locations, $\theta = \pi/2,\ \pi/8,\ \pi/2,\ 3\pi/8$, to solve for four symmetric A-coefficients.

Figure 10.42 presents the variation of δ with taper ratio for untwisted wings of various aspect ratio. It is worth noting that for an untwisted wing the induced drag has a minimum value at a taper ratio of around 0.35.

The Elliptical Wing

Rather than solving for the circulation distribution associated with a particular wing planform (the direct problem), an elliptical spanwise distribution of circulation can be assumed and the wing planform and characteristics then calculated (the indirect problem). An elliptical distribution of circulation is defined by

$$\Gamma(y) = \Gamma_r \sqrt{1 - \left(\frac{2y}{b}\right)^2} \quad \text{or} \quad \Gamma(\theta) = \tfrac{1}{2} Q_\infty c_r c_{l\alpha,r} A_1 \sin n\theta$$

Since only one A-coefficient has been stipulated the wing must have no aerodynamic twist. Substituting the elliptical value of A_1 into equations (10.22), (10.23), and (10.25) results in a constant spanwise induced angle of attack and a constant spanwise distribution of section lift coefficient,c_l, equal to that of the wing lift coefficient, C_L. Equation (10.24) thus states that the wing must have an untwisted elliptical spanwise variation in chord given by $c(\theta) = c_r \sin \theta$. Equation (10.26) gives $\delta = 0$, and thus an untwisted elliptical wing has the lowest induced drag coefficient.

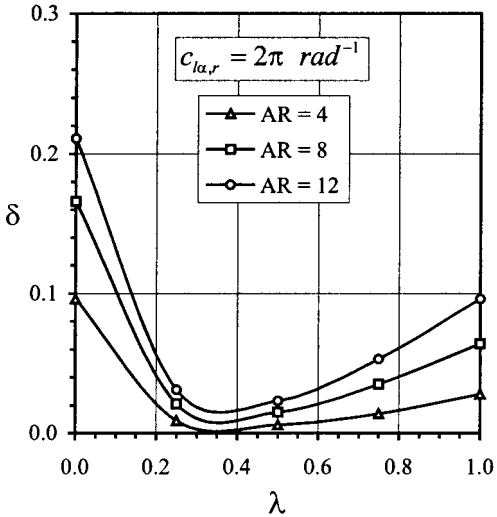


FIGURE 10.42 Variation of δ with taper ratio for an untwisted wing.

The Wing Lift Curve Slope

Consider a wing of arbitrary planform and aerodynamic twist. The wing life coefficient is given by

$$C_L = C_{l\alpha}(\alpha_r - \alpha_{L0})$$

and the root section lift coefficient is given by

$$c_{l,r} = c_{l\alpha,r}(\alpha_r - \alpha_{w,r} - \alpha_{l0,r})$$

Combining these two equations gives the following relationship:

$$C_{L\alpha} = \frac{c_{l\alpha,r}}{1 + \dfrac{c_{l\alpha,r}}{\pi A_R}(1 + \tau)} \tag{10.27}$$

where

$$(1 + \tau) = \frac{\pi A_R}{c_{l\alpha,r}}\left[\frac{c_{l,r}}{C_L} + \frac{c_{l\alpha,r}}{C_L}(\alpha_{w,r} + \alpha_{l0,r} - \alpha_{L0}) - 1\right]$$

For an untwisted, elliptic wing with constant airfoil section $c_{l\alpha,r} = c_{l\alpha}$ and $\tau = 0$ and equation (10.27) reduces to the form frequently used for incompressible flow.

Figure 10.43 displays the variations in τ with taper ratio for untwisted wings of various aspect ratio. As before, these data were obtained by solving Prandtl's lifting line equation using four spanwise locations.

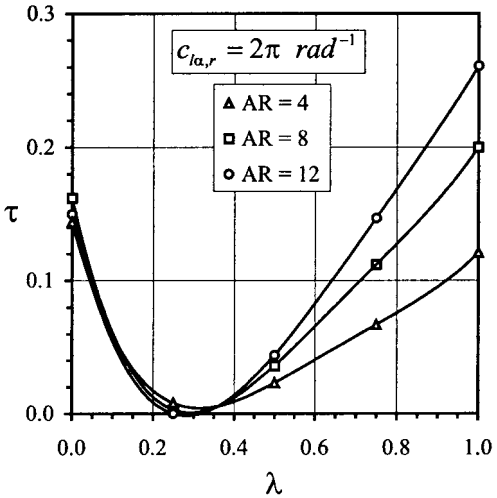


FIGURE 10.43 Variation of τ with taper ratio for an untwisted wing.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

The Monoplane Wing Equation

The general circulation distribution is sometimes given as

$$\Gamma(\theta) = 2Q_\infty b \sum_{n=1}^{N} A_n \sin n\theta \tag{10.28}$$

which results in the following form of Prandtl's lifting line equation:

$$\alpha(\theta) = \frac{4b}{c(\theta)c_{l_\alpha}(\theta)} \sum_{n=1}^{N} B_n \sin n\theta + \alpha_{l0}(\theta) + \sum_{n=1}^{N} \frac{nB_n \sin n\theta}{\sin \theta} \tag{10.29}$$

Equations (10.28) and (10.29) are equivalent mathematical statements of the lifting line model, as previously expressed by equations (10.20) and (10.21), where the A and B coefficients are related in the following manner:

$$B_n = \left(\frac{c_r c_{l_{\alpha,r}}}{4b}\right) A_n = d_r A_n$$

Equation (10.29) is occasionally written in the following way, which is known as the monoplane wing equation:

$$d(\theta)\{\alpha(\theta) - \alpha_{l0}(\theta)\} \sin \theta = \sum_{n=1}^{N} B_n \sin n\theta (nd(\theta) + \sin \theta)$$

where

$$d(\theta) = \frac{c(\theta)c_{l_\alpha}(\theta)}{4b}$$

Extended Lifting Line and Lifting Surface Theories

(*Applications—inviscid, incompressible, irrotational flow over swept, tapered, twisted, and yawed wings.*) Prandtl's lifting-line model assumes that the effects of camber and thickness are only governed by the local two-dimensional flow field around any particular wing section. This assumption is inappropriate when the wing is swept, and more sophisticated models have to be developed. As discussed in Subsection 10.11, the use of wing sweep brings many benefits during high-speed subsonic and supersonic flight, and thus its effect on low-speed flight has to investigated.

In general, the collection of swept-wing methods, known as extended lifting-line and lifting-surface models, all adopt the following solution methodology;

1. Distribute vorticity over the projection of the wing planform and the trailed wake onto the x–y plane.
2. Apply the Kutta condition in an appropriate form.
3. Use the Biot–Savart law to apply the surface flow tangency boundary condition at specified control points to produce a system of simultaneous algebraic equations for the unknown vorticity values.

4. Solve for the unknown vorticity values.

Differences in the various methods arise due to the following:

1. The manner of distributing the vorticity
2. The mathematical expression used to describe the vorticity distributions
3. The position and number of control points
4. The precise mathematical procedure used to obtain the solution

Thin airfoil theory (Subsection 10.7) shows that to account properly for the effects of airfoil camber, the vorticity must be distributed along the chord line. However, this theory also indicated that an approximate method could be used which calculated a value for the vorticity at the quarter chord based on the satisfaction of the surface flow tangency condition at the three-quarter chord point (the rear aerodynamic center).

As illustrated in Figure 10.44, extended lifting-line theory (also known as the three-quarter-point method) places a lifting line at the quarter-chord point, which produces a continuous sheet of trailing vortices. The surface flow tangency condition is then enforced at the rear aerodynamic center. Further details of this type of method are given by Weissinger (1947). An alternative to the extended lifting line methods is the vortex lattice model (Falkner 1943), where, as shown in Figure 10.45, the vorticity is distributed over the entire wing surface in the form of a finite number of elemental horseshoe vortices (a single vortex filament formed into a horseshoe shape). A chosen number of control points are now specified along the camber line rather than just one at the three-quarter chord point. The influence of all the elemental horseshoe vortices on each control point is found using the Biot–Savart law.

The vortex lattice method was the forerunner of the lifting surface model, which utilizes a continuous distribution of vorticity in both the chordwise and spanwise

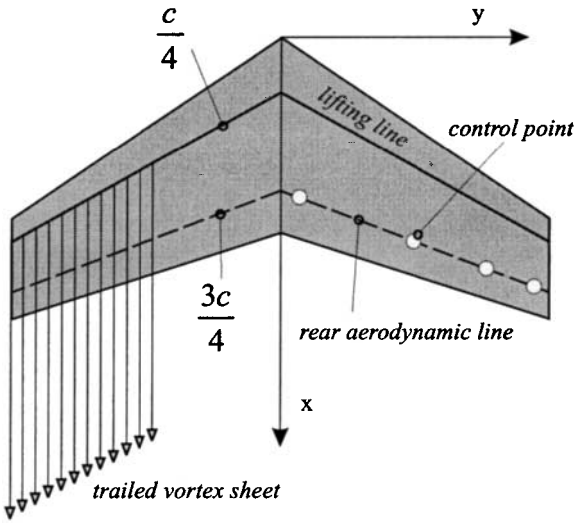


FIGURE 10.44 Weissinger's three-quarter chord method.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

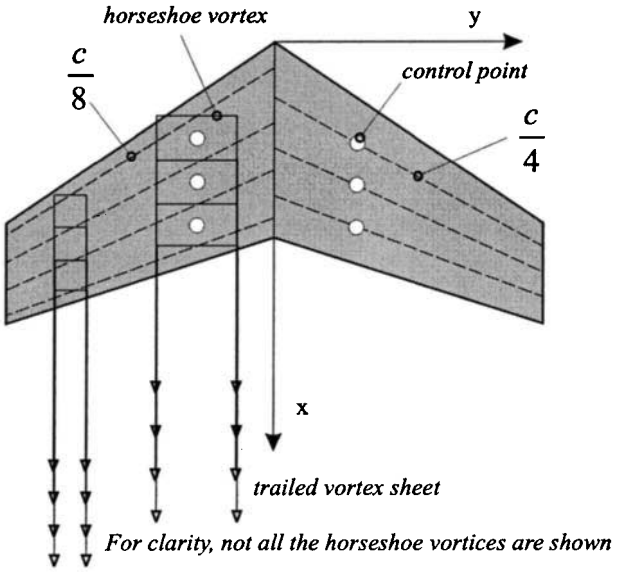


FIGURE 10.45 The vortex lattice method.

directions. Application of the Biot–Savart law yields a surface integral equation for the local downwash which has to be solved for the unknown distribution of vorticity subject to the surface flow tangency condition. For more information on vortex lattice and lifting surface methods Katz and Plotkin (1991) should be consulted.

Semiempirical Methods—Diederich's Method

(*Applications—compressible subcritical flow over swept, tapered, and twisted wings.*) Lifting-line, vortex lattice, lifting-surface, and panel methods all rely on the use of numerical algorithms to obtain the final solutions. A method is now described below which is highly amenable to spreadsheet analysis. This method is termed semiempirical because it utilizes correlations based on the results from the more sophisticated numerical models.

The following equations describe a semiempirical method developed by Diederich (1952). This method can be applied to subcritical compressible flow, which is described in more detail in Subsections 10.9, 10.10, and 10.11. It has been placed in this section for two reasons: first, because it is an excellent alternative to the numerical methods described above, and secondly, because it is used below to aid the discussion on the stalling characteristics of finite wings.

In general, the lift distribution of an arbitrary wing can be considered to be the superposition of two independent components:

1. An additional lift distribution, which depends on wing planform and angle of attack
2. A basic lift distribution, which gives zero wing lift and depends on camber and aerodynamic twist

This can be written in the basic form

$$c_l(y) = c_{la}(y) + c_{lb}(y) = \left(\frac{c_{la}(y)}{C_l}\right) C_L + c_{lb}(y)$$

$$= c_{la1}(y) C_{L.\alpha}(\alpha_L - \alpha_{L0}) + c_{lb}(y)$$

The lift distribution is often written in the following nondimensional parameters (Anderson 1936):

$$\frac{c_l(\hat{y})c(\hat{y})}{c_g} = L_a(\hat{y})C_L + L_b(\hat{y})\left(\frac{\varepsilon_{a,t}c_{l\alpha}}{E}\right)$$

where $\hat{y} = 2y/b$ and $L_a(y)$ and $L_b(y)$ are described below. The additional lift distribution is given by

$$L_a(\hat{y}) = \frac{c(\hat{y})c_{la}(\hat{y})}{c_g C_L} = C_1 \frac{c(\hat{y})}{c_g} + C_2 \frac{4}{\pi}\sqrt{1 - \hat{y}^2} + C_3 f(\hat{y})$$

The values of C_1, C_2, and C_3 depend on the aspect ratio and sweep of the wing and are given in Figure 10.46, while the function $f(\hat{y})$ is given in Figure 10.47 and depends only on the effective sweep (defined in Subsection 10.11) as

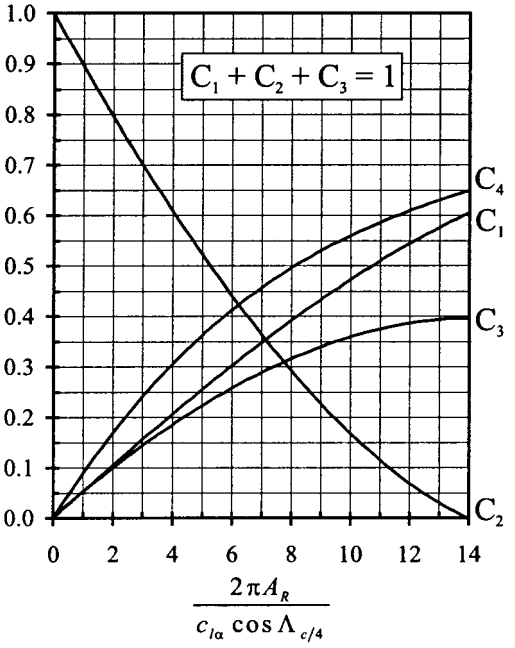


FIGURE 10.46 Factors in Diederich's method (data taken from Diederich 1952, courtesy of NASA).

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

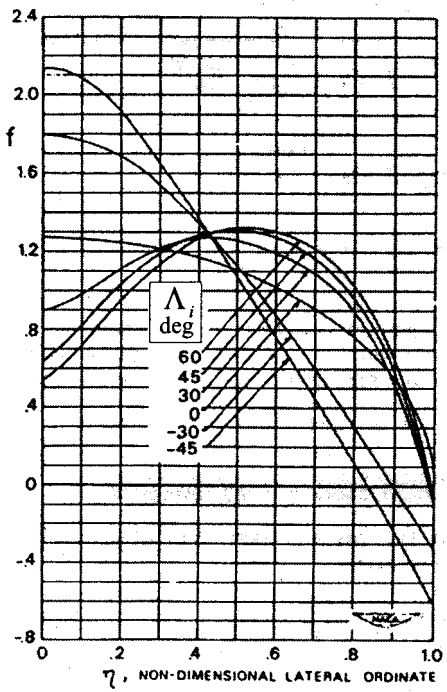


FIGURE 10.47 Function used in Diederich's method (Diederich 1952, courtesy of NASA).

$$\cos \Lambda_i = \frac{\beta_0 \cos \Lambda}{\sqrt{1 - M_\infty^2 \cos^2 \Lambda}}$$

where $\beta_0 = \sqrt{1 - M_\infty^2}$. The basic lift distribution is given by

$$L_b(\hat{y}) = \frac{c(\hat{y})c_{lb}(\hat{y})}{c_g}\left(\frac{E}{\varepsilon_{a,t}c_{l\alpha}}\right)$$

$$= C_4 E \beta_0 \cos \Lambda_\beta \left(\frac{\varepsilon_a(\hat{y})}{\varepsilon_{a,t}} + \alpha_{01}\right) L_a(\hat{y})$$

where C_4 is given in Figure 10.45 and

$$E = 1 + \frac{2\lambda}{A_R(1 + \lambda)}$$

which is known as Jones's edge factor (Jones 1941). The factor α_{01} is given by

$$\alpha_{01} = \frac{(\alpha_{L,0} - \alpha_{l0,r})}{\varepsilon_{at}} = -\int_0^1 \frac{\varepsilon_a(\hat{y})}{\varepsilon_{a,t}} L_a(\hat{y}) d\hat{y}$$

and, for unswept linearly tapered wings, has the simplified form of

$$\alpha_{01} = -C_1 \left(\frac{1 + 2\lambda}{3(1 + \lambda)} \right) - (C_2 + C_3) \frac{4}{3\pi}$$

Figure 10.48 illustrates distributions of $c_{la1}(\hat{y})$ and $c_{lb}(\hat{y})$ for a linearly tapered wing with the following characteristics: $A_R = 15$, $\lambda = 0.6$, $\Lambda = 0$, $\varepsilon_{a,t} = -2.2°$, and $\beta_0 \approx 1$. The zero lift angle of the wing (see Subsection 10.3) is given by

$$\alpha_{L0} = \alpha_{l0,r} + \alpha_{01}\varepsilon_{a,t}$$

and the lift curve slope is given by

$$C_{L\alpha} = \frac{c_{l\alpha} \cos \Lambda_i}{\sqrt{\dfrac{\beta_0^2}{\cos^2 \Lambda_i} + \left(\dfrac{c_{l\alpha}}{\pi A_R} \right)^2} + \dfrac{c_{l\alpha}}{\pi A_R}}$$

The wing lift coefficient can now be obtained from

$$C_L = C_{L\alpha}(\alpha_r - \alpha_{l0})$$

where the wing angle of attack, α_r, is defined as the angle between the root chord and the free-stream velocity.

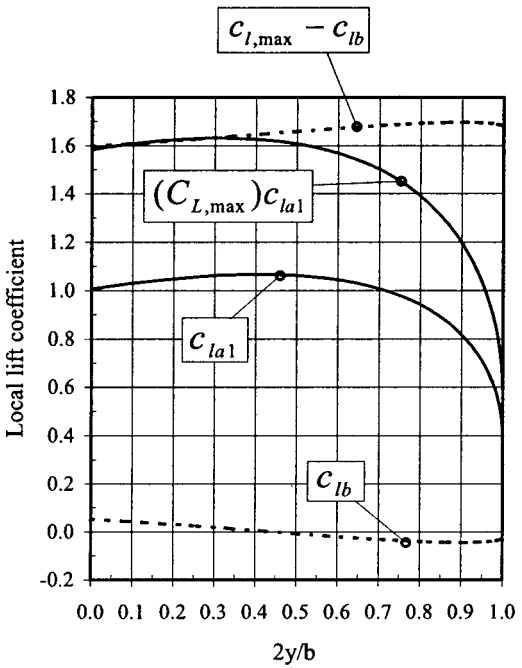


FIGURE 10.48 Spanwise lift distribution from Diederich's method.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

The wing pitching moment about the aerodynamic center of the wing (see Subsection 10.3), $C_{M,AC}$, is composed of two parts, one due to the spanwise distribution of section camber and one due to the combined effect of twist and sweep. Using the approximation that the section aerodynamic center lies at the quarter chord position, the wing pitching moment can be calculated from the following equation:

$$C_{M,AC} = C_{M,\text{camber}} + C_{M,\text{twist}}$$

where

$$C_{M,\text{camber}} = \frac{b}{S\bar{c}} \int_0^1 (c_{m,c/4}) c^2 d\hat{y}$$

$$C_{M,\text{twist}} = -\frac{b^2}{2S\bar{c}} \int_0^1 (c_{lb}) c\hat{y} \tan \Lambda_{c/4} d\hat{y}$$

The position of the wing aerodynamic center can be found using the method described in Subsection 10.3 and Figure 10.7.

The Stalling Characteristics of Finite Wings

The maximum wing lift coefficient will be reached when the local lift coefficient, at any spanwise position, reaches the maximum value appropriate to that location. When this occurs the wing is said to be stalled. Prior to the attainment of this condition an amount of trailing edge separation will have formed, on the wing upper surface, at this particular spanwise position. As the wing angle of attack is increased the trailing edge separation spreads over a region referred to as a stall cell. When considering the stability and control of an aircraft, it is important that this stall cell does not initially occur over any of the wing control surfaces, in particular the ailerons.

Figure 10.48 illustrates the spanwise variation of $(C_{l,\max} - c_{lb})$, which, for a given wing twist distribution, depends only on the airfoil sections which compose the wing. The maximum wing lift coefficient is given by the smallest value of $(c_{l,\max} - c_{lb})/c_{la1}$. The spanwise position at which this condition is satisfied identifies the location where the first stall cell will appear. The wing, used in this example, had a maximum wing lift coefficient of 1.54 and the stall cell first appeared at 26% of the span.

In general, a wing is tapered to reduce the wing root structural loads due to weight. However, as shown by Figure 10.40, taper increases the values of c_l towards the wing tips, which in turn increases the chance of flow separation over the ailerons. By introducing negative aerodynamic twist at the wing tip the tip values of c_l are reduced and the stall cell is forced inboard towards the wing root. In theory, this also makes the spanwise lift distribution more elliptical, which should reduce the induced drag. As discussed in Subsection 10.11, wing sweep is used to increase the drag divergence Mach number. For an aft-swept wing this has a similar effect on the spanwise distribution of c_l as increasing the taper. However, again this can be controlled by the use of aerodynamic twist. Chappell (1968) documents various empirical correlations which allow the stalling characteristics of various wings to be estimated.

10.9 SHOCK WAVE RELATIONSHIPS

Isentropic Flow

Combining the first law of thermodynamics for a closed-system noncyclic process with the definitions of entropy, work done, and a perfect gas gives two equations for the entropy change between two end states;

$$s_2 - s_1 = c_v \ln \frac{T_2}{T_1} + R \ln \frac{v_2}{v_1} \tag{10.30}$$

and

$$s_2 - s_1 = c_v \ln \frac{T_2}{T_1} - R \ln \frac{p_2}{p_1} \tag{10.31}$$

When the flow is regarded as both reversible and adiabatic there can be no change in entropy and the flow is said to be isentropic. Setting $s_2 - s_1$ in equation (10.31) results in the isentropic flow relationship

$$\frac{p_2}{p_1} = \left(\frac{T_2}{T_1}\right)^{\gamma/\gamma-1}$$

Many compressible flow fields can be regarded as isentropic since the effects of viscous diffusion and heat transfer are often small in the free stream. The opposite is true, however, within the boundary layer, which is a strong source of entropy generation. As indicated later, this is also the case for shock waves.

Speed of Sound and Mach Number

The speed of sound through an arbitrary fluid can be calculated from $a = \sqrt{\gamma RT}$ where T is the local fluid temperature. The Mach number is defined the ratio of the local fluid velocity to the speed of sound, i.e., $M = q/a$.

The Stagnation State

The stagnation value, of a fluid property is defined as the value it would ascertain if the flow were isentropically brought to rest without any work transfer. Under these conditions, the conservation of energy reduces to

$$h + \frac{q^2}{2} = \text{constant} = h_0$$

The stagnation temperature and pressure are respectively given by

$$\frac{T_0}{T} = 1 + \frac{\gamma - 1}{2} M^2 \quad \text{and} \quad \frac{p_0}{p} = \left(\frac{T_0}{T}\right)^{\gamma/\gamma-1}$$

Combining these two equations results in a useful relationship among the stagnation pressure, the static pressure, and the Mach number given by

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

$$\frac{p_o}{p} = \left(1 + \frac{\gamma - 1}{2} M^2\right)^{\gamma/\gamma-1} \tag{10.32}$$

Incompressible Flow Limit

The stagnation density of a moving perfect gas is given by the isentropic relation

$$\frac{\rho_o}{\rho} = \left(1 + \frac{\gamma - 1}{2} M^2\right)^{1/(\gamma-1)}.$$

For air, with $\gamma = 1.4$, the difference between the stagnation and the static densities is less than 5% for Mach numbers less than 0.32. As a result, it is generally accepted that the flow should be treated as compressible above Mach numbers of 0.3. Since the local maximum velocity over an arbitrary airfoil is approximately three times the free-stream value, the entire flow field can only be regarded as completely incompressible for free-stream Mach numbers of around 0.1. For low Mach numbers, equation (10.31) may be reduced to Bernoulli's equation using the binomial theorem, i.e.,

$$p_0 = p + \tfrac{1}{2}\gamma M^2 p, \quad \text{which gives} \quad p_0 = p + \tfrac{1}{2}\rho q^2$$

Mach Wave Formation

Any object moving through a fluid will cause pressure waves to propagate through the surrounding fluid at the local speed of sound. Figure 10.49 illustrates the spread

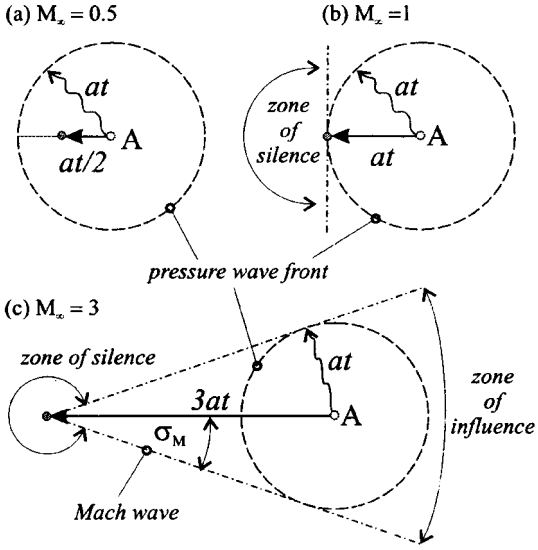


FIGURE 10.49 Mach wave formation.

of pressure waves from a small body traveling at subsonic and then supersonic speeds. For the supersonic case only the fluid that lies within the cone indicated is aware of the presence of the body. This is known as the zone of influence and the vertex angle, σ_M, is known as the Mach angle and is related to the Mach number as follows:

$$\sigma_M = \sin^{-1}(1/M_\infty)$$

Shock Wave Formation

To illustrate how a shock wave forms, consider the acceleration from rest of a piston within a long cylinder. The initial movement of the piston causes a Mach wave to travel downstream, which leaves the air behind with a slightly increased pressure and temperature. This, in turn, increases the local speed of sound so subsequent Mach waves, produced by the moving piston, travel faster downstream. A continual series of Mach waves will eventually merge to produce a strong shock wave across which the local pressure, temperature, velocity, and entropy will abruptly change. Although the flow through a shock wave is nonisentropic, the flow ahead and aft can often be considered to be isentropic.

Normal Shock Wave Relations

Applying the conservation of mass, momentum, and energy to a control volume which lies across a normal shock wave results in the following equations which are often referred to the Rankine–Hugoniot relationships.

$$M_2^2 = \frac{(\gamma - 1)M_1^2 + 2}{2\gamma M_1^2 - (\gamma - 1)}$$

$$\frac{p_2}{p_1} = \frac{2\gamma M_1^2 - (\gamma - 1)}{(\gamma + 1)}$$

$$\frac{p_{02}}{p_{01}} = \left[\frac{(\gamma + 1)M_1^2}{2 + (\gamma - 1)M_1^2}\right]^{\gamma/(\gamma-1)} \left[\frac{2\gamma M_1^2 - (\gamma - 1)}{(\gamma + 1)}\right]^{-1/(\gamma-1)}$$

$$\frac{T_2}{T_1} = \frac{[2\gamma M_1^2 - (\gamma - 1)](\gamma - 1)M_1^2 + 2]}{(\gamma + 1)^2 M_1^2}$$

$$\frac{\rho_2}{\rho_1} = \frac{(\gamma + 1)M_1^2}{2 + (\gamma - 1)M_1^2}$$

$$s_2 - s_1 = c_p \ln \left\{\frac{[2\gamma M_1^2 - (\gamma - 1)][(\gamma - 1)M_1^2 + 2]}{(\gamma + 1)^2 M_1^2}\right\} - R \ln \left\{\frac{2\gamma M_1^2 - (\gamma - 1)}{(\gamma + 1)}\right\} \tag{10.33}$$

Figure 10.50 graphically illustrates the ratios of the various flow variables as a function of the upstream Mach number. A number of sets of tables and graphs are

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

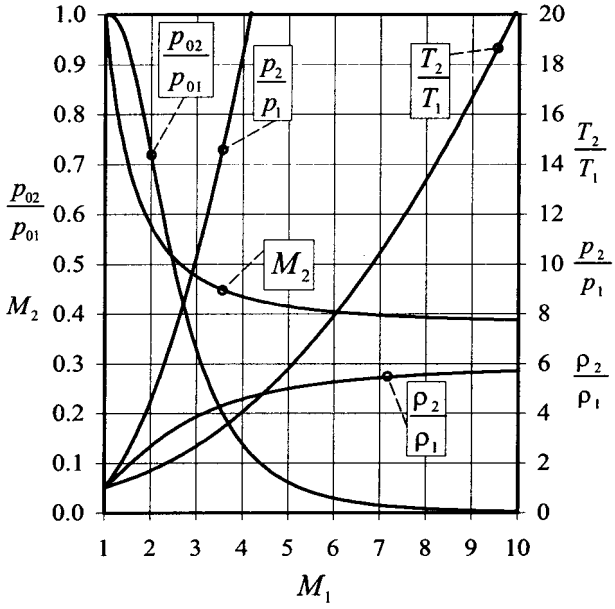


FIGURE 10.50 Normal shock wave properties ($\gamma = 1.4$).

available which give similar information e.g., Houghton and Brock 1978; Ames Research Staff 1947, 1953. Equation (10.33) indicates that the flow through a Mach wave ($M_1 = 1$) is isentropic while there is an increase in entropy through a shock wave ($M_1 > 1$).

Oblique Shock Wave Relations

Figure 10.51 defines the flow geometry associated with an oblique shock wave where σ is known as the shock wave angle and δ is referred to as the turning (or

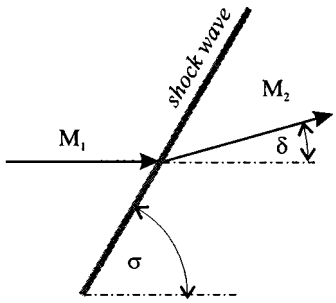


FIGURE 10.51 Oblique shock wave angles.

deflection) angle. Since there is only a change in the velocity normal to the oblique shock wave, we can use the normal shock wave relations defined above. This results in the following equations:

$$M_2^2 \sin^2(\sigma - \delta) = \frac{(\gamma - 1)M_1^2 \sin^2 \sigma + 2}{2\gamma M_1^2 \sin^2 \phi - (\gamma - 1)}$$

$$\frac{p_2}{p_1} = \frac{2\gamma M_1^2 \sin^2 \sigma - (\gamma - 1)}{(\gamma + 1)} \tag{10.34}$$

$$\frac{T_2}{T_1} = \frac{[2\gamma M_1^2 \sin^2 \sigma - (\gamma - 1)][(\gamma - 1)M_1^2 \sin^2 \sigma + 2]}{(\gamma + 1)^2 M_1^2 \sin^2 \sigma}$$

$$\frac{\rho_2}{\rho_1} = \frac{(\gamma + 1)M_1^2 \sin^2 \sigma}{2 + (\gamma - 1)M_1^2 \sin^2 \sigma}$$

It can be shown using the principle of increasing entropy that for an oblique shock $M_1 \sin \sigma \geq 1$. Equation (10.34) shows that the ratio of p_2/p_1 tends to unity as the shock wave angle tends to the Mach angle. Hence, the limits on the shock wave angle are $\sin^{-1}(1/M_1) \leq \sigma \leq 90°$. It should also be noted that the flow behind an oblique shock wave can be supersonic since $M_2 \sin(\sigma - \delta) \leq 1$. In order to utilize the above relationships, the relation between δ, σ, and M_1 has to known. From geometrical considerations, it can be shown that

$$\tan \delta = \frac{2(M_1^2 \sin^2 \sigma - 1)/\tan \sigma}{2 + M_1^2(\gamma + \cos 2\sigma)} \tag{10.35}$$

It is worth noting that the turning angle is 0 when $\sigma = 90°$ (normal shock) and also when $\sigma = \sin^{-1}(1/M_1)$ (Mach wave). A typical oblique shock wave solution procedure would be first to obtain σ for a given M_1 and δ and then to use the value of $M_1 \sin \sigma$ along with normal shock wave tables. The value of M_2 is given by the normal shock wave value divided by $\sin(\sigma - \delta)$.

Figures 10.52 and 10.53 show graphical representations of equation (10.35). For a given value of M_1, equation (10.35) may be differentiated, which gives the maximum turning angle to occur when

$$\sin^2 \sigma_{\max} = \frac{(\gamma + 1)M_1^2 - 4}{4\gamma M_1^2} + \frac{(\gamma + 1)^{1/2}}{4\gamma M_1^2}[16 + 8(\gamma - 1)M_1^2 + (\gamma + 1)M_1^4]^{1/2}$$

For a given Mach number $\sigma_{\max}$ can be calculated and thus the maximum turning angle, $\delta_{\max}$, can be obtained.

When a body forces the flow to deflect an amount greater than the maximum deflection angle, for the particular Mach number, a detached curved shock wave will form. For any deflection angle less than the maximum, there is a low and a high wave angle, which are respectively known as the weak and strong shock solutions. Experiments have shown that the observed shock angle nearly always corresponds to the weak solution.

Also indicated in Figure 10.52 is the locus of turning and wave angles which result in $M_2 = 1$, which are related by the following equation:

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

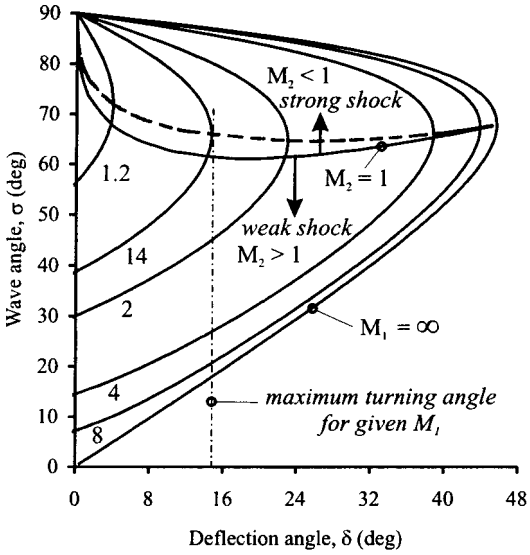


FIGURE 10.52 Oblique shock wave graph and definitions.

$$\sin^2 \sigma_s = \frac{(\gamma + 1)M_1^2 - 3 + \gamma}{4\gamma M_1^2} + \frac{(\gamma + 1)^{1/2}}{4\gamma M_1^2}[9 + \gamma + 2(\gamma - 3)M_1^2 + (\gamma + 1)M_1^4]^{1/2}$$

Expansion Waves (Prandtl–Meyer Flow)

Consider the flow through a Mach wave which has been produced by the flow turning through a small deflection, $d\delta$ (positive in anticlockwise direction). From geometrical considerations, and using the fact that the velocity component tangential to the wave remains constant, it can be shown that

$$\frac{dq}{q} = \frac{-d\delta}{\sqrt{M^2 - 1}} \tag{10.36}$$

Combining the steady flow energy equation with the speed of sound gives the following results:

$$\frac{dp}{p} = \frac{\gamma M^2 d\delta}{\sqrt{M^2 - 1}} \tag{10.37}$$

$$\frac{d\rho}{\rho} = \frac{M^2 d\delta}{\sqrt{M^2 - 1}} \tag{10.38}$$

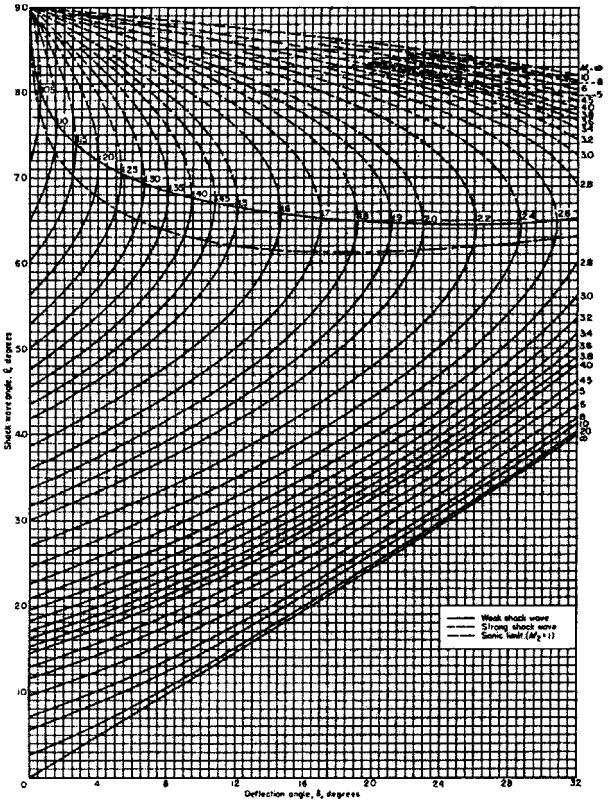


FIGURE 10.53 Standard oblique shock wave chart.

Using these differential changes in pressure and density with equations (10.30) and (10.31) results in the flow across the Mach wave being isentropic. Thus, a small deflection produces the following changes in the flow field:

$$dM \propto -d\delta, \quad dq \propto -d\delta, \quad dp \propto d\delta, \quad d\rho \propto d\delta, \quad \text{and} \quad ds = 0$$

Figure 10.54 shows the flow around a sharp convex corner to consist of an infinite number of Mach waves, each turning the flow through a differentially small angle. For negative changes in wall angle the Mach waves diverge and the flow remains isentropic throughout the entire fan. Such flows are called either centered expansion fans or Prandtl–Meyer flows. Differentiating the definition for the Mach number gives

$$\frac{dq}{q} = \frac{dM}{M} + \frac{da}{a} = \frac{dM}{M} + \tfrac{1}{2}\frac{dp}{p} - \tfrac{1}{2}\frac{d\rho}{\rho}$$

Using equations (10.35)–(10.37), the total turning angle is given from

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

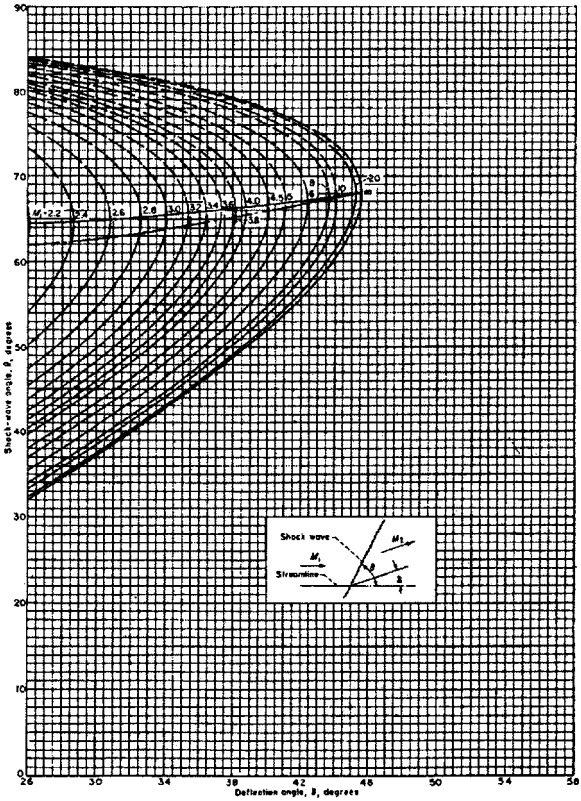


FIGURE 10.53 (*Continued*) Standard oblique shock wave chart.

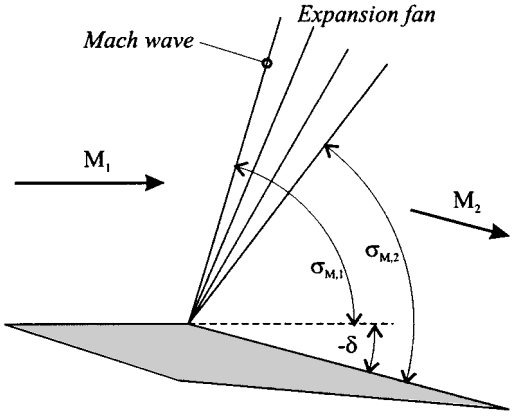


FIGURE 10.54 Prandtl–Meyer expansion fan.

$$-\int_0^{\delta} d\delta = \int_1^{M} \frac{2\sqrt{M^2 - 1}}{2 + (\gamma - 1)M^2} \frac{dM}{M}$$

Carrying out the integration gives the total turning angle equal to

$$-\delta = \sqrt{\frac{\gamma + 1}{\gamma - 1}} \tan^{-1} \left\{ \sqrt{\frac{\gamma + 1}{\gamma - 1} (M^2 - 1)} \right\} - \tan^{-1} \{ \sqrt{(M^2 - 1)} \} = v(M) \tag{10.36}$$

$v(M)$ is sometimes referred to as the Prandtl–Meyer function and is equal to 0 when the Mach number is unity. Knowing the upstream Mach number, M_1 and the deflection angle, δ, the downstream Mach number, M_2 can be calculated as follows.

1. From equation (10.36), Figure 10.55, or tables (Houghton and Brock 1978; Ames Research Staff 1947, 1953), obtain the value of δ_1 corresponding to M_1.
2. Calculate $\delta_2 = \delta - \delta_1$ and then find M_2 from tables, graphs, or equation.
3. Any other downstream property can be obtained from isentropic relationships or isentropic flow tables (noting that the stagnation pressure remains constant across the fan).
4. Calculate the fan boundaries using $\sigma_{M,1} = \sin^{-1}(1/M_1)$ and $\sigma_{M,2} = \sin^{-1}(1/M_2)$.

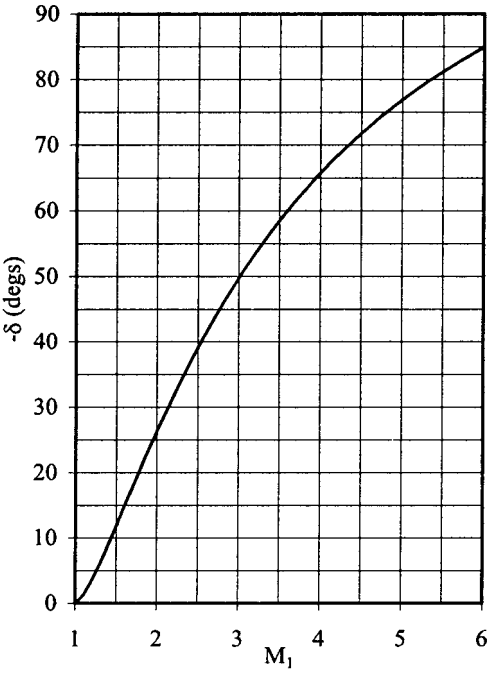


FIGURE 10.55 The Prandtl–Meyer function.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

Further Compressible Flow Phenomena

There are many more important and interesting compressible flow phenomena which have not been covered here. For further information Anderson (1990) and Oosthuizen and Carscallen (1997) should be consulted.

10.10 COMPRESSIBLE FLOW OVER AIRFOILS

Shock-Expansion Theory of Lift

(*Applications—two-dimensional, inviscid, isentropic, irrotational, supersonic compressible flow past a thin airfoil at low angles of attack. The airfoil must be made up of straight line segments and the deflection angles must be small enough not to induce a detached shock.*) Consider the supersonic flow past a flat plate at some angle of attack as depicted in Figure 10.56 As indicated in the figure, the upper and lower surfaces develop uniform pressure distributions which are due to the system of oblique shock waves and centered expansion fans. These pressures may be calculated using the relevant techniques described in Subsection 10.9 and thus the resultant aerodynamic force can be computed; see Anderson (1990) for numerical examples.

Linear Theory for Perturbated Compressible Flow

(*Applications—two-dimensional, inviscid, isentropic, irrotational, compressible flow past a thin airfoil at low angles of attack. The free-stream flow can be subsonic or supersonic, but not transonic or hypersonic.*) We define a disturbance (or perturbation) velocity potential as follows:

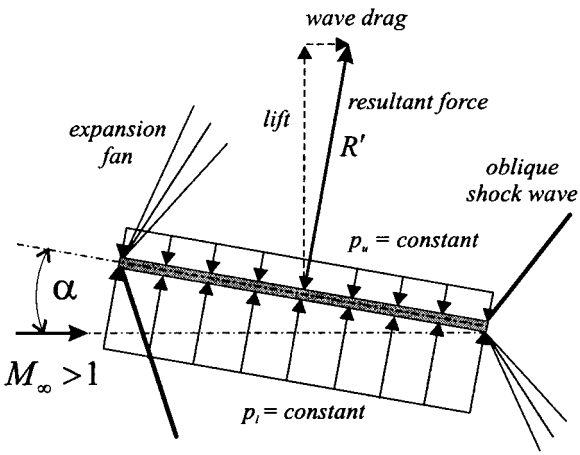


FIGURE 10.56 The shock expansion theory of lift.

$$u = Q_\infty + \hat{u} \quad v = 0 \quad w = \tilde{w}$$

$$\phi = Q_\infty x + \tilde{\phi} \quad u = \frac{\partial \phi}{\partial x} = Q_\infty + \frac{\partial \tilde{\phi}}{\partial x} \quad w = \frac{\partial \phi}{\partial z} = \frac{\partial \tilde{\phi}}{\partial z}$$

Starting from the continuity equation and utilizing Euler's equation (Subsection 10.4) and the isentropic speed of sound relation we obtain the disturbance velocity potential equation;

$$\left[a^2 - \left(Q_\infty + \frac{\partial \tilde{\phi}}{\partial x}\right)^2\right] \frac{\partial^2 \tilde{\phi}}{\partial x^2} + \left[a^2 - \left(\frac{\partial \tilde{\phi}}{\partial z}\right)^2\right] \frac{\partial^2 \tilde{\phi}}{\partial z^2} - 2\left(Q_\infty + \frac{\partial \tilde{\phi}}{\partial x}\right)\left(\frac{\partial \tilde{\phi}}{\partial z}\right) \frac{\partial^2 \tilde{\phi}}{\partial x \partial z} = 0$$

Expressing the disturbance potential in terms of the disturbance velocities, substituting in the energy equation, and excluding both the transonic flow ($0.8 < M_\infty < 1.2$) and hypersonic flow ($M_\infty > 5$) regimes, we obtain:

$$(1 - M_\infty^2) \frac{\partial^2 \tilde{\phi}}{\partial x^2} + \frac{\partial^2 \tilde{\phi}}{\partial z^2} = 0 \tag{10.39}$$

This linear partial differential equation can be solved with the following boundary conditions (also see Subsection 10.5):

1. Free-stream condition: $\tilde{u} = \tilde{w} = 0$ or $\tilde{\phi} = constant$.

2. Body flow tangency condition: $\tilde{w} = \partial \tilde{\phi} / \partial z = Q_\infty(\alpha - dz_s/dx)$

If the fluid is assumed to be a perfect gas, the linearized pressure coefficient is given by

$$C_p = \frac{p - p_\infty}{\frac{1}{2}\rho_\infty Q_\infty^2} = \frac{p - p_\infty}{\frac{1}{2}\gamma M_\infty^2 p_\infty} = -\frac{2\tilde{u}}{Q_\infty} \tag{10.40}$$

Subsonic Compressibility Correction Methods

(*Applications—two-dimensional, inviscid, isentropic, irrotational, subsonic compressible flow past a thin airfoil at low angles of attack.*) Subsonic compressibility correction methods relate the subsonic compressible flow past a particular airfoil to the incompressible flow past a second airfoil which is geometrically related to the first through an affine transformation. An affine transformation changes all the coordinates in a given direction by a uniform ratio. These methods generate expressions which are known as similarity laws. There are four methods in this class: Gothert's rule, the Prandtl–Glauert rule, Laitone's rule, and the Karman–Tsien rule. The elegance of the compressibility correction methods lies in the fact that compressible flow airfoil characteristics can be predicted by modifying the incompressible data obtained from either the methods described in Subsection 10.7 or from low-speed wind tunnel tests.

Gothert's rule essentially covers the fundamental transformation technique which is now described. Consider the following transformation:

$$\hat{x} = \zeta_x x, \quad \hat{z} = \zeta_z z, \quad \hat{\phi}(\hat{x},\hat{y}) = \zeta_\phi \tilde{\phi}(x,z)$$

Substituting these variables into equation (10.39) gives

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

$$\frac{\partial^2 \hat{\phi}}{\partial \hat{x}^2} + \frac{\partial^2 \hat{\phi}}{\partial \hat{z}^2} = 0$$

if we stipulate that

$$\frac{\zeta_z}{\zeta_x} = \sqrt{1 - M_\infty^2} = \beta_0$$

Since $\hat{\phi}$ is a solution to the Laplace equation, it must represent the disturbance velocity in incompressible flow. The incompressible and compressible airfoil profiles are geometrically related in the following manner:

$$(dz_s/dx)_{\text{inc}} = \beta_0 (dz_s/dx), \quad \hat{t}_{\text{inc}} = \beta_0 t \quad \text{and} \quad \hat{h}_{\text{inc}} = \beta_0 \hat{h}$$

The airfoil surface flow tangency condition on the incompressible plane is given by

$$\frac{\partial \hat{\phi}}{\partial \hat{z}} = Q_\infty \left\{ \left(\frac{\zeta_\phi}{\zeta_z} \right) (\alpha)_{\text{inc}} - \left(\frac{\zeta_\phi \zeta_x}{\zeta_z^2} \right) \frac{d\hat{z}_s}{d\hat{x}} \right\}$$

which means that the flow boundary conditions are satisfied on both the compressible and incompressible planes when

$$(Q_\infty)_{\text{inc}} = Q_\infty, \quad \alpha_{\text{inc}} = \beta_0 \alpha \quad \text{and} \quad \zeta_z^2 = \zeta_x \zeta_\phi$$

The compressible pressure coefficient is given by

$$C_p = \frac{p - p_\infty}{\frac{1}{2}\rho_\infty Q_\infty^2} = -\frac{2\tilde{u}}{Q_\infty} = -\frac{2}{Q_\infty} \frac{\partial \tilde{\phi}}{\partial x} = -\frac{\zeta_x}{\zeta_\phi} \frac{2}{Q_\infty} \frac{\partial \hat{\phi}}{\partial \hat{x}} = \frac{1}{\beta_0^2} (C_p)_{\text{inc}}$$

The Prandtl–Glauert rule utilizes the fact that for affinely related airfoils in incompressible flow the local pressure coefficient at corresponding points is approximately proportional to the thickness ratio, the camber ratio, and the angle of attack. Consider two airfoils in incompressible flow with the following geometric relationship:

$$\hat{t}_{\text{inc},2} = \frac{1}{\beta_0} \hat{t}_{\text{inc},1}, \quad \hat{h}_{\text{inc},2} = \frac{1}{\beta_0} \hat{h}_{\text{inc},1}, \quad \text{and} \quad \alpha_{\text{inc},2} = \frac{1}{\beta_0} \alpha_{\text{inc},1}$$

Then it follows that

$$(C_p)_{\text{inc},2} \approx \frac{1}{\beta_0} (C_p)_{\text{inc},1}$$

This result is now combined with Gothert's rule to relate the compressible and incompressible flow fields in the following manner

$$\frac{\hat{t}}{\hat{t}_{\text{inc},2}} = \frac{\hat{h}}{\hat{h}_{\text{inc},2}} = \frac{\alpha}{\alpha_{\text{inc},2}} = 1 \quad \text{and} \quad C_p = \frac{1}{\beta_0} (C_p)_{\text{inc},2}$$

Therefore, the compressible flow pressure coefficient can be obtained from the incompressible value for the same airfoil profile as

$$C_p = \frac{(C_p)_{\text{inc}}}{\sqrt{1 - M_\infty^2}} \tag{10.41}$$

where $(C_p)_{\text{inc}} = (C_p)_{\text{inc},2}$.

Since the lift and moment coefficients are obtained from the integrated pressure coefficient, we obtain

$$c_l = \frac{(c_l)_{\text{inc}}}{\sqrt{1 - M_\infty^2}} \quad \text{and} \quad c_m = \frac{(c_m)_{\text{inc}}}{\sqrt{1 - M_\infty^2}}$$

Laitone first hypothesized that the local Mach number (rather than the free-stream value) should be used to calculate β_0. The isentropic relations were then utilized to express the local Mach number in terms of the free stream value to obtain the following modification to the Prandtl–Glauert rule:

$$C_p = \frac{2\beta_0 (C_p)_{\text{incomp}}}{2\beta_0^2 + \tfrac{1}{2}\{2M_\infty^2 + (\gamma - 1)M_\infty^4\}(C_p)_{\text{incomp}}} \tag{10.42}$$

The Prandtl–Glauert rule assumes that the local speed of sound does not vary from point to point around the airfoil. This approximation was taken into account by Karman and Tsien, which resulted in

$$C_p = \frac{(C_p)_{\text{incomp}}}{\beta_0 + \tfrac{1}{2}\{1 - \beta_0\}(C_p)_{\text{incomp}}} \tag{10.43}$$

The two pressure coefficients in the Karman–Tsien rule do not strictly refer to the same airfoil profile since the transformation used distorts the geometry on the incompressible plane. However, this effect is small and the rule is commonly used for a fixed airfoil geometry.

Figure 10.57 displays a comparison between the three compressibility correction equations and wind tunnel data. In this example, the minimum experimental pressure coefficient was used at each Mach number and Laitones criterion gives a critical Mach number of just above 0.6 for this airfoil.

Critical Mach Number

The critical Mach number is defined as the free-stream Mach number at which the local Mach number on the airfoil surface becomes unity. The local pressure coefficient which coincides with this point is given, for isentropic flow by

$$C_{p,cr} = \frac{2}{\gamma M_{cr}^2}\left[\left(\frac{1 + \tfrac{1}{2}(\gamma - 1)M_{cr}^2}{1 + \tfrac{1}{2}(\gamma - 1)}\right)^{\gamma/\gamma - 1} - 1\right] \approx \frac{-2}{\gamma + 1}\frac{1 - M_{cr}^2}{M_{cr}^2} \tag{10.44}$$

Figure 10.57 shows that the critical Mach number may be estimated using either experimental data or one of the compressibility corrections along with equation (10.44). Laitone's correction intersected equation (10.44) at the lowest Mach number; which resulted in a critical Mach number of 0.6.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

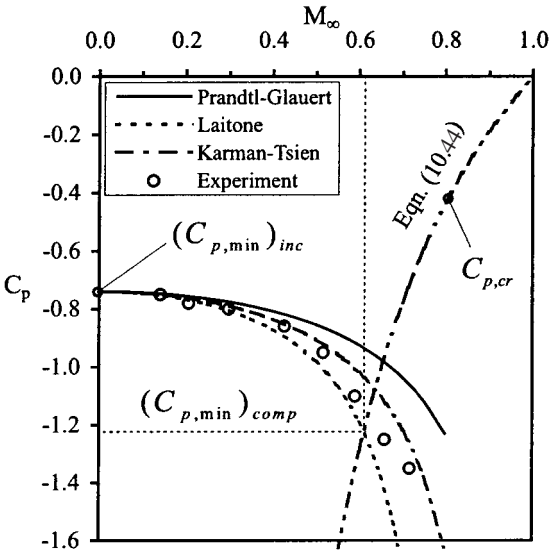


FIGURE 10.57 Compressibility correction methods (data from Stack et al. 1938, courtesy of NASA).

Linear Theory for Perturbated Supersonic Flow

(*Applications—two-dimensional, inviscid, isentropic, irrotational, supersonic compressible flow past a thin airfoil at low angles of attack.*) The following method was formulated by Ackeret (1925) and is often referred to as Ackeret's first order (or linear) theory. For Mach numbers between 1.2 and 5.0, equation (10.39) takes the form

$$\beta_1^2 \frac{\partial^2 \tilde{\phi}}{\partial x^2} - \frac{\partial^2 \tilde{\phi}}{\partial z^2} = 0$$

where

$$\beta_1 = \sqrt{M_\infty^2 - 1}$$

This equation has the classical wave equation form and thus has the general solution

$$\tilde{\phi} = \tilde{\phi}_1(x - \beta_1 z) + \tilde{\phi}_2(x + \beta_1 z)$$

By first considering $\tilde{\phi}_2 = 0$ and then $\tilde{\phi}_1 = 0$, it can be shown that lines of constant $\tilde{\phi}$ are in fact Mach waves respectively emanating from the upper and lower airfoil surfaces. Application of the flow tangency boundary condition to an arbitrary thin airfoil at a low angle of attack the velocity disturbances on the upper and lower surfaces are given by

$$\tilde{u} = \pm \frac{Q_\infty}{\beta_1}\left(\alpha - \frac{dz_s}{dx}\right) \quad \text{and} \quad \tilde{w} = \mp Q_\infty \left(\alpha - \frac{dz_s}{dx}\right)$$

where the upper sign applies to the upper surface. Thus, the local pressure coefficient is given from equation (10.40) as

$$C_p = -\frac{2\tilde{u}}{Q_\infty} = \mp \frac{2}{\beta_1}\left(\alpha - \frac{dz_s}{dx}\right) \tag{10.45}$$

If the upper and lower surface coordinates are written in terms of the camber line geometry and a thickness distribution (i.e., $\hat{z}_{u,l} = \hat{z}_c \pm \hat{z}_t$), the following relationships can be formulated:

$$\Delta C_p = \frac{p_l - p_u}{\tfrac{1}{2}\rho_\infty Q_\infty^2} = \frac{4}{\beta_1}\left(\alpha - \frac{d\hat{z}_c}{d\hat{x}}\right)$$

$$c_l = \int_0^1 \Delta C_p d\hat{x} = \frac{4\alpha}{\beta_1} \quad c_{l\alpha} = \frac{4}{\beta_1} \quad \alpha_{l0} = 0$$

$$c_{m,le} = -\int_0^1 \Delta c_p \hat{x} d\hat{x} = -\frac{2\alpha}{\beta_1} - \frac{4}{\beta_1}\int_0^1 \hat{z}_c d\hat{x}$$

$$c_{m,c/2} = c_{m,le} + \tfrac{1}{2}c_l = -\frac{4}{\beta_1}\int_0^1 \hat{z}_c d\hat{x}$$

$$\frac{x_{cp}}{c} \approx \frac{c_{m,le}}{c_l} = \frac{1}{2} - \frac{1}{2\alpha}\int_0^1 \hat{z}_c d\hat{x}$$

It is interesting to note that the following supersonic airfoil characteristics:

1. The lift coefficient is independent of airfoil shape.
2. The zero-lift angle is always zero.
3. The aerodynamic center is always located at the mid-chord position.

Also, drag is produced even though the flow is inviscid. It is known as wave drag, and the drag coefficient is given by

$$\begin{aligned} c_{d,w} &= \frac{4\alpha^2}{\beta_1} + \frac{4}{\beta_1}\left[\int_0^1 \left(\frac{d\hat{z}_c}{d\hat{x}}\right)^2 d\hat{x} + \int_0^1 \left(\frac{d\hat{z}_t}{d\hat{x}}\right)^2 d\hat{x}\right] \\ &= c_{d,i} + c_{d,t} \end{aligned} \tag{10.46}$$

The first term of equation (10.46) is known as the wave drag due to lift and the second two terms taken together are known as the wave drag due to thickness. Equation (10.46) shows that the flat plate has the lowest wave drag and can therefore be regarded as the best supersonic airfoil profile. For a diamond profile (double-wedge) the wave drag has the form

$$c_{d,w} = \frac{4\alpha^2}{\beta_1} + \frac{4f^2}{\beta_1}$$

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

Figure 10.58 illustrates a comparison between Ackeret's theory and wind tunnel data obtained for a biconvex airfoil. The difference between the prediction and the data can be addressed by using higher-order mathematical models such as that of Busemann; further information on higher-order supersonic theories can be found in Hilton (1951).

Drag Divergence Mach Number

As shown in Figure 10.57, it was relatively easy to calculate the critical Mach number for a given airfoil. Figure 10.59 illustrates the variation in drag coefficient, along with the changes in flow field around an airfoil, as the free-stream Mach number is increased past the critical value. For free-stream Mach numbers which lie between the critical and sonic values, the drag coefficient starts to increase rapidly (diverge) due to the formation of shock waves which terminate local pockets of supersonic flow on both upper and lower surfaces of the airfoil. The value of free-stream Mach number at which this phenomena occurs is known as the drag divergence Mach number. Nitzberg and Crandall (1949) observed that as M_∞ was further increased the shock waves grew in strength and eventually gave rise to boundary layer separation. This phenomenon was referred to as shock stall since the lift coefficient started to decrease rapidly at this point. As M_∞ approaches unity, the drag coefficient can easily increase by a factor of 10 or more. When the free-stream flow becomes supersonic (past the sound barrier) the drag coefficient drops due to formation of a system of shock waves which do not induce large amounts of flow separation.

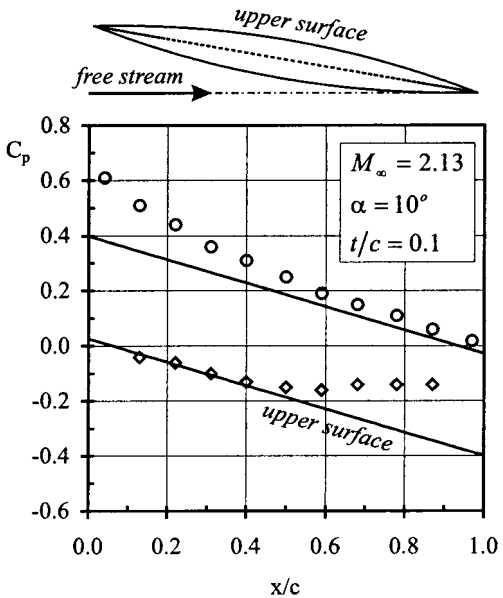


FIGURE 10.58 Ackert's linear theory (data from Ferri 1940, courtesy of NASA).

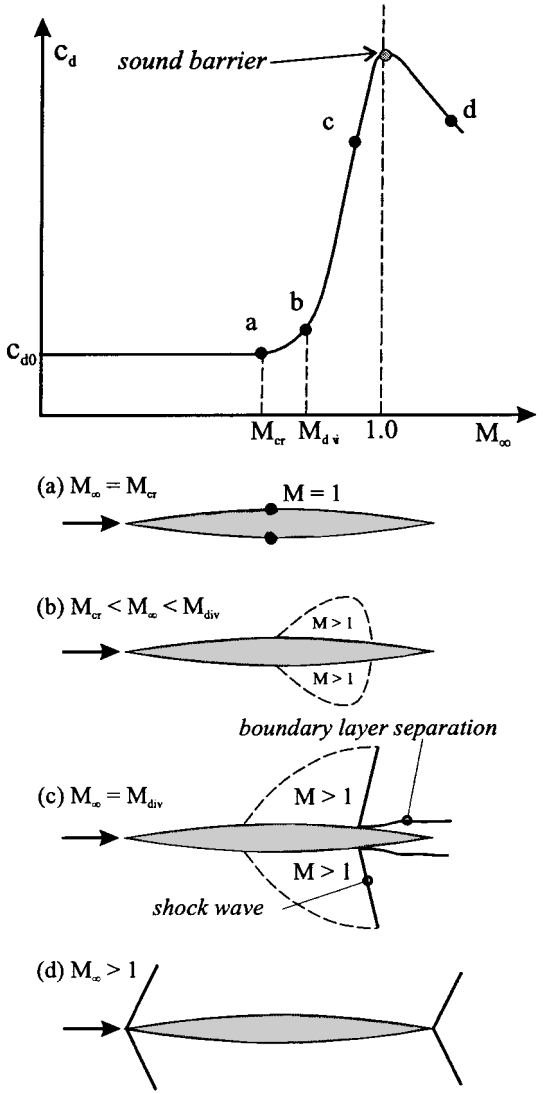


FIGURE 10.59 Drag divergence flow field changes.

The Supercritical Airfoil

Wind tunnel testing has shown that decreasing the airfoil thickness significantly increases M_{cr} (a similar effect can be achieved by using a swept wing, as discussed in Subsection 10.11). However, there are many wing design factors which prohibit the use of very thin airfoil sections. When $M_\infty < M_{cr}$, no shock waves form and the airfoil is termed subcritical. An airfoil profile which has been designed to pro-

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

duce a large upper surface region of low supersonic flow which can be terminated by a weak shock wave is referred to as supercritical and was originally developed by Whitcomb and Clark (1965). Although initially designed to encourage laminar flow, the NACA 6-series airfoils turned out to have high critical Mach numbers. These profiles have been successfully modified to produce families of supercritical airfoils, as demonstrated by von Doenhoff et al. (1947) and Loftin and Cohen (1948). Figure 10.60 shows that a typical supercritical airfoil has a low upper surface curvature and a concave lower surface at the trailing edge. Figure 10.61 illustrates that, at low angles of attack, the lift of a supercritical airfoil is produced by two mechanisms: a low pressure supersonic region spread over a large portion of the upper surface (known as a roof-top pressure distribution), and a high pressure region over the concave lower surface, (known as a rear loading).

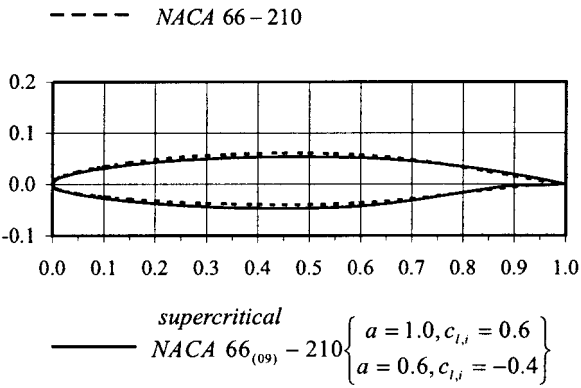


FIGURE 10.60 The supercritical airfoil.

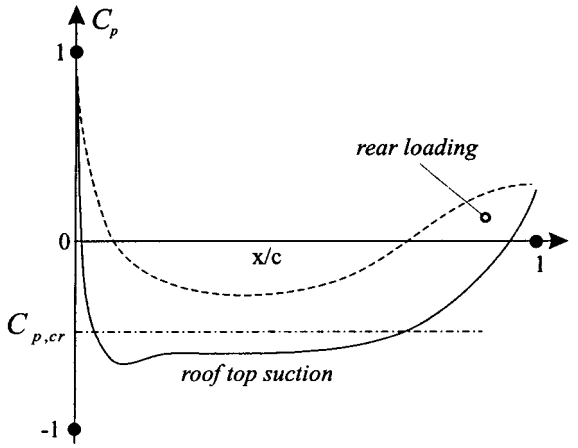


FIGURE 10.61 Typical supercritical pressure distribution.

10.11 COMPRESSIBLE FLOW OVER FINITE WINGS

Subsonic Compressibility Correction Methods

(*Applications—three-dimensional, inviscid, isentropic, irrotational, subsonic compressible flow past a finite wing of arbitrary aspect ratio, thin airfoil section and at low angles of attack.*) Gothert's rule, as described in Subsection 10.10, can be easily extended to three dimensions with the following affine transformation;

$$\hat{x} = \zeta_x x, \quad \hat{y} = \zeta_y y, \quad \hat{z} = \zeta_z z, \quad \hat{\phi}(\hat{x},\hat{y}) = \zeta_\phi \tilde{\phi}(x,z)$$

The governing equation for the linearized disturbance velocity potential becomes

$$\frac{\partial^2 \hat{\phi}}{\partial \hat{x}^2} + \frac{\partial^2 \hat{\phi}}{\partial \hat{y}^2} + \frac{\partial^2 \hat{\phi}}{\partial \hat{z}^2} = 0$$

if we stipulate that $\zeta_y = \zeta_z$ and $\zeta_z/\zeta_x = \sqrt{1 - M_\infty^2} = \beta_0$.

Flow boundary conditions are satisfied on both the compressible and incompressible planes when

$$(Q_\infty)_{\text{inc}} = Q_\infty, \quad \alpha_{\text{inc}} = \beta_0 \alpha \quad \text{and} \quad \zeta_z^2 = \zeta_x \zeta_\phi$$

Normally the transformation uses $\zeta_x = 1$, $\zeta_y = \zeta_z = \beta_0$, and $\zeta_\phi = \beta_0^2$. The compressible and incompressible wings are then geometrically related as follows;

$$\frac{\hat{t}_{\text{inc}}}{\hat{t}} = \frac{\hat{h}_{\text{inc}}}{\hat{h}} = \beta_0, \quad (c_t/c_r)_{\text{inc}} = c_t/c_r, \quad (A_R)_{\text{inc}} = \beta_0 A_R$$

$$\tan \Lambda_{\text{inc}} = \frac{1}{\beta_0} \tan \Lambda, \quad \text{and} \quad (A_R)_{\text{inc}} \tan \Lambda_{\text{inc}} = A_R \tan \Lambda$$

Shapiro (1953) points out that, for finite wing flows, it is not possible to formulate an equivalent three-dimensional version of the Prandtl–Glauert rule since the local pressure coefficients for two affinely related wings are not just proportional to the airfoil thickness ratio. Thus, the wings in the incompressible and compressible flows will not be of identical geometry. The pressure, lift, and moment coefficients are related as follows

$$\frac{C_p}{(C_p)_{\text{inc}}} = \frac{C_L}{(C_L)_{\text{inc}}} = \frac{C_m}{(C_m)_{\text{inc}}} = \frac{1}{\beta_0^2} \quad \text{and} \quad \frac{C_{L\alpha}}{(C_{L\alpha})_{\text{inc}}} = \frac{1}{\beta_0}$$

The wing lift curve slopes are related in the following manner (Hilton 1951):

$$\frac{(C_{L\alpha})}{(C_{L\alpha})_{\text{inc}}} = \frac{\pi A_R + (c_{l\alpha})_{\text{inc}}(1 + \tau)}{\pi A_R \beta_0 + (c_{l\alpha})_{\text{inc}}(1 + \tau)}$$

It is worth noting that the tail surfaces of an aircraft become less effective than the main wing at high Mach numbers, due to their smaller aspect ratio, and thus may have an impact on the stability of the aircraft.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

Mach Cone Flow Classification

As discussed in Subsection 10.9, a source of pressure disturbance in a supersonic flow can only influence the flow within a downstream conical volume extending from the source and whose curved surface is defined by Mach waves. In a similar fashion, the source of disturbance can only be influenced by the flow within a Mach cone facing upstream. Identifying the critical points on a wing from which Mach cones emanate is a important part of all supersonic finite wing theories. This method is used to identify whether a portion of the wing lies in either a subsonic or a supersonic flow and whether it behaves as though the flow is two dimensional or, as discussed below, conical (or cone symmetric). Figure 10.62, in conjunction with Table 10.7, documents examples of this technique.

Any edge of the wing perimeter can be classified as subsonic or supersonic based on whether the component of the free-stream velocity normal to the edge is respectively less than or greater than the speed of sound. Thus, subsonic edges will lie behind a Mach line, while supersonic edges will lie in front. Referring to Figure 10.62, the following edge number can be defined:

$$m = \frac{\tan \varepsilon}{\tan \sigma_M} = \beta_0 \tan \varepsilon$$

where subsonic edges have $m < 1$ and supersonic edges have $m > 1$. As illustrated in Figure 10.62, subsonic leading and trailing edges have high and zero lift loadings respectively, whilst supersonic edges have finite lift loadings.

The Method of Supersonic Singularities

(*Applications—three-dimensional, inviscid, isentropic, irrotational, supersonic compressible flow past a finite wing of arbitrary planform and aerodynamic twist, thin airfoil section and set at low angles of attack*) The linearized governing equations for the perturbation velocity potential in both subsonic and supersonic flow can be solved using Green's theorem along with the concept of source, doublet, and vortex potential functions. When the flow is supersonic, the influence a singularity has on the rest of the flow field is confined to the volume within the Mach cone. Some of the solutions contained within this section were obtained using these methods (see Heaslet and Lomax 1955 for further details).

Conical (or Cone-Symmetric) Flow

(*Applications—three-dimensional, inviscid, isentropic, irrotational, supersonic compressible flow past a finite wing whose planform perimeter consists of straight line segments, no aerodynamic twist, thin airfoil section and set at low angles of attack*) The conical flow field was originally conceived by Busemann (1947) and is defined as a flow in which all the flow properties are uniform along rays emanating from a single point. By using the conical flow model along with a suitable coordinate transformation, Busemann converted the linearized supersonic perturbation velocity potential equation into the Laplace equation in polar coordinates. As with the method of singularities, many of the solutions contained within this section were obtained using conical flow methods. Further details can be found in Lagerstrom (1950).

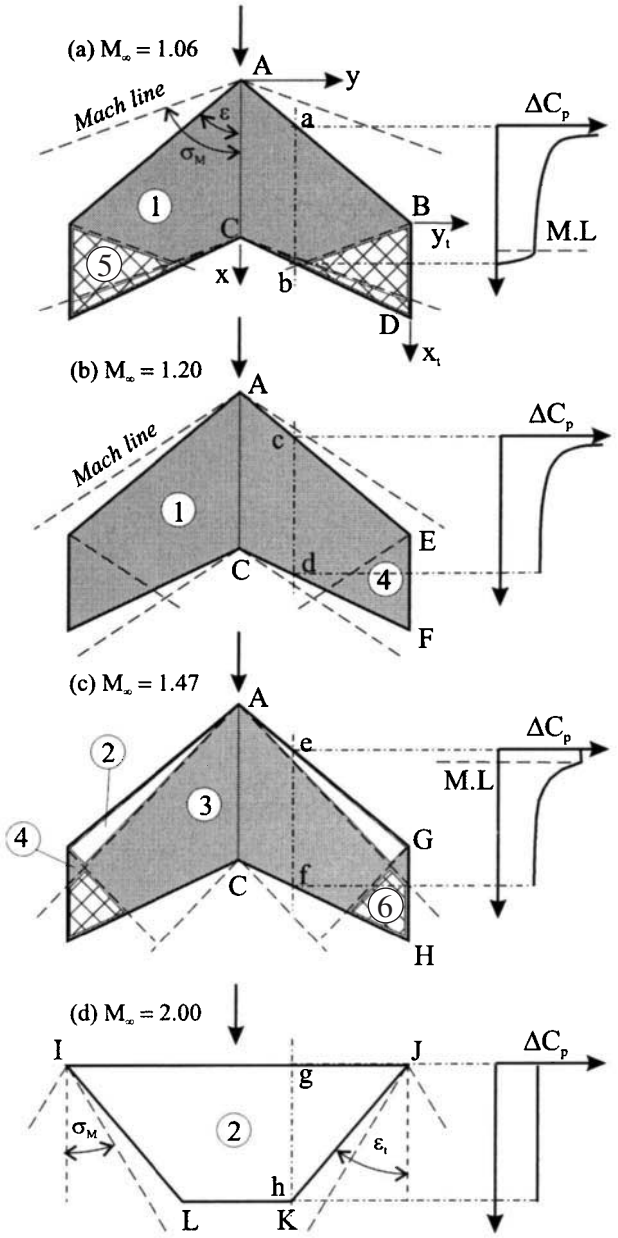


FIGURE 10.62 Examples of Mach cone zoning.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

TABLE 10.7 Mach Cone Zoning

Type of flow	Zones
Subsonic conical flow	1, 3, 4
2D supersonic flow	2
Superposed flow	5, 6

Effect of Wing Sweep

Figure 10.63 shows part of a infinite wing which has been swept back by an angle Λ. There are two types of swept-back wing: bent-back, where the wing is rotated such that the selected airfoil profile remains in a plane normal to the leading edge, and sheared-back, where the chosen profile remains parallel to the free-stream velocity. The effect of sweep relies on the fact that it is only the velocity component normal to the leading edge ($M_\infty \cos \Lambda$) which controls the pressure distribution over the wing section (Jones 1945). Thus, even though the free-stream Mach number may be equal to unity, the value of $M_\infty \cos \Lambda$ may be small enough to avoid shock wave formation. It is worth noting that shear-back increases the airfoil thickness

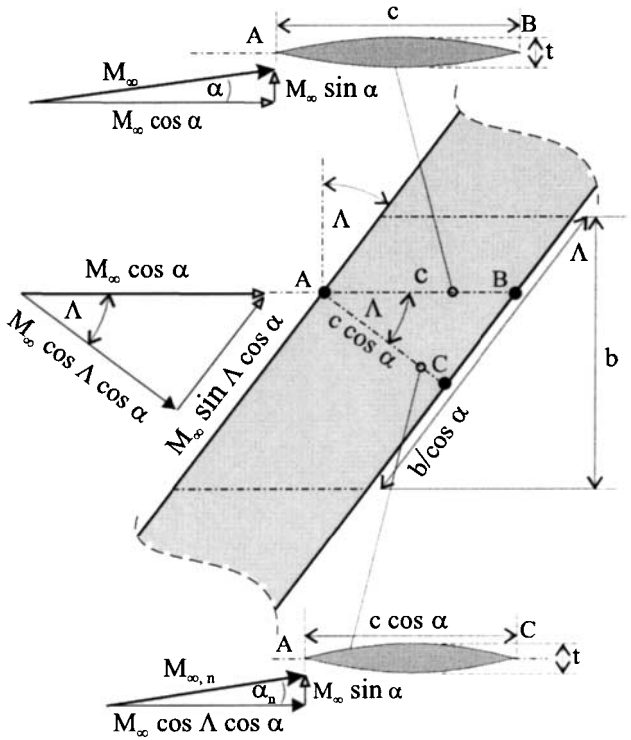


FIGURE 10.63 Airflow components over a swept wing.

normal to the leading edge and is thus not as effective as bend-back. Referring to Figure 10.63, the following relationships are obtained for an infinite wing:

$$M_{\infty,n} = M_\infty \sqrt{1 - \sin^2 \Lambda \cos^2 \alpha} \approx M_\infty \cos \alpha \tag{10.47}$$

$$\alpha_n = \arctan\left(\frac{\tan \alpha}{\cos \Lambda}\right) \approx \frac{\alpha}{\cos \Lambda} \tag{10.48}$$

$$c_l = (c_l)_n \left(\frac{M_{\infty,n}}{M_\infty}\right)^2 \approx (c_l)_n \cos^2 \alpha \tag{10.49}$$

$$c_{l\alpha} = (c_{l\alpha})_n \cos \alpha$$

$$c_d = (c_d)_n \cos \Lambda \left(\frac{M_{\infty,n}}{M_\infty}\right)^2 + c_{d,f} \tag{10.50}$$

$$\approx (c_{d,p})_n \cos^3 \alpha + c_{d,f}$$

$$\hat{t}_n = \hat{t}/\cos \Lambda \tag{10.51}$$

When the free-stream flow is subsonic, an equation can be derived (Schlichting and Truckenbrodt 1979) which includes the effects of both compressibility and sweep on the local pressure coefficient and is given by

$$C_p = \left(\frac{\cos \Lambda}{\sqrt{1 - M_\infty^2 \cos^2 \Lambda}}\right) (C_{p,n})_{\text{inc}} \tag{10.52}$$

where $(C_{p,n})_{\text{inc}}$ is the incompressible flow pressure coefficient obtained for the airfoil profile normal to the wing leading edge at an angle of attack equal to $\alpha/\cos \Lambda_i$ and referenced to Q_∞. Gothert's rule for a swept infinite wing gives the relationship

$$\cos \Lambda_i = \frac{\beta_0 \cos \Lambda}{\sqrt{1 - M_\infty^2 \cos^2 \Lambda}}$$

Using the Prandtl–Glauert compressibility rule along with equations (10.44) and (10.52) allows the effect of sweep to be clearly illustrated in Figure 10.64.

The effect of sweep when the flow is supersonic can be illustrated by considering an infinite sheared-back wing which has a diamond airfoil profile. Combining Ackeret's theory with equations (10.47)–(10.51) results in the following relationships:

$$c_l = \frac{4\alpha \cos \Lambda}{\sqrt{M_\infty^2 \cos^2 \Lambda - 1}}$$

$$c_d = \frac{4(\alpha^2 + \hat{t}^2) \cos \Lambda}{\sqrt{M_\infty^2 \cos^2 \Lambda - 1}} + c_{d,f}$$

Using these equations, the beneficial effect sweep has on the lift-to-drag ratio is clearly shown in Figure 10.65 where $c_{d,f} = 0.006$. The main advantage of wing sweep in supersonic flow is the reduction in wave drag due to the reduction in oblique shock wave strength.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

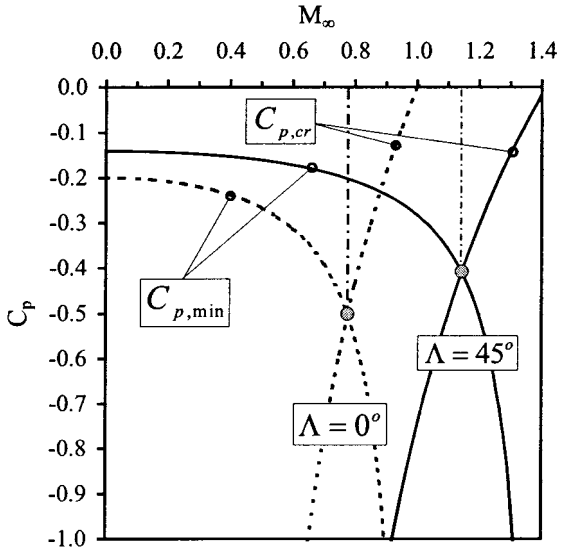


FIGURE 10.64 Swept wing critical Mach number.

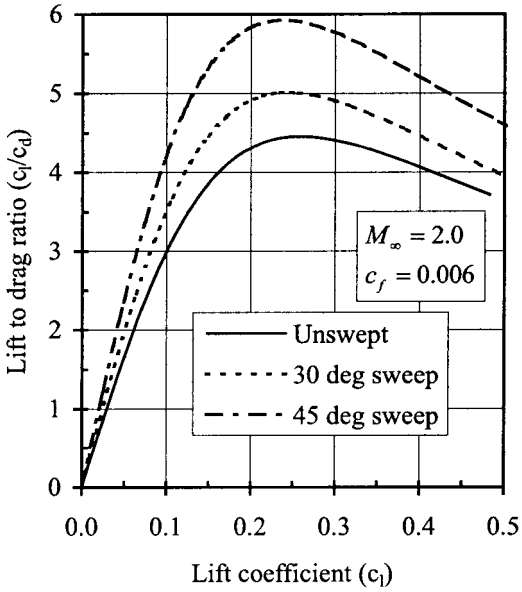


FIGURE 10.65 Effect of sweep on supersonic lift to drag ratio.

Surface Pressure Distributions—Zone 1 Subsonic Leading Edge ($m < 1$)

Figure 10.66 shows the flow over a wing with a flat plate profile and subsonic leading edges. The spanwise pressure distribution over the upper surface can be computed using conical flow theory and is given by

$$\frac{C_p}{C_{p,pl}} = \frac{m}{E(m)} \frac{1}{\sqrt{1 - r_0^2}}$$

where

$$r_0 = \frac{y}{x \tan \varepsilon} = \frac{\tan \theta}{\tan \varepsilon}$$

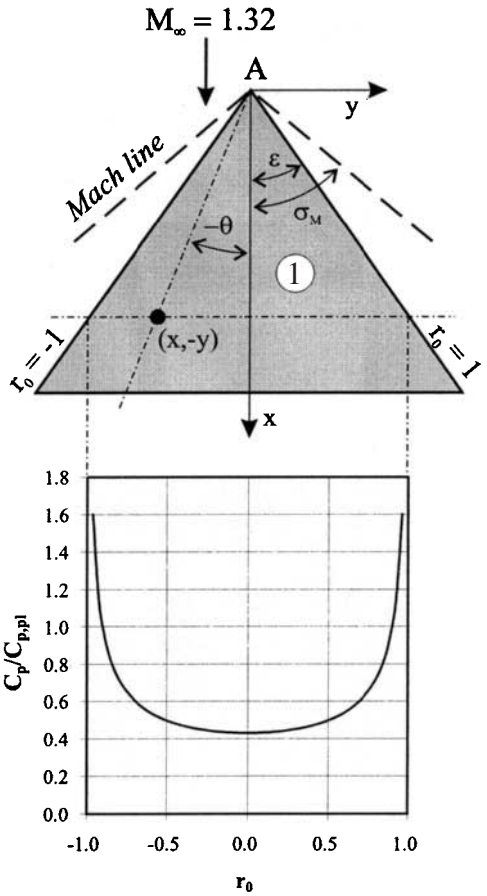


FIGURE 10.66 Subsonic leading edge spanwise distribution of pressure coefficient.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

and

$$C_{p,pl} = \mp \frac{2\alpha}{\beta_1} \quad \text{(negative denotes upper surface)}$$

$$E(m) = \int_0^{\pi/2} \sqrt{1 - (1 - m^2)\sin^2\theta}\, d\theta$$

$$= \frac{\pi}{2}\left\{1 - \left(\frac{1}{2}\right)^2 (1 - m^2)^2 - \left(\frac{1 \times 3}{2 \times 4}\right)^2 (1 - m^2)^4 - \cdots\right\}$$

$E(m)$ is known as an elliptic integral of the second kind. The mean value of pressure over the span is given by

$$\frac{\overline{C}_p}{C_{p,pl}} = \frac{m}{E(m)}\frac{\pi}{2}$$

Surface Pressure Distributions—Zone 2 Supersonic Leading Edge ($m > 1$)

Figure 10.67 illustrates the flow over a supersonic edge. The pressure distribution over a swept-back inclined flat plate can be computed directly from Ackeret's theory taking into account the sweep angle, i.e.,

$$C_p = \mp \frac{2 \sin \varepsilon}{\sqrt{M_\infty^2 \sin \varepsilon - 1}}\, \alpha$$

With reference to the flow over an unswept flat plate, the pressure distribution over a swept supersonic edge has the form

$$\frac{C_p}{C_{p,pl}} = \frac{m}{\sqrt{m^2 - 1}} \tag{10.53}$$

Surface Pressure Distributions—Zone 3 Supersonic Leading Edge, Region behind Mach Line ($m > 1$)

Figure 10.67 illustrates the variation in pressure coefficient after the Mach wave which is downstream of a supersonic leading edge. The variation is given by

$$\frac{C_p}{C_{p,pl}} = \frac{2}{\pi}\frac{m}{\sqrt{m^2 - 1}}\left(\arccos\sqrt{\frac{1 - r_1^2}{m^2 - r_1^2}}\right) \tag{10.54}$$

where

$$r_1 = \frac{y}{x \tan \sigma_M} = \frac{\tan \theta}{\tan \sigma_M}$$

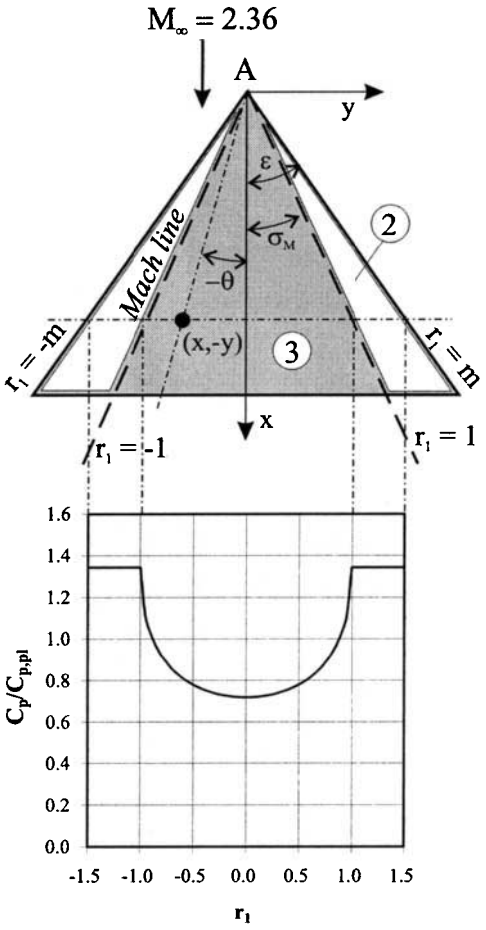


FIGURE 10.67 Supersonic leading edge spanwise distribution of pressure coefficient.

Surface Pressure Distributions—Zone 4 Subsonic Side Edge Bounded by a Supersonic Region

A side edge is defined as an edge, on the wing perimeter, which is parallel to the free-stream flow, e.g., side *BD* in Figure 10.62. When the wing leading edge (flat plate profile) is swept, the pressure coefficient is given by

$$\frac{C_p}{C_{p,pl}} = \frac{1}{\pi} \frac{m}{\sqrt{m^2 - 1}} \left\{ \arccos \left(1 + 2r_3 \frac{m + 1}{m - r_3} \right) \right\}$$

where

$$r_3 = \frac{y_t}{x_t \tan \sigma_M} = \frac{\tan \theta_t}{\tan \sigma_M}$$

When the side edge is part of an unswept wing, the pressure coefficient becomes

$$\frac{C_p}{C_{p,pl}} = \frac{1}{\pi} \arccos (1 + 2r_3)$$

Surface Pressure Distributions—Nonconical Zones

Zones which are influenced by more than one Mach cone are indicated by the crosshatched regions in Figure 10.62. These regions involve nonconical flow fields and thus are not covered by the solutions discussed above. However, solutions can be found using a superposition procedure as described by Jones and Cohen (1957).

Airloads—Delta Wing with Subsonic Leading Edge

Once the surface pressure distributions are known, the aerodynamic forces and moments can be calculated. For a delta wing with subsonic leading edges the wing lift coefficient is given by

$$C_L = \frac{m}{E(m)} \frac{2\pi}{\beta_0} \alpha$$

The drag of a the wing with a subsonic leading edge is composed of a component due to the lift and the suction force at the leading edge. The drag coefficient is given by

$$C_D = \frac{C_L^2}{\pi A_R} \{2E(m) - \sqrt{1 - m^2}\}$$

Airloads—Delta Wing with Supersonic Leading Edge

Interestingly, the mean spanwise pressure coefficient, as obtained from equations (10.53) and (10.54), turns out to be exactly $c_{p,pl}$, which results in both the lift and drag coefficients being equal to the two-dimensional results below:

$$C_L = c_l = \frac{4}{\beta_1} \alpha$$

$$C_D = c_d = \frac{4}{\beta_1} \alpha^2 = \frac{mC_L^2}{A_R}$$

Airloads—Rectangular and Trapezoidal Wings

Figure 10.62(d) illustrates a wing with a trapezium planform which is described as possessing raked tips. The values of lift and wave drag coefficients, expressed as a ratio of Ackeret's two-dimensional values, are given by

$$\frac{C_L}{c_l} = \frac{C_{D,w}}{c_{d,w}} = 1 - \frac{1 - r_4}{2A_R\beta_1}$$

where $r_4 = (\tan \varepsilon_t)/(\tan \sigma_M)$, $c_l = 4\alpha/\beta_1$, and $c_{d,w}$ is given by equation (10.54). This formula is only applicable if the two side edge Mach cones do not overlap, i.e., when $A_R\beta_1 > 2$.

REFERENCES

Abbott, I. H. and von Doenhoff, A. E. 1959. *Theory of Wing Sections,* Dover, New York.

Abbott, I. H., von Doenhoff, A. E. and Stivers, L. S. 1945. *Summary of Airfoil Data,* NACA Rpt. 824.

Ackeret, J. 1925. *Air Forces on Airfoils Moving Faster than Sound Velocity,* NACA TM-317.

Ames Research Staff. 1947. *Notes and Tables for Use in the Analysis of Supersonic Flow,* NACA TN-1428.

Ames Research Staff. 1953. *Equations, Tables and Charts for Compressible Flow,* NACA Rpt. 1135.

Anderson, J. D. 1990. *Modern Compressible Flow,* McGraw-Hill, New York.

Anderson, J. D., 1991. *Fundamentals of Aerodynamics,* McGraw-Hill, New York.

Anderson, R. F., 1936. *Determination of the Characteristics of Tapered Wings,* NACA Rpt. 572.

Busemann, A. 1947. *Infinitesimal Conical Supersonic Flow,* NACA TM-110.

Chappell, P. D., 1968. "Flow Separation and Stall Characteristics of Plane Constant-Section Wings in Subcritical Flow," *Journal of the Royal Aeronautical Society,* vol. 72.

Diederich, F. W., 1952. *A Simple Approximate Method for Calculating Spanwise Lift Distributions and Aerodynamic Influence Coefficients at Subsonic Speeds,* NACA TN-2751.

Eskinazi, S. 1967. *Vector Mechanics of Fluids and Magnetofluids,* Academic Press, New York.

Falkner, V. M. 1943. *The Calculation of Aerodynamic Loading on Surfaces of Any Shape,* R&M-1910, British A.R.C.

Ferri, A. 1940. *Experimental Results with Airfoils Tested in the High Speed Tunnel at Guidonia,* NACA TM-946.

Gault, D. E. 1957. *A Correlation of Low-Speed, Airfoil-Section Stalling Characteristics with Reynolds Number and Airfoil Geometry,* NACA TN-3963.

Heaslet, M. A. and Lomax, H. 1955. "Supersonic and Transonic Small Perturbation Theory," in *General Theory of High Speed Aerodynamics,* ed. W. R. Sears, High Speed Aerodynamics and Jet Propulsion 6, Princeton University Press, Princeton, NJ.

Hilton, W. F. 1951. *High Speed Aerodynamics,* Longmans, London.

Houghton, E. L. and Brock, A. E. 1977. *Aerodynamics for Engineering Students,* Edward Arnold, London.

Houghton, E. L. and Brock, A. E. 1979. *Tables for the Compressible Flow of Dry Air,* Edward Arnold, London.

Japikse, D. and Baines, N.C. 1994. *Introduction to Turbomachinery,* Concepts ETI Inc. White River Junction, VT, and Oxford University Press, Oxford.

Jones, R. T. 1941. *Correction of the Lifting-Line Theory for the Effect of the Chord,* NACA TN-817.

Jones, R. T. 1945. *Wing Plan Forms for High-Speed Flight,* NACA Rpt. 863.

Jones, R. T. and Cohen, D. 1957. "Aerodynamics of Wings at High Speeds," in High Speed Aerodynamics and Jet Propulsion 7, ed. A. F. Donovan and H. R. Lawrence, *Aerodynamic Components of Aircraft at High Speeds,* Princeton University Press, Princeton, NJ.

Katz, J. and A. Plotkin. 1991. *Low-Speed Aerodynamics: From Wing Theory to Panel Methods,* McGraw-Hill, New York.

Karamcheti, K. 1980. *Principles of Ideal-Fluid Aerodynamics,* Krieger, Malabar, FL.

Lagerstrom, P. A. 1950. *Linearized Supersonic Theory of Conical Wings,* NACA TN-1685.

Loftin, L. K., Jr. and Cohen, K. S. 1948. *An Evaluation of the Characteristics of a 10 Percent Thick NACA 66-series Airfoil Section with a Special Mean Camber Line Designed to Produce a High Critical Mach Number,* NASA TN-1633.

Lugt, H. J. 1995. *Vortex Flow in Nature and Technology,* Krieger, Malabar, FL.

Nitzberg, G. E. and Crandall, S. M. 1952. *A Comparative Examination of Some Measurements of Airfoil Section Lift and Drag at Supercritical Speeds,* NACA TN-2825.

Niven, A. J., 1988. "An Experimental Investigation into the Influence of Trailing-Edge Separation on an Aerofoil's Dynamic Stall Performance" (Ph.D Thesis, University of Glasgow).

McCullough, G. B. and Gault, D. E. 1951. *Examples of Three Representative Types of Airfoil-Section Stall at Low Speed,* NACA TN-2502.

Oosthuizen, P. H. and Carscallen, W. E. 1997. *Compressible Fluid Flow,* McGraw-Hill, New York.

Prandtl, L. and O. G. Tietjens. 1934. *Fundamentals of Hydro and Aerodynamics,* Dover, New York.

Schlichting, H. 1987. *Boundary Layer Theory.* McGraw-Hill, New York.

Schlichting, H. and Truckenbrodt, E. 1979. *Aerodynamics of the Airplane,* McGraw-Hill, New York.

Shapiro, A. H. 1953. *Compressible Fluid Flow,* Ronald Press, New York.

Sommerfeld, A. 1950. *Mechanics of Deformable Bodies,* Academic Press, London.

Stack, J., Lindsey, W. F., and Littell, R. E. 1938. *The Compressibility Burble and the Effect of Compressibility on Pressures and Forces Acting on an Airfoil,* NACA Rpt. 646.

Theodersen, T. 1931. *Theory of Wing Sections of Arbitrary Shape,* NACA Rpt. 411.

Theodorsen, T. and Garrick, I. E. 1932. *General Potential Theory of Arbitrary Wing Sections,* NACA Rpt. 452.

Vallentine, H. R. 1959. *Applied Hydrodynamics*, Butterworths, London.

Von Doenhoff, A. E., Stivers, L. S., and O'Connor, J. M. 1947. *Low-Speed Tests of Five NACA 66-series Airfoils Having Mean Lines Designed to Give High Critical Mach Numbers,* NASA TN-1276.

Ward, J. W., 1963. "The Behaviour and Effects of Laminar Separation Bubbles on Aerofoils in Incompressible Flow," *Journal of the Royal Aeronautical Society,* vol. 67.

Weissinger, J. 1947. *The Lift Distribution of Swept-Back Wings,* NACA TM-1120.

Whitcomb, R. T. and Clark, L. R. 1965. *An Airfoil Shape for Efficient Flight at Supercritical Mach Numbers,* NASA TMX-1109.

White, F. 1979. *Fluid Mechanics,* McGraw-Hill, New York.

White, F. 1991. *Viscous Fluid Flow,* McGraw-Hill, New York.

Wilcox, D. C., 1994. *Turbulence Modeling for CFD,* DCW Industries Inc., La Cañada, CA.

PART 2

AIRPLANE PERFORMANCE

Trevor Young

NOTATION

a	speed of sound
a	acceleration
a_c	centripetal acceleration
A_R	aspect ratio
c_t	specific fuel consumption of turbojet or turbofan airplane, defined in terms of mass flow rate
c_t'	specific fuel consumption of turbojet or turbofan airplane, defined in terms of weight flow rate
c_p	specific fuel consumption of piston or turboprop airplane, defined in terms of mass flow rate
c_p'	specific fuel consumption of piston or turboprop airplane, defined in terms of weight flow rate
C_L	lift coefficient
C_D	drag coefficient
C_{D_i}	lift-dependent drag coefficient
C_{D_0}	lift-independent drag coefficient (zero-lift drag coefficient)
CAS	calibrated airspeed
D	drag (force)
e	Oswald efficiency factor
EAS	equivalent airspeed
f_{acc}	acceleration factor (for climb or descent)
FAR	Federal Aviation Regulations
g	acceleration due to gravity
h	height
h_{SC}	screen height
H	geopotential height
ICAO	International Civil Aviation Organization
ISA	International Standard Atmosphere
JAA	Joint Aviation Authorities
JAR	Joint Aviation Requirements
K	lift-dependent drag factor
L	lapse rate
L	lift (force)
LRS	long-range speed
m	mass

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

M	Mach number
MRS	maximum range speed
MTOW	maximum takeoff weight
n	load factor
N	rotational speed of engine
p	pressure
P	power
P_D	drag power (power required)
P_S	shaft power
P_T	thrust power (power available)
q	dynamic pressure
Q	mass of fuel burned per unit time (fuel flow rate)
Q'	weight of fuel burned per unit time (fuel flow rate)
r	radius of turn
r_a	specific air range (SAR)
R	range
R	gas constant
ROC	rate of climb
ROD	rate of descent
ROS	rate of sink
s	ground distance
S	wing reference area
SAR	specific air range
SFC	specific fuel consumption
t	time
T	thrust
T	temperature
TAS	true airspeed
TOW	takeoff weight
V	true airspeed (TAS)
V_e	equivalent airspeed (EAS)
V_S	stall speed
V_1	decision speed in takeoff
W	weight (force)
x	still air distance

Greek

δ	relative pressure
ϕ	angle of bank

γ flight path angle
γ ratio of specific heats of air
λ ground effect factor
η_p propeller efficiency
μ_B coefficient of braking friction
μ_R coefficient of rolling friction
θ relative temperature
ρ density
σ relative density
Ω rate of turn

10.12 STANDARD ATMOSPHERE AND HEIGHT MEASUREMENT

International Standard Atmosphere (ISA)

The *International Standard Atmosphere* (ISO 2533, 1975) is an idealized model of the atmosphere which by international agreement is used for aircraft performance analysis and operation. The ISA describes a hypothetical vertical distribution of *temperature, pressure,* and *density*, which greatly simplifies numerical analysis. The ISA is identical to the International Civil Aviation Organization Standard Atmosphere (ICAO 1993) and the U.S. Standard Atmosphere for heights up 32 km. For aviation purposes, the ISA may be defined in two regions:

- The *troposphere* extends from the ISA datum height (ISA sea level) to the tropopause, at a geopotential height of 11,000 m (36,089.24 ft). The temperature is assumed to be exactly 15°C at the sea-level datum and to decrease linearly with altitude at a lapse rate of 6.5°C per 1000 m (exactly). The ISA datum is not defined by a geographical height, but rather by a reference pressure of 101,325 N/m^2 (2,116.217 lb/ft^2).
- The lower *stratosphere* is the region above the tropopause to a geopotential height of 20,000 m (65,616.80 ft) in which it is assumed that the temperature is a constant −56.5°C (exactly).

It is customary to define the air temperature (T), pressure (p), and density (ρ) as ratios of the standard sea-level (ISA datum) values and to use the subscript o to denote the standard sea-level conditions. Standard values are given in Table 10.8. The following definitions are used:

Relative temperature (temperature ratio):

$$\theta = \frac{T}{T_0} \quad \text{(absolute temperatures)} \tag{10.55}$$

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

TABLE 10.8 Standard Values of the ISA

Standard values	Symbol	SI units	Equivalent
Temperature at sea-level	T_0	288.15 K	518.67 R
Lapse rate in troposphere	L	6.5 K per 1,000 m	1.9812 K per 1,000 ft 3.56616 R per 1,000 ft
Temperature of tropopause	T^*	216.65 K	389.97 R
Height of tropopause	H^*	11,000 m	36,089.24 ft
Sea-level pressure	p_0	101,325 N/m^2 1013.25 hPa	29.92126 in. Hg 2,116.217 lb/ft^2
Sea-level density	ρ_0	1.225 kg/m^3	0.002376892 slugs/ft^3
Gravitational acceleration	g_0	9.80665 m/s^2	32.17405 ft/s^2
Speed of sound at sea-level	a_0	340.294 m/s	1,116.450 ft/s 661.4786 kt
Gas constant	R	287.05287 m^2/s^2 K	3,089.811 ft^2/s^2 K 1,716.562 ft^2/s^2 R
Ratio of specific heats of air	γ	1.40	

Relative pressure (pressure ratio):

$$\delta = \frac{p}{p_0} \tag{10.56}$$

Relative density (density ratio)

$$\sigma = \frac{\rho}{\rho_0} \tag{10.57}$$

Temperature, density, and pressure are not independent, as is evident from the perfect gas law, which for arbitrary conditions may be written as:

$$\frac{p}{\rho} = RT \tag{10.58}$$

where R is the gas constant. It follows that the ratios σ, δ, and θ are related by the following expression:

$$\sigma = \frac{\delta}{\theta} \tag{10.59}$$

Temperature, Pressure, and Density in the Standard Atmosphere

The pressure at any altitude may be determined by the following equation, as the sea-level conditions and the temperature are known:

$$\int \frac{1}{p}\, dp = -\int \frac{g}{RT}\, dh \tag{10.60}$$

The temperature is a linear function of height in the troposphere, and it has a constant value in the stratosphere. The integral is evaluated for these two regions by setting the gravitational acceleration equal to the standard sea-level value. For the troposphere, the integration is performed from sea level to altitude H, yielding the pressure ratio. In the stratosphere the integration is performed from the tropopause to the altitude H. The density ratio is determined from the perfect gas law (equation [10.58]). The resulting equations are given in Table 10.9. Tabulated values of the ISA are given in Table 10.10 as a function of geopotential height in feet, the unit of measure used internationally for aircraft operations. Note that the ISA is defined in terms of SI units and values in British units are obtained by conversion, which can lead to rounding error discrepancies.

Off-Standard Atmospheric Conditions

Whereas the ISA is essential for calculations of airplane performance prediction, flight tests and aircraft operations will generally not be conducted at atmospheric conditions that comply exactly with the ISA model. The evaluation of flight test data must therefore take into account actual flight conditions. Frequently it is required to determine the density ratio, given that the ambient temperature (T_{test}) was measured at a pressure height of H_{test}. The procedure is as follows:

1. From ISA tables, determine δ at the pressure height H_{test}.
2. Determine the relative temperature:

$$\theta = \frac{T_{\text{test}}}{T_0} \quad \text{(absolute temperatures)}$$

3. Determine σ from equation (10.59).

TABLE 10.9 Equations for the ISA

	In the troposphere	Tropopause	In the stratosphere
Relative temperature	$\theta = \frac{L}{T_0} H$	$\theta^* = 0.751865$	$\theta = \theta^*$
Relative pressure	$\delta = \left[1 - \frac{LH}{T_0}\right]^{g_0/RL}$	$\delta^* = 0.223361$	$\frac{\delta}{\delta^*} = e^{(-g_0/RT^*)(H-H^*)}$
Relative density	$\sigma = \left[1 - \frac{LH}{T_0}\right]^{(g_0/RL)-1}$ Alternatively, if θ and δ are known, use $\sigma = \frac{\delta}{\theta}$	$\sigma^* = 0.297076$	$\frac{\sigma}{\sigma^*} = e^{(-g_0/RT^*)(H-H^*)}$ Alternatively, if θ and δ are known, use $\sigma = \frac{\delta}{\theta}$

Note: The conditions at the tropopause are denoted by superscript asterisk (*).

TABLE 10.10 International Standard Atmosphere (−2,000 ft to 60,000 ft)

H ft	T K	θ	p N/m^2	δ	ρ kg/m^3	σ	a m/s
−2 000	292.11	1.013 75	108 866	1.074 42	1.298 3	1.059 85	342.63
−1 800	291.72	1.012 38	108 092	1.066 78	1.290 8	1.053 74	342.39
−1 600	291.32	1.011 00	107 322	1.059 19	1.283 4	1.047 66	342.16
−1 400	290.92	1.009 63	106 557	1.051 64	1.276 0	1.041 61	341.93
−1 200	290.53	1.008 25	105 797	1.044 13	1.268 6	1.035 59	341.69
−1 000	290.13	1.006 88	105 041	1.036 67	1.261 2	1.029 59	341.46
−800	289.73	1.005 50	104 289	1.029 25	1.253 9	1.023 62	341.23
−600	289.34	1.004 13	103 541	1.021 87	1.246 7	1.017 68	341.00
−400	288.94	1.002 75	102 798	1.014 54	1.239 4	1.011 76	340.76
−200	288.55	1.001 38	102 059	1.007 25	1.232 2	1.005 87	340.53
0	288.15	1.000 00	101 325	1.000 00	1.225 0	1.000 00	340.29
200	287.75	0.998 62	100 595	0.992 79	1.217 8	0.994 16	340.06
400	287.36	0.997 25	99 869	0.985 63	1.210 7	0.988 35	339.83
600	286.96	0.995 87	99 147	0.978 51	1.203 6	0.982 56	339.59
800	286.57	0.994 50	98 430	0.971 43	1.196 6	0.976 80	339.36
1 000	286.17	0.993 12	97 717	0.964 39	1.189 6	0.971 06	339.12
1 200	285.77	0.991 75	97 008	0.957 39	1.182 6	0.965 35	338.89
1 400	285.38	0.990 37	96 303	0.950 43	1.175 6	0.959 67	338.65
1 600	284.98	0.989 00	95 602	0.943 52	1.168 7	0.954 01	338.42
1 800	284.58	0.987 62	94 905	0.936 64	1.161 8	0.948 38	338.18
2 000	284.19	0.986 25	94 213	0.929 81	1.154 9	0.942 77	337.95
2 200	283.79	0.984 87	93 525	0.923 02	1.148 1	0.937 19	337.71
2 400	283.40	0.983 50	92 840	0.916 26	1.141 3	0.931 64	337.47
2 600	283.00	0.982 12	92 160	0.909 55	1.134 5	0.926 10	337.24
2 800	282.60	0.980 75	91 484	0.902 88	1.127 7	0.920 60	337.00
3 000	282.21	0.979 37	90 812	0.896 24	1.121 0	0.915 12	336.77
3 200	281.81	0.978 00	90 143	0.889 65	1.114 3	0.909 66	336.53
3 400	281.41	0.976 62	89 479	0.883 09	1.107 7	0.904 23	336.29
3 600	281.02	0.975 25	88 819	0.876 58	1.101 1	0.898 82	336.06
3 800	280.62	0.973 87	88 163	0.870 10	1.094 5	0.893 44	335.82
4 000	280.23	0.972 50	87 511	0.863 66	1.087 9	0.888 09	335.58
4 200	279.83	0.971 12	86 862	0.857 26	1.081 4	0.882 75	335.34
4 400	279.43	0.969 75	86 218	0.850 90	1.074 9	0.877 45	335.11
4 600	279.04	0.968 37	85 577	0.844 58	1.068 4	0.872 16	334.87
4 800	278.64	0.967 00	84 940	0.838 29	1.062 0	0.866 90	334.63
5 000	278.24	0.965 62	84 307	0.832 05	1.055 5	0.861 67	334.39
5 200	277.85	0.964 25	83 678	0.825 84	1.049 2	0.856 46	334.16
5 400	277.45	0.962 87	83 053	0.819 67	1.042 8	0.851 27	333.92
5 600	277.06	0.961 50	82 431	0.813 53	1.036 5	0.846 11	333.68
5 800	276.66	0.960 12	81 814	0.807 44	1.030 2	0.840 97	333.44
6 000	276.26	0.958 75	81 200	0.801 38	1.023 9	0.835 86	333.20
6 200	275.87	0.957 37	80 589	0.795 35	1.017 7	0.830 77	332.96
6 400	275.47	0.956 00	79 983	0.789 37	1.011 5	0.825 70	332.72
6 600	275.07	0.954 62	79 380	0.783 42	1.055 3	0.820 66	332.48
6 800	274.68	0.953 25	78 781	0.777 51	0.999 16	0.815 64	332.24

TABLE 10.10 International Standard Atmosphere (−2,000 ft to 60,000 ft) (*Continued*)

H ft	T K	θ	p N/m^2	δ	ρ kg/m^3	σ	a m/s
7 000	274.28	0.951 87	78 185	0.771 63	0.993 04	0.810 64	332.00
7 200	273.89	0.950 50	77 594	0.765 79	0.986 95	0.805 67	331.76
7 400	273.49	0.949 12	77 005	0.759 98	0.980 89	0.800 72	331.52
7 600	273.09	0.947 75	76 421	0.754 21	0.974 85	0.795 80	331.28
7 800	272.70	0.946 37	75 840	0.748 48	0.968 85	0.790 90	331.04
8 000	272.30	0.945 00	75 262	0.742 78	0.962 87	0.786 02	330.80
8 200	271.90	0.943 62	74 689	0.737 12	0.956 92	0.781 16	330.56
8 400	271.51	0.942 25	74 118	0.731 49	0.951 00	0.776 33	330.32
8 600	271.11	0.940 87	73 551	0.725 90	0.945 11	0.771 52	330.08
8 800	270.72	0.939 49	72 988	0.720 34	0.939 24	0.766 73	329.84
9 000	270.32	0.938 12	72 428	0.714 81	0.933 41	0.761 96	329.60
9 200	269.92	0.936 74	71 872	0.709 32	0.927 60	0.757 22	329.36
9 400	269.53	0.935 37	71 319	0.703 87	0.921 82	0.752 50	329.11
9 600	269.13	0.933 99	70 770	0.698 45	0.916 06	0.747 81	328.87
9 800	268.73	0.932 62	70 224	0.693 06	0.910 34	0.743 13	328.63
10 000	268.34	0.931 24	69 682	0.687 70	0.904 64	0.738 48	328.39
10 200	267.94	0.929 87	69 143	0.682 38	0.898 97	0.733 85	328.14
10 400	267.55	0.928 49	68 607	0.677 10	0.893 32	0.729 24	327.90
10 600	267.15	0.927 12	68 074	0.671 84	0.887 70	0.724 66	327.66
10 800	266.75	0.925 74	67 545	0.666 62	0.882 11	0.720 09	327.42
11 000	266.36	0.924 37	67 020	0.661 43	0.876 55	0.715 55	327.17
11 200	265.96	0.922 99	66 497	0.656 28	0.871 01	0.711 03	326.93
11 400	265.56	0.921 62	65 978	0.651 16	0.865 51	0.706 53	326.69
11 600	265.17	0.920 24	65 463	0.646 07	0.860 02	0.702 06	326.44
11 800	264.77	0.918 87	64 950	0.641 01	0.854 57	0.697 61	326.20
12 000	264.38	0.917 49	64 441	0.635 98	0.849 14	0.693 17	325.95
12 200	263.98	0.916 12	63 935	0.630 99	0.843 73	0.688 76	325.71
12 400	263.58	0.914 74	63 432	0.626 03	0.838 36	0.684 37	325.46
12 600	263.19	0.913 37	62 932	0.621 09	0.833 01	0.680 01	325.22
12 800	262.79	0.911 99	62 436	0.616 20	0.827 68	0.675 66	324.97
13 000	262.39	0.910 62	61 943	0.611 33	0.822 38	0.671 33	324.73
13 200	262.00	0.909 24	61 453	0.606 49	0.817 11	0.667 03	324.48
13 400	261.60	0.907 87	60 966	0.601 69	0.811 87	0.662 75	324.24
13 600	261.21	0.906 49	60 482	0.596 91	0.806 64	0.658 49	323.99
13 800	260.81	0.905 12	60 001	0.592 17	0.801 45	0.654 24	323.75
14 000	260.41	0.903 74	59 524	0.587 45	0.796 28	0.650 02	323.50
14 200	260.02	0.902 37	59 049	0.582 77	0.791 14	0.645 83	323.26
14 400	259.62	0.900 99	58 578	0.578 12	0.786 02	0.641 65	323.01
14 600	259.22	0.899 62	58 110	0.573 50	0.780 93	0.637 49	322.76
14 800	258.83	0.898 24	57 644	0.568 90	0.775 86	0.633 35	322.52
15 000	258.43	0.896 87	57 182	0.564 34	0.770 82	0.629 24	322.27
15 200	258.04	0.895 49	56 723	0.559 81	0.765 80	0.625 14	322.02
15 400	257.64	0.894 12	56 266	0.555 31	0.760 81	0.621 07	321.77
15 600	257.24	0.892 74	55 813	0.550 83	0.755 84	0.617 01	321.53
15 800	256.85	0.891 37	55 363	0.546 39	0.750 90	0.612 98	321.28

TABLE 10.10 International Standard Atmosphere (−2,000 ft to 60,000 ft) (*Continued*)

H ft	T K	θ	p N/m^2	δ	ρ kg/m^3	σ	a m/s
16 000	256.45	0.889 99	54 915	0.541 97	0.745 98	0.608 96	321.03
16 200	256.05	0.888 62	54 471	0.537 58	0.741 09	0.604 97	320.78
16 400	255.66	0.887 24	54 029	0.533 23	0.736 22	0.600 99	320.53
16 600	255.26	0.885 87	53 590	0.528 90	0.731 37	0.597 04	320.29
16 800	254.87	0.884 49	53 155	0.524 60	0.726 55	0.593 11	320.04
17 000	254.47	0.883 12	52 722	0.520 32	0.721 76	0.589 19	319.79
17 200	254.07	0.881 74	52 292	0.516 08	0.716 99	0.585 30	319.54
17 400	253.68	0.880 36	51 864	0.511 86	0.712 24	0.581 42	319.29
17 600	253.28	0.878 99	51 440	0.507 67	0.707 52	0.577 57	319.04
17 800	252.88	0.877 61	51 019	0.503 51	0.702 82	0.573 73	318.79
18 000	252.49	0.876 24	50 600	0.499 38	0.698 14	0.569 91	318.54
18 200	252.09	0.874 86	50 184	0.495 28	0.693 49	0.566 12	318.29
18 400	251.70	0.873 49	49 771	0.491 20	0.688 87	0.562 34	318.04
18 600	251.30	0.872 11	49 360	0.487 15	0.684 26	0.558 58	317.79
18 800	250.90	0.870 74	48 952	0.483 12	0.679 68	0.554 84	317.54
19 000	250.51	0.869 36	48 548	0.479 13	0.675 13	0.551 12	317.29
19 200	250.11	0.867 99	48 145	0.475 16	0.670 59	0.547 42	317.04
19 400	249.71	0.866 61	47 746	0.471 21	0.666 08	0.543 74	316.79
19 600	249.32	0.865 24	47 349	0.467 30	0.661 60	0.540 08	316.54
19 800	248.92	0.863 86	46 955	0.463 41	0.657 13	0.536 44	316.28
20 000	248.53	0.862 49	46 563	0.459 54	0.652 69	0.532 81	316.03
20 200	248.13	0.861 11	46 174	0.455 71	0.648 28	0.529 20	315.78
20 400	247.73	0.859 74	45 788	0.451 89	0.643 88	0.525 62	315.53
20 600	247.34	0.858 36	45 404	0.448 11	0.639 51	0.522 05	315.28
20 800	246.94	0.856 99	45 023	0.444 35	0.635 16	0.518 05	315.02
21 000	246.54	0.855 61	44 645	0.440 61	0.630 83	0.514 97	314.77
21 200	246.15	0.854 24	44 269	0.436 90	0.626 53	0.511 45	314.52
21 400	245.75	0.852 86	43 896	0.433 22	0.622 25	0.507 96	314.26
21 600	245.36	0.851 49	43 525	0.429 56	0.617 99	0.504 98	314.01
21 800	244.96	0.850 11	43 157	0.425 93	0.613 76	0.501 02	313.76
22 000	244.56	0.848 74	42 791	0.422 32	0.609 54	0.497 58	313.50
22 200	244.17	0.847 36	42 428	0.418 73	0.605 35	0.494 16	313.25
22 400	243.77	0.845 99	42 068	0.415 18	0.601 18	0.490 76	312.99
22 600	243.37	0.844 61	41 709	0.411 64	0.597 03	0.487 37	312.74
22 800	242.98	0.843 24	41 354	0.408 13	0.592 91	0.484 00	312.48
23 000	242.58	0.841 86	41 001	0.404 64	0.588 80	0.480 65	312.23
23 200	242.19	0.840 49	40 650	0.401 18	0.584 72	0.477 32	311.97
23 400	241.79	0.839 11	40 301	0.397 74	0.580 66	0.474 01	311.72
23 600	241.39	0.837 74	39 956	0.394 33	0.576 62	0.470 71	311.46
23 800	241.00	0.836 36	39 612	0.390 94	0.572 60	0.467 43	311.21
24 000	240.60	0.834 99	39 271	0.387 57	0.568 61	0.464 17	310.95
24 200	240.20	0.833 61	38 932	0.384 23	0.564 63	0.460 92	310.70
24 400	239.81	0.832 24	38 596	0.380 91	0.560 68	0.457 70	310.44
24 600	239.41	0.830 86	38 262	0.377 62	0.556 75	0.454 49	310.18
24 800	239.02	0.829 49	37 930	0.374 34	0.552 84	0.451 29	309.93

TABLE 10.10 International Standard Atmosphere (−2,000 ft to 60,000 ft)

H ft	T K	θ	p N/m^2	δ	ρ kg/m^3	σ	a m/s
25 000	238.62	0.828 11	37 601	0.371 09	0.548 95	0.448 12	309.67
25 200	238.22	0.826 74	37 274	0.367 86	0.545 08	0.444 96	309.41
25 400	237.83	0.825 36	36 949	0.364 66	0.541 23	0.441 82	309.15
25 600	237.43	0.823 99	36 627	0.361 48	0.537 40	0.438 69	308.90
25 800	237.04	0.822 61	36 307	0.358 32	0.533 59	0.435 59	308.64
26 000	236.64	0.821 23	35 989	0.355 18	0.529 81	0.432 50	308.38
26 200	236.24	0.819 86	35 673	0.352 07	0.526 04	0.429 42	308.12
26 400	235.85	0.818 48	35 360	0.348 97	0.522 30	0.426 37	307.86
26 600	235.45	0.817 11	35 049	0.345 90	0.518 57	0.423 33	307.61
26 800	235.05	0.815 73	34 740	0.342 85	0.514 87	0.420 30	307.35
27 000	234.66	0.814 36	34 433	0.339 83	0.511 19	0.417 30	307.09
27 200	234.26	0.812 98	34 129	0.336 82	0.507 52	0.414 30	306.83
27 400	233.87	0.811 61	33 826	0.333 84	0.503 88	0.411 33	306.57
27 600	233.47	0.810 23	33 526	0.330 88	0.500 26	0.408 37	306.31
27 800	233.07	0.808 86	33 228	0.327 94	0.496 65	0.405 43	306.05
28 000	232.68	0.807 48	32 932	0.325 02	0.493 07	0.402 51	305.79
28 200	232.28	0.806 11	32 639	0.322 12	0.489 51	0.399 60	305.53
28 400	231.88	0.804 73	32 347	0.319 24	0.485 96	0.396 70	305.27
28 600	231.49	0.803 36	32 058	0.316 38	0.482 44	0.393 83	305.01
28 800	231.09	0.801 98	31 770	0.313 55	0.478 93	0.390 97	304.75
29 000	230.70	0.800 61	31 485	0.310 73	0.475 45	0.388 12	304.48
29 200	230.30	0.799 23	31 202	0.307 94	0.471 98	0.385 29	304.22
29 400	229.90	0.797 86	30 921	0.305 16	0.465 54	0.382 48	303.96
29 600	229.51	0.796 48	30 642	0.302 41	0.465 11	0.379 68	303.70
29 800	229.11	0.795 11	30 365	0.299 67	0.461 70	0.376 90	303.44
30 000	228.71	0.793 73	30 090	0.296 96	0.458 31	0.374 13	303.17
30 200	228.32	0.792 36	29 817	0.294 27	0.454 94	0.371 38	302.91
30 400	227.92	0.790 98	29 546	0.291 59	0.451 59	0.368 65	302.65
30 600	227.53	0.789 61	29 277	0.288 94	0.448 26	0.365 93	302.38
30 800	227.13	0.788 23	29 010	0.286 30	0.444 95	0.363 22	302.12
31 000	226.73	0.786 86	28 745	0.283 69	0.441 65	0.360 53	301.86
31 200	226.34	0.785 48	28 482	0.281 09	0.438 38	0.357 86	301.59
31 400	225.94	0.784 11	28 220	0.278 51	0.435 12	0.355 20	301.33
31 600	225.54	0.782 73	27 961	0.275 96	0.431 88	0.352 66	301.07
31 800	225.15	0.781 36	27 704	0.273 42	0.428 66	0.349 93	300.80
32 000	224.75	0.779 98	27 449	0.270 90	0.425 46	0.347 31	300.54
32 200	224.36	0.778 61	27 195	0.268 40	0.422 28	0.344 72	300.27
32 400	223.96	0.777 23	26 944	0.265 92	0.419 11	0.342 13	300.01
32 600	223.56	0.775 86	26 694	0.263 45	0.415 97	0.339 56	299.74
32 800	223.17	0.774 48	26 447	0.261 01	0.412 84	0.337 01	299.47
33 000	222.77	0.773 11	26 201	0.258 58	0.409 73	0.334 47	299.21
33 200	222.37	0.771 73	25 957	0.256 17	0.406 63	0.331 95	298.94
33 400	221.98	0.770 36	25 715	0.253 78	0.403 56	0.329 44	298.68
33 600	221.58	0.768 98	25 474	0.251 41	0.400 50	0.326 94	298.41
33 800	221.19	0.767 61	25 236	0.249 06	0.397 46	0.324 46	298.14

TABLE 10.10 International Standard Atmosphere (−2,000 ft to 60,000 ft) (*Continued*)

H ft	T K	θ	p N/m^2	δ	ρ kg/m^3	σ	a m/s
34 000	220.79	0.766 23	24 999	0.246 72	0.394 44	0.321 99	297.87
34 200	220.39	0.764 85	24 764	0.244 40	0.391 44	0.319 54	297.61
34 400	220.00	0.763 48	24 531	0.242 10	0.388 45	0.317 10	297.34
34 600	219.60	0.762 10	24 300	0.239 82	0.385 48	0.314 68	297.07
34 800	219.20	0.760 73	24 070	0.237 55	0.382 53	0.312 27	296.80
35 000	218.81	0.759 35	23 842	0.235 30	0.379 60	0.309 87	296.54
35 200	218.41	0.757 98	23 616	0.233 07	0.376 68	0.307 49	296.27
35 400	218.02	0.756 60	23 392	0.230 86	0.373 78	0.305 13	296.00
35 600	217.62	0.755 23	23 169	0.228 66	0.370 90	0.302 77	295.73
35 800	217.22	0.753 85	22 948	0.226 48	0.368 03	0.300 43	295.46
36 000	216.83	0.752 48	22 729	0.224 32	0.365 18	0.298 11	295.19
36 089	216.65	0.751 87	22 632	0.223 36	0.363 92	0.297 08	295.07
36 200	216.65	0.751 87	22 512	0.222 17	0.361 98	0.295 50	295.07
36 400	216.65	0.751 87	22 296	0.220 05	0.358 52	0.292 67	295.07
36 600	216.65	0.751 87	22 083	0.217 94	0.355 09	0.289 87	295.07
36 800	216.65	0.751 87	21 872	0.215 86	0.351 70	0.287 10	295.07
37 000	216.65	0.751 87	21 663	0.213 79	0.348 33	0.284 35	295.07
37 200	216.65	0.751 87	21 455	0.211 75	0.345 00	0.281 63	295.07
37 400	216.65	0.751 87	21 250	0.209 72	0.341 70	0.278 94	295.07
37 600	216.65	0.751 87	21 047	0.207 72	0.338 43	0.276 27	295.07
37 800	216.65	0.751 87	20 846	0.205 73	0.335 19	0.273 63	295.07
38 000	216.65	0.751 87	20 646	0.203 76	0.331 98	0.271 01	295.07
38 200	216.65	0.751 87	20 449	0.201 81	0.328 81	0.268 42	295.07
38 400	216.65	0.751 87	20 253	0.199 88	0.325 66	0.265 85	295.07
38 600	216.65	0.751 87	20 059	0.197 97	0.322 55	0.263 30	295.07
38 800	216.65	0.751 87	19 867	0.196 08	0.319 46	0.260 79	295.07
39 000	216.65	0.751 87	19 677	0.194 20	0.316 41	0.258 29	295.07
39 200	216.65	0.751 87	19 489	0.192 34	0.313 38	0.255 82	295.07
39 400	216.65	0.751 87	19 303	0.190 50	0.310 38	0.253 37	295.07
39 600	216.65	0.751 87	19 118	0.188 68	0.307 41	0.250 95	295.07
39 800	216.65	0.751 87	18 935	0.186 87	0.304 47	0.248 55	295.07
40 000	216.65	0.751 87	18 754	0.185 09	0.301 56	0.246 17	295.07
40 200	216.65	0.751 87	18 574	0.183 32	0.298 67	0.243 81	295.07
40 400	216.65	0.751 87	18 397	0.181 56	0.295 82	0.241 48	295.07
40 600	216.65	0.751 87	18 221	0.179 83	0.292 99	0.239 17	295.07
40 800	216.65	0.751 87	18 046	0.178 10	0.290 18	0.236 88	295.07
41 000	216.65	0.751 87	17 874	0.176 40	0.287 41	0.234 62	295.07
41 200	216.65	0.751 87	17 703	0.174 71	0.284 66	0.232 37	295.07
41 400	216.65	0.751 87	17 533	0.173 04	0.281 93	0.230 15	295.07
41 600	216.65	0.751 87	17 366	0.171 39	0.279 24	0.227 95	295.07
41 800	216.65	0.751 87	17 200	0.169 75	0.276 57	0.225 77	295.07
42 000	216.65	0.751 87	17 035	0.168 12	0.273 92	0.223 61	295.07
42 200	216.65	0.751 87	16 872	0.166 51	0.271 30	0.221 47	295.07
42 400	216.65	0.751 87	16 711	0.164 92	0.268 70	0.219 35	295.07
42 600	216.65	0.751 87	16 551	0.163 34	0.266 13	0.217 25	295.07
42 800	216.65	0.751 87	16 392	0.161 78	0.263 59	0.215 17	295.07

TABLE 10.10 International Standard Atmosphere (−2,000 ft to 60,000 ft) (*Continued*)

H ft	T K	θ	p N/m^2	δ	ρ kg/m^3	σ	a m/s
43 000	216.65	0.751 87	16 236	0.160 23	0.261 07	0.213 11	295.07
43 200	216.65	0.751 87	16 080	0.158 70	0.258 57	0.211 08	295.07
43 400	216.65	0.751 87	15 927	0.157 18	0.256 09	0.209 06	295.07
43 600	216.65	0.751 87	15 774	0.155 68	0.253 64	0.207 06	295.07
43 800	216.65	0.751 87	15 623	0.154 19	0.251 22	0.205 08	295.07
44 000	216.65	0.751 87	15 474	0.152 71	0.248 81	0.203 11	295.07
44 200	216.65	0.751 87	15 326	0.151 25	0.246 43	0.201 17	295.07
44 400	216.65	0.751 87	15 179	0.149 81	0.244 08	0.199 25	295.07
44 600	216.65	0.751 87	15 034	0.148 37	0.241 74	0.197 34	295.07
44 800	216.65	0.751 87	14 890	0.146 95	0.239 43	0.195 45	295.07
45 000	216.65	0.751 87	14 748	0.145 55	0.237 14	0.193 58	295.07
45 200	216.65	0.751 87	14 607	0.144 16	0.234 87	0.191 73	295.07
45 400	216.65	0.751 87	14 467	0.142 78	0.232 62	0.189 90	295.07
45 600	216.65	0.751 87	14 328	0.141 41	0.230 40	0.188 08	295.07
45 800	216.65	0.751 87	14 191	0.140 06	0.228 19	0.186 28	295.07
46 000	216.65	0.751 87	14 056	0.138 72	0.226 01	0.184 50	295.07
46 200	216.65	0.751 87	13 921	0.137 39	0.223 85	0.182 73	295.07
46 400	216.65	0.751 87	13 788	0.136 08	0.221 71	0.180 98	295.07
46 600	216.65	0.751 87	13 656	0.134 77	0.219 59	0.179 25	295.07
46 800	216.65	0.751 87	13 525	0.133 49	0.217 48	0.177 54	295.07
47 000	216.65	0.751 87	13 396	0.132 21	0.215 40	0.175 84	295.07
47 200	216.65	0.751 87	13 268	0.130 94	0.213 34	0.174 16	295.07
47 400	216.65	0.751 87	13 141	0.129 69	0.211 30	0.172 49	295.07
47 600	216.65	0.751 87	13 015	0.128 45	0.209 28	0.170 84	295.07
47 800	216.65	0.751 87	12 891	0.127 22	0.207 28	0.169 21	295.07
48 000	216.65	0.751 87	12 767	0.126 00	0.205 30	0.167 59	295.07
48 200	216.65	0.751 87	12 645	0.124 80	0.203 33	0.165 99	295.07
48 400	216.65	0.751 87	12 524	0.123 60	0.201 39	0.164 40	295.07
48 600	216.65	0.751 87	12 404	0.122 42	0.199 46	0.162 82	295.07
48 800	216.65	0.751 87	12 286	0.121 25	0.197 55	0.161 27	295.07
49 000	216.65	0.751 87	12 168	0.120 09	0.195 66	0.159 72	295.07
49 200	216.65	0.751 87	12 052	0.118 94	0.193 79	0.158 20	295.07
49 400	216.65	0.751 87	11 937	0.117 80	0.191 94	0.156 68	295.07
49 600	216.65	0.751 87	11 822	0.116 68	0.190 10	0.155 18	295.07
49 800	216.65	0.751 87	11 709	0.115 56	0.188 28	0.153 70	295.07
50 000	216.65	0.751 87	11 597	0.114 46	0.186 48	0.152 23	295.07
50 200	216.65	0.751 87	11 486	0.113 36	0.184 70	0.150 77	295.07
50 400	216.65	0.751 87	11 376	0.112 28	0.182 93	0.149 33	295.07
50 600	216.65	0.751 87	11 268	0.111 20	0.181 18	0.147 90	295.07
50 800	216.65	0.751 87	11 160	0.110 14	0.179 45	0.146 49	295.07
51 000	216.65	0.751 87	11 053	0.109 08	0.177 73	0.145 09	295.07
51 200	216.65	0.751 87	10 947	0.108 04	0.176 03	0.143 70	295.07
51 400	216.65	0.751 87	10 843	0.107 01	0.174 35	0.142 32	295.07
51 600	216.65	0.751 87	10 739	0.105 98	0.172 68	0.140 96	295.07
51 800	216.65	0.751 87	10 636	0.104 97	0.171 03	0.139 61	295.07

TABLE 10.10 International Standard Atmosphere (−2,000 ft to 60,000 ft) (*Continued*)

H ft	T K	θ	p N/m^2	δ	ρ kg/m^3	σ	a m/s
52 000	216.65	0.751 87	10 534	0.103 97	0.169 39	0.138 28	295.07
52 200	216.65	0.751 87	10 434	0.102 97	0.167 77	0.136 95	295.07
52 400	216.65	0.751 87	10 334	0.101 99	0.166 16	0.135 64	295.07
52 600	216.65	0.751 87	10 235	0.101 01	0.164 57	0.134 35	295.07
52 800	216.65	0.751 87	10 137	0.100 04	0.163 00	0.133 06	295.07
53 000	216.65	0.751 87	10 040	0.099 087	0.161 44	0.131 79	295.07
53 200	216.65	0.751 87	9 943.9	0.098 139	0.159 90	0.130 53	295.07
53 400	216.65	0.751 87	9 848.8	0.097 200	0.158 37	0.129 28	295.07
53 600	216.65	0.751 87	9 754.6	0.096 270	0.156 85	0.128 04	295.07
53 800	216.65	0.751 87	9 661.3	0.095 349	0.155 35	0.126 82	295.07
54 000	216.65	0.751 87	9 568.8	0.094 437	0.153 86	0.125 60	295.07
54 200	216.65	0.751 87	9 477.3	0.093 534	0.152 39	0.124 40	295.07
54 400	216.65	0.751 87	9 386.6	0.092 639	0.150 93	0.123 21	295.07
54 600	216.65	0.751 87	9 296.8	0.091 753	0.149 49	0.122 03	295.07
54 800	216.65	0.751 87	9 207.9	0.090 875	0.148 06	0.120 87	295.07
55 000	216.65	0.751 87	9 119.8	0.090 005	0.146 64	0.119 71	295.07
55 200	216.65	0.751 87	9 032.6	0.089 144	0.145 24	0.118 56	295.07
55 400	216.65	0.751 87	8 946.1	0.088 292	0.143 85	0.117 43	295.07
55 600	216.65	0.751 87	8 860.6	0.087 447	0.142 48	0.116 31	295.07
55 800	216.65	0.751 87	8 775.8	0.086 610	0.141 11	0.115 19	295.07
56 000	216.65	0.751 87	8 691.8	0.085 782	0.139 76	0.114 09	295.07
56 200	216.65	0.751 87	8 608.7	0.084 961	0.138 43	0.113 00	295.07
56 400	216.65	0.751 87	8 526.3	0.084 148	0.137 10	0.111 92	295.07
56 600	216.65	0.751 87	8 444.8	0.083 343	0.135 79	0.110 85	295.07
56 800	216.65	0.751 87	8 364.0	0.082 546	0.134 49	0.109 79	295.07
57 000	216.65	0.751 87	8 284.0	0.081 756	0.133 20	0.108 74	295.07
57 200	216.65	0.751 87	8 204.7	0.080 974	0.131 93	0.107 70	295.07
57 400	216.65	0.751 87	8 126.2	0.080 200	0.130 67	0.106 67	295.07
57 600	216.65	0.751 87	8 048.5	0.079 432	0.129 42	0.105 65	295.07
57 800	216.65	0.751 87	7 971.5	0.078 672	0.128 18	0.104 64	295.07
58 000	216.65	0.751 87	7 895.2	0.077 920	0.126 95	0.103 64	295.07
58 200	216.65	0.751 87	7 819.7	0.077 174	0.125 74	0.102 64	295.07
58 400	216.65	0.751 87	7 744.9	0.076 436	0.124 54	0.101 66	295.07
58 600	216.65	0.751 87	7 670.8	0.075 705	0.123 34	0.100 69	295.07
58 800	216.65	0.751 87	7 597.4	0.074 981	0.122 16	0.099 726	295.07
59 000	216.65	0.751 87	7 524.7	0.074 263	0.121 00	0.098 772	295.07
59 200	216.65	0.751 87	7 452.7	0.073 553	0.119 84	0.097 827	295.07
59 400	216.65	0.751 87	7 381.4	0.072 849	0.118 69	0.096 891	295.07
59 600	216.65	0.751 87	7 310.8	0.072 152	0.117 56	0.095 964	295.07
59 800	216.65	0.751 87	7 240.9	0.071 462	0.116 43	0.095 046	295.07
60 000	216.65	0.751 87	7 171.6	0.070 778	0.115 32	0.094 137	295.07

Height Scales

In airplane performance work it is necessary to distinguish between different height scales that are used.

***Geometric Height* (h).** Geometric height is the true vertical distance of a point from a datum plane, which in this case is the mean sea-level datum. It is used to define the height of buildings and terrain, for example.

***Geopotential Height* (H).** Geopotential height is the height in a hypothetical uniform gravitational field which would give the same potential energy as the point under consideration in the actual, variable gravitational field. Thus, if the standard value of gravity is used in conjunction with geopotential height, the variation of gravity with height is automatically taken into account. Note that the integration of equation (10.60) to give the pressure and density variations with height in the ISA was performed using the standard value of g. The difference between geometric and geopotential height for typical airplane cruise altitudes is small and is usually ignored (for example, at 40,000 ft the difference is less than 0.2%).

***Pressure Height* (h_p).** The ISA defines a unique relationship between pressure and geopotential height, as is evident from the equations in Table 10.9. The height in the standard atmosphere can thus be considered as a scale of pressure. By definition, the pressure height at a point in any atmosphere (standard or off-standard) is the height in the ISA which has the same pressure.

***Flight Level* (FL).** For airplane operations, altitude is specified in terms of a flight level, which is pressure height expressed in hundreds of feet. FL 350 for example, represents a pressure height of 35,000 ft. The standard sea-level setting of 1013.2 hPa (29.92 in. mercury) is used as the reference pressure.

***Density Height* (h_ρ).** This height scale is not as widely used as pressure height. *Density height* is the equivalent height for a given density in the ISA.

Altimeter Settings

An altimeter is essentially a pressure gauge that works on the principle of differential pressure between the inside and outside of a sealed chamber. The relationship between pressure and height is defined by equations based on those given in Table 10.9, where the constants correspond to those of the ISA. An altimeter thus provides the pilot with a reading of pressure altitude. For an altimeter to be a useful instrument for day-to-day operations, where varying atmospheric pressure conditions would be encountered, it must be possible to adjust the zero height reading. The pilot accomplishes this by rotating the subscale knob on the instrument, in effect changing the reference pressure used to determine the height. For standard aircraft operations, altitude is measured in feet. There are three useful settings that may be used by the pilot.

Standard Setting. By agreement, all aircraft traffic above a specified transition height use the standard sea-level setting of 1013.2 hPa (29.92 in. mercury) as the reference pressure. Because all aircraft operate by the same rules, there is no danger

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

of collision even if the ambient conditions depart substantially from standard ISA conditions. This is the standard adopted for flight level operation.

QNH Setting. If the pilot selects a reference pressure that results in the altimeter correctly reading the elevation of the local airport (when the aircraft is on the runway), this is called a QNH setting. It is used for takeoff and climb-out to the transition altitude (at which point the pilot switches over to the standard setting), as well as for landing (where the setting is based on the destination airport). For operations below the transition height, QNH may be used throughout the flight.

QFE Setting. The pilot sets the subscale knob such that the altimeter reads zero on the ground, irrespective of the airport's actual elevation, for a QFE setting. During flight the altimeter will indicate the height above that airport.

Further Reading

Further details of the Standard Atmosphere can be found in ISO 2533 (1975), ESDU 68046 (1992), ESDU 72018 (1972), and ICAO (1993), including tables with height in meters and off-standard atmospheric properties. Swatton (2000) and Lowry (1999) provide more information on altimeters and the use of pressure height for flight operations.

10.13 AIRSPEED AND AIRSPEED MEASUREMENT

Speed of Sound (a)

The speed of sound in the atmosphere is a function of the ambient conditions and is given by:

$$a = \sqrt{\gamma RT} \tag{10.61}$$

where γ is the ratio of specific heats of air, R is the gas constant, and T is the ambient temperature (absolute). This equation may also be written as a function of the temperature ratio and the standard sea-level value of the speed of sound, i.e.,

$$a = a_0\sqrt{\theta} \tag{10.62}$$

Note that equation (10.62) is correct for any temperature, corresponding to a standard or off-standard day. In the ISA the temperature decreases linearly with altitude to the tropopause and thus the speed of sound will also decrease, although not linearly, but as a function of $\sqrt{\theta}$. In the stratosphere the speed of sound is constant.

True Airspeed (V)

The true airspeed (TAS) is the speed of the airplane relative to the surrounding air mass.

Mach Number (M)

An airplane's flight Mach number is defined as the ratio of its TAS to the speed of sound of the ambient air.

$$M = \frac{V}{a} \tag{10.63}$$

Dynamic Pressure (q)

The *dynamic pressure* is defined as:

$$q = \frac{1}{2}\rho V^2 \tag{10.64}$$

For high-speed flight Mach number and not TAS is used as the measure of speed. For this reason it is convenient to express the dynamic pressure in terms of Mach number and pressure ratio, i.e.,

$$q = \frac{1}{2}p_0\delta\gamma M^2 \tag{10.65}$$

Airspeed Indication (Incompressible Flow)

The Bernoulli equation for incompressible flow states that the sum of the static pressure and dynamic pressure along a streamline is constant. It may be represented as:

$$p_t = p + \frac{1}{2}\rho V^2 = \text{constant} \tag{10.66}$$

where p_t is the total pressure and p is the static pressure.

A *Pitot static* system is used for measuring airspeed. It has two pressure openings, a *total pressure* port and a *static pressure* port. A small Pitot tube is aligned approximately in the direction of the incoming air; this port senses the total pressure of the air as it momentarily comes to rest. The second port senses the static air pressure at the side of the tube, or on the side of the fuselage. A mechanical measuring device is used to measure the difference in air pressure between the two ports. This, by the application of Bernoulli's equation, is a measure of the dynamic pressure, and, provided air density is known, may be used to indicate the airspeed for incompressible flow. For airspeeds above about 200 kt, compressibility effects become important and the form of the Bernoulli expression given above in equation (10.66) is not satisfactory. The practical measurement of airspeed on actual aircraft introduces specific difficulties, which are discussed below under Calibrated Airspeed.

Ground Speed (V_g)

The *ground speed* is the actual speed of the airplane relative to the ground. It is given by the sum of the TAS vector and the wind velocity.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

Equivalent Airspeed (V_e)

The *equivalent airspeed* (EAS) is the equivalent speed which the airplane would have at sea level (air density ρ_0) if it developed the same *dynamic pressure* as it does moving at its TAS at the altitude concerned (air density ρ). Mathematically this can written as:

$$q = \frac{1}{2}\rho_0 V_e^2 = \frac{1}{2}\rho V^2$$

Thus,

$$V_e = \sqrt{\frac{\rho}{\rho_0}}\, V = \sqrt{\sigma}\, V \tag{10.67}$$

Equivalent airspeed is a very useful parameter for engineering analysis. However, for flight operations reference is made to *calibrated airspeed*.

Calibrated Airspeed (V_C)

The *calibrated airspeed* (CAS)—very occasionally called the *rectified airspeed*—is the airspeed reading on a calibrated airspeed indicator connected to a Pitot static system, that is assumed to be entirely free of error. It is common practice in aircraft operations to write KCAS for *knots calibrated airspeed*. A better representation of the operation of an airspeed indicator than that given above is provided by the following equation, which applies to *compressible* airflow:

$$\left(\frac{p_t}{p}\right)_{\text{compressible}} = \left(1 + \frac{\gamma - 1}{2} M^2\right)^{\gamma/(\gamma-1)} \tag{10.68}$$

As the Pitot static system provides a measurement of $(p_t - p)$, this term will be isolated on the left-hand side of the equation, and noting that $\gamma = 1.4$ and $M = V/a$, the equation becomes:

$$(p_t - p) = p\left\{\left[1 + 0.2\left(\frac{V}{a}\right)^2\right]^{3.5} - 1\right\} \tag{10.69}$$

Equation (10.69) may be turned into a useful expression for determining airspeed. This is done by selecting standard sea-level values for pressure and speed of sound on the right hand side of the equation and by defining the resulting velocity as calibrated airspeed. The left-hand side of equation (10.69) is provided by the pitot static system. From this equation it is evident that an airspeed indicator may be correctly calibrated for any Mach numbers at sea level, where CAS will always equal EAS. However due to the use of sea-level values, CAS will be greater than EAS by a small amount that increases with Mach number and altitude.

Compressibility Correction for CAS

The difference between CAS and EAS is called the *compressibility correction* and is designated as ΔV_C. By definition:

$$V_C = V_e + \Delta V_C \tag{10.70}$$

and the magnitude of ΔV_C is given by:

$$\Delta V_C = a_0\{\sqrt{5\{[\delta\{[1 + 0.2M^2]^{3.5} - 1\} + 1]^{1/3.5} - 1\}} - M\sqrt{\delta}\} \tag{10.71}$$

The compressibility correction factor is negligibly small for operations below about 10,000 ft and 200 kt CAS. A useful equation for converting CAS to Mach number, accounting for compressibility correction, is:

$$M = \sqrt{5\left\{\left[\frac{1}{\delta}\left\{\left[1 + 0.2\left(\frac{V_c}{a_0}\right)^2\right]^{3.5} - 1\right\} + 1\right]^{1/3.5} - 1\right\}} \tag{10.72}$$

Indicated Airspeed (V_I)

The indicated airspeed (IAS) is the reading of an actual airspeed indicator. Most frequently the speed is indicated in knots (nautical miles per hour), but units of miles per hour and km per hour are also used. During flight the IAS may differ from the CAS because of an instrument calibration error or an error arising from the inability of the pitot static system to measure the correct total and static pressures accurately. The error associated with an individual instrument may be corrected using charts supplied by the manufacturer; modern instruments are very accurate and the instrument error is usually negligible. Total pressure recovery and static pressure errors are very small on well-designed systems, and the instrument can be calibrated to reduce these errors. On a modern airliner the correction is taken into account by the air data computer, and for all practical purposes the IAS will then equal the CAS.

Further Reading

Consult Eshelby (2000), Lowry (1999), or Lan and Roskam (1981) for further details on airspeed measurement.

10.14 DRAG AND DRAG POWER (POWER REQUIRED)

Drag Components

For the purposes of performance analysis the airplane's overall drag coefficient can be divided into two components, i.e.,

$$C_D = C_{D_0} + C_{D_i} \tag{10.73}$$

where C_{D_0} is the *zero-lift* drag coefficient (i.e., the drag coefficient when C_L equals zero) and C_{D_i} is the *lift-dependent* drag coefficient (induced drag). C_{D_i} is mainly wing *trailing vortex drag,* which is proportional to C_L^2 but also includes additional lift-dependent interference drag due to the effect of the fuselage, nacelles, etc. on

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

the wing planform, and a small contribution due to the increase of boundary layer (profile) drag with angle of attack.

Drag Polar

The characteristic C_D versus C_L relationship for an airplane is commonly referred to as the drag polar. It is convenient to represent the drag polar by a mathematical model. To an acceptable approximation the lift-dependent drag coefficient of the whole aircraft may be taken to be proportional to C_L^2 and the drag polar written as:

$$C_D = C_{D_0} + KC_L^2 \tag{10.74}$$

where K is the *lift-dependent* drag factor. For initial performance evaluation of an airplane, it is often useful to express K in terms of a span efficiency factor. The equation is then written as:

$$C_D = C_{D_0} + \frac{1}{\pi A_R e} C_L^2 \tag{10.75}$$

where A_R is the wing aspect ratio and e is the Oswald efficiency factor, which is generally in the range of $0.65 \leq e \leq 0.90$. High aspect ratio wings are seen from this equation to result in low induced drag. The use of this idealized *parabolic* drag relationship greatly simplifies calculations and is sufficiently accurate for most performance work, providing the following factors are noted.

- *Mach number:* At a speed known as the *drag rise Mach number* (M_{DR}) the drag starts to rise rapidly due to compressibility effects (Figure 10.68). This drag increase is called *wave drag*. The drag rise Mach number thus sets an upper limit to the validity of the low-speed drag polar.
- *Aircraft configuration:* During flight the pilot may change the drag characteristics of the aircraft in several ways, such as by deploying air-brakes or flaps or lowering the undercarriage. Each of these factors will change the drag polar.
- *Ground effect:* Operation of the aircraft in close proximity to the ground results in a change in the trailing vortex sheet and a reduction in the lift-dependent drag. This effect depends on the aircraft type and its proximity to the ground.

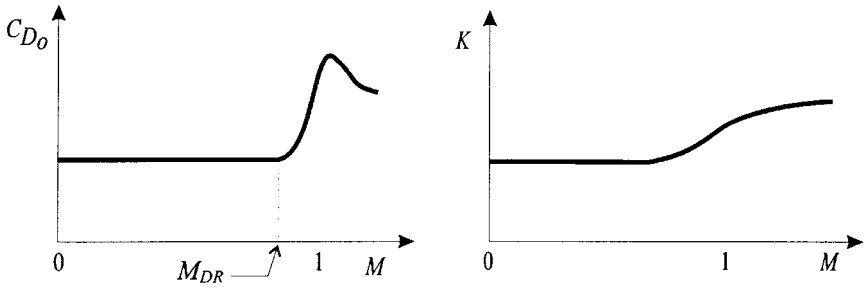


FIGURE 10.68 Influence of Mach number on drag parameters.

Actual Drag Polars

Drag polars derived from experimental results will usually differ slightly from the idealized parabolic form given by equation (10.74) above. The effect of substantial camber on a wing for example, will result in the minimum drag ($C_{D_{\min}}$) point occurring not at $C_L = 0$, as would be expected from equation (10.74), but at some small positive value of the lift coefficient ($C_{L_{\min\ drag}}$). The drag polar in such cases may be approximated by:

$$C_D = C_{D_{\min}} + K(C_L - C_{L_{\min\ drag}})^2 \tag{10.76}$$

For most calculations, however, the difference between $C_{D_{\min}}$ and C_{D_0} is small and equation (10.74) may be used.

High-Speed Drag Polars

For any airplane it is possible to represent its low-speed flight characteristics in terms of a single drag polar. For aircraft that operate in the transonic flight regime, it is necessary to take into account the compressibility drag rise. It is seen from Figure 10.68 that above M_{DR} there is a unique drag polar for each Mach number (Figure 10.69).

Drag versus EAS Relationship

It is convenient in many applications to write the drag in terms of V_e, based on the parabolic drag polar:

$$D = \left[\frac{C_{D_0}\rho_0 S}{2}\right] V_e^2 + \left[\frac{2\ KW^2}{\rho_0 S}\right] \frac{1}{V_e^2} \tag{10.77}$$

or alternatively:

$$D = A_1 V_e^2 + B_1 V_e^{-2} \tag{10.78}$$

where

$$A_1 = \frac{C_{D_0}\rho_0 S}{2} \quad \text{and} \quad B_1 = \frac{2KW^2}{\rho_0 S}$$

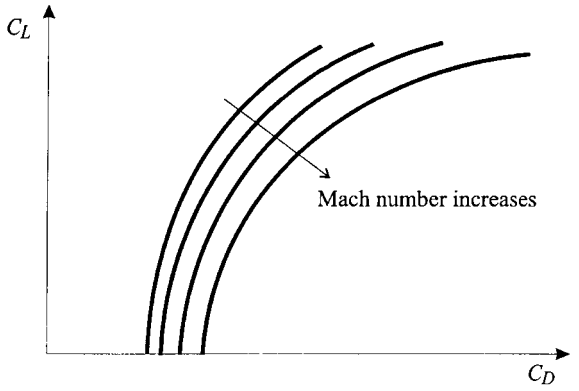


FIGURE 10.69 High-speed drag polars.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

The substitution has introduced constants A_1 and B_1 to simplify the mathematics. It is important to note that A_1 and B_1 will change if the airplane configuration or weight changes. Equation (10.77) shows that there are two contributions to the drag of the airplane. The first, the lift-independent drag (D_0) contribution, is proportional to V_e^2, whereas the second contribution, the lift-dependent drag (D_i) term, is inversely proportional to V_e^2. At low speeds D_i is the dominant part, whereas at high speeds D_0 is dominant as shown in Figure 10.70.

Minimum Drag Condition

The speed at which the airplane's drag will be a minimum may be determined by differentiating equation (10.77) with respect to V_e and setting the result equal to zero.

$$V_{e_{md}} = \left[\frac{2W}{\rho_0 S}\sqrt{\frac{K}{C_{D_0}}}\right]^{1/2} \tag{10.79}$$

At this speed there are equal contributions to the drag from D_0 and D_i and the minimum drag is given by:

$$D_{\min} = 2W\sqrt{KC_{D_0}} \tag{10.80}$$

Lift-to-Drag Ratio

The lift-to-drag ratio is a measure of an aircraft's *aerodynamic efficiency*. For an aircraft in *steady level flight,* as shown in Figure 10.71, the thrust (T) will equal the drag (D) and the lift (L) will equal the weight (W). By combining these equations it is seen that the thrust required to sustain steady (i.e., unaccelerated) level flight for a given airplane weight depends on the lift-to-drag ratio, i.e.,

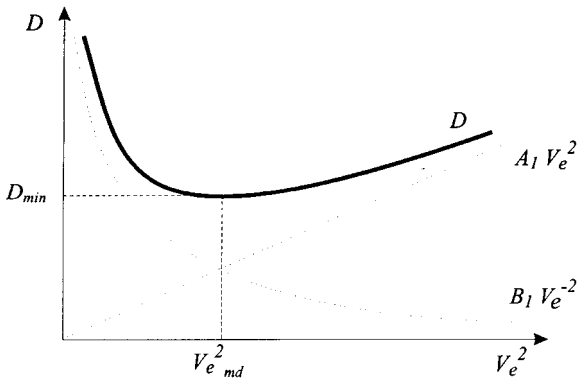


FIGURE 10.70 Drag function.

where: $A_1 = \dfrac{C_{D_0}\,\rho_0 S}{2}$ and $B_1 = \dfrac{2\,KW^2}{\rho_0 S}$

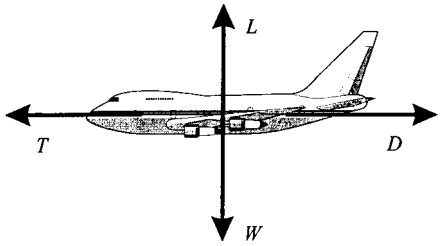


FIGURE 10.71 Steady level flight.

$$T = \frac{1}{(L/D)} W \tag{10.81}$$

An aerodynamically efficient airplane, with a high (L/D) ratio, will require a lower thrust to maintain steady level flight than a comparable airplane of the same weight.

Maximum Lift-to-Drag Ratio

Although the value of (L/D) will change during flight, each aircraft has a maximum value which it cannot exceed. The maximum lift-to-drag ratio is a figure of merit, widely used to assess aerodynamic efficiency. The value of $(L/D)_{\max}$ can be determined graphically by drawing a line tangent to the drag polar through the origin. Based on the parabolic drag polar, the maximum lift-to-drag ratio occurs at the conditions associated with *minimum drag*. The minimum drag lift coefficient $(C_{L_{md}})$ for a parabolic drag polar is:

$$C_{L_{md}} = \sqrt{\frac{C_{D_0}}{K}} \tag{10.82}$$

and the lift-to-drag ratio corresponding to this condition is:

$$\left(\frac{L}{D}\right)_{\max} = \frac{1}{2}\sqrt{\frac{1}{C_{D_0}K}} \tag{10.83}$$

Power Required and Power Available

The product of an airplane's total drag and its speed is a useful concept for performance analysis. By definition:

$$\text{Drag power} = P_D = DV \tag{10.84}$$

Because the thrust produced by the aircraft's engine(s) is required to overcome the drag, the drag power may be considered as the power required for flight. The thrust power is by definition the net propulsive force multiplied by the speed; it is thus the power available to propel the aircraft due to the engine(s). By definition:

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

$$\text{Thrust power} = P_T = TV \tag{10.85}$$

These two quantities, the drag power (power required) and the thrust power (power available), are equal in steady (i.e., unaccelerated) level flight, as the thrust is equal to the drag (see Figure 10.71). However, in climbing, descending, or accelerated flight this will generally not be true. An excess of *thrust power* to *drag power* will enable the aircraft to perform a steady climb and a deficit will result in the aircraft losing height.

Minimum Power Condition

The drag power (power required) may be expressed as a function of the EAS based on the parabolic drag polar. Using equations (10.67) and (10.77), it follows that:

$$P_D = \left[\frac{C_{D_0}\rho_0 S}{2\sqrt{\sigma}}\right] V_e^3 + \left[\frac{2KW^2}{\sqrt{\sigma}\rho_0 S}\right]\frac{1}{V_e} \tag{10.86}$$

This function has a minimum at a speed that is a little slower than the minimum drag speed. The flight condition is identified by the subscript *mp* for *minimum power.* At the speed for minimum power ($V_{e_{mp}}$) the drag will be slightly greater than $D_{\min}$. Expressions for the EAS, drag, lift-to-drag ratio, and C_L corresponding to the minimum power condition, derived using the parabolic drag polar, are given in Table 10.11.

Further Reading

Consult Anderson (1999), Mair and Birdsall (1992), or Ojha (1995) for further details on aircraft drag. Methods for estimating drag polars for airplanes are provided by Lan and Roskam (1981), Torenbeek (1982), and Raymer (1989).

10.15 ENGINE (POWERPLANT) PERFORMANCE

Installation Considerations

The analysis of an aircraft's performance is based on the *net installed engine thrust.* For a jet engine this takes into account the inlet pressure recovery, power and bleed air extraction, and drag contributions associated with the propulsion system. The airplane's drag is a function of the entire flowfield around the airplane, and this flowfield is influenced by the engine inlet and exhaust stream-tubes. For example, if the pilot throttles back, spillage will occur around the lip of the inlet, resulting in an increment in drag. It is conventional to regard these components of drag, whose magnitudes depend on the position of the throttle, as decrements of thrust rather than as increments of airframe drag. The accurate prediction of an aircraft's performance thus requires a consistent definition of what constitutes *propulsion system thrust* and what constitutes *propulsion system drag*. A general rule that is often used regards all fore and aft components of force that depend on throttle setting as increments or decrements of thrust.

TABLE 10.11 Summary of Performance Parameters Based on the Parabolic Drag Polar

Drag:	Drag power:
$D = \left[\frac{C_{D_0}\,\rho_0 S}{2}\right] V_e^2 + \left[\frac{2\,KW^2}{\rho_0 S}\right] \frac{1}{V_e^2}$	$P_D = DV = D\,\frac{V_e}{\sqrt{\sigma}} = \left[\frac{C_{D_0}\,\rho_0\,S}{2\sqrt{\sigma}}\right] V_e^3 + \left[\frac{2\,KW^2}{\sqrt{\sigma}\,\rho_0\,S}\right] \frac{1}{V_e}$
At the minimum drag condition:	At minimum drag power condition:
$C_{D_i} = C_{D_0}$	$C_{D_i} = 3\,C_{D_0}$
$D_{\min} = 2\,W\sqrt{KC_{D_0}}$	$D_{mp} = \frac{4}{\sqrt{3}}\,W\sqrt{KC_{D_0}}$
$D_{\min} = \frac{1}{(L/D)_{\max}}\,W$	$D_{mp} = \frac{2}{\sqrt{3}}\,D_{\min} = 1.15\,D_{\min}$
$C_{L_{md}} = \sqrt{\frac{C_{D_0}}{K}}$	$C_{L_{mp}} = \sqrt{\frac{3C_{D_0}}{K}}$
	$C_{L_{mp}} = \sqrt{3}\,C_{L_{md}} = 1.73\,C_{L_{md}}$
$V_{e_{md}} = \left[\frac{2\,W}{\rho_0 S}\sqrt{\frac{K}{C_{D_0}}}\right]^{1/2}$	$V_{e_{mp}} = \left[\frac{2\,W}{\rho_0 S}\sqrt{\frac{K}{3C_{D_0}}}\right]^{1/2}$
	$V_{e_{mp}} = \frac{1}{\sqrt[4]{3}}\,V_{e_{md}} = 0.760\,V_{e_{md}}$
$\left(\frac{L}{D}\right)_{\max} = \frac{1}{2}\sqrt{\frac{1}{KC_{D_0}}}$	$\left(\frac{L}{D}\right)_{mp} = \frac{1}{4}\sqrt{\frac{3}{KC_{D_0}}}$
	$\left(\frac{L}{D}\right)_{mp} = \frac{\sqrt{3}}{2}\left(\frac{L}{D}\right)_{\max} = 0.866\left(\frac{L}{D}\right)_{\max}$

Turbofan Thrust Variation

The following functional relationship describes the net thrust in terms of the dominant parameters:

$$\frac{T}{\delta} = f_1\left(\frac{N}{\sqrt{\theta}}, M\right) \tag{10.87}$$

where δ and θ are the ambient pressure and temperature ratios respectively and N is the rotational speed of the engine (ESDU 70020 1970). In the case of multiple-shaft engines, the speed is usually defined as the speed of the fan or low-speed compressor. For a particular engine setting (i.e., N is constant) the thrust is a function of the atmospheric conditions and the Mach number. It is useful to consider the effect of these variables for various flight conditions. For an airliner the cruise will generally be at a constant Mach number; if the height does not change very much, then, from equation (10.87), it is seen that the thrust will be constant. In the stratosphere, where temperature is constant, the thrust for a given Mach number will decay linearly with pressure as the height increases, i.e.,

$$\frac{T}{\delta} = \text{constant} \tag{10.88}$$

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

In the troposphere there will be a steadily decreasing thrust as the air pressure is reduced. Mattingly et al. (1987) indicate that for turbofan engines with high bypass ratios, the thrust will decay as an approximate power law function of the air density, for any constant Mach number up to 0.9, i.e.,

$$\frac{T}{T_{SL}} \propto \sigma^n \tag{10.89}$$

where T_{SL} is the sea-level reference thrust. The exponent $n = 0.6$ has been shown to provide a reasonable approximation to actual engine data (Mair and Birdsall 1992).

Thrust Ratings

Engine manufacturers will use one of two alternative ways of setting the thrust on the engine. The first method uses the engine speed as the index, and the second, more widely used method uses the *engine pressure ratio* (EPR) as the thrust setting variable. The EPR is essentially the ratio of the pressure in the exhaust flow to the pressure of the flow just ahead of the compressor. To avoid exceeding engine design limitations and to achieve the maximum life of the turbine, the pilot should adhere to specified thrust ratings during each stage of the flight.

Specific Fuel Consumption for Turbojet / Turbofan Engine

The rate at which fuel is burnt in the engine is usually expressed as a specific fuel consumption (SFC) rather than in absolute terms. The SFC of turbojet and turbofan engines is commonly known as the *thrust specific fuel consumption* (TSFC). It may be defined in two ways—a source of much confusion. SFC is either the *mass* of fuel (m_f) burned per unit time, divided by the thrust (convenient when working in SI units), or the *weight* of fuel (W_f) burned per unit time, divided by the thrust (convenient when working in British units). The SFC is a figure of merit used to assess an engine's efficiency in converting fuel into thrust. The symbol that is widely used is c. Various subscripts/superscripts are used to distinguish between different expressions of SFC. Using the first definition:

$$\text{SFC} = c_t = \frac{-(dm_f/dt)}{T} = \frac{Q}{T} \quad \text{(convenient for SI units)} \tag{10.90}$$

where Q is the mass fuel flow. The reason for the minus sign is that the rate of change of aircraft fuel mass is negative but SFC is positive. In consistent SI units SFC will be measured in $\text{kgN}^{-1}\text{s}^{-1}$; however, traditional engineering practice has been to quote the SFC in terms of $\text{kgN}^{-1}\text{h}^{-1}$ (or $\text{mgN}^{-1}\text{s}^{-1}$). Alternatively, SFC may be expressed in terms of weight flow, i.e.,

$$\text{SFC} = c_t' = \frac{-(dW_f/dt)}{T} = \frac{Q'}{T} \quad \text{(convenient for British units)} \tag{10.91}$$

Here the consistent units are $\text{lb lb}^{-1}\text{s}^{-1}$; however, the most widely used units in the industry are $\text{lb lb}^{-1}\text{h}^{-1}$.

The following functional relationship (ESDU 70020 1970) describes the fuel flow in terms of the dominant parameters:

$$\frac{Q}{\delta\sqrt{\theta}} = f_2\left(\frac{N}{\sqrt{\theta}}, M\right) \tag{10.92}$$

Combining this equation with (10.87) and (10.90) leads to the functional relationship for SFC:

$$c_t = \frac{Q}{T} = \sqrt{\theta}\, f_3\left(\frac{N}{\sqrt{\theta}}, M\right) \tag{10.93}$$

This equation indicates that the SFC will depend on height and Mach number for a given engine speed. Note that for an airplane cruising in the stratosphere the SFC will depend only on Mach number for a given engine speed.

SFC Models for Turbojet / Turbofan Engine

It is usually not possible to obtain expressions for the functions f_1, f_2, and f_3, and as a result simple algebraic expressions are used which allow the SFC to be represented with reasonable accuracy over a limited speed range. Several methods are described in ESDU 73019 (1982).

1. The variation of SFC over segments of the cruise is often small, and the use of a mean constant value will usually yield satisfactory results for cruise analysis.
2. A more accurate method is to assume a law of the form:

$$c_t = c_1\sqrt{\theta}\, M^n \tag{10.94}$$

   With suitably chosen values of the constants c_1 and n, this expression is reported to provide an accurate approximation to measured SFC figures for turbofans within the limited range of N, θ, and M values associated with subsonic cruising flight.
3. A third model that takes into account the variation of SFC with M is:

$$c_t = c_2 + c_3 M \tag{10.95}$$

   This can provide a reasonable approximation to the manufacturers' data (at constant height and engine speed).

Turbojet / Turbofan Idealizations

For turbojet/turbofan engines, the following idealizations are often made to simplify the cruise analysis and enable performance problems to be solved analytically:

1. Thrust in the cruise is assumed to be independent of speed.
2. The thrust is assumed to be proportional to ambient pressure for altitude variations.
3. The SFC is assumed to be independent of speed and altitude.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

Turboprop Engine Performance

In a turboprop engine a gas turbine drives the propeller, which generates the thrust. The product of the propeller thrust (T_p) and the forward velocity (V) is by definition equal to the thrust power (power available). A propeller is never 100% efficient in converting the shaft power (P_S) to *thrust power* and it is thus necessary to introduce a propeller efficiency (η_p), i.e.,

$$\eta_p P_S = T_p V \tag{10.96}$$

The analysis of a turboprop engine is complicated by the fact that there is a small amount of jet thrust provided by the residual energy in the exhaust gases. The total thrust (T) must include the residual jet thrust (T_J), i.e.,

$$T = T_p + T_J = T_p \left(1 + \frac{T_J}{T_p}\right) \tag{10.97}$$

It is convenient to rate turboprop engine output in terms of an *equivalent power,* rather than in terms of thrust. The equivalent power (P_e) is equal to the shaft power plus the equivalent power contribution of the direct jet thrust. In other words, the equivalent power is the hypothetical power that would be needed to drive the propeller (at the same propeller efficiency) to produce the same total thrust, i.e.,

$$\eta_p P_e = TV \tag{10.98}$$

With this definition and referring to equations (10.96) and (10.97), the equivalent power can be written as:

$$P_e = P_S \left(1 + \frac{T_J}{T_p}\right) \tag{10.99}$$

Although the ratio T_J/T_p does change during flight, the impact of this variation on the ratio P_e/P_S is small and for most calculations the ratio P_e/P_S may be assumed to be constant.

Power Models for Turboprop

The shaft power of a *flat rated* engine at either the climb or cruise power rating is essentially constant over a range of speeds, from sea level up the rated height, typically more than 10,000 ft. Above this height, the power output will decay with height in much the same way as described above for the turbojet/turbofan engine. The ratio of maximum shaft power to the sea-level reference power $(P_S)_{SL}$ is given by:

$$\frac{P_S}{(P_S)_{SL}} \propto \sigma^n \tag{10.100}$$

which is analogous to equation (10.89). Furthermore, the shaft power will tend to increase with Mach number because of an increase in ram pressure in the inlet. The variation of power with speed at a fixed height has been shown (Mair and Birdsall 1992) to follow an approximate power law relation:

$$\frac{P_S}{(P_S)_{SL}} = AM^n \tag{10.101}$$

where A and the exponent n (between 0 and 1) are selected to suit the engine data.

Specific Fuel Consumption for Turboprop

There are a number of ways of defining SFC for a turboprop engine. It is given the symbol c_p, and when working in SI units it is convenient to define SFC as the *mass* of fuel burned per unit time (Q) divided by the equivalent power (P_e).

$$\text{SFC} = c_p = \frac{\left(-\dfrac{dm_f}{dt}\right)}{P_e} = \frac{Q}{P_e} \quad \text{(convenient for SI units)} \tag{10.102}$$

From a theoretical perspective it is correct to base SFC on the *equivalent power.* However, engine specifications are frequently based on *shaft power,* and it is thus necessary to check this before using the data. In consistent SI units SFC for a turboprop engine will be expressed in terms of kg $W^{-1}s^{-1}$. However, common engineering practice is to use kg $W^{-1}h^{-1}$ or μgJ^{-1}. The alternative definition of SFC is *weight* of fuel burned per unit time divided by the equivalent power, i.e.,

$$\text{SFC} = c_p' = \frac{\left(-\dfrac{dW_f}{dt}\right)}{P_e} = \frac{Q'}{P_e} \quad \text{(convenient for British units)} \tag{10.103}$$

Engineering practice for many years has been to express the SFC of turboprop engines in terms of the inconsistent units of lb $hp^{-1}h^{-1}$. If the SFC is based on shaft power, it is usually written as lb $shp^{-1}h^{-1}$.

Turboprop Idealizations

For turboprops in the range of speeds used for subsonic cruising, shaft power varies with Mach number and altitude, and simple idealizations are not possible. The variation of SFC with height and Mach number is very small and for most calculations can be ignored.

Piston Engine Performance

The power produced by a reciprocating piston engine is directly proportional to the *mass flow* of the air in the intake manifold. Two factors influence the mass flow: the *air density* and the *manifold pressure*. For normally aspirated engines (i.e., not supercharged) the intake manifold pressure will be equal to ambient pressure, or a little greater than ambient at high speed due to the influence of ram air pressure in the inlet. As modern piston engine aircraft fly at low Mach numbers, the ram effect

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

is small and can usually be neglected. A piston engine suffers a considerable drop in power as it climbs. To prevent this loss of power, the intake manifold pressure may be boosted by means of a mechanical air compressor. Both superchargers and turbochargers are capable of maintaining sea-level pressure in the intake manifold to altitudes of about 20,000 ft.

Piston Engine Thrust Power

The product of the propeller thrust and the forward velocity is equal to the thrust power (power available). The engine power (P) is related to the thrust power (P_T) by the propeller efficiency.

$$P_T = \eta_P P \tag{10.104}$$

Specific Fuel Consumption for Piston Engine

The SFC of a piston engine is defined in a similar way to that of a turboprop engine, and the same notation is used. In SI units it is convenient to define SFC as the *mass* of fuel burned per unit time (Q) divided by the engine power (P), while in British units it is common to use the *weight* of fuel burned per unit time divided by the power. Thus,

$$\text{SFC} = c_p = \frac{\left(-\dfrac{dm_f}{dt}\right)}{P} = \frac{Q}{P} \quad \text{(convenient for SI units)} \tag{10.105}$$

or

$$\text{SFC} = c_p' = \frac{\left(-\dfrac{dW_f}{dt}\right)}{P} = \frac{Q'}{P} \quad \text{(convenient for British units)} \tag{10.106}$$

Piston Engine Idealizations

The following idealizations are often made in order to analyze the airplane's cruise performance numerically.

1. For a given throttle setting and altitude, the power is assumed to be independent of speed (for the range of speeds normally used for cruising).
2. The shaft power is assumed to be proportional to ambient density for altitude variations. (For engines that are supercharged or turbocharged, sea-level power will be maintained to the rated height.)
3. The SFC is assumed to be independent of speed and altitude.

Further Reading

For further information on turbine engine performance consult Mattingly et al. (1987) and Mair and Birdsall (1992). Lan and Roskam (1981) and Lowry (1999) describe the performance of piston engines; both references provide details on propeller performance and efficiency.

10.16 LEVEL FLIGHT PERFORMANCE

Level Flight Turbojet/Turbofan Performance

In this analysis the airplane is considered to be flying straight (i.e., not turning) and level (i.e., not climbing or descending) at constant velocity. A very useful analysis technique is to superimpose on a single graph the thrust (T) and drag (D) variations as functions of V_e. For the idealized turbojet/turbofan engine (see Subsection 10.15) operating at a set throttle position, the thrust decreases with altitude and is independent of V_e. The family of T curves is drawn as horizontal lines, with the greatest thrust at sea level and reducing with increasing height. The D curve for a given aircraft weight is independent of altitude if the function is plotted against EAS rather than TAS. A single curve will thus represent the drag at any altitude. The drag acting on the airplane may be modelled by equation (10.77), derived from the parabolic drag polar. The intersection points of the T and D curves represent a series of steady state level flight conditions (Figure 10.72). A low thrust line on the graph will have two points of intersection with the drag function, one at a speed less than minimum drag speed ($V_{e_{md}}$) and one at a speed greater than $V_{e_{md}}$. The thrust functions for which this will be true correspond to flight at high altitude where the thrust has been substantially reduced, or alternatively at any altitude, but with a reduced throttle setting.

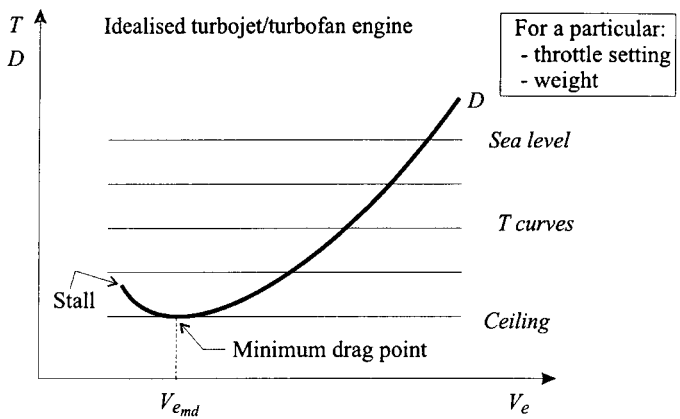


FIGURE 10.72 Thrust and drag for the idealized turbojet/turbofan engine airplane.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

Level Flight Piston Engine Performance

For the idealized piston engine the T curves are rectangular hyperbolas, as thrust power (P_T) is constant. For a set throttle position, the thrust variation with altitude is shown superimposed on the drag versus EAS graph in Figure 10.73. The lowest thrust curve that will intersect the drag curve, giving the absolute ceiling, does so not at the $V_{e_{md}}$, as was the case for the turbojet/turbofan engine, but at the minimum power speed ($V_{e_{mp}}$). From a practical perspective, it is not convenient to represent piston engine performance in terms of thrust. Piston engines are rated in terms of power, and changes in the throttle setting will result in changes in the thrust power. The thrust may be replaced by an equivalent group of terms which includes the engine power and propeller efficiency. Based on equations (10.96) and (10.67) the thrust may be expressed as:

$$T = \frac{\eta_p P \sqrt{\sigma}}{V_e} \tag{10.107}$$

Maximum Level Speed

It is apparent that the construction of T and D versus V_e graphs for the maximum throttle setting will give, at their intersection, the maximum level speed condition at any altitude. For a piston engine aircraft (flying at speeds below any compressibility effects), the *maximum level speed* corresponding to a maximum power ($P_{\max}$) may be determined using the parabolic drag polar. As the aircraft is in steady level flight, thrust power equals drag power and hence from equation (10.86), it can be deduced that:

$$\eta_p P = \frac{C_{D_0} \rho_0 S}{2\sqrt{\sigma}} V_e^3 + \frac{2KW^2}{\rho_0 S \sqrt{\sigma}} V_e^{-1} \tag{10.108}$$

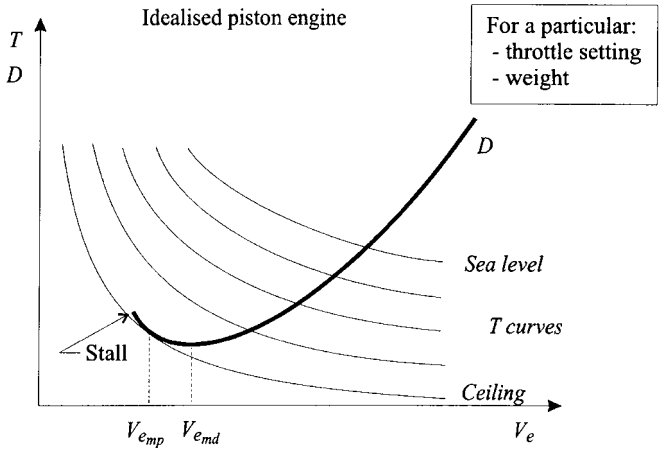


FIGURE 10.73 Thrust and drag for the idealized piston engine airplane.

Equation (10.108) applies to any steady level flight condition. Note that the equation does not provide a closed form solution for V_e and must be solved by iteration. For the maximum level speed, it is possible to obtain an approximate solution by noting that the second of the two terms is a function of V_e^{-1} and this term becomes very small (in comparison to the first term), as speed increases. Hence,

$$V_{e_{\max}} \approx \sqrt[3]{\frac{2\eta_p P_{\max}\sqrt{\sigma}}{C_{D_0}\,\rho_0 S}} \tag{10.109}$$

For turbojet/turbofan-powered aircraft the maximum speed is usually within the drag rise Mach regime, and the low-speed drag polar is not applicable, rendering the approach given above rather inaccurate due to the difficulty in accounting for the wave drag.

Speed Stability—Turbojet/Turbofan Airplane

In Figure 10.74 the thrust for a selected throttle setting of a jet airplane is superimposed on the drag relationship. If the aircraft is flying straight and level at a speed V_{e_1}, it will be operating at a condition of *equilibrium*. If the aircraft speeds up a little (without the pilot changing the throttle position) to a speed V_{e_2}, then T will be less than D and the aircraft will tend to slow down to the original speed. By a similar argument, if the speed decreased to V_{e_3}, this would cause the aircraft to accelerate, as the drag would have decreased and T will be greater than D. At the slow speed equilibrium point V_{e_4} the situation is different. If the aircraft slows down, the drag will increase, causing the aircraft to slow down further. It can thus be concluded that a jet airplane is unstable with regard to speed changes for operation below the minimum drag speed ($V_{e_{md}}$). This region is called the back end of the drag curve. If the aircraft slows down from $V_{e_{md}}$ there will be a thrust deficit and the aircraft will start to sink, requiring a corrective action by the pilot.

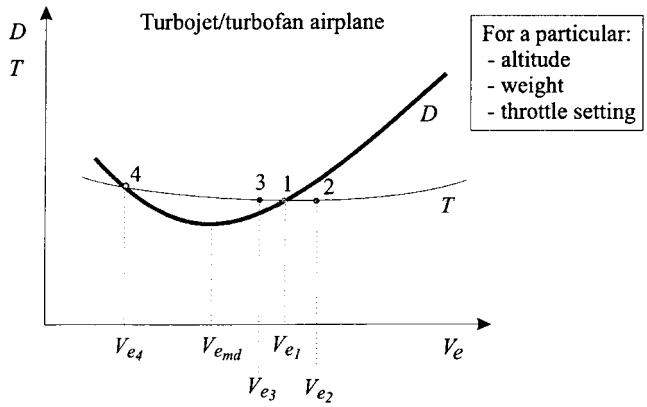


FIGURE 10.74 Speed stability for turbojet/turbofan engine airplane.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

Speed Stability—Piston Engine Airplane

For piston engine aircraft, a similar but not identical deduction to that presented for the jet airplane can be made. Although the form of the drag function is unchanged, the thrust curves are rectangular hyperbolas (for the idealized engine). The curve corresponding to the lowest power setting that will permit steady level flight touches the drag curve at the minimum power speed ($V_{e_{mp}}$). For speeds less than $V_{e_{mp}}$ the aircraft will be statically unstable with regard to speed changes. From a performance perspective, there are two important differences regarding speed instability between piston engine and jet aircraft: the first is that because $V_{e_{mp}} = 0.76V_{e_{md}}$, the speed instability affects a smaller speed range for the piston engine airplane, and the second is that because piston engines produce high thrust at low speed, the thrust deficit is comparatively small. For these reasons, pilots of piston engine light aircraft are usually not aware of the condition.

Absolute Ceiling—Turbojet/Turbofan

As the thrust produced by a jet engine decays with altitude, there will be a maximum altitude (for every throttle setting) beyond which it will not be possible for the airplane to climb. The absolute ceiling is defined as the maximum altitude at which level flight can be maintained with maximum available thrust. An aircraft can fly higher than its absolute ceiling by a zoom maneuver, in which kinetic energy is exchanged for potential energy, but these altitudes cannot be sustained. The absolute ceiling is given by the intersection of the lowest thrust curve with the lowest point on the drag curve on Fig 10.72. Steady flight is possible only at one speed; based on the parabolic drag polar and the idealized thrust relationship, this speed is $V_{e_{md}}$. Flight at the absolute ceiling is difficult to sustain and largely of theoretical interest.

Service Ceiling—Turbojet/Turbofan

The service ceiling is a practical upper limit and is the greatest altitude at which a maximum rate of climb of 100 ft/min (0.508 m/s) can be achieved. The cruise ceiling (used in military applications) is the greatest altitude at which a rate of climb of 300 ft/min (1.53 m/s) can be achieved (Raymer 1989).

Single-Engine Inoperative Service Ceiling

In the event of an engine failure of a multiple-engine aircraft, the service ceiling will be significantly reduced as it depends on the net thrust. En route flight planning over mountainous terrain is based on an airplane's single-engine inoperative service ceiling, defined in terms of a minimum gradient that must be maintained by the remaining engines (set at maximum continuous thrust) following an engine failure, i.e., 1.6% for four engine aircraft, 1.4% for a three engine aircraft, and 1.1% for a two engine aircraft (FAR Part 25.123). Airlines plan flights in such a way that following an engine failure at any stage during the flight, the airplane will drift down to its single-engine inoperative service ceiling, clearing all obstacles along a predefined corridor by at least 2000 ft.

Ceiling—Piston Engine

For the idealized piston engine aircraft the absolute ceiling is achieved by flying at V_{emp}. For piston engine airplanes the absolute ceiling is largely of theoretical significance because most of these aircraft are unpressurized and will usually not be flown higher than about 10,000 ft (a height at which oxygen is required).

Further Reading

For further details on level flight performance, speed stability, and aircraft ceilings, consult Raymer (1989), Mair and Birdsall (1992), Ojha (1995), Anderson (1999), Lowry (1999), or Eshelby (2000).

10.17 CLIMBING AND DESCENDING FLIGHT

Climb Speed Schedule

The initial part of a typical climb for an airliner will be at constant CAS. The implication of this is that the TAS will increase as the air density drops (Figure 10.75). The combined effect of an increase in TAS and a reduction in the speed of sound results in a rapid increase in the Mach number. To avoid a substantial increase in drag as the speed approaches the drag rise Mach number, a change to the climb schedule will usually be required at some point. The climb speed schedule may, for example, require the pilot to climb at 300 kt (CAS) until Mach 0.82 is reached and then to hold the Mach number constant for the remainder of the climb. Climbing at constant Mach number implies that there will be a slight decrease in TAS up to the tropopause, but in the stratosphere it will be at constant TAS.

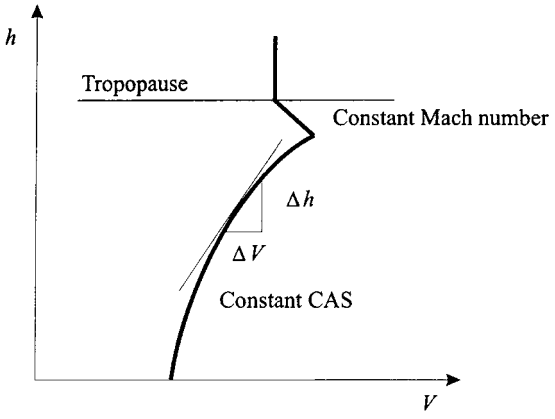


FIGURE 10.75 Typical airliner climb schedule: constant CAS, followed by constant Mach number climb.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

Climb Analysis

Figure 10.76 shows an aircraft performing a climb in still air with a climb angle of γ. For a *steady* climb the aircraft is at a constant TAS and is in a state of equilibrium at all points along the flight path. The sum of the forces acting along the flight path in this case is zero. In flight this is seldom the situation and very often the aircraft will *accelerate* along the flight path, as is the case for a constant CAS climb. Under these flight conditions the airplane's flight path is not exactly straight but will have a slight curve. The curvature, however, is very small and the centripetal acceleration is approximately zero and, for a typical climb analysis, can be ignored. For most problems involving climbing or descending flight, the angle between the thrust line and the flight path is relatively small and is ignored.

Angle of Climb

By summing the forces acting on the aircraft in Figure 10.76, it can be shown that the climb angle (γ) relates to the thrust (T), weight (W), and drag (D) by the following relationship:

$$\sin \gamma = \frac{\dfrac{T}{W} - \dfrac{D}{W}}{\left(1 + \dfrac{V}{g}\dfrac{dV}{dh}\right)} \tag{10.110}$$

where h is the height. This equation is valid for a climb (γ positive) or descent (γ negative) and is applicable to an aircraft performing an accelerated climb or descent. For steady flight the rate of change of speed with respect to height is zero and the equation may be simplified, i.e.,

$$\sin \gamma = \frac{T}{W} - \frac{D}{L}\cos \gamma \tag{10.111}$$

The angle of climb for the average steady climb is small, permitting a small angle approximation to be used:

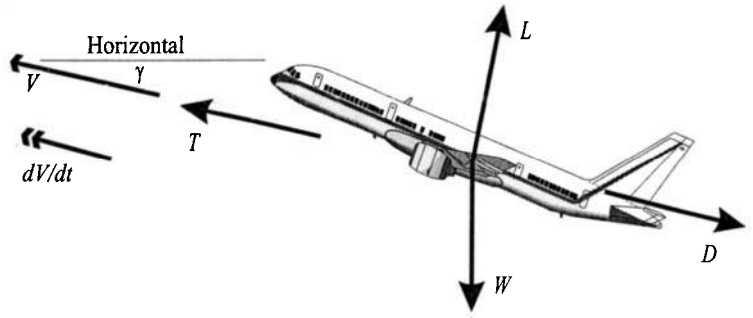


FIGURE 10.76 Climbing flight.

$$\gamma \approx \frac{T}{W} - \frac{D}{L} \tag{10.112}$$

Equation (10.112) provides a graphical method for the determination of the climb angle. A plot of the ratios (T/W) and (D/L) for a selected aircraft weight, superimposed on the same graph, as a function of V_e is prepared. The angle of climb (measured in radians) is given by the difference between the two curves. This is illustrated for the idealized turbojet/turbofan in Figure 10.77.

Climb Gradient

The climb gradient is given by tan γ. The climb gradient represents the ratio of the gain in height to the horizontal distance flown and is usually expressed as a percentage.

Best Angle of Climb Speed for Turbojet/Turbofan Aircraft

The flight conditions that will give the maximum angle of climb are of interest in clearing obstacles after takeoff. For the idealized turbojet/turbofan it is evident from Figure 10.77 that the best angle of climb speed, at any altitude is $V_{e_{md}}$.

Best Angle of Climb Speed for Piston Engine Aircraft

The expression for the angle of climb is:

$$\sin \gamma = \frac{\eta_p P \sqrt{\sigma}}{W V_e} - \frac{D}{W} \tag{10.113}$$

In the case of the idealized piston engine, the thrust curves are rectangular hyperbolas and the speed that will achieve the maximum angle of climb will be between the stall speed and $V_{e_{mp}}$, but will depend on the altitude. It is possible to obtain an expression for the best angle of climb speed for a piston engine aircraft. However the expression does not have a closed-form solution and must be solved by iteration. Using the parabolic drag polar, the equation that will give the best angle of climb speed for a piston engine aircraft is:

$$C_{D_0} \rho_0 S V_e^4 + \eta_p P \sqrt{\sigma} V_e - \frac{4KW^2}{\rho_0 S} = 0 \tag{10.114}$$

The solution predicted by this equation is based on the parabolic drag polar and the assumptions of the idealized piston engine/propeller combination. At the very low speeds associated with this flight condition, the accuracy of these idealizations is reduced, resulting in a poor speed estimation. If actual data of power, drag and propeller efficiency are available, then a superior approach is to use equation (10.113) and solve for the optimum speed graphically.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

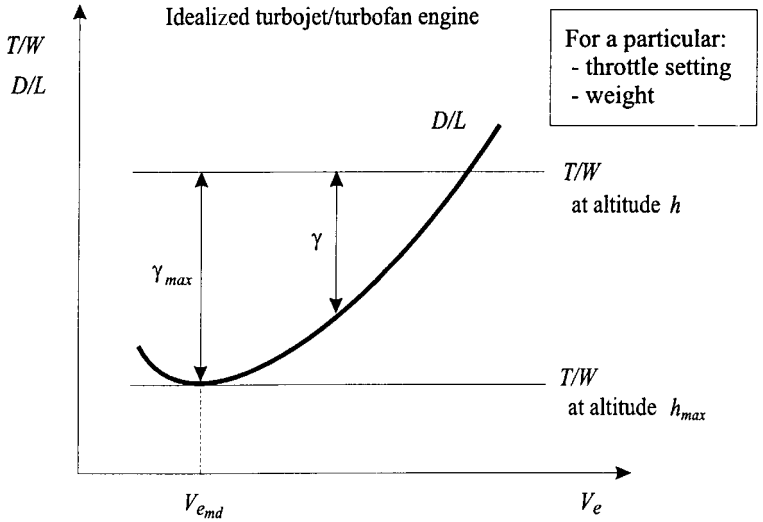


FIGURE 10.77 Angle of climb for the idealized turbojet/turbofan engine airplane.

General Equation for Rate of Climb (ROC)

The rate of climb is often written as R/C, but this notation has been the source of some confusion, particularly when written in an equation, and for this reason the abbreviation ROC will be used here. The rate of climb in the absence of updrafts is the change of height with respect to time. From equation (10.110) the ROC may be deduced:

$$\text{ROC} = \frac{dh}{dt} = V \sin \gamma = \frac{\left(\frac{T}{W} - \frac{D}{W}\right) V}{\left(1 + \frac{V}{g}\frac{dV}{dh}\right)} \tag{10.115}$$

The usual way to analyze equation (10.115) is to introduce an acceleration factor, defined as:

$$f_{acc} = \frac{V}{g}\frac{dV}{dh} \tag{10.116}$$

hence,

$$\text{ROC} = \frac{\left(\frac{T}{W} - \frac{D}{W}\right) V}{(1 + f_{acc})} \tag{10.117}$$

The significance of this factor is illustrated in Figure 10.75. The slope of the curve is (dh/dV), which changes with height. For a constant CAS climb the *acceleration factor* will increase as the altitude increases. This is evident from the fact that the TAS increases and (dV/dh), which is the reciprocal of the slope of the line, also

increases for a constant CAS climb. The acceleration factor depends on the Mach number, the height, and the climb speed condition. Equations to determine f_{acc} are given in Table 10.12. For high-speed aircraft, it is preferable to express the ROC in terms of Mach number and divide both numerator and denominator by the pressure ratio (δ) to get the thrust in a form consistent with the usual presentation of thrust data for a jet engine, i.e.,

$$\text{ROC} = \frac{\left(\frac{T}{\delta} - \frac{D}{\delta}\right) Ma_0\sqrt{\theta}}{(1 + f_{acc})\frac{V}{\delta}} \tag{10.118}$$

Rate of Climb (ROC) at Constant TAS

For a steady unaccelerated climb (i.e., constant TAS) f_{acc} is zero and equation (10.117) reduces to:

$$\text{ROC} = \frac{(T - D)V}{W} = \frac{P_T - P_D}{W} \tag{10.119}$$

The steady rate of climb is thus proportional to the excess of thrust power (power available) to drag power (power required).

Maximum Rate of Climb Speed for Turbojet/Turbofan

The optimum climb speed that will reduce the total trip fuel is very close to the maximum rate of climb speed. (Airliners are most efficient during cruise, so it is desirable to climb as fast as possible.) The maximum rate of climb is achieved at a speed higher than the minimum drag speed. An approximation of the best rate

TABLE 10.12 Acceleration Factor for ISA Conditions

The acceleration factor is by definition: $f_{\text{acc}} = \frac{V}{g}\frac{dV}{dh}$

and its magnitude is given by: $f_{\text{acc}} = 0.7\, M^2\psi$

Where for:

Constant Mach no.	$\psi = -C$
Constant EAS	$\psi = 1 - C$
Constant CAS	$\psi = \frac{\{[1 + 0.2\, M^2]^{3.5} - 1}{0.7\, M^2\, [1 + 0.2\, M^2]^{2.5}} - C$

And where:

$C = 0.190263$ (below tropopause)

$C = 0$ (above tropopause)

Based on Boeing 1989.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

of climb speed for the idealized turbojet/turbofan at a particular altitude and throttle setting may be derived using the parabolic drag expression.

$$V_e = \left[\frac{W}{3C_{D_0}\rho_0 S}\left\{\frac{T}{W} \pm \sqrt{\left(\frac{T}{W}\right)^2 + 12KC_{D_0}}\right\}\right]^{1/2} \tag{10.120}$$

Alternatively, a graphical method may be used to determine the speed for the best rate of climb. The latter method is particularly suitable for problems where the parabolic drag polar or the idealized thrust function is not valid.

Maximum Rate of Climb Speed for Piston Engine Aircraft

In the case of the idealized piston engine aircraft $(T/W)V_e$ lines are horizontal, as thrust power is independent of speed (Figure 10.78). It is seen that the maximum rate of climb for the idealized piston engine aircraft is at the minimum power speed. The comment made earlier (in the angle of climb analysis) regarding the validity of the parabolic drag polar and power idealization at low speeds is also valid here. A graphical approach to determine the best rate of climb speed is better if data are available.

Time to Climb

The time to climb from height h_1 to height h_2 is given by:

$$t = \int_{h_1}^{h_2} \frac{1}{\text{ROC}}\, dh \tag{10.121}$$

where the ROC is given by equation (10.117) or (10.119). The integration of equa-

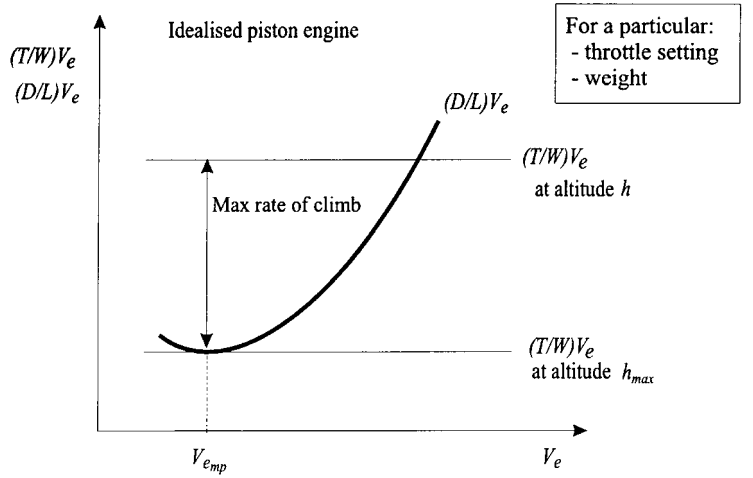


FIGURE 10.78 Rate of climb for the idealized piston engine airplane.

tion (10.121) for a general problem is not easy because of the interdependency of the many variables. The best approach for determining the time to climb is to divide the climb into intervals, each interval corresponding to a change of height (Δh). The ROC is then determined at the start of the ith interval (ROC_i) and at end of the interval (ROC_{i+1}) based on the aircraft weight at the start of the interval (W_i). By making the assumption that the ROC will change linearly across the interval, equation (10.121) may be integrated to give the increment in time (Δt_i) for the ith interval. The weight at the end of the interval (W_{i+1}) is now determined from the fuel flow and the time Δt_i. To improve the accuracy of the calculation, the ROC at the end of the interval is then recalculated using W_{i+1} and a second iteration of the increment in time is obtained. Based on a revised weight for the end of the interval, the next interval is considered.

Effect of Wind on Climb Performance

In Figure 10.79 the aircraft is shown to be flying in an air mass and the entire air mass is moving with a speed V_w. The still air angle of climb is γ. The ground speed is given by a sum of the aircraft's TAS vector and the wind velocity. For a tailwind the absolute climb gradient is reduced, but the rate of climb is unchanged.

Angle of Descent and Descent Gradient

The angle of climb and the climb gradient expressions derived above are also valid for the descent; the only difference is that the angle γ is negative. The optimum decent conditions for airline operations are those that result in the lowest trip fuel. As the descent is usually at idle thrust (lowest fuel burn), the optimum flight plan is one that results in the longest distance in descent, which implies that the descent speed should be close to the flight condition for maximum lift-to-drag ratio. For

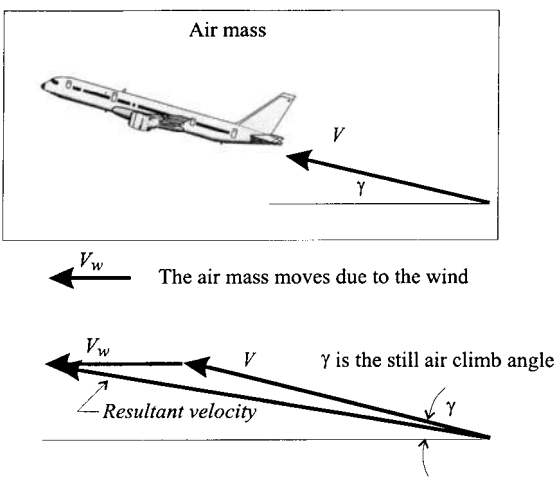


FIGURE 10.79 Effect of wind on climb performance.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

transport aircraft operations the descent may be determined by the rate of re-pressurization of the cabin or air traffic control requirements.

Rate of Descent (ROD)

The rate of descent, often written as R/D, may be determined directly from equations (10.117) and (10.119) derived for the rate of climb. The only difference is that the ROD is positive when (dh/dt) is negative.

Glide Angle for Unpowered Flight

Multiple engine failures on airliners are a very rare occurrence. There have been a few reported cases where the engines have failed simultaneously due to fuel starvation or the ingestion of volcanic dust. For light aircraft the incidence rate is much higher. The better an airplane can glide, the more time the pilot will have to restart the engine or find a suitable place to land. It can be shown that an aerodynamically efficient airplane with a high lift-to-drag ratio will glide very well. For a steady descent (constant TAS) with zero thrust, equation (10.112) may be rewritten as:

$$\gamma \approx -\frac{1}{L/D} \tag{10.122}$$

where γ is defined positive for a climb. The smallest glide angle (which will produce the maximum range in still air) will be achieved at the flight condition of maximum lift-to-drag ratio, which implies that the aircraft should be flown at the minimum drag speed ($V_{e_{md}}$).

Sailplane Performance

For unpowered flight the term *rate of sink* (ROS) is preferable to *rate of descent*. In essence the two terms are synonymous. For steady (constant TAS flight) the ROS for unpowered flight is given by:

$$\text{ROS} = \left(\frac{D}{W}\right) V = \frac{P_D}{W} \tag{10.123}$$

In order to achieve the maximum duration in a glide from a given altitude, the aircraft must fly at a slower speed than that required for the best glide angle. From equation (10.123) it can be deduced that in still air the lowest ROS will occur at the airplane's minimum power speed ($V_{e_{mp}}$). For this class of airplane use is made of the term *glide ratio* as a figure of merit to characterize performance. The glide ratio is the horizontal distance covered divided by the loss of height, hence,

$$\text{glide ratio} = -\tan\gamma \approx \frac{L}{D} \tag{10.124}$$

This has a maximum at the same condition as the minimum glide angle. Based on the parabolic drag polar

$$\left(\frac{L}{D}\right)_{\max} = \frac{1}{2}\sqrt{\frac{\pi A_R e}{C_{D_0}}}$$

hence good gliding performance is associated with aircraft of low C_{D_0} and superbly efficient, high aspect ratio (A_R) wings.

Further Reading

For further details on climbing and descending flight, including energy methods, consult Mair and Birdsall (1992), Ojha (1995), Anderson (1999), and Eshelby (2000). Lowry (1999) deals with light aircraft and glider performance, expanding the basic analysis (presented herein) to take into account the effect of wind.

10.18 TURNING PERFORMANCE

Load Factor

The load factor (n) is by definition equal to the ratio of lift to weight. In straight and level flight, n equals 1; however, in a turn or a pull-up maneuver, n will be greater than 1. By definition,

$$n = \frac{L}{W} \tag{10.125}$$

During maneuvers the load factor may not exceed the structural limits of the airplane. The limiting envelope is called a V–n diagram and is a plot of allowable load factor versus airspeed.

Sustained Level Turn

For a sustained level turn the aircraft must maintain both speed and height in the turn. The additional lift (compared to level flight) required to provide the centripetal acceleration results in an increase in the drag. If the pilot progressively increases the angle of bank, tightening the turn, the required thrust to sustain the turn will increase. This may continue until a limiting condition is reached; the limit may be imposed by the maximum available thrust, the maximum load factor, or the maximum lift coefficient. It may not always be possible to sustain a constant speed in a level turn. If the speed is permitted to drop, the resulting instantaneous turn rate could be higher than the comparative sustained turn rate.

Angle of Bank

In a correctly banked (coordinated turn) the airplane sideslips neither inwards (overbanked) nor outwards (underbanked), so that the lift force lies in the aircraft's plane of symmetry. The component of lift acting towards the center of the turn provides the required centripetal force to accelerate the aircraft of mass m in a

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

circular flight path of radius r, as shown in Figure 10.80. In the analysis presented below it is assumed that the thrust axis is approximately aligned with the flight direction.

Resolving the forces vertically and horizontally provides the following two equations:

$$L \cos \phi = W = mg \tag{10.126}$$

$$L \sin \phi = ma_c = m \frac{V^2}{r} = mV\Omega \tag{10.127}$$

where a_c = the centripetal acceleration
Ω = the rate of turn (units of radians per second)

In a turn the load factor may be obtained from equations (10.125) and (10.126).

$$n = \frac{L}{W} = \frac{1}{\cos \phi} \tag{10.128}$$

As the angle of bank is increased, it is evident that the load factor will increase. The maximum angle of bank in a sustained turn may therefore be restricted by the maximum allowable load factor.

Turn Rate in Sustained Level Turn

The turn rate is obtained from equations (10.126)–(10.128).

$$\Omega = \frac{q}{V} \sqrt{n^2 - 1} \tag{10.129}$$

For a given speed the maximum rate of turn will occur at the maximum allowable load factor. Conversely, for a given load factor, the absolute maximum rate of turn will occur at the lowest possible speed, the turn stall speed, provided that there is sufficient thrust to maintain the turn. Note that for most flights the pilot will not maneuver the airplane at turn rates anywhere near the limiting conditions. Under normal conditions pilots often execute a so-called rate one turn, which implies a turn rate of 180° in one minute, or 3°/s.

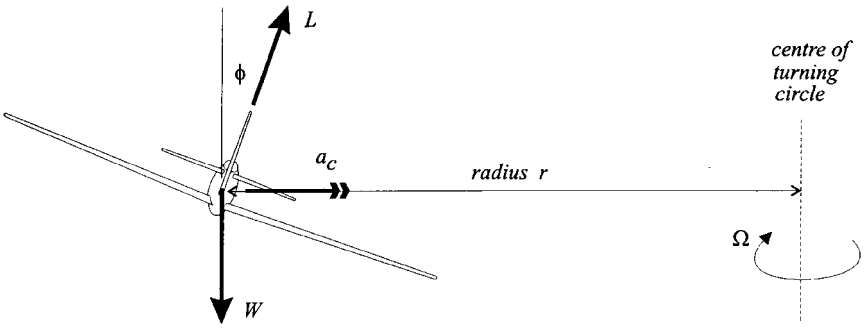


FIGURE 10.80 Turning performance.

Radius of Turn in Sustained Level Turn

The radius of turn (r) is given by:

$$r = \frac{V^2}{g\sqrt{n^2 - 1}} = \frac{V^2}{\text{g tan } \phi} \tag{10.130}$$

The radius of turn is an important parameter for obstacle clearance after takeoff. Routine flight planning for aircraft operations out of airports with mountains in the vicinity must take into account the impact on the flight path due to the loss of thrust resulting from an engine failure on a multiple-engine aircraft.

Further Reading

For further details on turn performance and other maneuvers consult Mair and Birdsall (1992), Ojha (1995), Anderson (1999), Lowry (1999), and Eshelby (2000). Maneuver load factor design limitations are described by Torenbeek (1982) and Raymer (1989).

10.19 STALL AND SPIN

Stall Condition

An airplane stalls when the angle of attack exceeds the critical stall angle of attack. To enter a stall from level flight, the pilot would reduce the airspeed and, to compensate for the loss of lift, simultaneously pull the stick/yoke back, increasing the angle of attack. The stall speed is the lowest speed at which steady controllable flight can be maintained. This is preceded in most aircraft by buffeting, associated with the initial separation of airflow from the upper wing surface, striking the tailplane. A further increase in the angle of attack results in substantial separation of the flow on the upper wing surface, a reduction in lift, and an increase in drag. The aircraft loses height, and for most designs the strong nose-down pitching moment associated with the stall will rapidly reduce the angle of attack, reattaching the airflow on the wing. After initially relaxing the stick/yoke, the pilot will raise the nose and apply power to restore steady flight. It is important that the stall speeds be correctly calculated as they impact directly on the operational safety of the airplane. Apart from the obvious desire for the pilot not to inadvertently stall the aircraft without sufficient height for recovery, the importance of the stall speeds are linked to their use as reference speeds during takeoff and landing.

Level Flight Stall Speed

The C_L corresponding to the stall speed (V_S) is the maximum lift coefficient ($C_{L_{\max}}$), which depends on the airplane's configuration, or more precisely on the position of the high-lift devices. The purpose of leading- and trailing-edge devices, such as flaps and slats, is to increase the value of $C_{L_{\max}}$ and to delay the stall. In level flight, the lift is equal to the weight and the corresponding stall speed, designated as V_{S_1}, is given by:

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

$$V_{S_1} = \sqrt{\frac{W}{\frac{1}{2}\rho SC_{L_{\max}}}} \tag{10.131}$$

Strictly speaking the V_{S_1} speed is defined by a 1 kt/s deceleration with the power off.

Maneuver Stall Speed

For an airplane performing a maneuver, such as a pull-up or a turn, the stall speed will be *higher* than the stall speed determined by equation (10.131). As the aircraft is not flying straight and level, the lift is equal to the weight multiplied by the load factor, hence,

$$V_S = \sqrt{\frac{nW}{\frac{1}{2}\rho SC_{L_{\max}}}} = \sqrt{n}V_{S_1} \tag{10.132}$$

Factors Influencing Stall Speed

It is evident that an airplane does not have a single stalling speed, as this depends on air density, weight, aircraft configuration, and load factor. Furthermore, a change in the stall characteristics will occur when the aircraft is stalled with power on. At high angles of attack the thrust will have a significant vertical component. This will produce a small increment to the total lift, which will reduce the stall speed. For propeller-driven aircraft, with engines mounted ahead of the wings, the influence of the slipstream on the air flowing over the wings is to delay flow separation and reduce the stall speed.

Spin

Due to an asymmetry of the airflow over the wings, it is possible that one wing will stall before the other. This will result in a rolling moment because of the reduction of lift on the stalled wing, and a yawing moment because of a local increase in drag. The result is that the airplane may enter into a tight spiral or spin. A pilot may deliberately put an airplane into a spin by progressively raising the nose, simultaneously reducing the engine power, and then, at the onset of the stall, deliberately yawing the aircraft using the rudder. This will result in the nose dropping abruptly, with one wing falling faster than the other and a rapid yawing of the aircraft. Depending on the airplane's aerodynamic characteristics, mass distribution, and direction of propeller motion, it may continue to spin, or it may stop yawing and just descend in a stall. If the spin becomes established, the airplane will continue to yaw at a steady rate, rapidly losing height in a motion that involves roll and sideslip. Recovery from an established spin is initiated by the pilot stopping the yawing motion by applying maximum rudder deflection in the opposite direction to the spin. The pilot then recovers from the stall condition by initially allowing

the nose to drop a little, until airflow is reattached over the wings, and then pulling back on the stick/yoke, raising the nose and applying power to sustain the pull-out from the dive. For some aircraft designs the relative position of the horizontal tailplane results in a wash of stalled air striking the fin in a spin, making it very difficult, and in some cases impossible, for the pilot to stop the yawing motion, particularly when the center of gravity is located in an aft position. Flight testing may require the installation of a spin parachute to assist in the recovery. Inverted spin recovery is not possible for many aircraft.

Further Reading

Consult Stinton (1996).

10.20 RANGE AND ENDURANCE

Fuel Consumption Definitions

The *range* relates to the *distance* that an aircraft can fly on a given quantity of fuel, while the *endurance* relates to the *time* for which the aircraft can fly on that fuel quantity. Both parameters depend on the *rate* at which the fuel is consumed. As explained in Subsection 10.15, the fuel flow may be defined as either the *mass* of fuel consumed per unit time, which is convenient for SI units (and is given the symbol Q), or the *weight* of fuel consumed per unit time, which is convenient for British units (given the symbol Q'). In this subsection the range and endurance equations are derived from fuel flow definitions based on the mass flow rate. The alternative expressions based on the weight flow rate are summarized in Table 10.13. For a jet engine it is evident from equation (10.90) that the fuel flow is the

TABLE 10.13a Summary of Range and Endurance Expressions (Mass Flow Basis)

	Range	Endurance
Governing equation	$R = \int_{m_2}^{m_1} \frac{V}{Q}\, dm$	$t = \int_{m_2}^{m_1} \frac{1}{Q}\, dm$
Turbojet/turbofan airplane	$Q = c_t\, g \left(\frac{D}{L}\right) m$	
	$R = \frac{V}{c_t g}\left(\frac{L}{D}\right) \ln\left(\frac{m_1}{m_2}\right)$ for constant c_t, V, and C_L	$t = \frac{1}{c_t g}\left(\frac{L}{D}\right) \ln\left(\frac{m_1}{m_2}\right)$ for constant c_t and C_L
Piston-propeller airplane	$Q = \frac{c_p\, g\, m\, V}{\eta_P}\left(\frac{D}{L}\right)$	
	$R = \frac{\eta_p}{c_p g}\left(\frac{L}{D}\right) \ln\left(\frac{m_1}{m_2}\right)$ for constant η_p,c_p, and C_L	$t = \frac{\eta_p}{c_p g V}\left(\frac{L}{D}\right) \ln\left(\frac{m_1}{m_2}\right)$ for constant η_p, c_p, V and C_L

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

TABLE 10.13b Summary of Range and Endurance Expressions (Weight Flow Basis)

	Range	Endurance
Governing equation	$R = \int_{W_2}^{W_1} \frac{V}{Q'}\, dW$	$t = \int_{W_2}^{W_1} \frac{1}{Q'}\, dW$
Turbojet/turbofan airplane	$Q' = c_t' \left(\frac{D}{L}\right) W$	
	$R = \frac{V}{c_t'} \left(\frac{L}{D}\right) \ln \left(\frac{W_1}{W_2}\right)$ for constant c_t', V, and C_L	$t = \frac{1}{c_t'} \left(\frac{L}{D}\right) \ln \left(\frac{W_1}{W_2}\right)$ for constant c_t' and C_L
Piston-propeller airplane	$Q' = \frac{c_p' W V}{\eta_P} \left(\frac{D}{L}\right)$	
	$R = \frac{\eta_p}{c_p'} \left(\frac{L}{D}\right) \ln \left(\frac{W_1}{W_2}\right)$ for constant η_p, c_p', and C_L	$t = \frac{\eta_p}{c_p V} \left(\frac{L}{D}\right) \ln \left(\frac{W_1}{W_2}\right)$ for constant η_p, c_p', V, and C_L

product of SFC and thrust, while for a piston or turboprop engine, it is the product of SFC and *power*, as is evident from equations (10.102) and (10.105).

Specific Air Range (SAR)

The specific air range (r_a), also referred to as the *specific range* or the *fuel mileage,* is defined as the distance travelled per unit fuel mass consumed. Thus,

$$r_a = \left(-\frac{dx}{dm_f}\right) = \frac{\left(\frac{dx}{dt}\right)}{\left(-\frac{dm_f}{dt}\right)} = \frac{V}{Q} \tag{10.133}$$

where x = still air distance
m_f = onboard fuel mass
Q = net *mass* flow of fuel to all engines.

The reason for the minus sign in equation (10.133) is that the change of fuel dm_f is a negative quantity and SAR is a positive quantity.

Cruise Speeds for Jet Airplanes

The greatest possible range that an airplane may achieve (for a fixed fuel quantity) is obtained by flying at all times at the flight condition for maximum SAR. This is called the maximum range speed (MRS). The MRS decreases as fuel is burnt (at a set altitude). In practice, airlines usually fly faster than this, sacrificing a small increase in fuel to obtain a shorter cruise time. A portion of an airline's cost is proportional to the flight time, and flying faster will reduce this. The speed that

will give the lowest total trip cost for a particular set of operating costs, is called the economy (econ) speed. This can be difficult to calculate without complete cost data, and a simpler approach is to fly at a fixed percentage faster than the MRS. The so-called long-range speed (LRS) is typically set about 2% to 4% faster than the MRS and has a 1% reduction in SAR, as shown in Figure 10.81.

Turbojet / Turbofan Airplane Range Equation

The change in airplane mass is equal to the change in total onboard *fuel mass*. (This is obviously true for all commercial aircraft operations, but not for military operations where weapons are released.) The still air range, R, for an airplane with initial mass m_1 and final mass m_2 is given by:

$$R = -\int_{\text{start}}^{\text{end}} r_a dm = -\int_{m_1}^{m_2} \frac{V}{Q}\, dm \tag{10.134}$$

For a jet airplane in level flight, the fuel flow may be written as:

$$Q = c_t T = c_t \left(\frac{D}{L}\right) mg \tag{10.135}$$

Hence,

$$R = -\int_{m_1}^{m_2} \frac{V}{c_t g m} \left(\frac{L}{D}\right) dm \tag{10.136}$$

Breguet Solution for the Turbojet / Turbofan Airplane

To evaluate equation (10.136) it is necessary to describe the variables c, V, and (L/D) during the cruise. The most widely used solution to this equation is obtained

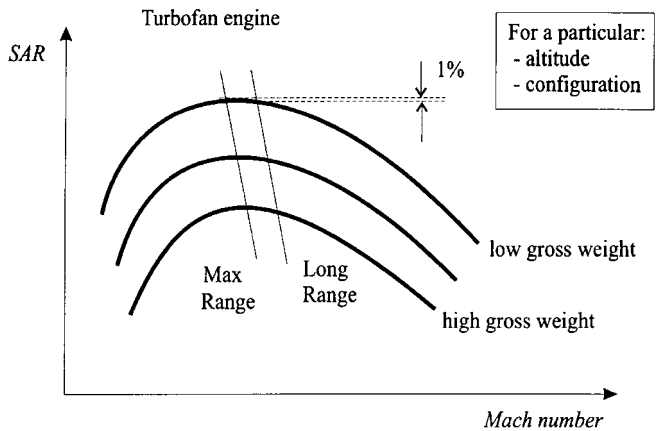


FIGURE 10.81 Specific air range versus Mach number.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

by assuming that SFC, TAS, and C_L are constant. Note that the condition of constant C_L implies that (L/D) is constant. The integration is thus straightforward.

$$R = \frac{V}{c_t g}\left(\frac{L}{D}\right)\ln\left(\frac{m_1}{m_2}\right) \tag{10.137}$$

This expression is widely referred to as the Breguet range equation (although the original equation was derived for a piston engine airplane). A number of other solutions exist to the range integral given by equation (10.136), resulting from the aircraft being flown under different constraints. Eshelby (2000) and Mair and Birdsall (1992) present solutions for other range scenarios that assume constant C_L and altitude, and constant Mach number and altitude. The Breguet range equation is the simplest solution, and because the deviation between the results of this method and other expressions is usually not significant, it is most often used for performance estimation. For a jet airplane the Breguet range equation can be written in a slightly different way:

$$R = \frac{a_o\sqrt{\theta}}{c_t g}\left(\frac{ML}{D}\right)\ln\left(\frac{m_1}{m_2}\right) \tag{10.138}$$

It is evident from this equation that if the SFC is assumed to be constant, then the flight condition at any height that will give the greatest range for a given fuel load occurs when (ML/D) is a maximum.

Cruise-Climb

For the Breguet range equation to be valid, C_L and V must be held constant during flight. This implies that the airplane is flown in a way that ensures that the ratio W/σ remains constant. This is possible if the airplane is allowed to climb very slowly so that the relative air density decreases proportionally to the decrease in weight. In the stratosphere, the thrust will automatically decrease as the aircraft climbs, without the throttle setting being altered. Thus, the pilot's instructions are simply to maintain a constant Mach number, allowing the aircraft to drift up as the flight progresses. This process is called a cruise-climb as altitude is not constant, and it is in fact an elegant solution for obtaining the maximum possible range. A small increase in thrust is required to maintain the climb angle, but this can be neglected for range estimation. In a cruise-climb the flight parameters for the starting condition (given the subscript 1) and the final condition (subscript 2) are related as follows:

$$\frac{W_1}{W_2} = \frac{\sigma_1}{\sigma_2} \tag{10.139}$$

Step Climb

The cruise-climb gives the greatest possible range, however its practical use is limited by air traffic control. As a result, aircraft often fly a stepped approximation of the cruise-climb, climbing to a higher cruise altitude as fuel is burnt.

Integrated Range Method

If SAR values can be determined for the cruise, then a simple numerical integration may be performed to determine the range. The technique, usually called the integrated range method, follows directly from equation (10.134). A graph of SAR versus aircraft mass is prepared. The range is the area under the graph between the points, representing the end of the cruise (lowest weight) and the start of the cruise (highest weight).

Piston Engine Airplane Range

For a piston engine aircraft the fuel flow may be determined using equations (10.104) and (10.105) for steady level flight, i.e.,

$$Q = c_p P = \frac{c_p mgV}{\eta_p}\left(\frac{D}{L}\right) \tag{10.140}$$

The range equation (10.134) is applicable to any airplane, hence the still air range for a piston engine airplane is given by:

$$R = -\int_{m_1}^{m_2} \frac{\eta_p}{c_p g m}\left(\frac{L}{D}\right) dm \tag{10.141}$$

Breguet Solution for Piston Engine Airplane

For an idealized piston engine the SFC and propeller efficiency are both constant and may be taken out of the integral equation (10.141). This assumption is acceptable for most applications as their variation during cruise is small. For the flight condition where C_L is constant, the integral yields:

$$R = \frac{\eta_p}{c_p g}\left(\frac{L}{D}\right) \ln\left(\frac{m_1}{m_2}\right) \tag{10.142}$$

By inspection it is seen that the maximum range will be achieved if the aircraft flies throughout the cruise at the flight condition of $(L/D)_{\max}$). Equation (10.142) is the Breguet range equation for piston engine aircraft. The equation is valid for flight schedules of either constant C_L and altitude, or constant C_L and airspeed. If the altitude is constant then airspeed must be reduced to maintain a constant C_L as fuel is burned. If, on the other hand airspeed is constant, then the aircraft must fly a cruise-climb.

Payload Range Diagram

Figure 10.82 is a typical payload versus range graph for an airliner. With the *maximum* allowable payload, the amount of fuel that can be taken onboard will usually be limited not by the size of the fuel tanks, but by the allowable takeoff weight (TOW). Under standard conditions the allowable TOW will be the maximum takeoff weight (MTOW) and the aircraft will have certain nominal range. If the nominal

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

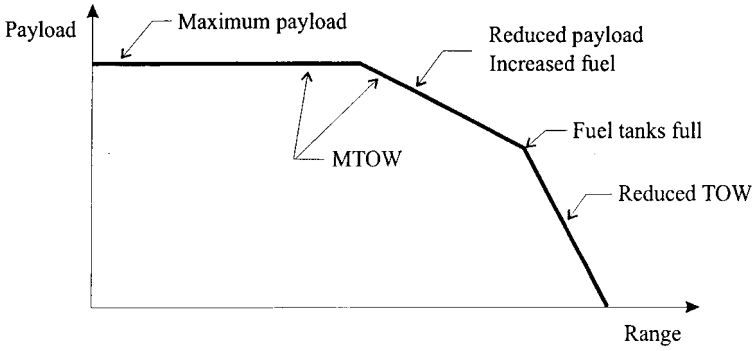


FIGURE 10.82 Typical payload-range graph for an airliner.

range is inadequate for the planned mission, then it will be necessary to reduce the payload in order to take on more fuel, but without exceeding the MTOW. Progressively longer mission lengths may be achieved by trading payload for fuel. When the point is reached when the fuel tanks are full, the only way the range can be increased is by further reducing the payload. The greatest possible range will correspond to zero payload.

Maximum Endurance for Turbojet / Turbofan Airplane

For some applications it is necessary that the airplane remain in the air for as long as possible on a given fuel load; for example, an aircraft on coastal patrol duties or an airliner holding at its destination, awaiting clearance to land. It is desirable that the airplane fly during these times at the speed for lowest fuel consumption per unit time. From equation (10.135) it is evident that if the SFC is assumed to be constant, then this occurs when (L/D) is a maximum. The airplane must therefore fly at V_{emd} to achieve the greatest endurance time. A plot of Q taking into account actual engine characteristics, as opposed to idealized characteristics that assume that SFC is constant, shows that the speed for minimum fuel flow is a little slower than V_{emd}. However, this speed is in the speed instability region (see Subsection 10.16) and so for an airplane without an auto-throttle function, the speed schedule usually chosen for holding is at or very close to V_{emd}.

Turbojet / Turbofan Airplane Endurance

The rate of change of airplane mass is equal to the fuel mass burned per unit time. Hence, if the initial mass is m_1 and final mass is m_2, the endurance time (t) is given by:

$$t = -\int_{m_1}^{m_2} \frac{1}{Q}\, dm = -\int_{m_1}^{m_2} \frac{1}{c_t g m} \left(\frac{L}{D}\right) dm \tag{10.143}$$

For constant c_t and C_L, the endurance time is:

$$t = \frac{1}{c_t g}\left(\frac{L}{D}\right)\ln\left(\frac{m_1}{m_2}\right) \tag{10.144}$$

This equation is valid only if the airplane is flown at a constant C_L. Thus, for flight at constant altitude, the pilot must reduce airspeed to compensate for the reduction in weight. Alternatively, if the aircraft is permitted to fly a cruise-climb, then it is possible to maintain a constant airspeed; however, this is not possible in a hold where the pilot must keep the airplane at a given altitude.

Piston Engine Airplane Endurance

It may be deduced from equation (10.140) that the lowest fuel consumption occurs when the aircraft is flown at the condition for minimum power. The endurance time for a piston engine is given by:

$$t = -\int_{m_2}^{m_2}\frac{1}{Q}\,dm = -\int_{m_1}^{m_2}\frac{\eta_p}{c_p g m V}\left(\frac{L}{D}\right)dm \tag{10.145}$$

With the assumptions that η_p, c_p, V, and C_L are all constant, the integral yields the following expression:

$$t = \frac{\eta_p}{c_p g V}\left(\frac{L}{D}\right)\ln\left(\frac{m_1}{m_2}\right) \tag{10.146}$$

Because the assumptions given above do not correspond to the flight condition for minimum power, the solution given by equation (10.146) does not give the greatest possible endurance time.

Further Reading

Consult Mair and Birdsall (1992) and Eshelby (2000) for alternative solutions to the basic range and endurance integral equations (resulting from the airplane being flown under different constraints). Anderson (1999) derives solutions to the range and endurance integrals with fuel flow based on weight rather than mass, as done here. Lowry (1999) describes practical performance analysis of piston engine aircraft, while Smetana (2001) covers methods for assessing the en route performance of new designs.

10.21 TAKEOFF AND LANDING PERFORMANCE

Takeoff

A schematic of the takeoff is shown in the Figure 10.83. The aircraft accelerates from rest to a speed that will provide sustained controllable flight, at which point the pilot will pull the stick/yoke back, causing the airplane to rotate as the tail moves downwards. A few seconds later it lifts off and climbs to clear an imaginary screen height. This height (h_{sc}) is generally 50 ft (15.2 m) for military or light

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

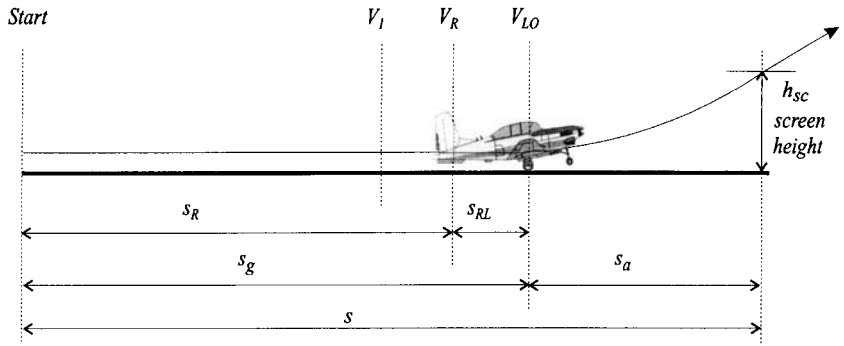


FIGURE 10.83 Takeoff profile.

aircraft and 35 ft (10.7 m) for commercial aircraft. The total takeoff distance (s) consists of a ground segment (s_g) and an air segment (s_a). The ground segment, called the *ground roll* or *ground run*, may be divided into two elements, the distance taken from rest to the point where the aircraft rotates (s_R) and the distance from the point of rotation to liftoff (s_{RL}). At speed V_R the aircraft rotates, increasing its angle of attack and lift; shortly afterwards, at a speed V_{LO}, where the aircraft has reached sufficient forward speed to generate the required lift, liftoff occurs. Methods to evaluate the takeoff distance follow.

Forces Acting on the Airplane During Takeoff

The forces acting on an aircraft during takeoff are shown in Figure 10.84. The runway has a gradient of $\tan\gamma_G$, where a positive angle of γ_G will be used to indicate an uphill takeoff. The rolling coefficient of friction is μ_R. The thrust (T) of the engine (or propeller) accelerates the aircraft. Resistance to forward motion comes from the aerodynamic drag (D), rolling friction of the tires, and, for an inclined runway, the component of weight acting parallel to the runway. The net acceleration force is given by:

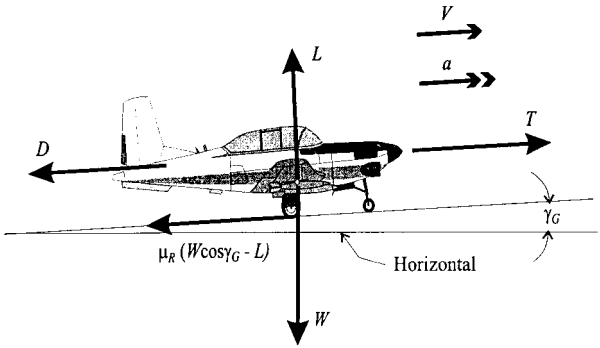


FIGURE 10.84 Forces acting on an aircraft during takeoff.

$$\sum \text{forces along runway} = T - \mu_R (W \cos \gamma_G - L) - D - W \sin \gamma_G \quad (10.147)$$

Before equation (10.147) can be used to evaluate the takeoff distance, it is necessary to describe the forces acting on the airplane during the takeoff. The weight is reduced by only a very small amount during the takeoff and may be regarded as constant, but the other forces will change as the speed increases.

Lift and Drag. Because changes in the angle of attack can only result from the differential extension and contraction of the nose and main landing gear, the lift coefficient will be essentially constant up to the point of rotation. The drag coefficient will also be essentially constant. The lift and drag forces will thus vary as functions of V^2.

Ground Effect. Performance calculations in close proximity to the ground require a correction to the drag polar determined away from the ground. During the ground run the aircraft's lift-dependent drag (C_{D_i}) is reduced by a ground effect factor (λ) as a result of a reduction in the trailing vortex drag. The magnitude of λ essentially depends on the wing span and height of the wing above the ground. Torenbeek (1982) provides data to estimate the impact of this effect.

Flaps and Undercarriage. With the flaps set for takeoff and the undercarriage extended, the applicable drag polar must include terms that correct for these factors. Whereas the undercarriage would increase the value of the clean aircraft C_{D_0}, the flaps would change both C_{D_0} and C_{D_i}. Typical values may be obtained from Torenbeek (1982) or Raymer (1989).

Thrust. In general, the thrust from the engine (or propeller) depends on the atmospheric conditions and the airspeed and will vary during the takeoff run.

Rolling Friction. The value of μ_R is dependent on the tire pressure and the runway surface type and does change a little during the takeoff. However, the influence of these considerations on the ground run is very small and a mean value may be used. Because the rolling resistance on a hard dry surface is small in comparison to the other forces in equation (10.133) an approximate value of μ_R may be used without this parameter significantly affecting the calculated takeoff distance. For a hard, dry surface, the most usual value for μ_R that is used is 0.02 (ESDU 85029 1985). Other values suggested for dry concrete are 0.025 (Mair and Birdsall 1992) and 0.015 (Boeing 1989).

Analytical Evaluation of the Ground Distance (Zero Wind)

In the absence of wind, the distance to the point of rotation (s_R) is given by:

$$s_R = \int_0^{V_R} \frac{V}{a}\, dV \quad (10.148)$$

where s is the ground distance, V is the airspeed (equal to the ground speed in the absence of wind), and a is the acceleration. At any instant during the ground run, the acceleration may be obtained from equation (10.147) by applying Newton's second law. Because the runway gradient is always a small quantity, the approximations $\cos \gamma_G \approx 1$ and $\gamma_G \approx \gamma_G$ may be introduced. Using the parabolic drag polar, corrected for ground effect, the acceleration may be written as:

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

$$a = \frac{1}{m}\left[T - \mu_R(W - \tfrac{1}{2}\rho V^2 S C_L) - \tfrac{1}{2}\rho V^2 S(C_{D_0} + \lambda K C_L^2) - \gamma_G W\right] \quad (10.149)$$

Mean Thrust. The analysis is simplified by assuming that the thrust is equal to $\overline{T}$, a mean constant value selected to give a good approximation of the takeoff distance. It has been shown (Boeing 1989) that for a jet airplane, the acceleration varies approximately linearly with V^2 from zero speed to V_R. The thrust and acceleration may thus be calculated at the speed

$$V = 1/\sqrt{2}\ V_R \approx 0.71 V_R$$

For a propeller-driven aircraft a better estimate of the ground run is obtained if the propeller thrust is calculated at a speed of $V = 0.74V_R$ (Mair and Birdsall 1992). With T taken as constant and all other variables written as functions of V^2, the integral expression (10.148) may be evaluated to give:

$$s_R = \frac{1}{2gB_2} \ln\left(\frac{A_2 + B_2 V_R^2}{A_2}\right) \quad (10.150)$$

where

$$A_2 = \frac{\overline{T}}{W} - \mu_R - \gamma_G \quad \text{and} \quad B_2 = \frac{\rho}{2(W/S)}(\mu_R C_L - C_{D_0} - \gamma C_L^2)$$

Mean Acceleration. A popular and relatively simple method for estimating the takeoff run is based on the use of a mean acceleration ($\overline{a}$). The approach may be summarized as follows:

1. Determine the mean thrust ($\overline{T}$) at $V = 0.71V_R$ (for a jet engine) or $V = 0.74V_R$ (piston engine) from engine data.
2. Calculate $\overline{a}$ for $T = \overline{T}$ from equation (10.149) where $V = 0.71V_R$ (jet engine) or $V = 0.74V_R$ (piston engine).
3. Estimate the ground distance (s_R) from the equation for uniform acceleration, i.e.,

$$S_R = \frac{1}{2}\frac{V_R^2}{\overline{a}} \quad (10.151)$$

Effect of Wind on the Ground Distance

The component of wind acting along the runway is designated as V_w. By convention, V_w is positive for a headwind and negative for a tailwind. (Note that this is opposite to the convention usually adopted for the cruise.) At the start of the ground run, the aircraft is stationary, but the presence of the wind implies that the airspeed is equal in magnitude to V_w. In this situation equation (10.148) may be written as:

$$S_R = \int_{V_w}^{V_R} \frac{V - V_w}{a}\, dV \quad (10.152)$$

This expression may be evaluated as described above. The mean acceleration may be determined, as before, from equation (10.149). Based on a mean acceleration, the ground distance is:

$$s_R = \frac{1}{2\bar{a}} (V_R - V_w)^2 \tag{10.153}$$

The significance of a headwind on reducing the takeoff distance is evident from this equation.

Numerical Evaluation of the Ground Run

Equation (10.152) can be integrated numerically by dividing the takeoff run into n segments. Using the trapezoidal rule, the ground distance is given by:

$$S_R = \tfrac{1}{2} \sum_{i=0}^{n} \left(\frac{(V)_{i+1} - V_w}{(a)_{i+1}} + \frac{(V)_i - V_w}{(a)_i} \right) \Delta V \tag{10.154}$$

Estimation of the Rotation Distance

The rotation distance (s_{RL}) is usually small in comparison to s_R but is difficult to estimate accurately due to the changes in C_L. At the point of rotation, the pilot will pull the stick/yoke back, raising the nose and increasing the angle of attack. The time that it takes for the airplane to rotate depends on the rate that the pilot pulls the stick/yoke back and on the type of aircraft. The duration is of the order of 1 to 3 seconds; small light aircraft may rotate in 1 second or less, with large transport aircraft taking longer. An estimate of the distance s_{RL} may be obtained by multiplying the rotation time by the mean ground speed. The assumption of zero acceleration from rotation to liftoff is reasonable. In the absence of substantive data, the expression $V_{LO} = 1.2V_{S_1}$ may be used to estimate the liftoff speed.

Climb-out to Screen Height

After the liftoff there is a transition phase in which the flight path is curved, the lift is greater than the weight, and there is a small increase in speed. After the transition the airplane will climb at an approximate constant climb angle. The point at which the screen height is reached may be either before or after the end of the curved transition. The air segment (s_a) is difficult to calculate accurately due to the variation of the governing parameters and the influence of varying pilot technique. The simplest method is to multiply an average time by the average ground speed. The time is best determined from experimental data. It would typically be between 2 and 8 seconds and is largely a function of the thrust-to-weight ratio of the airplane. An alternative approach to estimating s_a assumes that the flight path is a circular arc and the distance is then be calculated directly. The method is described by Mair and Birdsall (1992) and Anderson (1999).

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

Landing Procedure

The landing segment is shown in Figure 10.85. The aircraft descends initially along a straight glide path, which is typically at an angle of 3° to the horizontal. At the threshold (screen height), usually taken to be 50 ft (15.2m), the approach speed is required to be not less than 1.3 times the stalling speed (in the landing configuration), according to the Airworthiness Regulations. The pilot reduces the vertical component of the airplane's velocity by a flare. Depending on pilot technique, there may be a short hold-off period where he or she permits the airplane to float a little, allowing the speed to reduce before it touches down. Because the landing distance depends substantially on pilot technique, analytical estimates often compare poorly with actual test data. It is common practice in theoretical analysis to assume that there is no float and that the touchdown occurs with zero vertical velocity. The speed at the touchdown (V_T) will typically be about 5% to 15% higher than the stall speed (in the landing configuration). At touchdown the nose wheel should still be well above the runway and the pilot will then allow it to descend gently onto the runway. A delay of a few seconds is typical before the pilot applies the brakes. The aircraft is brought to rest by use of the wheel brakes, sometimes assisted by lift-dumpers or spoilers and reverse thrust from the engines.

Braking Force

The analytical evaluation of the landing distance can be undertaken in an almost identical manner to that presented for the takeoff. The one significant difference is that a braking force replaces the rolling resistance. It is possible to estimate the maximum braking force, albeit with some difficulty. Under normal operations the braking force would be substantially lower than the maximum design force, which would only be required under emergency conditions. The braking force of the wheels is given by:

$$\text{Braking force} = \mu_B R_{MW} = \mu_B\,(kW - L) \tag{10.155}$$

where μ_B is the coefficient of braking friction, R_{MW} is the reaction force of the main wheels on the runway, and k is the portion of the airline's weight carried by the

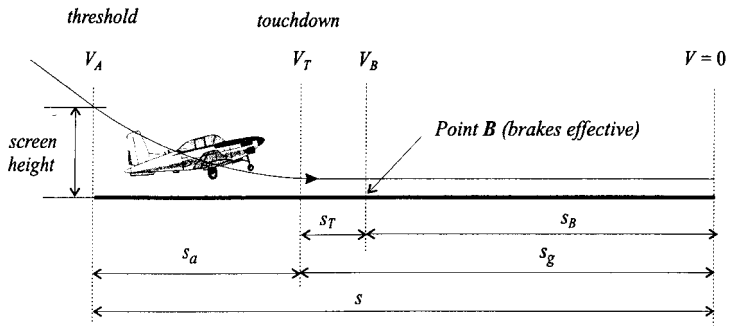


FIGURE 10.85 Landing profile.

main wheels (typically between 0.8 and 0.9). The effect of spoilers (which destroy the lift) in increasing the braking force is seen from this equation. The braking coefficient of friction is not constant and will increase as the airplane slows down. When the runway is dry, the increase is fairly small, but with a wet surface there is a large nonlinear increase. For hard, dry runways the maximum braking coefficient ($\mu_{B_{max}}$) typically increases from about 0.7 at 100 kt to about 0.8 as the speed decreases to zero, while for wet conditions $\mu_{B_{max}}$ is about 0.2 (at 100 kt) increasing to about 0.7. Maximum braking coefficient values for various runway surfaces and tyre pressures are provided by ESDU 71026 (1995). Icy slush or wet snow is a particular danger because μ_B could be reduced to less than 0.05 at the speeds associated with touchdown. The value of μ_B achieved in practice is a function of the amount of slip taking place between the tires and the runway. If the wheels are permitted to roll freely, the coefficient will equal μ_R, but as the brakes are applied the coefficient increases rapidly and then starts to reduce if the tires slip. If the brakes are manually controlled, the mean effective braking force will be about 30% to 50% of the theoretical maximum braking force (ESDU 71026 1995). Antiskid cycling by automatic braking systems protect the wheels from locking; in these cases the effective braking can be as high as 80% to 90% of the maximum value. Once a theoretical braking force has been determined, it is necessary to check if this results in the maximum permissible brake torque, or maximum brake system pressure, being exceeded. The actual braking force could thus be substantially less than that determined from a simple calculation based on a theoretical value of μ_B. For initial estimations Raymer (1989) suggests typical mean μ_B values of 0.3–0.5 for dry concrete or asphalt surfaces.

Landing Distance

The total landing distance is given by:

$$s = s_a + s_T + s_B \tag{10.156}$$

An estimate of the airborne distance (s_a) can be obtained by multiplying the average time by the average ground speed. After touchdown there is a slight delay before the wheel brakes become effective, usually about 2 to 3 seconds, during which time the speed falls by a few percent. The distance s_T can be estimated from the delay time and the touchdown speed. The equations to be used for calculating the length of the ground run after the point B are essentially the same as those used for the takeoff. Equation (10.152) may be rewritten to give the braking distance, i.e.,

$$s_B = \int_{V_B}^{V_W} \frac{V - V_w}{-a}\, dV \tag{10.157}$$

where a is negative and is given by equation (10.149). The following three points are important:

1. The braking coefficient (μ_B) will replace the rolling friction coefficient (μ_R).
2. The drag coefficient will be greater than that used for the takeoff because of the greater flap angle used for landing.
3. In cases where reversed thrust can be used, it is typically applied after the spoilers and brakes become effective. When thrust reversers are not available,

the engines are run at idling speed and the thrust is usually small enough to be neglected.

Because the braking force cannot be represented as a function of V^2, it is often necessary to evaluate the integral by step-by-step computation. An estimate of the distance s_B may be obtained by determining the mean acceleration ($\bar{a}$) calculated for $V = 0.71V_B$. Equation (10.153), used for the takeoff analysis, may be rewritten for the landing distance:

$$s_B = \frac{-1}{2\bar{a}} (V_B - V_w)^2 \tag{10.158}$$

The determination of the required runway distance for actual aircraft operations is discussed in Subsection 10.22.

Further Reading

Further information on the takeoff and landing distance calculation is presented by Mair and Birdsall (1992), Ojha (1995), Eshelby (2000), and Lowry (1999); the last reference focuses on light aircraft performance. Details on the coefficient of rolling friction (μ_R) and braking coefficient (μ_B) are contained in ESDU 85029 (1985) and ESDU 71026 (1995), respectively. Stinton (1996) and Swatton (2000) describe the takeoff and landing performance from the pilot's perspective, discussing the regulatory requirements.

10.22 AIRPLANE OPERATIONS

Regulations and Requirements

Regulations and requirements have been established to ensure that all airplanes engaged in public transport flights meet a minimum standard of safety deemed appropriate to the operation. Two complementary sets of measures contain specific details regarding the required performance of these aircraft.

The first is concerned with the *operation* of the airplane. The most important are:

- FAR 121 (Federal Aviation Regulation Part 121), *Operating requirements: Domestic, flag, and supplemental operations.*
- JAR OPS 1 (Joint Airworthiness Requirement OPS Part 1), *Commercial Air Transportation* (*Aeroplanes*).

The second set of measures pertains to the certification of new airplanes, as described in the Airworthiness Regulations/Requirements. These include:

- FAR 23 (Federal Aviation Regulation Part 23), *Airworthiness standards: Normal, utility, acrobatic, and commuter category airplanes.*
- FAR 25 (Federal Aviation Regulation Part 25), *Airworthiness standards: Transport category airplanes.*

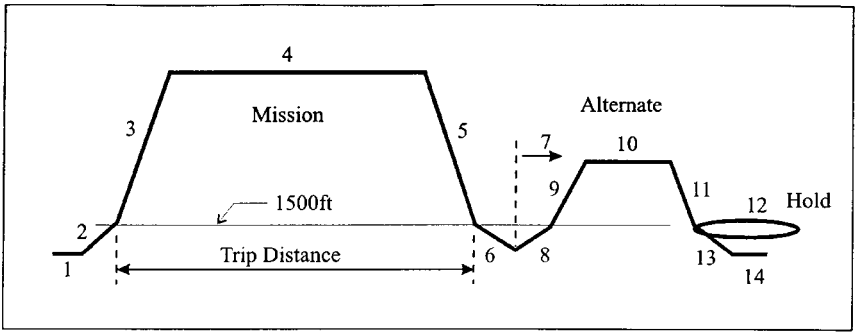
- JAR 23 (Joint Aviation Requirement Part 23), *Normal, Utility, Aerobatic, and Commuter Category Aeroplanes.*
- JAR 25 (Joint Aviation Requirement Part 25), *Large Aeroplanes.*

En Route Flight Planning—Fuel Required

The en route flight profile is divided into several parts for the purpose of flight planning, as illustrated in Figure 10.86. Specific requirements exist for the determination of the required fuel for the mission. These depend on the operator (flag or foreign), the type of airplane, the route (domestic or international), and the availability of alternate airports (if the airplane cannot land at the destination airport for any reason). For example, U.S. flag and supplemental operations on international routes where an alternate airport is specified must comply with FAR 121.645. It is stated that no person may release for takeoff a turbine-engine powered airplane (not including a turboprop airplane) unless, considering wind and other weather conditions expected, it has enough fuel:

1. To fly to and land at the airport to which it is released
2. After that, to fly for a period of 10% of the total time required to fly from the airport of departure to, and land at, the airport to which it was released
3. After that, to fly to and land at the most distant alternate airport specified in the flight release
4. After that, to fly for 30 minutes at holding speed at 1,500 feet above the alternative airport under standard temperature conditions.

The fuel required, time, and distance for each segment are determined. It is usual to calculate the trip fuel and time from brake release at the departure aerodrome to



Index

1	Engine start-up, taxi	8	Climb to 1,500 ft
2	Takeoff and climb to 1,500 ft	9	Climb to alternate cruise altitude
3	Climb to cruise altitude	10	Alternate cruise
4	Cruise	11	Descent to 1,500 ft
5	Descent to 1,500 ft	12	Hold for 30 minutes
6	Approach and land	13	Approach and land
7	En route reserve	14	Taxi and shutdown

FIGURE 10.86 Typical flight profile for fuel planning.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

touchdown at the destination aerodrome. The trip distance calculation may in some cases ignore the climb and descent below 1,500 ft (as shown in Figure 10.86). The block fuel and time includes engine start-up and taxi, and the taxi after landing.

Takeoff Reference Speeds

The takeoff is one of the critical parts of any flight. Commercial airlines are required to operate their aircraft under strict safety regulations. Definitions of the important reference speeds dealing with the takeoff operation of multiple-engine transport airplanes are given in FAR Part 25.107 and Part 25.149. The most important of these speeds are described below.

Minimum Control Speed—Ground **(V_{MC_G}).** This is the minimum speed on the ground at which, when the critical engine suddenly becomes inoperative and with the remaining engine(s) operating at full takeoff thrust, it is possible to recover control of the airplane with the use of primary aerodynamic controls alone (without the use of nose wheel steering) to enable the takeoff to be safely continued using normal piloting skill. The critical engine is the outboard engine that results in the most severe consequence for the takeoff.

Minimum Control Speed—Air **(V_{MC_A}).** This is the airspeed at which, when the critical engine suddenly becomes inoperative and with the remaining engine(s) operating at maximum available takeoff thrust, it is possible to recover control of the airplane using normal piloting skill and maintain straight flight with an angle of bank of not more than 5°. V_{MV_A} may not exceed 1.2 V_S, where V_S is determined for the maximum takeoff weight.

Takeoff Decision Speed **(V_1).** This is the speed at which a multiple-engine airplane must continue the takeoff, even if one engine fails (completely). Thus, during the takeoff, up to the V_1 speed, the pilot will be able to bring the airplane safely to a stop if there is an engine failure. If there is an engine failure after V_1, the pilot shall have sufficient thrust (from the remaining engines) and sufficient remaining runway to take off safely and clear the specified screen height. The exact definition of the V_1 speed accounts for the reaction time of the pilot and is the speed of the airplane at the instant the pilot has recognized and reacted to the engine failure. The V_1 speed may not be less than V_{MC_G}.

Takeoff Rotation Speed **(V_R).** This is the speed at which rotation is initiated. V_R must not be less than 1.05 times the minimum control speed (air) nor less than V_1.

Minimum Unstick Speed **(V_{MU}).** This is the minimum speed at which the airplane can be made to lift off the ground and continue the takeoff without displaying any hazardous characteristics. V_{MU} speeds are determined for the all-engines-operating and the one-engine-inoperative conditions.

Liftoff Speed **(V_{LOF}).** The liftoff speed is closely associated with the V_R speed. The all engines operating liftoff speed must not be less than 110% of V_{MU}, assuming maximum practicable rotation rate. The one engine inoperative liftoff speed must not be less than 105% of V_{MU}.

***Takeoff Safety Speed* (V_2).** This is a reference speed used to determine the climb performance of the airplane during the initial climb-out, with one engine inoperative. V_2 is equal to the actual speed at the 35 ft (10.7 m) height as demonstrated in flight and must be equal to or greater than 120% of the stall speed in the takeoff configuration, or 110% of V_{MC_A}. For turbopropeller and reciprocating engine powered airplanes with more than three engines, and for certain turbojet powered airplanes, the 120% stall speed requirement is reduced to 115%.

Operational Field Length for Takeoff

The operational field length for a given airplane gross weight, airport elevation, ambient temperature, and wind is equal to the longest of the following three calculated distances:

1. The one-engine-out takeoff distance, which is the distance required to reach a height of 35 ft (10.7 m) over the runway, assuming a complete engine failure at V_1, rotation at V_R, and achieving a speed of V_2 no later than the 35 ft point.
2. The accelerate-stop distance, which is the distance required to accelerate to V_1 and then to bring the airplane to a complete stop, using the wheel brakes only (i.e., no thrust reverse).
3. The normal all-engine takeoff distance plus a margin of 15%. This is the FAR takeoff field length for all engines operating.

Balanced Field Length

The first two distances (above) are functions of V_1. By increasing the selected V_1 speed, the calculated one-engine-out takeoff distance will decrease, but the *accelerate-stop* distance will increase. There exists a unique V_1 speed, where the two distances are equal, called the balanced takeoff field length. The determination of the V_1 speed for a balanced runway is performed for a given aircraft gross weight, airport elevation, ambient temperature, and wind. There are cases where some of the takeoff requirements are not met with a balanced field length and the operation is planned using an unbalanced field length, i.e., the one-engine-out and accelerate-stop distances will not be equal. This may result from the use of clearways and stopways (see below).

Unbalanced Field Length—Clearways and Stopways

If the airplane is certified for an unbalanced field length takeoff and the runway has a *clearway* and/or a *stopway*, then it is possible to increase the airplane's takeoff weight for the given runway length. A clearway is an area of prescribed width beyond the end of the runway under the control of the airport authority. Instead of reaching the height of 35 ft (10.7 m) at the end of the runway, the pilot may lift off farther down the runway and use the clearway to climb to the 35 ft (10.7 m) screen height. However, the accelerate-stop distance must still equal the available runway distance. Achieving the increased weight but still retaining the same accelerate-stop distance implies that the V_1 speed (unbalanced) will be less than the V_1 speed (balanced). A stopway is a hard surfaced area, aft of the runway,

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

that may be used for braking. When a stopway is available, the additional braking distance may be taken into account to determine the accelerate-stop distance for the increased takeoff weight.

Climb-out Gradient Requirements

The flight path after liftoff is divided into three segments (Figure 10.87). It is required that the airplane be capable of maintaining specified minimum climb gradients, with one engine inoperative during each segment. The first segment is from liftoff to the point of complete gear retraction. The airplane will have the gear extended, the flaps set for takeoff, and the throttle position set for takeoff. The second segment starts at the point of complete retraction of the gear and ends at a height 400 ft (122 m) above the runway. In this segment the airplane climbs at a speed of no less than V_2 with the gear retracted, flaps set for takeoff. The third segment, called the final segment, extends from the end of the second segment to a height of at least 1,500 ft (457 m). During this segment the flaps are retracted and the airplane accelerates to the en route climb speed, which is required to be 25% higher than the stall speed. The thrust will be at takeoff (or maximum continuous) setting. The minimum climb gradients apply after the airplane has cleaned up and accelerated along a horizontal flight path.

Obstacle Clearance Requirements

In addition to the climb gradient requirements, the airplane must be operated within the safety regulations dealing with obstacle clearance at specific airports under actual flight conditions. Buildings, trees, communication towers, etc. are all potential dangers for an aircraft that suffers an engine failure during takeoff. Furthermore, due to tailwinds, its actual flight path may be lower than that predicated by the calculated still air climb gradient. When an operator flies out of a specific airport, the net flight path must clear all obstacles within a defined flight corridor by at least 35 ft (10.7 m). The net flight path is a conservative definition that requires the actual one-engine-inoperative flight path (determined for actual headwind or

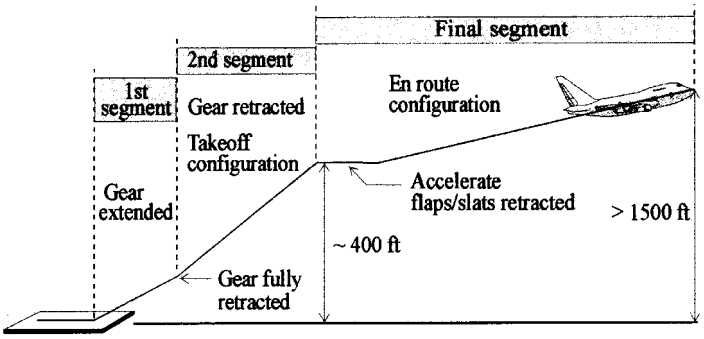


FIGURE 10.87 Climb out profile.

tailwind conditions) to be reduced by a fixed percentage. The amount is 0.8% for two-engine aircraft, 0.9% for three-engine aircraft, and 1.0% for four-engine aircraft (FAR 25.115). This ensures that minor errors in loading, optimistic approximations in performance predictions, or changes in wind speed or direction will not result in catastrophe when an engine failure occurs on takeoff.

Climb Requirements Following an Overshoot on Landing

One of the potentially dangerous situations that an operator has to take into account during flight planning is the possibility of an aborted landing with one engine inoperative. The airplane will approach the runway, then, after an overshoot, will need to climb, clearing all obstacles. The resulting still air climb gradient must exceed a minimum specified value for the airplane type. The minimum climb gradient is 2.1% for a two-engine airplane, 2.4% for a three-engine airplane, and 2.7% for a four-engine airplane. A second climb requirement is specified for aborted landings when all engines are operating. With the airplane in the landing configuration, all airplane types must be able to maintain a climb gradient of 3.2%.

Required Runway Length for Landing

The demonstrated landing distance is based on the airplane crossing the 50 ft (15.2 m) threshold at a speed 30% above the stall speed. After the touchdown, the airplane will be brought to a stop by means of the wheel brakes only. This conservative assumption ensures that if thrust reversers are used, the airplane will stop within the calculated landing distance. The distance from the threshold until the airplane stops is the measured landing distance. The required runway (dry) for jet airplane operations is determined by multiplying the measured landing distance by an operational reserve factor of 1.667. This factor implies that the airplane should ideally require only 60% of the runway, with the remaining 40% regarded as an *operational reserve*. For turboprop aircraft the percentage is 70%, with 30% regarded as the operational reserve. For wet runways, various policies are in use. Airlines may typically add a fixed distance (say 300 m) to the required runway (dry) to get the required runway (wet). In the case of adverse runway conditions of ice or snow contamination, additional allowances will be made to cover the reduced braking capability of the airplane.

Further Reading

Consult the FARs and JARs for complete details on the regulations. Stinton (1996) and Swatton (2000) provide useful information on the flight implications of these regulations.

REFERENCES

Anderson, J. D., Jr. 1999. *Aircraft Performance and Design,* McGraw-Hill, New York.

Boeing. 1989. *Jet Transport Performance Methods,* Boeing Commercial Airplane Company.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

ESDU 68046. 1992. *Atmospheric Data for Performance Calculations, amendment* (*d*), ESDU, London.

ESDU 70020. 1970. *Non-dimensional Approach to Engine Thrust and Airframe Drag for the Analysis of Measured Performance Data: Aircraft with Turbo-jet and Turbo-fan Engines,* ESDU, London.

ESDU 71026. 1995. *Frictional and Retarding Forces on Aircraft Tyres, Part II, amendment* (*d*), ESDU, London.

ESDU 72018. 1972. *International Standard Atmosphere* (*−2000 ft to 105,000 ft, data in SI units*), ESDU, 27 London.

ESDU 73019. 1982. *Approximate Methods for Estimating Cruise Range and Endurance: Aeroplanes with Turbo-jet and Turbo-fan engines, amendment* (*c*), ESDU, London.

ESDU 85029. 1985. *Calculation of Ground Performance in Take-off and Landing,* ESDU, 27 London.

Eshelby, M. E. 2000. *Aircraft Performance: Theory and Practice,* Edward Arnold, London.

Federal Aviation Regulation Part 23 (FAR 23). *Airworthiness Standards: Normal, Utility, Acrobatic, and Commuter Category Airplanes*. Federal Aviation Administration (FAA), Washington, DC.

Federal Aviation Regulation Part 25 (FAR 25). *Airworthiness Standards: Transport Category Airplanes,* Federal Aviation Administration (FAA), Washington, DC.

Federal Aviation Regulation Part 121 (FAR 121). *Operating Requirements: Domestic, Flag, and Supplemental Operations,* Federal Aviation Administration (FAA), Washington, DC.

ICAO. 1993. *Manual of the ICAO Standard Atmosphere,* Doc. 7488/1, International Civil Aviation Organization (ICAO), Montreal.

ISO 2533. 1975. *Standard Atmosphere,* International Organisation for Standardization.

Joint Aviation Requirement Part 23 (JAR 23). *Normal, Utility, Aerobatic, and Commuter Category Aeroplanes,* Joint Aviation Authorities (JAA) of Europe, Hoofdorp, The Netherlands.

Joint Aviation Requirement Part 25 (JAR). *Large Aeroplanes,* Joint Aviation Authorities (JAA) of Europe, Hoofdorp, The Netherlands.

Joint Airworthiness Requirement OPS Part 1 (JAR OPS 1). *Commercial Air Transportation* (*Aeroplanes*), Joint Aviation Authorities (JAA) of Europe, Hoofdorp, The Netherlands.

Lan, C.-T. E. and Roskam, J. 1981. *Airplane Aerodynamics and Performance,* Roskam Aviation and Engineering Corp.

Lowry, J. T., 1999. *Performance of Light Aircraft,* AIAA, Reston, VA.

Mair, W. A. and Birdsall, D. L. 1992. *Aircraft Performance,* Cambridge University Press, Cambridge.

Mattingly, J. D., Heiser, W. H. and Daley, D. H. 1987. *Aircraft Engine Design,* AIAA, Reston, VA.

Ojha, S. K. 1995. *Flight Performance of Aircraft,* AIAA, Reston, VA.

Raymer, D. 1989. *Aircraft Design: A Conceptual Approach,* AIAA, Reston, VA.

Smetana, F. O. 2001. *Flight Vehicle Performance and Aerodynamic Control,* AIAA, Reston, VA.

Stinton, D. 1996. *Flying Qualities and Flight Testing of the Aeroplane*, Blackwell Science, Oxford.

Swatton, P. J. 2000. *Aircraft Performance Theory for Pilots,* Blackwell Science, Oxford.

Torenbeek, E. 1982. *Synthesis of Subsonic Airplane Design,* Delft University Press, Delft.

PART 3

AIRCRAFT STABILITY AND CONTROL

Douglas G. Thomson

NOTATION

$\mathbf{A}$	system matrix
a_t	tailplane lift curve slope, per radian
a_e	elevator effectiveness, per radian
a_F	lift-curve slope of fin, per radian
a_r	rudder effectiveness, per radian
$\mathbf{B}$	control matrix
b	wing semispan, m
$\mathbf{C}$	output matrix
$\mathbf{D}$	direct matrix
$\mathbf{E}$	gust influence matrix
C_D, C_L, C_T	coefficients of drag, lift, and thrust
C_m	pitching moment coefficient
C_{D0}, $C_{D\alpha}$	coefficients of empirical equation for drag coefficient
C_{L_0}, C_{L_α} $C_{L_{\delta e}}$	coefficients of empirical equation for lift coefficient
C_{M_0}, $C_{M\alpha}$, $C_{M_{\delta e}}$, C_{M_q}	coefficients of empirical equation for pitching moment coefficient
$\bar{c}$	mean aerodynamic chord, m
D	aircraft drag, N
g	acceleration due to gravity, m/s^2
h_T	offset of thrust-line from aircraft x_b body axes, m
I_{xx}, I_{yy}, I_{zz}	aircraft moments of inertia about x_b, y_b, z_b body axes, kg m^2
I_{xz}	product of inertia about aircraft y_b axis, kg m^2
L	aircraft lift, N
L, M, N	external moments, Nm
l_F	distance between fin aerodynamic center and aircraft c.g., m
l_t	tailplane lever arm, m
M	Mach number
m	aircraft mass, kg
P, Q, R	angular velocities in direction of x_b, y_b, z_b body axes, rad/s
p, q, r	perturbations in angular velocities in direction of x_b, y_b, z_b body axes, rad/s
p_d	dynamic pressure, N/m^2
S	wing area, m^2
S_f	fin area, m^2

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

S_t	tailplane area, m^2
T	period of oscillation, s
T	aircraft thrust, N
$\mathbf{T}$	Euler transformation matrix
t_{half}	time to half amplitude, s
t_{double}	time to double amplitude, s
$\mathbf{u}$	control vector
U, V, W	translational velocities in direction of x_b, y_b, z_b body axes, m/s
u, v, w	perturbations in translational velocities in direction of x_b, y_b, z_b body axes, m/s
V_f	aircraft flight velocity, m/s
V_g	gust velocity vector
V_H	tailplane volume ratio
V_v	fin volume ratio
X, Y, Z	external forces, N
$\mathbf{x}$	state vector
z_F	height of fin mean aerodynamic center above x-axis, m
x_b, y_b, z_b	body axes
$x\text{E}, y_E, z_E$	earth axes
$\mathbf{y}$	output vector

Greek

α, β	incidence angles (attack and sideslip), rad
δ	control deflection, rad
ε	tailplane downwash angle, rad
Φ, Θ, Ψ	euler angles (roll, pitch, and yaw), rad
ϕ, θ, ψ	perturbations in Euler angles (roll, pitch, and yaw), rad
λ	eigenvalue
η, ζ, ξ	elevator, rudder, and aileron deflections, rad
ρ	air density, kg/m^3
σ	sidewash angle, rad

Subscripts

b	body axes set
dr	dutch roll mode
E	earth axes set
e	equilibrium condition (trim state)

e, a, r	elevator, aileron, rudder
g	gust
ph	phugoid mode
r	roll mode
s	spiral mode
sp	short-period mode

10.23 MATHEMATICAL MODELING AND SIMULATION OF FIXED WING AIRCRAFT

Aircraft Nonlinear Equations of Motion

The aircraft, in its most basic form, has 6 degrees of freedom, as summarized in Table 10.14.

These consist of translational motions in the directions of the axes set fixed in the aircraft, and three rotations about these axes (Figure 10.88). The six aircraft states (or *state variables*) are (U, V, W, P, Q, R). The *body fixed frame of reference* (x_b, y_b, z_b) has its origin at the center of gravity of the aircraft, with the x_b axis pointing forwards, usually down the centerline of the fuselage, the z_b axis downwards, and the y_b axis in the starboard direction. The forces and translational velocities are positive in these directions. The positive direction for angular quantities is determined by the right-hand rule (the right thumb is pointed in the positive direction of the axis and the direction of curl of the fingers gives the positive direction for angular quantities). Hence, a positive roll rate gives starboard wing down, a positive pitch rate gives nose up, and a positive yaw rate gives nose right.

As the aircraft is a rigid body translating and rotating in 3D space it is appropriate to apply the *Euler equations* to its motion:

$$m(\dot{U} + QW - RV) = X - mg \sin \Theta \tag{10.159}$$

$$m(\dot{V} + RU - PW) = Y + mg \cos \Theta \sin \Phi \tag{10.160}$$

$$m(\dot{W} + PV - QU) = Z + mg \cos \Theta \cos \Phi \tag{10.161}$$

$$I_{xx}\dot{P} - I_{xz}\dot{R} + QR(I_{zz} - I_{yy}) - PQI_{xz} = L \tag{10.162}$$

$$I_{yy}\dot{Q} + RP(I_{xx} - I_{zz}) + (P^2 - R^2)I_{xz} = M \tag{10.163}$$

$$I_{zz}\dot{R} - I_{xz}\dot{P} + PQ(I_{yy} - I_{xx}) - QRI_{xz} = N \tag{10.164}$$

TABLE 10.14 State Variable and Parameters for 6-Degree of Freedom Aircraft Model

Aircraft axis	Translational			Angular			
	Motion	Velocity	Force	Motion	Displacement	Rate	Moment
x_b	Fore/aft	U	X	Roll	Φ	P	L
y_b	Sideward	V	Y	Pitch	Θ	Q	M
z_b	Heave	W	Z	Yaw	Ψ	R	N

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

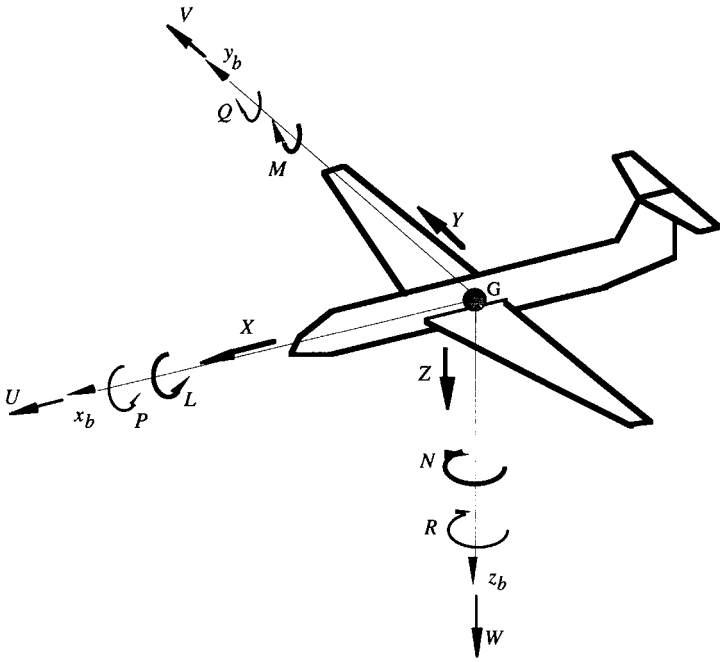


FIGURE 10.88 State variables as referred to the body axes set.

Equations (10.159)–(10.161) are the *translational equations of motion,* derived by consideration of linear momentum. In effect they may be simply expressed as:

$$ma_x = F_x;, \quad \text{where } a_x = \dot{U} + QW - RV, \quad \text{and} \quad F_x = X - mg \sin \Theta \tag{10.165}$$

$$ma_y = F_y, \quad \text{where } a_y = \dot{V} + RU - PW, \quad \text{and} \quad F_y = Y + mg \cos \Theta \sin \Phi \tag{10.166}$$

$$ma_z = F_z, \quad \text{where } a_z = \dot{W} + PV - QU, \quad \text{and} \quad F_z = Z + mg \cos \Theta \cos \Phi \tag{10.167}$$

The component accelerations (a_x, a_y, a_z) are the *absolute* (or *inertial*) accelerations of the center of gravity; recall that the terms QW etc. occur as the frame of reference rotates as it translates. The forces F_x, F_y, and F_z are the total *external forces,* which are composed of the gravitational terms (mg sinΘ, etc.) and the aerodynamic and propulsive terms (X, Y, Z).

The rotational equations of motion are derived from the principles of angular momentum. The *external moments* (L, M, N) are due to aerodynamic and propulsive loads.

Axes Sets and the Euler Transformation. The Euler equations are referred to a frame of reference with origin located at the center of gravity of the system. As previously mentioned, this frame of reference is known as the *body fixed axes set*. This axes set moves with the aircraft and is of practical use only when referred to

an *inertial frame* of reference, i.e., *the earth fixed axes set* (x_E, y_E, z_E). The origin of this axes set is nominal, but the normal convention for directions is that the x_E axis points north, the y_E axis to the east, and the z_E axis down towards the center of the earth. In practical terms the origin of this axes set is often taken as the position of the aircraft at the initiation of a simulation, with the x body and earth axes coincident. The aircraft's position and orientation in terms of the Euler or *attitude* angles (Φ, Θ, Ψ) is given relative to this axes set (Figure 10.89).

The transformation from earth axes (O, x_E, y_E, z_E) to body axes (O, x_b, y_b, z_b) may be achieved through the action of three consecutive rotations (Figure 10.90):

1. A rotation of Ψ (the heading or azimuth angle) about Oz_E to give the intermediate frame (O, x_1, y_1, z_E)

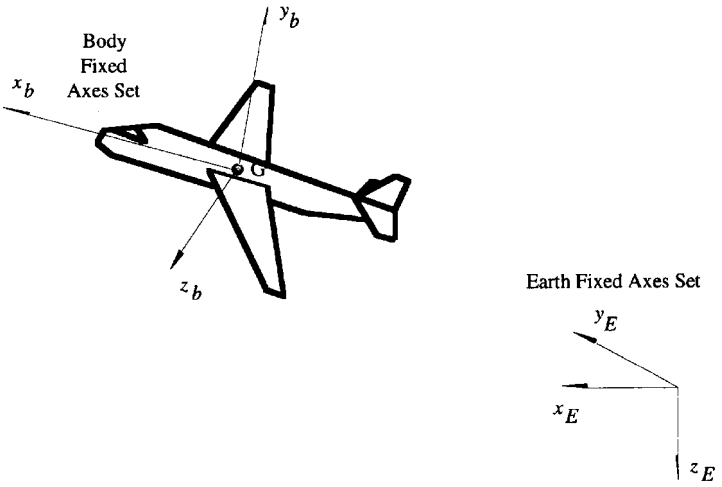


FIGURE 10.89 Earth and body fixed frames of reference.

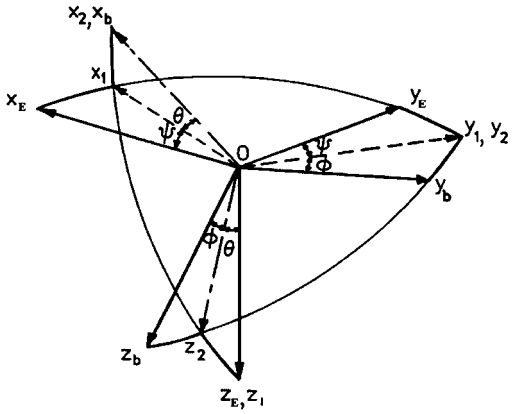


FIGURE 10.90 The Euler angle transformation.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

2. A rotation of Θ (the pitch angle) about Oy_1 to give the intermediate frame (O, x_2, y_1, z_2)
3. A rotation of Φ (the roll or bank angle) about Ox_2 to give the body fixed frame (O, x_b, y_b, z_b)

Defining unit vectors in the x_b, y_b, z_b directions as (i_b, j_b, k_b) for the body axes and unit vectors in the directions x_E, y_E, z_E for the earth axes as (i_E, j_E, k_E), the transformation from the body to the earth axes frame is given by:

$$\begin{bmatrix} i_E \\ j_E \\ k_E \end{bmatrix} = \begin{bmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{bmatrix} \begin{bmatrix} i_b \\ j_b \\ k_b \end{bmatrix} \tag{10.168}$$

where $l_1 = \cos\Theta\cos\Psi$
$l_2 = \cos\Theta\sin\Psi$
$l_3 = -\sin\Theta$
$m_1 = \sin\Phi\sin\Theta\cos\Psi - \cos\Phi\sin\Psi$
$m_2 = \sin\Phi\sin\Theta\sin\Psi + \cos\Phi\cos\Psi$
$m_3 = \sin\Phi\cos\Theta$
$n_1 = \cos\Phi\sin\Theta\cos\Psi + \sin\Phi\sin\Psi$
$n_2 = \cos\Phi\sin\Theta\sin\Psi - \sin\Phi\cos\Psi$
$n_3 = \cos\Phi\cos\Theta$

or, for a general earth fixed axes vector, λ_E, for an earth to body axes transformation, we may write:

$$\lambda_E = \mathbf{T}\lambda_b \tag{10.169}$$

and to transform from body to earth axes the transpose of the matrix is used:

$$\lambda_b = \mathbf{T}^T\lambda_E \tag{10.170}$$

The matrix $\mathbf{T}$ is known as the *Euler angle transformation matrix* while its elements $l_1 \ldots n_3$ are termed the *direction cosines*.

Hence, for translational velocities where

$$\lambda_b = \begin{bmatrix} U \\ V \\ W \end{bmatrix} \quad \text{and} \quad \lambda_E = \begin{bmatrix} \dot{X}_E \\ \dot{Y}_E \\ \dot{Z}_E \end{bmatrix}$$

we would have:

$$\begin{bmatrix} \dot{X}_E \\ \dot{Y}_E \\ \dot{Z}_E \end{bmatrix} = \begin{bmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{bmatrix} \begin{bmatrix} U \\ V \\ W \end{bmatrix} \tag{10.171}$$

and

$$\begin{bmatrix} U \\ V \\ W \end{bmatrix} = \begin{bmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \\ n_1 & n_2 & n_3 \end{bmatrix} \begin{bmatrix} \dot{X}_E \\ \dot{Y}_E \\ \dot{Z}_E \end{bmatrix} \tag{10.172}$$

Note that (X_E, Y_E, Z_E) is in effect the position of the aircraft relative to the earth fixed frame of reference and hence $(\dot{X}_E, \dot{Y}_E, \dot{Z}_E)$ are the component velocities in the directions of the axes. It is possible also to transform the angular velocities such that the earth fixed frame-related Euler angle rates $(\dot{\Phi}, \dot{\Theta}, \dot{\Psi})$ may be expressed in terms of their body fixed equivalents (P, Q, R):

$$P = \dot{\Phi} - \dot{\Psi} \sin \Theta \tag{10.173}$$

$$Q = \dot{\Theta} \cos \Phi + \dot{\Psi} \sin \Phi \cos \Theta \tag{10.174}$$

$$R = \dot{\Psi} \cos \Phi \cos \Theta - \dot{\Theta} \sin \Phi \tag{10.175}$$

and these expressions may be inverted to give:

$$\dot{\Phi} = P + Q \sin \Phi \tan \Theta + R \cos \Phi \tan \Theta \tag{10.176}$$

$$\dot{\Theta} = Q \cos \Phi - R \sin \Phi \tag{10.177}$$

$$\dot{\Psi} = Q \sin \Phi \sec \Theta + R \cos \Phi \sec \Theta \tag{10.178}$$

Control Variables. For a basic aircraft there are three primary flight controls: elevator, ailerons, and rudder. In mathematical terms the "δ" notation can be used, giving the control variables: δ_e, elevator; δ_r rudder; δ_a, aileron. Note that the symbols η, ζ, ξ are often adopted for elevator, rudder, and aileron.

Deflections of these three control surfaces effectively cause changes in the angular rates about the three axes of the aircraft. The mechanism and sign convention for each of them is given below.

- *Elevator* (pitch control): stick forward gives positive δ_e, elevator is depressed, thereby increasing the effective camber of the tailplane, increasing its lift, and thus producing a pitch down moment about the center of gravity (i.e., $+ve\ \delta_e$ gives $-ve\ Q$).
- *Rudder* (yaw control): Left pedal forward denotes positive δ_r, rudder is displaced to left when viewed from above (i.e., towards the port wing), the fin becomes cambered producing an increase in sideforce towards the starboard side, which produces a negative yawing moment about the center of gravity and turns the nose to the left (i.e., $+\ ve\ \delta_r$ gives $-ve\ R$).
- *Aileron* (roll control): Stick right gives positive δ_a, the port aileron is depressed increasing this wing's camber and hence lift, while the starboard aileron is raised, reducing this wing's camber and hence lift. The resulting positive rolling moment causes the aircraft to bank to the right (i.e., $+\ ve\ \delta_a$ gives $+\ ve\ P$).

Simulation of Longitudinal Motion of a Fixed-Wing Aircraft

In this subsection a basic simulation of an aircraft is developed. The starting point is to compile a set of expressions for the external forces and moments, i.e., develop the mathematical model. For convenience the equations of motion are often split into two sets: longitudinal and lateral-directional. Here, a longitudinal simulation is presented in some detail. A full simulation is simply an extension of what is presented.

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

The Mathematical Model. Longitudinal motions occur in the aircraft xz-plane where the state variables are (U, W, Q) and the control variable is the elevator angle, δ_e. In effect longitudinal motions cover fore and aft motions (i.e., accelerations), climbing flight, and pitching flight. The equations used are then (10.159), (10.161), (10.163), and (10.172), with the lateral/directional variables (V, P, R, Φ, Ψ) set to zero.

$$m(\dot{U} + QW) = X - mg \sin \Theta \tag{10.179}$$

$$m(\dot{W} - QU) = Z + mg \cos \Theta \tag{10.180}$$

$$I_{yy}\dot{Q} = M \tag{10.181}$$

$$\dot{\Theta} = Q \tag{10.182}$$

In modelling the longitudinal dynamics of the aircraft it is necessary to calculate X, Z, and M as functions of the state and control variables. From Figure 10.91 we can readily see that these loads are given by:

$$X = T - D \cos \alpha + L \sin \alpha \tag{10.183}$$

$$Z = -L \cos \alpha - D \sin \alpha \tag{10.184}$$

$$M - M_A + Th_T \tag{10.185}$$

where L and D are the total aircraft lift and drag, T is the total engine thrust assumed to act a distance h_T below the x_b axis, and M_A is the aerodynamic pitching moment derived below.

The angle of attack, α, is obtained from

$$\tan \alpha = \frac{W}{U} \tag{10.186}$$

Calculation of the thrust T is dependent on the powerplant, while L, D, and M_A are obtained as follows.

First, the forces are nondimensionalized by division by $\frac{1}{2}\rho V_f^2 S$ such that equations (10.174) and (10.175) become:

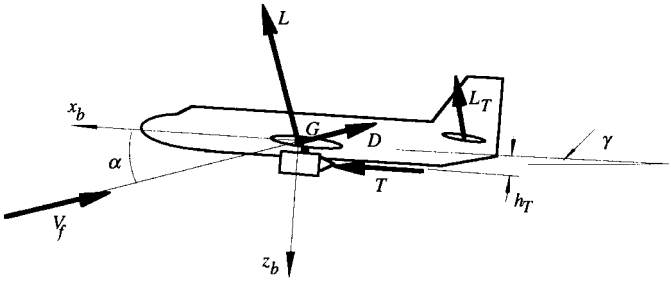


FIGURE 10.91 External forces on aircraft—longitudinal motion.

$$X = \tfrac{1}{2}\rho V_f^2 S(C_T - C_D \cos \alpha + C_L \sin \alpha) \tag{10.187}$$

$$Z = \tfrac{1}{2}\rho V_f^2 S(-C_L \cos \alpha - C_D \sin \alpha) \tag{10.188}$$

Equation (10.176) is divided by $\tfrac{1}{2}\rho V_f^2 S\bar{c}$ (where $\bar{c}$ is the mean aerodynamic chord) to give:

$$M = \tfrac{1}{2}\rho V_f^2 S\bar{c}\left(C_{M_A} + \frac{h_T}{\bar{c}} C_T\right) \tag{10.189}$$

The aerodynamic coefficients are usually obtained either from wind tunnel data or by semiempirical methods. In general terms one might write: $C_L = f(\alpha, \delta_e, M)$, where M = Mach number. Wind tunnel data may be presented in the form of a look-up table, and at any point in the simulation where α, δ_e, and M are known, C_L, C_D, and C_M are found by linear interpolation. Using semiempirical methods, typical expressions for the coefficients are:

$$C_L = C_{L_0} + C_{L_\alpha}\alpha + C_{L_{\delta e}}\delta_e \tag{10.190}$$

$$C_D = C_{D_0} + C_{D_\alpha}\alpha \tag{10.191}$$

$$C_M = C_{M_0} + C_{M_\alpha}\alpha + C_{M_{\delta e}}\delta_e + C_{M_q}\hat{q} \tag{10.192}$$

where the nondimensional pitching velocity, $\hat{q}$, is given by:

$$\hat{q} = \frac{Q\bar{c}}{V_f} \tag{10.193}$$

and

C_{D_0}, C_{L_0} = lift and drag coefficients at zero angle of attack

$C_{L_\alpha} = \dfrac{\partial C_L}{\partial \alpha}$ = lift-curve slope

$C_{L_{\delta e}} = \dfrac{\partial C_L}{\partial \delta_e}$ = change in lift coefficient due to displacement of elevator

$C_{D_\alpha} = \dfrac{\partial C_D}{\partial \alpha}$.

Simulation Procedure. The nonlinear differential equations of motion (10.174–10.177) may be solved numerically, using a Runge–Kutta scheme, for example, to give time histories of the state variables, U, W, Q, Θ in response to a deflection in the elevator angle, δ_e. The convention is to solve the equations from some initial condition representing a *trim state* of the aircraft.

Calculation of a Trim State. The longitudinal trim of an aircraft is usually defined by setting the accelerations and angular velocities to zero. Equations (10.174)–(10.176) become:

$$X - mg \sin \Theta = 0 \tag{10.194}$$

$$Z + mg \cos \Theta = 0 \tag{10.195}$$

$$M = 0 \tag{10.196}$$

and hence there are three equations to satisfy for three unknowns. For a given flight velocity, V_f, altitude (hence air density, ρ) and climb angle, γ, the unknowns are the required thrust and elevator angle and the resulting fuselage pitch attitude, T, δ_e, and Θ, respectively. Hence, substituting equations (10.178)–(10.180) into (10.189)–(10.191) we have:

$$g_1(\Theta, T, \delta_e) = T - D \cos \alpha + L \sin \alpha - mg \sin \Theta = 0 \tag{10.197}$$

$$g_2(\Theta, T, \delta_e) = -L \cos \alpha - D \sin \alpha + mg \cos \Theta = 0 \tag{10.198}$$

$$g_3(\Theta, T, \delta_e) = M_A + Th_T = 0 \tag{10.199}$$

We therefore obtain a system of three nonlinear, algebraic equations g_1, g_2, and g_3, to be solved for three unknowns, T, δ_e, and Θ, which is usually solved using a Newton–Raphson (or similar) iterative scheme.

Calculation of Response to Controls. Noting that $\Phi = \Psi = 0$, for longitudinal motion and rewriting the Euler transformation (10.166) accordingly to give equations (10.199) and (10.200), and recasting equations (10.174)–(10.177) provides a set of six coupled nonlinear differential equations:

$$\dot{U} = -QW + \frac{X}{m} - g \sin \Theta \tag{10.200}$$

$$\dot{W} = QU + \frac{Z}{m} + g \cos \Theta \tag{10.201}$$

$$\dot{Q} = \frac{M}{I_{yy}} \tag{10.202}$$

$$\dot{\Theta} = Q \tag{10.203}$$

$$\dot{X}_E = U \cos \Theta + W \sin \Theta \tag{10.204}$$

$$\dot{Z}_E = -U \sin \Theta + W \cos \Theta \tag{10.205}$$

which can be solved simultaneously for the six states: (U, W, Q, Θ, X_E, Z_E) in response to inputs of elevator δ_e. The elevator input feeds into the equations of motion through the lift and pitching moment (10.185) and (10.187).

10.24 DEVELOPMENT OF THE LINEARIZED EQUATIONS OF MOTION

Although the methods of solving the nonlinear equations computationally are well established and understood, simplified *linearized models* are far more appropriate in order to establish the stability characteristics of an aircraft.

Small-Disturbance Theory—Basic Concept

In small-disturbance theory, the aircraft's motion consists of small deviations from some reference steady flight state (a *trim state*). This assumption is valid for all of the most common flight conditions, and it is only in gross maneuvering flight (e.g., high angle of attack, high-speed maneuvering of fighter aircraft) where the linearized, small-disturbance equations are invalid and the full nonlinear equations must be applied.

Using small-disturbance theory, we assume that the instantaneous total value of each of the state and control variables is composed of two components:

$$U = U_e + u, \quad \Theta = \Theta_e + \theta, \quad \delta_a = \delta_{a_e} + \delta_a', \text{ etc.} \tag{10.206}$$

where the subscript e denotes the reference trim or equilibrium state of the vehicle and the lowercase denotes a perturbation from the reference state. Note that the prime notation is used for perturbations of control variables. In a similar way, it is assumed that the aerodynamic force and moments have two components: the reference value, still denoted by subscript e, and a perturbation, this time denoted by Δ:

$$X = X_e + \Delta X, \text{ etc.}$$

There are three major limitations on the use of the linearized equations of motion:

1. The linearized equations of motion are valid only for small disturbances from the reference trim state. This is a consequence of the small-disturbance assumption, and it implies that calculation of only a few seconds of disturbed (from trim) flight using the linearized equations may be valid.
2. The equations are derived for symmetrical aircraft only. This may not seem too much of a problem, but it does exclude helicopters, which are not symmetrical due to the necessity of having a tail rotor.
3. The equations are derived assuming a rigid aircraft. For small aircraft (even fighters) flying at subsonic speeds this assumption is valid because aeroelastic effects are minimal.

The Reference Trim State

It is convenient both mathematically and physically to refer a dynamic analysis of aircraft motion to a reference trim state. The following assumptions are commonly used:

1. There are no resultant accelerations on the aircraft ($\dot{U}_e = \dot{V}_e = \ldots \dot{R}_e = 0$).
2. The aircraft has no angular velocity ($P_e = Q_e = R_e = 0$).
3. The aircraft is assumed to be in wings-level ($\Phi_e = 0$), symmetric flight ($V_e = 0$).

Choice of Axes Set—Stability Axes

The most commonly used axes set for analysis are the *stability axes*. In this frame the x-axis is fixed in the aircraft in the direction of motion, i.e., the x-body axis is aligned with the relative wind, such that

$$U_e = V_f \quad \text{and} \quad W_e = 0$$

Choosing this axes set has the advantage that it simplifies the calculation of the external forces Figure 10.92(a)) in trimmed flight to

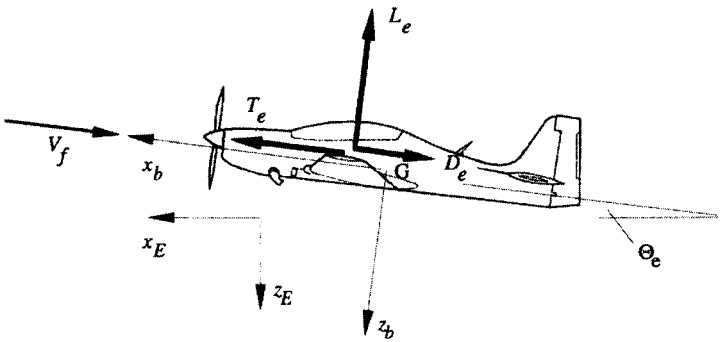
$$X_e = T_e - D_e \tag{10.207}$$

$$Z_e = -L_e \tag{10.208}$$

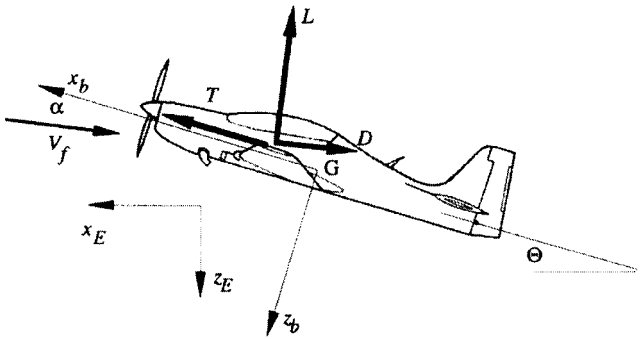
Further, as $W_e = 0$, it is clear from equation (10.186) that $\alpha_e = 0$. In disturbed flight the angle between the relative wind and the x-body axis is the angle of attack, α. As the lift and drag act parallel and perpendicular to the relative wind, the external forces are obtained by resolving lift and drag through the angle α (Figure 10.92(b)):

$$X = T - D\cos\alpha + L\sin\alpha \tag{10.209}$$

$$Z = -L\cos\alpha - D\sin\alpha \tag{10.210}$$



(a) Trimmed Flight



(b) General Disturbed Flight

FIGURE 10.92 Definition of stability axes.

The disadvantage of this choice is that because the aircraft will adopt a different angle of attack (and pitch attitude) for each trimmed flight speed, the x-axis will be oriented with respect to a geometrical datum at each flight speed. Consequently, the values of the moments of inertia, I_{xx}, I_{zz}, etc. will vary with reference flight speed. This is usually considered a minor effect because the angle of attack variation over the speed range may only be a few degrees.

Procedure for Linearizing the Nonlinear Equations

The process of linearization can be summarized as follows:

1. Replace full nonlinear variables values by the reference (trim) plus small perturbation value (i.e., $\dot{U} = \dot{U}_e + \dot{u}$, etc.).
2. Apply appropriate trim values as listed above (i.e., $\dot{U}_e = 0$, $\dot{U} = \dot{u}$, etc.).
3. Make small-angle assumption for attitude perturbations (i.e., $\cos\theta \approx 1$, $\sin\theta \approx \theta$ etc.).
4. Eliminate products of perturbations (for example, it is assumed that $\dot{u} >> (qw - vr)$
5. Eliminate the trim value of the external force or moment (for example, applying equation [10.159] at the trim state gives $0 = X_e - mg \sin\Theta_e$, etc.).

The Linearized Equations of Motion in Basic Form

Applying the procedure described above to the nonlinear equations, (10.159)–(10.164), produces the following set of linearized equations

$$m\dot{u} = \Delta X - mg\theta\cos\Theta \tag{10.211}$$

$$m\dot{v} = \Delta Y - U_e r + mg\phi\cos\Theta_e \tag{10.212}$$

$$m\dot{w} = \Delta Z + mU_e q - mg\theta\sin\Theta_e \tag{10.213}$$

$$I_{xx}\dot{p} - I_{xz}\dot{r} = \Delta L \tag{10.214}$$

$$I_{yy}\dot{q} = \Delta M \tag{10.215}$$

$$I_{zz}\dot{r} - I_{xz}\dot{p} = \Delta N \tag{10.216}$$

The expressions for the body rates in terms of the Euler angle rates, (10.173)–(10.175) are also linearized to give:

$$p = \dot{\phi} - \dot{\psi}\sin\Theta_e \tag{10.217}$$

$$q = \dot{\theta} \tag{10.218}$$

$$r = \dot{\psi}\cos\Theta_e \tag{10.219}$$

Linear Expressions for the Aerodynamic Force and Moment Perturbations

Linear expressions for the aerodynamic and propulsive force and moment perturbations are obtained by assuming that the external forces and moments are functions

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

of the instantaneous values of the disturbance velocities, control angles and their time derivatives, i.e.,

$$X = f_1(u, \dot{u}, v, \dot{v}, w, \dot{w}, p, \dot{p}, q, \dot{q}, r, \dot{r}, \delta_e, \dot{\delta}_e, \delta_r, \dot{\delta}_r, \delta_a, \dot{\delta}_a)$$

$$\vdots$$

$$N = f_6(u, \dot{u}, v, \dot{v}, w, \dot{w}, p, \dot{p}, q, \dot{q}, r, \dot{r}, \delta_e, \dot{\delta}_e, \delta_r, \dot{\delta}_r, \delta_a, \dot{\delta}_a)$$

The method normally used to linearize the external forces and moments is to represent them by a Taylor series expansion. The Taylor series expansion for a multivariable problem can be applied to the external forces and moments to give for the *X*-force:

$$\Delta X = \frac{\partial X}{\partial u} u + \frac{\partial X}{\partial \dot{u}} \dot{u} + \frac{\partial X}{\partial v} v + \ldots + \frac{\partial X}{\partial \dot{\delta}_a} \dot{\delta}_a$$

where the higher order derivatives have been neglected.

It should also be noted that the derivatives have to be evaluated at the point about which the expansion was derived—the equilibrium trim state—and hence the derivatives must be calculated using the state values from trim, denoted by the subscript *e*. The full set of external force and moment perturbation linearizations is:

$$\Delta X = \left(\frac{\partial X}{\partial u}\right)_e u + \left(\frac{\partial X}{\partial \dot{u}}\right)_e \dot{u} + \left(\frac{\partial X}{\partial v}\right)_e v + \ldots + \left(\frac{\partial X}{\partial \dot{\delta}_a}\right)_e \dot{\delta}_a$$

$$\vdots$$

$$\vdots$$

$$\Delta N = \left(\frac{\partial N}{\partial u}\right)_e u + \left(\frac{\partial N}{\partial \dot{u}}\right)_e \dot{u} + \left(\frac{\partial N}{\partial v}\right)_e v + \ldots + \left(\frac{\partial N}{\partial \dot{\delta}_a}\right)_e \dot{\delta}_a$$

These derivatives are known as the *stability derivatives* or the *aerodynamic derivatives*. In their full form as shown above there are six states plus three controls and their derivatives, giving 18 aerodynamic derivatives to represent each external force or moment. Clearly, if all 18 derivatives in all six equations ($6 \times 18 = 108$) were used, the equations would become large and difficult to manipulate. Fortunately, for a wide range of flight states many of the derivatives are small and may be neglected. The rationale behind neglecting certain derivatives is as follows.

1. For any condition of symmetric flight (in the *xz*-plane) the asymmetric forces and moments (Y, L, N) are zero. It then follows that the derivatives of the asymmetric forces and moments with respect to the symmetric variables (u, w, q, δ_e) will be zero, i.e.,

$$\frac{\partial Y}{\partial u} = \frac{\partial Y}{\partial w} = \frac{\partial Y}{\partial q} = \frac{\partial Y}{\partial \delta_e} = \cdots = \frac{\partial L}{\partial u} = \cdots \frac{\partial N}{\partial \delta_e} = 0$$

2. Similarly, the derivatives of the symmetric forces and moments (X, Z, M) with respect to the asymmetric variables (v, p, r, δ_r, δ_a) are zero, i.e.,

$$\frac{\partial X}{\partial v} = \frac{\partial X}{\partial p} = \frac{\partial X}{\partial r} = \frac{\partial X}{\partial \delta_r} = \frac{\partial X}{\partial \delta_a} = \cdots = \frac{\partial Z}{\partial v} = \cdots \frac{\partial M}{\partial \delta_a} = 0$$

3. It has also been found through experiment and experience that the derivatives with respect to acceleration are all negligible except $(\partial M)/(\partial \dot{w})$, i.e.,

$$\frac{\partial X}{\partial \dot{u}} = \frac{\partial X}{\partial \dot{v}} = \frac{\partial X}{\partial \dot{w}} = \cdots = \frac{\partial N}{\partial \dot{r}} = 0$$

4. The control rate derivatives are all negligible:

$$\frac{\partial X}{\partial \dot{\delta}_e} = \frac{\partial Y}{\partial \dot{\delta}_e} = \frac{\partial Z}{\partial \dot{\delta}_e} = \cdots = \frac{\partial N}{\partial \dot{\delta}_a} = 0$$

5. Again, through experiment and experience, the following derivatives may also be neglected

$$\frac{\partial X}{\partial q} = \frac{\partial X}{\partial \delta_e} = \frac{\partial Y}{\partial p} = \frac{\partial Y}{\partial r} = \frac{\partial Y}{\partial \delta_a} = 0$$

The Linearized Equations of Motion

The simplified expressions for perturbation aerodynamic forces and moments, ΔX, . . . , ΔN, can be substituted into equations (10.211)–(10.216) to give:

$$m\dot{u} = \frac{\partial X}{\partial u} u + \frac{\partial X}{\partial w} w - mg \cos \Theta_e \theta \tag{10.220}$$

$$m\dot{v} = \frac{\partial Y}{\partial v} v - mU_e r + mg \cos \Theta_e \phi + \frac{\partial Y}{\partial \delta_r} \delta_r \tag{10.221}$$

$$m\dot{w} = \frac{\partial Z}{\partial u} u + \frac{\partial Z}{\partial w} w + \frac{\partial Z}{\partial q} q + mU_e q - mg \sin \Theta_e \theta + \frac{\partial Z}{\partial \delta_e} \delta_e \tag{10.222}$$

$$I_{xx}\dot{p} - I_{xz}\dot{r} = \frac{\partial L}{\partial v} v + \frac{\partial L}{\partial p} p + \frac{\partial L}{\partial r} r + \frac{\partial L}{\partial \delta_r} \delta_r + \frac{\partial L}{\partial \delta_a} \delta_a \tag{10.223}$$

$$I_{yy}\dot{q} = \frac{\partial M}{\partial u} u + \frac{\partial M}{\partial w} w + \frac{\partial M}{\partial \dot{w}} \dot{w} + \frac{\partial M}{\partial q} q + \frac{\partial M}{\partial \delta_e} \delta_e \tag{10.224}$$

$$I_{zz}\dot{r} - I_{xz}\dot{p} = \frac{\partial N}{\partial v} v + \frac{\partial N}{\partial p} p + \frac{\partial N}{\partial r} r + \frac{\partial N}{\partial \delta_r} \delta_r + \frac{\partial N}{\partial \delta_a} \delta_a \tag{10.225}$$

Linearized Equations in Compact Form

On examination, it is clear that there are two sets of equations—longitudinal equations, (10.220), (10.222), (10.224), where u, w, q are controlled by δ_e, and the lateral/directional set, (10.221), (10.223), (10.225) where v, p, r are controlled by δ_r, and δ_a. The longitudinal and lateral/directional dynamics of the aircraft may be

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

treated separately. The above equations may be manipulated and written in a more compact form.

Including the kinematic expression (10.218), and substituting for $\dot{w}$ in (10.224) using (10.222), the longitudinal linearized equations of motion may be written in matrix form as:

$$\begin{bmatrix} \dot{u} \\ \dot{w} \\ \dot{q} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} X_u & X_w & 0 & -g\cos\Theta_e \\ Z_u & Z_w & Z_q + U_e & -g\sin\Theta_e \\ M_u^* & M_w^* & M_q^* & M_\theta^* \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} u \\ w \\ q \\ \theta \end{bmatrix} + \begin{bmatrix} 0 \\ Z_{\delta_e} \\ M_{\delta_e} \\ 0 \end{bmatrix} [\delta_e] \qquad (10.226)$$

where

$$X_u = \frac{1}{m}\frac{\partial X}{\partial u},\ X_w = \frac{1}{m}\frac{\partial X}{\partial w},\ Z_u = \frac{1}{m}\frac{\partial Z}{\partial u},\ Z_w = \frac{1}{m}\frac{\partial Z}{\partial w},\ Z_q = \frac{1}{m}\frac{\partial Z}{\partial q},\ Z_{\delta_e} = \frac{1}{m}\frac{\partial Z}{\partial \delta_e},$$

$$M_u^* = M_u + M_{\dot{w}} Z_u, \quad M_w^* = M_w + M_{\dot{w}} Z_w, \quad M_q^* = M_q + M_{\dot{w}} U_e,$$

$$M_\theta^* = -M_{\dot{w}} g\sin\Theta_e, \quad M_{\delta_e}^* = M_{\delta_e} + M_{\dot{w}} Z_{\delta_e}, \quad \text{and} \quad M_u = \frac{1}{I_{yy}}\frac{\partial M}{\partial u}, \text{ etc.}$$

Decoupling equations (10.223) and (10.225) and including the kinematic expressions (10.217) and (10.219) the full set of linearized lateral/directional equations may be written in matrix form as:

$$\begin{bmatrix} \dot{v} \\ \dot{p} \\ \dot{r} \\ \dot{\phi} \\ \dot{\Psi} \end{bmatrix} = \begin{bmatrix} Y_v & 0 & -U_e & g\cos\Theta_e & 0 \\ L_v^* & L_p^* & L_r^* & 0 & 0 \\ N_{sv}^* & N_p^* & N_r^* & 0 & 0 \\ 0 & 1 & \tan\Theta_e & 0 & 0 \\ 0 & 0 & \sec\Theta_e & 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ p \\ r \\ \phi \\ \Psi \end{bmatrix} + \begin{bmatrix} Y_{\delta_r} & 0 \\ L_{\delta_r}^* & L_{\delta_a}^* \\ N_{\delta_r}^* & N_{\delta_a}^* \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \delta_r \\ \delta_a \end{bmatrix} \qquad (10.227)$$

where

$$Y_v = \frac{1}{m}\frac{\partial Y}{\partial v}, \quad Y_{\delta_r} = \frac{1}{m}\frac{\partial Y}{\partial \delta_r},$$

$$L_v^* = \frac{I_{xx}I_{zz}}{I_{xx}I_{zz} - I_{xz}^2}\left(L_v + \frac{I_{xz}}{I_{xx}} N_v\right), \quad L_p^* = \frac{I_{xx}I_{zz}}{I_{xx}I_{zz} - I_{xz}^2}\left(L_p + \frac{I_{xz}}{I_{xx}} N_p\right), \text{ etc.},$$

$$L_v = \frac{1}{I_{xx}}\frac{\partial L}{\partial v}, \quad L_p = \frac{1}{I_{xx}}\frac{\partial L}{\partial p}, \text{ etc.},$$

and

$$N_v^* = \frac{I_{xx}I_{zz}}{I_{xx}I_{zz} - I_{xz}^2}\left(N_v + \frac{I_{xz}}{I_{zz}} L_v\right), \quad N_p^* = \frac{I_{xx}I_{zz}}{I_{xx}I_{zz} - I_{xz}^2}\left(N_p + \frac{I_{xz}}{I_{zz}} L_p\right), \text{ etc.},$$

$$N_v = \frac{1}{I_{zz}}\frac{\partial N}{\partial v}, \quad N_p = \frac{1}{I_{zz}}\frac{\partial N}{\partial p}$$

The Incidence Angles Expressed in Linear Form

The nonlinear expressions for angle of attack and angle of sideslip are:

$$\tan \alpha = \frac{W}{U} \quad \text{and} \quad \sin \beta = \frac{V}{V_f}$$

By making small-angle assumptions and applying the appropriate trim information, the following linear expressions are derived:

$$\alpha = \frac{w}{U_e} \quad \text{and} \quad \beta = \frac{v}{U_e}$$

The Nondimensional Linearized Equations of Motion

The equations of motion are often used in their nondimensional form. The main advantage of this is that the derivatives (which become coefficients) of different aircraft can be directly compared. The nondimensionalizing quantities used are listed in Table 10.15. Applying the nondimensionalizing factors given in Table 10.15 to the longitudinal set of equations (10.226) gives the following set of nondimensional equations:

$$\begin{bmatrix} 2\mu & 0 & 0 & 0 \\ 0 & 2\mu & 0 & 0 \\ 0 & -C_{m_{\dot{\alpha}}} & i_{yy} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} D\hat{u} \\ D\alpha \\ D\hat{q} \\ D\theta \end{bmatrix}$$

$$= \begin{bmatrix} (C_{x_u} + 2C_{L_0} \tan \theta_e) & C_{x_\alpha} & 0 & -C_{L_0} \\ (C_{z_u} - 2C_{L_0}) & C_{z_\alpha} & 2\mu & -C_{L_0} \tan \theta_e \\ C_{m_u} & C_{m_\alpha} & C_{m_q} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \hat{u} \\ \alpha \\ \hat{q} \\ \theta \end{bmatrix} + \begin{bmatrix} 0 \\ C_{z_{\delta_e}} \\ C_{m_{\delta_e}} \\ 0 \end{bmatrix} [\delta_e]$$

Similarly, the lateral/directional equations in nondimensional form are:

$$\begin{bmatrix} 2\mu & 0 & 0 & 0 & 0 \\ 0 & i_{xx} & -i_{xz} & 0 & 0 \\ 0 & -i_{xz} & i_{zz} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} D\beta \\ D\hat{p} \\ D\hat{r} \\ D\phi \\ D\Psi \end{bmatrix}$$

$$= \begin{bmatrix} C_{y_\beta} & 0 & -2\mu & C_{L_0} & 0 \\ C_{l_\beta} & C_{l_p} & C_{l_r} & 0 & 0 \\ C_{n_\beta} & C_{n_p} & C_{n_r} & 0 & 0 \\ 0 & 1 & \tan \theta_e & 0 & 0 \\ 0 & 0 & \sec \theta_e & 0 & 0 \end{bmatrix} \begin{bmatrix} \beta \\ \hat{p} \\ \hat{r} \\ \phi \\ \Psi \end{bmatrix} + \begin{bmatrix} 0 & C_{y_{\delta r}} \\ C_{l_{\delta a}} & C_{l_{\delta r}} \\ C_{n_{\delta a}} & C_{n_{\delta r}} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \delta_a \\ \delta_r \end{bmatrix}$$

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

TABLE 10.15 Nondimensionalizing Factors

Dimensional quantity	Divisor	Nondimensional quantity
X, Y, Z	$\frac{1}{2}\rho V^2 S$	C_x, C_y, C_z
L, M, N	$\rho V^2 S l$	C_l, C_m, C_n
U, V, W	U_e	$\hat{u}, \beta, \alpha$
P, Q, R	$\frac{1}{t^*}$	$\hat{p}, \hat{q}, \hat{r}$
$\dot{\alpha}, \dot{\beta}$	$\frac{1}{t^*}$	$D\alpha, D\beta$
$\dot{\delta}_e, \dot{\delta}_a, \dot{\delta}_r$	$\frac{1}{t^*}$	$D\delta_e, D\delta_a, D\delta_r$
m	$\rho S l$	μ
$I_{xx}, I_{yy}, I_{zz}, I_{xz}$	$\rho S l^3$	$i_{xx}, i_{yy}, i_{zz}, i_{xz}$
t	t^*	$\hat{t}$

$D = \frac{d}{d\hat{t}}$, $t^* = \frac{l}{U_e}$, and $l = \frac{\bar{c}}{2}$ (longitudinal equations), and $l = \frac{b}{2}$ (lat/dir equations).

The nondimensional form of the derivatives are for

$$C_{x_u} = \frac{\partial C_x}{\partial \hat{u}}, \text{ etc.}$$

Note that the relationship between the compact form of the derivatives given in equations (10.226) and (10.227) do not have a direct mapping to their nondimensional counterparts.

The Equations of Motion in State Space Form

It is clear from equations (10.226) and (10.227) that the linearized equations of may be written in state space form:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{10.228}$$

where $\mathbf{x}$ = the state vector (n)
$\mathbf{u}$ = the control vector (m)
$\mathbf{A}$ = the system matrix ($n \times n$)
$\mathbf{B}$ = the control matrix ($n \times m$)

for a system with n states and m controls. For the longitudinal equations in state space form:

$$\mathbf{A} = \begin{bmatrix} X_u & X_w & 0 & -g\cos\theta_e \\ Z_u & Z_w & (Z_q + U_e) & -q\sin\theta_e \\ M_u^* & M_w^* & M_q^* & M_\theta^* \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ Z_{\delta_e} \\ M_{\delta_e}^* \\ 0 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} u \\ w \\ q \\ \theta \end{bmatrix}, \quad \mathbf{u} = [\delta_e]$$

and for the lateral equations in state space form

$$\mathbf{A} = \begin{bmatrix} Y_v & 0 & -U_e & g\cos\theta_e & 0 \\ L_v^* & L_p^* & L_r^* & 0 & 0 \\ N_v^* & N_p^* & N_r^* & 0 & 0 \\ 0 & 1 & \tan\theta_e & 0 & 0 \\ 0 & 0 & \sec\theta_e & 0 & 0 \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} Y_{\delta_r} & 0 \\ L_{\delta_r}^* & L_{\delta_a}^* \\ N_{\delta_r}^* & N_{\delta_a}^* \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} v \\ p \\ r \\ \phi \\ \Psi \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} \delta_r \\ \delta_a \end{bmatrix}$$

The response of the aircraft in terms of variables other than the state variables can be obtained by use of an *output equation* of the form

$$\mathbf{y} = \mathbf{Cx} + \mathbf{Du} \tag{10.229}$$

where $\mathbf{y}$ = the output vector (p)
$\mathbf{C}$ = the output matrix ($p \times n$)
$\mathbf{D}$ = the direct matrix ($p \times m$)

for an output response with p states. Further, it is possible to obtain a transfer function relating an output, y_i, from the output vector, $\mathbf{y}$, and a single control, u_j, from the control vector, $\mathbf{u}$, using the expression:

$$\frac{Y_i(s)}{U_j(s)} = \mathbf{C_i}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B_j} + d_{ij} \tag{10.230}$$

where $\mathbf{B_j}$ is the column of the $\mathbf{B}$ matrix associated with the control u_j (the jth column of $\mathbf{B}$), $\mathbf{C}_i$ is the row of the $\mathbf{C}$ matrix corresponding to the output y_i (the ith row of $\mathbf{C}$), and d_{ij} is the element of $\mathbf{D}$ associated with output y_i and control u_j.

10.25 CALCULATION OF AERODYNAMIC DERIVATIVES

The linearized equations of motion given by (10.226) and (10.227) are valid for any aircraft. The main factor that determines the dynamic characteristics of a particular aircraft will be the values of its aerodynamics or stability derivatives. These derivatives can be obtained experimentally by wind tunnel testing or extracted from data recorded in flight trials. There is, of course, a need to be able to estimate the value of derivatives from basic configurational information. The following analyt-

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

ical expressions for derivatives are based on readily obtainable aircraft data and given in dimensional and nondimensional coefficient form.

Force Coefficients

Referring to Figure 10.92, the aircraft X and Z external forces may be expressed as:

$$X = T - D + L\alpha \qquad Z = -L - D\alpha$$

assuming that the angle of attack, α, is small. Noting that if stability axes are used the angle of attack in the trim state is zero, the trim values of the X and Z forces are:

$$X_e = -L_e \qquad Z_e = -L_e$$

and nondimensionalizing gives:

$$C_x = C_T - C_D + \alpha C_L \qquad C_Z = -C_L = -\alpha C_D$$

and

$$C_{x_e} = -C_{T_e} - C_{D_0} \qquad C_{z_e} = -C_{L_0}$$

where C_{L_0} and C_{D_0} refer to the values of these coefficients at zero angle of attack.

The Longitudinal Derivatives

1. The u-derivatives:

$$X_w = \frac{\rho S U_e}{2m} C_{x_\alpha} \qquad C_{x_u} = -2(C_{D_0} + C_{L_0} \tan \Theta_e) - M_e \frac{\partial C_D}{\partial M}$$

where M_e = the Mach number in equilibrium flight.

$$Z_u = \frac{\rho S U_e}{2m} (C_{Z_u} + 2C_{Z_e}) \qquad C_{Z_u} = -M_e \frac{\partial C_L}{\partial \hat{u}}$$

$$M_u = \frac{\rho S \bar{c} U_e}{2I_{yy}} C_{m_u} \qquad C_{M_u} = M_e \frac{\partial C_M}{\partial \hat{u}} + \rho U_e^2 \frac{\partial C_M}{\partial p_d}$$

where p_d = dynamic pressure.

2. The w-derivatives:

$$X_w = \frac{\rho S U_e}{2m} C_{x_\alpha} \qquad C_{x_\alpha} = C_{L_0} - C_{D_\alpha}$$

$$Z_w = \frac{\rho S U_e}{2m} C_{z_\alpha} \qquad C_{z_\alpha} = -C_{L_\alpha} - C_{D_0}$$

where

$$C_{L_\alpha} = \frac{\partial C_L}{\partial \alpha} \quad \text{and} \quad C_{D_\alpha} = \frac{\partial C_D}{\partial \alpha}$$

$$M_w = \frac{\rho S U_e \bar{c}}{2I_{yy}} C_{m_\alpha} \qquad C_{m_\alpha} = C_{L_\alpha}(h - h_n)$$

3. The q-derivatives—tailplane contribution only:

$$M_q = \frac{\rho S \bar{c}^2 U_e}{4I_{yy}} C_{m_q} \qquad C_{m_q} = \frac{-2V_H a_t l_t}{\bar{c}}$$

$$Z_q = \frac{\rho S \bar{c} U_e}{2m} C_{z_q} \qquad C_{z_q} = -2V_H a_t$$

where $V_H = \dfrac{S_t l_t}{S\bar{c}}$ = tailplane volume ratio
S_t = tailplane area
l_t = tailplane moment arm
a_t = tailplane lift curve slope = $(\partial C_{L_t})/(\partial \alpha_t)$

4. The $\dot{w}$-derivative—tailplane contribution only:

$$M_{\dot{w}} = \frac{\rho S \bar{c}^2}{4I_{yy}} C_{m_{\dot{\alpha}}} \qquad C_{M_{\dot{\alpha}}} = -2a_t V_H \frac{l_t}{\bar{c}} \frac{\partial \varepsilon}{\partial \alpha}$$

where ε = the tailplane downwash angle.

5. The δ_e-derivatives:

$$Z_{\delta_e} = \frac{\rho S U_e^2}{2m} C_{z_{\delta e}} \qquad C_{z_{\delta e}} = -a_e \frac{S_t}{S}$$

$$M_{\delta_e} = \frac{\rho S \bar{c} U_e^2}{2I_{yy}} C_{m_{\delta e}} \qquad C_{m_{\delta e}} = -a_e V_H$$

where $a_e = (\partial C_{L_t})/(\partial \delta_e)$.

The Lateral / Directional Derivatives

1. The v-derivatives—fin contribution only:

$$Y_v = \frac{\rho S U_e}{2m} C_{y_\beta} \qquad C_{y_\beta} = -a_F \left(1 - \frac{\partial \sigma}{\partial \beta}\right) \frac{S_F}{S}$$

$$L_v = \frac{\rho S b U_e}{2I_{xx}} C_{l_\beta} \qquad C_{l_\beta} = -a_F \left(1 - \frac{\partial \sigma}{\partial \beta}\right) \frac{S_F}{S} \frac{z_F}{b}$$

$$N_v = \frac{\rho S b U_e}{2I_{zz}} C_{n_\beta} \qquad C_{n_\beta} = a_F V_v \left(1 - \frac{\partial \sigma}{\partial \beta}\right)$$

where a_F = lift curve slope of the fin
σ = sidewash angle
S_F = fin area
z_F = height of fin mean aerodynamic center above x-axis
V_v = fin volume ratio = $(S_F l_F)/(Sb)$

2. The p-derivatives—fin contribution only:

$$Y_p = \frac{\rho U_e S b}{4m} C_{y_p} \qquad C_{y_p} = -a_F \frac{S_F}{S}\left(2\frac{z_F}{b} - \frac{\partial\sigma}{\partial\hat{p}}\right)$$

$$N_p = \frac{\rho U_e S b^2}{4I_{zz}} C_{n_p} \qquad C_{n_p} = a_F V_v \left(2\frac{z_F}{b} - \frac{\partial\sigma}{\partial\hat{p}}\right)$$

3. The r-derivatives—fin contribution only:

$$Y_r = \frac{\rho U_e S b}{4m} C_{y_r} \qquad C_{y_r} = -a_F \frac{S_F}{S}\left(2\frac{l_F}{b} - \frac{\partial\sigma}{\partial\hat{r}}\right)$$

$$L_r = \frac{\rho U_e S b^2}{4I_{xx}} C_{l_r} \qquad C_{y_r} = a_F \frac{S_F}{S}\frac{z_F}{b}\left(2\frac{l_F}{b} - \frac{\partial\sigma}{\partial\hat{r}}\right)$$

$$N_r = \frac{\rho U_e S b^2}{4I_{zz}} C_{n_r} \qquad C_{n_r} = -a_F V_v \left(2\frac{l_F}{b} - \frac{\partial\sigma}{\partial\hat{r}}\right)$$

where l_F = distance between fin aerodynamic center and aircraft center of gravity.

4. The δ_r-derivatives:

$$Y_{\delta_r} = \frac{\rho S U_e^2}{2m} C_{y_{\delta_r}} \qquad C_{y_{\delta_r}} = a_r \frac{S_F}{S}$$

$$L_{\delta_r} = \frac{\rho S U_e^2 b}{2I_{xx}} C_{l_{\delta_r}} \qquad C_{l_{\delta_r}} = a_r \frac{S_F}{S}\frac{z_F}{b}$$

$$N_{\delta_r} = \frac{\rho S U_e^2 b}{2I_{zz}} C_{n_{\delta_r}} \qquad C_{l_{\delta_r}} = -a_r V_v$$

where a_r = rudder effectiveness = $(\partial C_{Y_F})/(\partial\delta_r)$

5. The δ_a-derivatives: Simple expressions for these derivatives are not available; they are normally estimated from wind tunnel tests.

10.26 AIRCRAFT DYNAMIC STABILITY

Prediction of Stability—General Theory

The stability of any dynamic system is obtained by consideration of its free (unforced) motion. For an aircraft, unforced motion implies that there should be no control inputs (i.e., $\mathbf{u} = 0$), such that the controls remain fixed at their trim value. The state space equation (10.180) becomes:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \tag{10.231}$$

Consider the general case where the aircraft has n degrees of freedom. Equation (10.223) will have a general solution:

$$\mathbf{x} = \mathbf{x}_0 e^{\lambda t}$$

Substitution of this solution into equation (10.231) yields the familiar eigenvalue problem:

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{x_0} = 0 \tag{10.232}$$

In this expression $(\lambda\mathbf{I} - \mathbf{A})$ is an $(n \times n)$ matrix, while $\mathbf{x}_0$ is a vector of dimension (n). Equation (10.232) has a trivial solution ($\mathbf{x_0} = 0$), or the more useful solution that the determinant of $(\lambda\mathbf{I} - \mathbf{A})$ should be zero, giving the *characteristic equation:*

$$|\lambda\mathbf{I} - \mathbf{A}| = 0 \tag{10.233}$$

the solution of which yields n eigenvalues, λ_i, $(i = 1, n)$. If these values of λ are substituted into equation (10.232), then for the eigenvalue λ_i:

$$(\lambda_i\mathbf{I} - \mathbf{A})\mathbf{x_0} = 0 \tag{10.234}$$

Equation (10.234) can be solved for each eigenvalue, λ_i, to give a vector of amplitudes x_{0_i}, $i = 1, n$. This is known as the eigenvector, and its value can assist in the determining which state variables are influenced by each eigenvalue.

Equation (10.233) is a polynomial function of λ which when solved gives the system eigenvalues which can have real or imaginary values upon which the stability of the system is dependent.

Imaginary Eigenvalues: $\lambda = n \pm \omega i$. The general solution for a state i takes the form:

$$x_i = x_{0_i} e^{\lambda t} = x_{0_i} e^{(n \pm \omega i)t} = x_{0_i} e^{nt} e^{\pm \omega i t} = x_{0_i} e^{nt}(\cos \omega t \mp i \sin \omega t)$$

The response is an oscillation with angular frequency ω (the imaginary part of the eigenvalue). The amplitude of the oscillation will decrease provided the real part is negative. In this case the system is said to be *dynamically stable*. Conversely, should the real part be positive, then the amplitude of the oscillations will increase and the system will be *dynamically unstable*.

Real Eigenvalues. Now the general solution for a state i takes the form:

$$x_i = x_{0_i} e^{\lambda t}$$

If λ is negative then the response will be an exponential decay and the system is then *statically stable.* If the λ is positive then the response will be an exponential growth and the system is statically unstable.

Period and Time to Half or Double Amplitude. From above it is clear that for an imaginary eigenvalue, $\lambda = n \pm \omega i$, the imaginary part, ω, determines the period of the oscillation:

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

$$T = \frac{2\pi}{\omega} s$$

or, more generally,

$$T = \frac{2\pi}{\text{Im}(\lambda)} s$$

The damping of the mode is usually measured by the *time to half amplitude* in the case of a convergent mode or *time to double amplitude* in the case of a divergent mode. It can be shown that:

$$t_{\text{half}} \quad \text{or} \quad t_{\text{double}} = \frac{0.69}{|n|} s$$

or, more generally,

$$t_{\text{half}} \quad \text{or} \quad t_{\text{double}} = \frac{0.69}{|\text{Re}(\lambda)|}$$

Note that the expressions given above are valid where λ has been calculated from the dimensional equations of motion. When the nondimensional equations are being used they must be multiplied by the factor t^* to obtain values in seconds.

Prediction of Aircraft Stability

It is possible to predict aircraft stability using the methods described at the start of this subsection. For almost all fixed-wing aircraft the longitudinal and lateral/directional modes are uncoupled and so can be treated separately. This simplifies the problem as two independent lower-order systems may be analyzed. For aircraft such as helicopters that generate nonsymmetrical loads and exhibit heavy coupling between longitudinal and lateral directional modes, much higher-order system matrices are generated.

Aircraft Longitudinal Dynamic Stability

The longitudinal stability properties of an aircraft with the controls fixed (i.e., the elevator does not move, $\delta_e = 0$) can be determined by expressing equation (10.226) in the form:

$$\begin{bmatrix} \dot{u} \\ \dot{w} \\ \dot{q} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} X_u & X_w & 0 & -q\cos\Theta_e \\ Z_u & Z_w & Z_q + U_e & -g\sin\Theta_e \\ M_u^* & M_w^* & M_q^* & M_\theta^* \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} u \\ w \\ q \\ \theta \end{bmatrix}$$

This is in the same form as equation (10.231) and hence the characteristic equation can be obtained from (10.233). If it is assumed that $\Theta_e = 0$ (level trimmed flight) and that $U_e >> Z_q$, then the characteristic equation may be written in the form:

$$\lambda^4 + A\lambda^3 + B\lambda^2 + C\lambda + D = 0 \tag{10.235}$$

where $A = -(M_q^* + X_u + Z_w)$
$B = [-M_w^* U_e + M_q^*(Z_w + X_u) + X_u Z_w - X_w Z_u]$
$C = [gM_u^* + U_e\,(M_w^* X_u - M_u^* X_w) + M_q^*(X_w Z_u - X_u Z_w)]$
$D = g(M_w^* Z_u - M_u^* Z_w)$

Attempting to factorize to find λ analytically is unrealistic. In general, the longitudinal equations will give two oscillatory modes, with the characteristic equation factorizing to give:

$$(\lambda^2 + 2\zeta_{ph}\omega_{n_{ph}}\lambda + \omega_{n_{ph}}^2)(\lambda^2 + 2\zeta_{sp}\omega_{n_{sp}}\lambda + \omega_{n_{sp}}^2) = 0 \tag{10.236}$$

Most aircraft exhibit these two classical longitudinal modes: the phugoid (ph) and the short-period (sp) modes. These modes are now discussed individually.

The Phugoid Mode. The phugoid mode is characterized by lightly damped oscillations in altitude and airspeed. The period can typically be from 10 s to around 2 minutes in the case of a large airliner. Inspection of the corresponding eigenvector reveals that this mode influences the aircraft speed, u, more than its angle of attack, α. The phugoid mode is therefore a pitching oscillation at almost constant angle of attack, often lightly damped and occasionally unstable.

Although an exact analytical representation of the phugoid characteristic equation is not practical, it is possible to obtain an approximation. Assuming that the pitch acceleration ($\dot{q}$) is small, the pitching moment equation becomes one of static balance and it can be shown that the phugoid characteristic equation can be approximated by:

$$\lambda^2 + \left[-X_u - \frac{M_u^*}{M_w^*}\left(\frac{g}{U_e} - X_w\right)\right]\lambda + \frac{g}{U_e}\left(-Z_u + Z_w\frac{M_u^*}{M_w^*}\right) = 0$$

The Short-Period Mode. The short-period mode is a very fast and heavily damped oscillation in pitch. Its period can be less than a second for highly maneuverable aircraft and no more than a few seconds for larger vehicles. The variable most influenced by the short-period oscillation is angle of attack with little or no change in airspeed. This is readily confirmed by inspection of the eigenvector.

Although an exact analytical representation of the short-period characteristic equation is not practical, it is possible to obtain an approximation. Assuming that the velocity changes very little during the short-period motion, it is possible to neglect the x-equation of motion, and it can be shown that the short-period characteristic equation can be approximated by:

$$\lambda^2 - (Z_w + M_q^*)\lambda + (Z_w M_q^* - U_e M_w^*) = 0$$

Aircraft Lateral/Directional Stability

The stick-fixed stability is investigated by assuming that the aileron and rudder controls are locked ($\delta_r = \delta_a = 0$). The lateral/directional equations of motion (10.227) then become

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

$$\begin{bmatrix} \dot{v} \\ \dot{p} \\ \dot{r} \\ \dot{\phi} \\ \dot{\Psi} \end{bmatrix} = \begin{bmatrix} Y_v & 0 & -U_e & g\cos\Theta_e & 0 \\ L_v^* & L_p^* & L_r^* & 0 & 0 \\ N_v^* & N_p^* & N_r^* & 0 & 0 \\ 0 & 1 & \tan\Theta_e & 0 & 0 \\ 0 & 0 & \sec\Theta_e & 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ p \\ r \\ \phi \\ \Psi \end{bmatrix}$$

This is in the same form as equation (10.231), and hence the characteristic equation can be obtained from (10.233). If it is assumed that $\Theta_e = 0$ (level trimmed flight), then the characteristic equation may be written in the form:

$$\lambda(\lambda^4 + 4\lambda^3 + B\lambda^2 + C\lambda + D) = 0 \tag{10.237}$$

where $A = -(L_p^* + N_r^* + Y_v)$
$B = [N_v^* U_e + Y_v(L_p^* + N_r^*) + L_p^* N_r^* - L_r^* N_p^*]$
$C = [-gL_v^* + U_e(L_v^* N_p^* - L_p^* N_v^*) + Y_v(L_r^* N_p^* - L_p^* N_r^*)]$
$D = g(L_v^* N_r^* - L_r^* N_v^*)$

Attempting to factorize to find λ analytically is unrealistic. In general, the lateral/directional equations factorize to give three modes, one of which is oscillatory. The factorized characteristic equation takes the form:

$$\lambda(\lambda - \lambda_s)(\lambda - \lambda_r)(\lambda^2 + 2\zeta_{dr}\omega_{n_{dr}}\lambda + \omega_{n_{dr}}^2) = 0 \tag{10.238}$$

Most aircraft exhibit these three classical modes: the spiral mode (s), the roll mode (r), and the Dutch roll mode (dr). These modes are now discussed individually.

The Spiral Mode. This mode can be described as a true banked turn (i.e., without sideslip). It can be stable or unstable and is usually very slow, with time to half or double amplitude typically many seconds. When the spiral mode is stable, the turn has increasing radius, and effectively a heading change occurs. When unstable, the radius decreases and a spiral motion occurs.

Because the spiral mode is effectively a true banked turn, there is little or no sideslip. In attempting to approximate the spiral mode eigenvalue, it is therefore possible to disregard the sideforce equation of motion. Expressing the angle of bank, ϕ, in terms of the turn rate, r (for a true banked turn it can be shown that $r/\phi = g/U_e$), an approximation for the spiral mode eigenvalue is found to be:

$$\lambda_s = -\frac{\dfrac{g}{U_e}(L_v^* N_r^* - N_v^* L_r^*)}{(L_v^* N_p^* - N^* v L_p^*)}$$

The Roll Mode. This stable and relatively fast mode heavily influences the aircraft's roll degree of freedom. Anaytically, only the eigenvector elements associated with p and ϕ are usually of significance. This mode is often modeled as a single-degree of freedom rotation about the x-axis:

$$\dot{p} = L_p^* p$$

Making the substitution $p = p_0 e^{\lambda t}$ yields the common approximation:

$$\lambda_r = L_p^*$$

The Dutch Roll Mode. This motion consists essentially of sideslip, yaw, and rolling motions in combination. As the aircraft sideslips in one direction it yaws in the other, thus maintaining an almost linear flight path. This yawing/sideslipping motion causes the aircraft to roll in the same direction as the yaw. Generally this motion is stable and relatively heavily damped. Occasionally the mode can be unstable, causing serious handling deficiencies.

By using the knowledge that sideslip and yaw mirror one another in this mode (i.e., $\Psi = -\beta$) it is possible to reduce the full lateral/directional characteristic equation to:

$$\lambda^3 - (N_r^* + L_p^*)\lambda^2 + (N_\beta^* + L_p^* N_r^* - N_r^* L_r^*)\lambda + (N_p^* L_\beta^* - L_p^* N_\beta^*) = 0$$

where $L_\beta^* = U_e L_v^*$ and $N_\beta^* = U_e N_v^*$. Note that the roll mode is a factor of this cubic.

10.27 AIRCRAFT RESPONSE TO CONTROLS AND ATMOSPHERIC DISTURBANCES

Response Calculation—A General Approach

From the state space representation of the aircraft dynamics it is possible to derive transfer functions relating output states to controls. The Laplace transform of the state equation:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{10.239}$$

gives the general form of the transfer function:

$$\frac{X(s)}{U(s)} = (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} \tag{10.240}$$

for a state X and control U.

Longitudinal Response to Elevator

For longitudinal response the entries into equation (10.240) are:

$$\mathbf{x} = \begin{bmatrix} u \\ w \\ q \\ \theta \end{bmatrix}, \quad \mathbf{u} = [\delta_e], \quad \mathbf{A} = \begin{bmatrix} X_u & X_w & 0 & -g\cos\Theta_e \\ Z_u & Z_w & Z_q + U_e & -g\sin\Theta_e \\ M_u^* & M_w^* & M_q^* & M_\theta^* \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ Z_{\delta_e} \\ M_{\delta_e}^* \\ 0 \end{bmatrix}$$

which on solution will give the four transfer functions:

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

$$\frac{u(s)}{\delta_e(s)}, \quad \frac{w(s)}{\delta_e(s)}, \quad \frac{q(s)}{\delta_e(s)}, \quad \frac{\theta(s)}{\delta_e(s)}$$

which can be used to obtain the state response to inputs of elevator. These expressions are too cumbersome to present here, but it can be appreciated by consideration of equation (10.235) that, for example, they take the form:

$$\frac{u(s)}{\delta_e(s)} = \frac{f(s)}{s^4 + As^3 + Bs^2 + Cs + D}$$

where f is a polynomial function of s (of order 4 or less) with coefficients dependent on the derivatives X_u . . . $M^*_{\dot{\delta}_e}$. The form of the response becomes apparent on consideration of equation (10.236), which allows us to deduce that:

$$\frac{u(s)}{\delta_e(s)} = \frac{f(s)}{(s^2 + 2\zeta_{ph}\omega_{n_{ph}}s + \omega^2_{n_{ph}})(s^2 + 2\zeta_{sp}\omega_{n_{sp}}s + \omega^2_{n_{sp}})}$$

It is apparent from above that the response will have two components, a short-term response related to the short-period mode and a longer-duration response related to the phugoid mode. The two characteristic motions are superimposed on one another, but because the short-period mode is very fast and heavily damped while the phugoid is of a very much longer period, the effect of the short-period motion disappears quickly, leaving only the effect of the phugoid.

Because the phugoid and short period modes are widely separated in terms of their frequencies, it is possible to decouple them to obtain approximations.

Phugoid Response—An Approximation. Extending the approximate method detailed in Subsection 10.26, it is possible to derive approximate transfer functions which define the low-frequency phugoid response:

$$\frac{u(s)}{\delta_e(s)} = \frac{\dfrac{g}{U_e}Z_{\delta_e} - \dfrac{M^*_{\delta_e}}{M^*_w}\left[S\left(X_u - \dfrac{g}{U_e}\right) + \dfrac{g}{U_e}\right]}{s^2 + \left[-X_u - \dfrac{M^*_u}{M^*_w}\left(\dfrac{g}{U_e} - X_w\right)\right]s + \dfrac{g}{U_e}\left(-Z_u + Z_w\dfrac{M^*_u}{M^*_w}\right)}$$

$$\frac{w(s)}{\delta_e(s)} = \frac{-\dfrac{g}{U_e}\dfrac{M^*_u}{M^*_w}Z_{\delta_e} - \dfrac{M^*_{\delta_e}}{M^*_w}\left[s^2 - sX_u - \dfrac{g}{U_e}Z_u\right]}{s^2 + \left[-X_u - \dfrac{M^*_u}{M^*_w}\left(\dfrac{g}{U_e} - X_w\right)\right]s + \dfrac{g}{U_e}\left(-Z_u + Z_w\dfrac{M^*_u}{M^*_w}\right)}$$

$$\frac{q(s)}{\delta_e(s)} = \frac{-\dfrac{Z_{\delta_e}}{U_e}\left(s - X_u + \dfrac{M^*_u}{M^*_w}X_w\right) - \dfrac{M^*_{\delta_e}}{U_eM^*_w}[s^2 - (X_u + Z_w)s - X_wZ_u + X_uZ_w]}{s^2 + \left[-X_u - \dfrac{M^*_u}{M^*_w}\left(\dfrac{g}{U_e} - X_w\right)\right]s + \dfrac{g}{U_e}\left(-Z_u + Z_w\dfrac{M^*_u}{M^*_w}\right)}$$

Short-Period Response—An Approximation. Extending the approximate method detailed in Subsection 10.26, it is possible to derive approximate transfer functions which define the higher-frequency short-period response:

$$\frac{w(s)}{\delta_e(s)} = \frac{(s - M_q^*)Z_{\delta_e} + U_e M_{\delta_e}^*}{s^2 - (Z_w + M_q^*)s + (Z_w M_q^* - U_e M_w^*)}$$

and

$$\frac{q(s)}{\delta_e(s)} = \frac{(s - Z_w)M_{\delta_e}^* + M_w^* Z_{\delta_e}}{s^2 - (Z_w + M_q^*)s + (Z_w M_q^* - U_e M_w^*)}$$

Recall that the assumption is that there is no change in forward velocity during this mode ($u = 0$) and as $q = \dot{\theta}$, $q(s) = s\,\theta(s)$.

Lateral/Directional Response

For lateral/directional response the entries into equation (10.240) are:

$$\mathbf{x} = \begin{bmatrix} v \\ p \\ r \\ \phi \\ \Psi \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} \delta_r \\ \delta_a \end{bmatrix},$$

$$\mathbf{A} = \begin{bmatrix} Y_v & 0 & -U_e & g\cos\Theta_e & 0 \\ L_v^* & L_p^* & L_r^* & 0 & 0 \\ N_v^* & N_p^* & N_r^* & 0 & 0 \\ 0 & 1 & \tan\Theta_e & 0 & 0 \\ 0 & 0 & \sec\Theta_e & 0 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} Y_{\delta_r} & 0 \\ L_{\delta_r}^* & L_{\delta_a}^* \\ N_{\delta_r}^* & N_{\delta_a}^* \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

which on solution will give 10 transfer functions:

$$\frac{v(s)}{\delta_r(s)}, \ldots, \frac{\psi(s)}{\delta_a(s}$$

which can be used to obtain the state response to inputs of rudder and aileron. These expressions are too cumbersome to present here, but it can be appreciated by consideration of equation (10.237) that, for example, they take the form:

$$\frac{v(s)}{\delta_r(s)} = \frac{f(s)}{s(s^4 + As^3 + Bs^2 + Cs + D)}$$

where f is a polynomial function of s (of order 5 or less) with coefficients dependent on the derivatives $Y_v \ldots N_{\delta_a}^*$. The form of the response becomes apparent on consideration of equation (10.238), which allows us to deduce that:

$$\frac{u(s)}{\delta_e(s)} = \frac{f(s)}{s(s - \lambda_s)(s - \lambda_r)(s^2 + 2\zeta_{dr}\omega_{n_{dr}}s + \omega_{n_{dr}}^2) = 0}$$

Again we can see that the response of the aircraft will be made up of the component modes spiral, roll, and Dutch roll.

Roll Response to Aileron. As mentioned in Subsection 10.26, the roll motion of an aircraft can often be treated as a single-degree of freedom system. The rolling equation of motion becomes

$$\dot{p} = L_p^* p + L_{\delta_a}^* \delta_a$$

from which the roll rate can be estimated from:

$$p(t) = \frac{L_{\delta_a}^*}{L_p^*}(1 - e^{-t/\tau_r})$$

where

$$\tau_r = -\frac{1}{L_p^*} = \text{the roll time constant}$$

Dutch Roll Response to Rudder. As in Subsection 10.26, making the assumption that the aircraft center of gravity follows a straight line in a Dutch roll motion, it is possible to ignore the y-equation of motion and derive the following approximate transfer functions relating roll rate, p, and sideslip angle, β, to rudder deflection, δ_r:

$$\frac{p(s)}{\delta_r(s)} = \frac{(s^2 - N_r^* s + N_\beta^*)L_{\delta_r}^* - N_{\delta_r}^*(L_\beta^* - sL_r^*)}{s^3 - (N_r^* + L_p^*)s^2 + (N_r^* L_p^* - L_r^* N_p^* + N_\beta^*)s + (N_p^* L_\beta^* - L_p^* N_\beta^*)}$$

$$\frac{\beta(s)}{\delta_r(s)} = \frac{-N_p^* L_{\delta_r}^* - N_{\delta_r}^*(s - L_p^*)}{s^3 - (N_r^* + L_p^*)s^2 + (N_r^* L_p^* - L_r^* N_p^* + N_\beta^*)s + (N_p^* L_\beta^* - L_p^* N_\beta^*)}$$

Response of Aircraft to Atmospheric Disturbances

Consider an aircraft suddenly immersed completely in a gust of constant velocity V_g with components $-u_g$ and $-w_g$ (the negative sign simply denotes that the gust is a headwind with an upward component). (See Figure 10.93.) We can write the equations of motion in the form

$$\begin{bmatrix} \dot{u} \\ \dot{w} \\ \dot{q} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} X_u & X_w & 0 & -g\cos\theta_e \\ Z_u & Z_w & u_e & -g\sin\theta_e \\ M_u^* & M_w^* & M_q^* & M_\theta^* \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} u - u_g \\ w - w_g \\ q \\ \theta \end{bmatrix} + \begin{bmatrix} 0 \\ Z_{\delta_e} \\ M_{\delta_e}^* \\ 0 \end{bmatrix} [\delta_e]$$

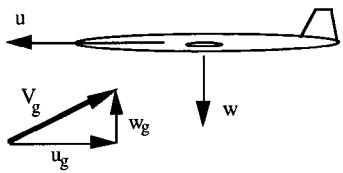


FIGURE 10.93 Aircraft acted upon by a gust.

or

$$\begin{bmatrix} \dot{u} \\ \dot{w} \\ \dot{q} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} X_u & X_w & 0 & -g\cos\theta_e \\ Z_u & Z_w & u_e & -g\sin\theta_e \\ M_u^* & M_w^* & M_q^* & M_\theta^* \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} u \\ w \\ q \\ \theta \end{bmatrix} + \begin{bmatrix} 0 \\ Z_{\delta_e} \\ M_{\delta_e}^* \\ 0 \end{bmatrix} [\delta_e] + \begin{bmatrix} -X_u & -X_w \\ -Z_u & -Z_w \\ -M_u^* & -M_w^* \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_g \\ w_g \end{bmatrix}$$

which may be written as

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{E}V_g$$

where

$$V_g = \begin{bmatrix} u_g \\ w_g \end{bmatrix} = \text{gust vector}$$

and

$$\mathbf{E} = \begin{bmatrix} -X_u & -X_w \\ -Z_u & -Z_w \\ -M_u^* & -M_w^* \\ 0 & 0 \end{bmatrix} = \text{gust influence matrix}$$

For a sidewind of velocity $-v_g$ (i.e., from the starboard direction) we would have

$$\begin{bmatrix} \dot{v} \\ \dot{p} \\ \dot{r} \\ \dot{\phi} \\ \dot{\Psi} \end{bmatrix} = \begin{bmatrix} Y_v & 0 & -u_e & g\cos\theta_e & 0 \\ L_v^* & L_p^* & L_r^* & 0 & 0 \\ N_v^* & N_p^* & N_r^* & 0 & 0 \\ 0 & 1 & \tan\theta_e & 0 & 0 \\ 0 & 0 & \sec\theta_e & 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ p \\ r \\ \phi \\ \Psi \end{bmatrix} + \begin{bmatrix} Y_{\delta_r} & 0 \\ L_{\delta_r}^* & L_{\delta_a}^* \\ N_{\delta_r}^* & N_{\delta_a}^* \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \delta_r \\ \delta_a \end{bmatrix} + \begin{bmatrix} -Y_v \\ -L_v^* \\ -N_v^* \\ 0 \\ 0 \end{bmatrix} [v_g]$$

Note that for a large aircraft the distributed effect of the gust might impose rotations on the aircraft as well as translational velocities such that the V_g vector may include terms such as p_g and q_g.

FURTHER READING

Cook, M. V., *Flight Dynamics Principles,* Edward Arnold, London (1997).

Etkin, B., and Reid, L. D., *Dynamics of Flight Stability and Control,* 3d ed., John Wiley & Sons, New York (1996).

AERODYNAMICS, PERFORMANCE AND STABILITY AND CONTROL

Hancock, G. J., *An Introduction to the Flight Dynamics of Rigid Aeroplanes,* Ellis Horwood, Chichester (1995).

McCormick, B., W., *Aerodynamics, Aeronautics and Flight Mechanics,* John Wiley & Sons, New York (1979).

McLean, D. *Automatic Flight Control Systems,* Prentice Hall, London (1990).

McRuer, D., Ashkenas, I., and Graham, D. *Aircraft Dynamics and Automatic Control,* Princeton University Press, Princeton (1973).

Pallett, E. H. J., and Coyle, S. *Automatic Flight Control,* 4th ed., Blackwell Science, Oxford (1973).

Rolfe, K. M., and Staples, K. J. *Flight Simulation,* Cambridge University Press, Cambridge (1986).

SECTION 11
AVIONICS AND ASTRIONICS

Section Editor: Marc Pelegrin

PART 1

THE ELECTROMAGNETIC SPECTRUM

Florent Christophe

Radio waves propagate in the vacuum or through the complex media surrounding the Earth—or other planets—and carry information from the source to the receiver. We will focus in this subsection on radio waves involving man-made transmitters for avionics or astrionics applications—i.e., communications, navigation, surveillance, and radiolocation—but the case of the transmitter as a natural source may also be considered for radioastronomy or Earth observation from space.

11.1 RADIOWAVES IN A VACUUM

Radio waves, inferred from Maxwell equations, first observed by Hertz in 1886, and applied for long range by Marconi in 1906, are a combination of an electric field E and a magnetic field H of periodic time variations, produced by electric charge displacements or electric currents.

In a vacuum, far enough from those electric sources, E and H fields appear as plane waves following expressions:

$$E(r,t) = E_0 \exp(\omega t - k \cdot r) \tag{11.1}$$

$$H(r,t) = H_0 \exp(\omega\, t - k \cdot r) \tag{11.2}$$

where E_0 and H_0 are orthogonal vectors, their modules being linked by

$$E_0/H_0 = 120\pi \tag{11.3}$$

The power density transported by such waves is given by:

$$D = E_0 H_0 \quad \text{or} \quad D = E_0^2/120\pi \quad \text{or} \quad D = 120\pi H_0^2 \tag{11.4}$$

The direction of vector E_0 defines the linear polarization of the wave; ω is its frequency, and k is the wave vector indicating the direction of wave propagation, orthogonal to both E_0 and H_0. They are linked through

$$k = c/\omega \tag{11.5}$$

where c is the velocity of light in a vacuum, equal to $3 \cdot 10^8$ m/s. Vector r defines the point in space where the fields are considered. The time period and frequency can be easily derived

$$T = 2\pi/\omega \qquad f = 1/T = \omega/2\pi \tag{11.6}$$

as well as the spatial period or wavelength:

TABLE 11.1 From Extremely Low to Extremely High Frequencies

Frequency (MHz)	3.10^{-4}	3.10^{-3}	0.03	0.3	3	30	300	3.10^{3}	3.10^{4}
Wavelength (m)	10^{6}	10^{5}	10^{4}	10^{3}	100	10	1	0.1	0.01
Denomination	ELF	VLF	LF	MF	HF	VHF	UHF	SHF	EHF

$$\lambda = 2\pi/k = c/f \tag{11.7}$$

Table 11.1 illustrates equation (11.7) and indicates the usual denomination of frequency bands, following a logarithmic scale.

Beyond the upper EHF limit (300 GHz), atmospheric attenuation (discussed below) makes most terrestrial applications impractical up to about 10 THz (30 μm wavelength), where the far infrared region of optics begins. The low frequencies (LF band and below) are restricted to submarine communications, where they are useful for their ability to penetrate saltwater despite their limited bandwidth and impractical wavelength-sized antennas.

Most avionics or astrionics systems make use of wavelengths from a meter to a centimeter, i.e., VHF to SHF bands, but the selection of a frequency band for a given application is governed by bandwidth requirements and depends on antenna and wave propagation, which will be discussed now.

11.2 ANTENNAS AND POWER BUDGET OF A RADIO LINK

An antenna is a transducer that transforms an electric current from a transmission line into a radio wave (on transmit) or a radio wave into an electric current (on receive). This transformation should be done with maximum efficiency, but attention must be paid to the direction and polarization of the radiated wave.

We will first introduce wire antennas, a classical example of which is the half-wavelength dipole. It has been shown that most of the power coming from a transmission line (like a coaxial cable often used at UHF or a twin wire line for VHF and below) is radiated when it is connected to two thin collinear wires at a resonance frequency for which each wire has a length close to a quarter of the wavelength. The modulus of the electric field radiated by such an antenna is maximal in a plane orthogonal to the wires, without angular dependency in this plane, the electric field being parallel to the wires. The power efficiency of such a transducer being fair only very close to the resonance frequency, practical implementations where relative bandwidths of at least a few percent are required are based on thick dipoles.

For another classical transmission line at SHF and EHF such as the rectangular waveguide, the antenna has to adapt the guided wave progressively to a free space propagating wave (or vice versa) without major discontinuities that would make detrimental wave reflections. A natural solution is to widen the section of the waveguide progressively, making a pyramidal horn. In its aperture (the base of the pyramid) the wave is weakly coupled to the walls and thus ready to be launched to free space. Such a wave will have as its preferred beam direction the axis of the

pyramid. The width of this beam in each plane (expressed in radians) may be related to the aperture of the horn antenna by approximate formulas:

$$\delta\phi = \lambda/a, \quad \delta\theta = \lambda/b \tag{11.8}$$

where a and b are the sides of the rectangular radiating aperture. For other aperture antennas, such as reflector antennas, for which a parabolic reflector may be illuminated by a small horn (or primary feed) located near its focus, the same formulas apply, a and b now being the dimensions of the reflector itself.

If we now consider how the radiated power P is distributed around the antenna, we may write the power density D at distance R from the antenna as if the antenna were isotropic, times a power-concentrating factor, g, i.e.:

$$D = Pg/4\pi R^2 \tag{11.9}$$

This power-concentrating factor g is known as the antenna gain* and is given by the approximate formula:

$$g = 4\pi S/\lambda^2 \tag{11.10}$$

where S is the surface of the aperture (i.e., $a \cdot b$), therefore:

$$g = 4\pi/\delta\phi\delta\theta \tag{11.11}$$

From the power density, the electric and magnetic fields at distance R may be derived making use of formulas (11.4) or directly from the power available at the output of a second antenna of section S' (and gain g' given by equation (11.10)) as $P' = S'D$, or:

$$P' = Pgg'(\lambda/4\pi R)^2 \tag{11.12}$$

This formula governing the power transfer from one antenna to another is known as the radio-communication equation. Compared to the equivalent thermal noise at the input of the receiver, it allows the signal-to-noise ratio, i.e., the quality of the radio link, to be inferred.

Further information concerning antennas† may be found in Rudge et al. (1982, 1983).

11.3 RADIOWAVE PROPAGATION IN THE TERRESTRIAL ENVIRONMENT

Depending on the wavelength, many effects are likely to occur when a radio wave interacts with the Earth's surface or with the atmosphere. These effects, which are

*The notion of antenna gain, which was introduced with regard to aperture radiating antennas, also applies to wire antennas. The gain g of a wire antenna is limited to a few units (1.5 for a half-wavelength dipole), whereas it exceeds 1,000 for a 10 $\lambda \times 10\ \lambda$ aperture.

†The pencil beam is always associated with *side lobes* (either *near lobes*, far lobes, or even *back lobes*, with reference to their angular distance to the main beam), which may produce detrimental coupling effects. Parasitic side lobes might be reduced through various techniques, but at the expense of widening the main beam and reducing the gain. Those techniques, which include windowing, are similar to frequency-filtering techniques and lead to the concept of angular filtering.

presented below, are summarized in Table 11.2. More detailed information is available in the ITU-R Handbooks (see, e.g., ITU 1991, 1996, 1997).

Interactions with the Earth's Surface

If we first consider the ground as an homogeneous flat medium, its electrical properties (dielectric constant and conductivity) may be reduced to a refractive index ranging from 2 to 9, depending on frequency, soil composition, and moisture content. Applying Fresnel's laws of optics to an electromagnetic wave incident upon the plane interface demonstrates a specularly reflected wave and a refracted wave below the surface.*

The reflected wave may produce an interference with the directly transmitted one, a classical situation being grazing angle at the interface where both waves with similar elevation angles cannot be separated by the antenna beam, whereas the reflection coefficient is close to –1. The resulting interference pattern exhibits at range R of a transmitter h_t above ground successive nulls at height h_n, given by:

$$h_n = R\lambda(1 + 2n)/4h_t \tag{11.13}$$

Earth Curvature and Relief Effects. Between a transmitter and a receiver close to the Earth's surface the direct wave may be obstructed either by obstacles either coming from the local relief or by the curvature of the Earth. The limiting situation in this last case is the horizon plane tangent to the Earth's sphere, which a receiver at height h_r crosses at range R_h for a transmitter at height h_t, r_e being the Earth's radius:†

$$R_h = \sqrt{2r_e h_t} + \sqrt{2r_e h_r} \tag{11.14}$$

Such an obstruction does not occur at a sharp cutoff when crossing the horizon plane, due to the diffraction effect. For radio waves at wavelengths comparable to the radius of curvature of the obstacle, this diffraction may compensate most of the shadowing effect. Electromagnetic waves at MF and below are therefore suitable for propagating hundreds or thousands of kilometers near the Earth's surface.§

TABLE 11.2 Summary of the Interactions of a Radio Wave with Earth's Environment: Reduced (−), Limited (+), or Strong (++)

Wavelength	>10 m	10 m–10 cm	<10 cm
Ground effects	+	+	++
Tropospheric effects	−	−	+
Ionospheric effects	++	+	−

* This refracted wave usually vanishes rapidly due to the finite conductivity of the soil

† In fact, the refraction of radio waves propagating quasihorizontally in the stratified lower troposphere bends the ray downwards. In standard troposphere, this bending is usually compensated for by modifying in the horizon formula the real earth radius of 6,400 km into a value of 8,500 km.

§ For completeness, a creeping wave guided near the interface must also be considered in such conditions.

Soil Roughness Effects. Small-scale random features on the surface (either of natural origin or produced by agriculture on bare soils, but vegetation can be also accounted for, as well as wind-driven waves on the sea surface or swell) disturb the previously mentioned specular reflection, reducing the corresponding wave and creating more and more diffuse scattering as the ratio of the typical size of the irregularities to the wavelength is augmented.*

Interactions with the Troposphere

In this lower part of the atmosphere, below 12 km, gaseous molecules of sufficient density, and hydrometeors (the liquid water or ice particles inside clouds, rain, snow, hail, etc.) may interact with radio waves. Molecular absorption is due to the exchange of energy between the wave and quantified levels of vibrations or rotations and appears as absorption lines. In the upper SHF and EHF bands, water vapor creates such an absorption line around 22 GHz and a rapidly increasing continuum beyond 50 GHz, whereas oxygen creates a strong absorbing continuum around 58 GHz. Those effects limit long-range surface-based applications to around 45 GHz, but radiometers take advantage of such molecular interactions for remote sensing of the atmosphere.

The major interaction with hydrometeors may be described as Rayleigh scattering from individual particles that are small with respect to the wavelength; the scattering cross-section (ratio of scattered power to incident power density) is then written for spherical water droplets of diameter d as:

$$\sigma = 190 d^6/\lambda^4 \tag{11.15}$$

This relationship shows a strong increase with frequency and drop size. In a given volume of interaction where droplets of different sizes (larger than 0.1 mm in diameter, up to 7 mm for thunderstorms) are falling at different velocities, their distribution may be derived from the rain rate observed at ground level. This droplet distribution may then be used for estimating the overall scattering cross-section per volume unit, which happens to be the attenuation coefficient per length unit.

A further effect at EHF due to interaction with the lower troposphere is scintillation (rapid fluctuations of amplitude and phase) due to the transit of the wave through the time-varying heterogeneities of refractive index caused by turbulence.

Interactions with the Ionosphere

The ionosphere is mostly created, at altitudes higher than 80 km, by UV radiation and cosmic rays from the Sun, ionizing the low-density air molecules. The electron density exhibits a strong dependency with solar flare activity (varying on an 11-year cycle basis), hour of the day and season, latitude (the auroral oval is the place where charged cosmic rays penetrate the high atmosphere), and altitude. The maximum electron density is observed around 350 km and varies from $N = 3 \cdot 10^{10}$ m^{-3} at night for a low-activity Sun to more than $N = 10^{12}$ m^{-3} at noon for an active Sun.

*Part of this energy is backscattered towards the transmitter, creating surface clutter for radar. See Part 5.

The propagation of a radio wave in such cold plasma, neglecting collisions, is mainly ruled by the equation giving the refractive index:

$$n(f) = 1 - \tfrac{1}{2}(f_p/f)^2 \tag{11.16}$$

for a frequency f which is large with respect to the plasma frequency f_p, given by (for N in m^{-3} and f_p in Hz):

$$f_p = 9\sqrt{N} \tag{11.17}$$

The maximum value for f_p is about 10 MHz. An electromagnetic wave at a frequency below or close to this value would be refracted downwards before reaching the region of the maximum ionization; this effect is used for establishing over-the-horizon radio links in the HF band. Waves at frequencies beyond that value are able to transit through the whole ionosphere, with effects mostly due to an augmentation of the time delay. With respect to propagation in a vacuum, the path is apparently augmented by:

$$\Delta R = 40.5\ \mathrm{TEC}/f^2 \tag{11.18}$$

TEC (total electron content) is the integral of the electron density along the path. Maximum values of 10^{18} m^{-2} may be encountered for a slanting path through the whole ionosphere, giving an augmented path of 450 m at 300 MHz down to 4.5 m at 3 GHz. Such ionospheric delays cause the major errors in the budget of satellite navigation systems, most of them being corrected in dual-frequency systems (see also Part 3, section 11). Additional effects for transionospheric systems operating at UHF and SHF comes from scintillation caused by the crossing of heterogeneities.*

11.4 MANAGEMENT OF THE ELECTROMAGNETIC SPECTRUM

To avoid interference between systems, international regulations have been agreed upon for allocating frequency bands to a single application with possible secondary applications (ITV). World Radiocommunications Conferences update these allocations every three years, following technical discussions under the auspices of the International Telecommunications Union. National authorities must negotiate with their public or private users the best ways to match the limited spectral resource to an increasing demand.

Among solutions for overcoming such frequency congestion bottlenecks, the design of adaptive systems more robust against interference, as well as better prediction of the effects of interference, would allow the overall number of users in the same frequency band to be increased. Also considered, despite the adverse propagation conditions in the troposphere, is the EHF band, which has large available bandwidths.

* Anisotropy of the ionospheric plasma must also be considered due to the magnetic field of the Earth, which principally affects the polarization of transionospheric electromagnetic waves.

REFERENCES

International Telecommunication Union (ITU). Radio Regulations, ITU, Geneva.

International Telecommunication Union (ITU). 1991. *Curves for Radiowave Propagation over the Surface of the Earth*, ITU, Geneva.

International Telecommunication Union (ITU). 1996. *Radiometeorology*, ITU, Geneva.

International Telecommunication Union (ITU). 1997. *Ionosphere and Its Effects on Terrestrial and Earth-to-Space Radiowave Propagation from VLF to SHF*, ITU, Geneva.

Rudge, A. W., Milne, K. Olver, A. D., and Knight, P. 1982, 1983. *The Handbook of Antenna Design*, vol. 1 (1982), vol. 2 (1983), Peregrinus, London.

PART 2

THE SPACECRAFT ENVIRONMENT

D. Boscher

11.5 INTRODUCTION TO THE SPACE ENVIRONMENT

The natural space environment in the solar system has several components (Hargreaves 1992; Suess and Tsurutani 1998):

1. Meteoroids, created in the early solar system. They reach the vicinity of the Sun with nearly a constant flux that is, however, significantly intensified during known meteor showers (e.g., Leonids).
2. Cosmic rays, charged particles with high energy. These originate from both inside and outside the galaxy. Moreover, so-called anomalous cosmic rays are produced by the interaction between the interstellar neutrals and the heliosphere.
3. Solar energetic particles escaping sporadically from the Sun. Nearly all elements are found but with an abundance and charge state different in cosmic rays.
4. Neutrons ejected by the Sun. These neutrons decay rapidly in the interplanetary medium and only a few of them remain in the Earth vicinity (but they can be much more abundant near the Sun).
5. Photons, from the γ rays to the radio frequency, coming essentially from the Sun. The maximum of the spectrum is in the visible, but the UV rays have a great influence on the near-ground environment of planets and moons.
6. The solar wind, consisting of plasma and magnetic field, originating in the Sun. At low heliospheric latitudes its density is high but its mean velocity is about 400 km/s with sporadic effects of coronal mass ejections; at high latitudes, as it comes from coronal holes, its velocity is higher (800 km/s) but its density is lower.

Moreover, in the vicinity of planets and moons, there are other components:

7. Dust, created by the interaction of meteoroids and small moons. Usually this dust is attracted by the moon that created it, but it can form rings like those around Saturn or Jupiter.
8. Neutral atoms and molecules forming atmospheres, the size and composition of which depend on the size of the planet and its proximity to the Sun.
9. Interaction between cosmic rays/solar energetic particles and the atmosphere creates many secondaries, including pions, muons, and neutrons. Some of these neutrons can escape the planet and decay in its vicinity, producing protons and electrons.
10. Reflection of the solar spectrum by planets (ground + atmosphere) and moons is known as the albedo. Moreover, as the planets (and moons) are heated by the Sun, they emit in the infrared part of the spectrum.

11. The interaction between UV rays and atmosphere-created ionospheres, which are plasmas made of ions of planetary origin and electrons. The temperature is very low as compared to the solar wind, for example.
12. The planets and moons can have a magnetosphere, which is formed by the interaction of the solar wind with the quasi-dipolar internal magnetic field of the body. The magnetosphere produces radiation belts, energetic particles trapped by the magnetic field, whose origins are the solar wind or the neutrons escaping from the atmospheres. The magnetosphere modifies also the electromagnetic spectrum of the waves in the radio frequency range.

Finally, human beings are responsible for two more components of the space environment:

13. An electromagnetic environment. Man-made VLF transmitters have interacted with particles in the radiation belt since before space flight.
14. Debris from satellites and launchers, an increasing space environment problem around the Earth.

Particular attention must be paid to magnetospheres because they imply special processes which lead to special effects. Magnetospheres are created by the interaction of the solar wind with the quasi-dipolar magnetic field of the body. This interaction leads to different regions (Kivelson and Russell 1995):

1. The bow shock, where the supersonic and superalfvenic solar wind velocity is decreased by collisionless interaction
2. The magnetosheath, a region with high turbulence, for velocity as well as for the magnetic field
3. The magnetopause, where the solar wind pressure is counterbalanced by the magnetic pressure due to the internal magnetic field
4. The magnetotail, in the anti-solar direction, where the plasma sheet separates it into two lobes
5. The internal magnetosphere, with the cusps surrounded by the auroral ovals, and the radiation belt region

In these different regions the amplitude and direction of the magnetic field can be different, and the plasma density and temperature can be as well.

11.6 EFFECTS OF THE SPACE ENVIRONMENT ON SPACECRAFT, LAUNCHERS AND AIRPLANES

Most of the components listed above have an impact on satellites and launchers, depending on the energy of the component. Low-energy components interacting with the surface of the satellites (thermal coatings, solar cells, antennas, optical systems) are:

- Photons, from UV to radio
- The solar wind
- The atmosphere and its different atoms and molecules (oxygen atoms are very corrosive due to chemical reactions)

- The ionosphere
- Plasma in the magnetosphere
- Light dust, micrometeoroids, and microdebris

The interaction depends on the location of the satellite/launcher in space. Low-altitude satellites and launchers are impacted by all these components (except solar wind), while interplanetary missions can only interact with meteoroids, photons, solar wind, and eventually plasmas in the planetary magnetotails.

As the energy increases, particles and photons can penetrate the satellite body. In this category are:

- Heavy dust, light meteoroids, and debris (millimetric)
- Energetic particles from cosmic ray, solar, or radiation belt origin, including neutrons
- High-energy photons, from γ-rays to X-rays

These may affect astronauts, particularly during extra-vehicular activities (EVAs). The worst are meteoroids and heavy debris, whose impact can definitely damage the satellite, launcher, or human, though the probability of impact is very low (even so, the French satellite CERISE suffered such a fatal impact).

As far as effects on satellites are concerned, two different approaches must be considered. The first one deals with the determination of the lifetime of the satellite. Long-term effects must be well modeled and predicted. They are related to statistical properties of the space environment. Total dose effects and aging of the satellite belong in this category, but sporadic effects must also be considered by their mean properties. For such an approach, models already exist, though they generally have to be improved, for example the radiation belt environment model. The second approach deals with satellite operations and requires a better knowledge of the real-time environment (the space weather). Some components of the space environment vary greatly and rapidly, in particular the plasma and the energetic particle flux. Major events can endanger the life of the satellite, or at least require observation by satellite operators or on-board computers. Worst cases depend strongly on the architecture of the satellite, and it is difficult to quantify the impact of a worst-case event on a particular satellite subsystem. We will review the most important effects of the space environment on different satellite subsystems, distinguishing between long-term effects and short-term (space weather) ones.

Neutrals from the atmosphere have two effects on spacecraft. They can induce drag on spacecraft, and they can induce chemical reactions, as there are oxidizing molecules and atoms, in particular atomic oxygen, but also NO, NO_3, O_3, HO, etc. The drag decreases the apogee of a spacecraft and consequently also the perigee. Maneuvers are needed to get the spacecraft back in its nominal orbit. The drag is also important in the calculation of the entry of planetary probes or reentry of large vehicles. It is also an important effect to take into account in the opening of the launcher-protecting shroud. It is a space weather effect as the heating of high-altitude neutrals is modified by UV rays which are Sun-flare dependent, and the atmosphere can be modified in hours. As for oxidization, it is rather a cumulative effect on satellites. This induces erosion of the surface but also contamination of the near-satellite environment, and the glowing of that environment by interaction with ambient plasmas.

Photons from the sun have a spectral distribution similar to that of a black body at a temperature $T \approx 5600$ K. The maximum of the emission is in the visible range and the mean total flux is around 1370 W/m^2. The first effect of photons emitted

by the Sun is the heating of illuminated surfaces of the satellites. Because unlit surfaces are not heated, this effect produces temperature gradients in the satellite, which are controlled using highly reflective thermal coatings. Thermal control of spacecraft is a major problem in space, but as the solar flux is nearly constant, it is not affected much by space weather.

At the two edges of the spectrum (UV and radio) the photon flux is more variable and is related to solar flares. Because the UV photons are more energetic, they interact with the spacecraft surface materials (thermal control coatings as well as antenna and optics). They principally induce colored centers and chemical reactions which modify the reflectivity of those materials and therefore the thermal properties of the satellite. It is a long-term effect which is important to take into account for the aging of satellites, but apart from a solar cycle influence it is not important for space weather applications.

Photons can also increase the background noise of photon detectors with multiple diffusions. It is a space weather effect only for the variable parts of the Sun spectrum (γ-rays, X-rays, UVs, radio). Radio waves emitted from the Sun can also perturb the transmissions between the spacecraft and the ground (in particular for satellites at the Lagrangian L1 point and interplanetary probes).

Plasmas from solar wind or ionospheric origin are composed of ions and electrons with temperatures up to tens of keV. With their relatively low energy, these particles cannot penetrate the satellite more than a few microns. They produce surface effects like those produced by UV light (creation of colored centers and chemical reactions), but as they are charged they also induce charging effects, specially on the dielectrics or insulated conductors. Moreover, as they interact with different materials, differential potentials are set up and discharges may appear. Charging is always present on a satellite, but normally the potential levels are not high enough to create discharges (apart from high-insulating materials). The potentials are higher in the wake of low-altitude satellites, during periods of precipitating electrons, or near geosynchronous orbit in the midnight region, during substorm injection. As low-altitude satellites fly in the ionosphere they produce a wake at the trailing edge.

Energetic particles produce numerous effects. They induce:

- Aging of electronics, optics, and materials
- Single-event effects
- Internal charging and electrostatic breakdown and discharges
- Background of electronics and Cerenkov effects

When energetic ions interact with sensitive electronic components, a modification of electronic characteristics is recorded. This depends on the total dose and the dose rate, but it is rather a weak process. It is very important to take this aging into account for the lifetime of the satellite, but this effect is not a space weather one, though it depends on the solar cycle.

The energetic ions can also produce single-event effects (SEEs) in electronics (single hard errors, single-event upsets, latchups, burnouts, gate and dielectric ruptures). These effects are normally due to heavy ions from galactic or solar cosmic rays, but particles as light as protons or neutrons can produce the same effects because they generate secondary heavy ions through nuclear reactions with the silicon inside the electronics (in the future, due to microelectronics evolution, protons will be able to directly induce SEEs). These effects, because they are sporadic, are of major concern for space weather. Nevertheless, most of these effects are diminished when hardened or tolerant electronics and redundant and rewriting com-

puters are used. Some of them are permanent, either directly (e.g., latchup, burnout) or in the case of modified allocation tables. Launchers (and even aircraft) are also affected by SEEs as solar proton events create neutrons at low altitudes and can directly interact at higher altitudes, especially for GTO and SSO launches.

The same kind of effects are induced in human beings. Dose effects affect all cells, especially those which are not renewed or rapidly renewed. Single-energetic particles can also break the DNA chain in the cell nucleus, producing chromosome aberrations, translocations, and tumor induction. They can also induce cell mutation, which can have effects on the genetics. It is essential to know the real-time space weather during EVAs and for future planetary missions.

Energetic ions and electrons also increase locally dark currents in detectors. This effect is clearly visible from imagers on board Earth or Sun observatories. These energetic ions and electrons also produce atom displacements in solar cells, decreasing their output powers. Energetic electrons, as they penetrate inside the spacecraft, produce internal charging and electrostatic discharges. The effect of discharges can be direct through line destruction but usually creates electromagnetic pulses that produce signals understood as false commands by on-board computers. Apart from dose effects, all the energetic particles produce space weather effects, either by single particles (the probability of interaction then depends on the flux) or by cumulative effects (the case of charging, for example).

Debris or meteoroids, as they impact parts of the satellite, can induce destructive effects on it. The effect depends on their mass. Light particles induce cumulative effects, but as the weight increases effects are less probable but more dangerous. Heavy particles, by momentum transfer, can also induce problems in attitude control. In any case, the impacts produce secondaries which can in turn impact other parts of the satellite. Space weather in this field is concerned only with meteoroid clouds like Leonids, and perhaps the differential orbit of debris relative to satellite when heating of the atmosphere is recorded. For spacecraft control, it is important to track big debris. This is done through radar in the United States.

In fact, all these components are never alone in the space environment, and synergistic effects appear which are not necessarily worse than the effects taken separately. This synergistic effect is well known with regard to materials in low-earth orbit satellites. The combination of UV, particles, and atomic oxygen gives less degradation than applied separately, as oxidation and sputtering strip the surface, which then is clean and pure. But other synergistic effects are worse, such as debris and plasmas. Debris, by its impact on the surface, increases its roughness and produces sharpened edges, which helps the discharges to be triggered (plasmas can also be ejected during the impact, thus increasing the discharge probability). In this case, even if debris effects are not directly space weather effects, the combined effects are.

There is also another component of the space environment, the magnetic fields, but these cannot produce direct effects on satellites unless they are field spin stabilized. Nevertheless, these fields have indirect effects as they modify the densities of charged particles and plasmas.

We can sum up the effects of the space environment in Table 11.3. In this table, direct effects which are concerned by Space Weather studies are set in bold type.

11.7 THE ENERGETIC CHARGED PARTICLE ENVIRONMENT

Most of these parts of the environment are normally taken into account in the conception of a spacecraft, using for example highly reflective materials for thermal

TABLE 11.3 The Components of the Space Environment and Their Effects on Spacecraft and Their Parts

Components of the space environment	Effects, **space weather effects**	Direct problems	Induced problems
Neutrals	**Drag**	**Orbit control**	
	Oxidization	Temperature control	
		Erosion	Contamination, glow
Photons	Heating	Temperature control	Contamination
	Surface aging	Temperature control	
	Background noise increase	Signal/noise ratio	
Plasmas	Surface aging	Temperature control	
	Surface charging	**ESDs**	Contamination
Particles	Aging	Temperature control	
		Electronic characteristics changes	
	Atom displacements	**Solar cell power decrease**	
		Electronic problems	
	Background noise increase	**Detector background**	
	Internal charging	**ESDs**	
	SEEs	**Electronic problems**	
	Human effects	**DNA/cell damage**	Genetics, cancer
		Cell destruction	Death
Debris, meteoroids	**Impact**	Partial destruction	
		Attitude control problem	
		Erosion	ESDs, Contamination

Space weather effects are set in bold type.

control, or potential continuity for differential charging avoidance. Nevertheless, in the first case the aging of the materials (due to the combined effect of UV and particles) leads to thermal problems for long-duration missions, and in the second case, it will be impossible or at least extremely difficult to avoid this problem in the solar cells. In any case, the most worrisome problems are related to the presence of charged energetic particles, including:

- Cosmic rays
- Solar energetic particles
- Radiation belts

The charged particles in a space where a magnetic field is present do not have free motion. Due to the Lorentz force, particles have a gyration motion around field lines. The gyration frequency depends on its velocity (kinetic energy) and on the magnitude of the field:

$$\nu = \frac{qB}{2\pi m}$$

where q is the charge of the particle (C), m is its (relativistic) mass (kg), B the field

magnitude (T), ν thus being in Hz. As the field is locally uniform, the particle trajectory is along a helix with a radius (R), named the *gyroradius*, when projected on to a plane perpendicular to the field:

$$R = \frac{mv_{\perp}}{qB}$$

where v is the particle velocity (m/s) and $v_{\perp}$ is its perpendicular component with respect to the field. The velocity component parallel to the field gives the mean motion of the particle (the guiding center; see Roederer 1970). In a quasi-dipole field, such as near the Earth or near other planets with a strong magnetic field, such as Jupiter, Saturn, Uranus, and Neptune, particles are trapped by the magnetic field. Their motion can be divided into three different motions: the gyration around the field line, the bounce between mirror points along the field line, and a drift motion around the planet, due to the gradients and curvature of the field. Each of these motions being periodic, the total one is called quasi-periodic, and particles gyrate, bounce, and drift around the planet for days and even years, though their velocities approach that of light for relativistic particles (electrons in the hundreds of keV range, protons in the hundreds of MeV range, heavy ions in the tens of GeV range). Particle periods associated with gyration, bounce, and drift are given for different planets in Table 11.4 for 1 MeV equatorially mirroring trapped electrons.

Nevertheless, where (or when) the gyroradius is higher than the gradients of the magnetic field, particles cross the field lines and penetrate magnetospheres. This effect is well known around the Earth where high-energy cosmic rays are detected at sea level in mid-latitude regions. In fact, the latitude reached by a particle depends on its energy and charge state. This is known as magnetospheric shielding and was extensively studied by Smart and Shea (1985) using different magnetic field models. Table 11.5 gives the minimum latitude reached by protons for various energies at sea level.

TABLE 11.4 Periods of Gyration, Bounce, and Drift of 1 MeV Electrons Mirroring at the Equator of Various Planets at an Altitude Equal to the Planet's Radius

Periods	Earth	Jupiter	Saturn
Gyration	30 μs	0.2 μs	40 μs
Bounce	0.1 s	0.7 s	1.3 s
Drift	2,000 s (30 mn)	$6.8.10^6$ s (79 days)	$1.3.10^5$ s (1.5 days)

TABLE 11.5 Minimum Latitude Reached by Protons Coming from Outside the Magnetosphere at Sea Level

Energy	Latitude
1 MeV	62°
10 MeV	59°
100 MeV	52°
1 GeV	44°
10 GeV	5°

Magnetospheric shielding gives the boundary where cosmic rays, as well as solar energetic particles, penetrate the magnetosphere. It also gives an idea of the external boundary of the radiation belts. At synchronous distance from the Earth, nearly all ions with energy greater than 1 MeV penetrate, so it is easy to detect solar energetic particle events there.

Due to the relation between the particle motion and the magnetic field, the flux of energetic charged particles in the heliosphere is subject to variations. First, because they are guided by the magnetic field, they are subject to field fluctuations. In the solar wind, these variations can be due to the magnetic field of the sun itself (the different polarities of the two hemispheres, with a nonplanar equator), plus its fluctuations, due to coronal mass ejections (CMEs), interplanetary shocks, or corotating interaction regions (CIRs). Near the planets, these fluctuations imply substorms and magnetic storms, which are seen, for example, on Earth as auroral intensification. Superposed on these fluctuations is a slow variation due to solar activity, with a period of 10.5 years. Activity is maximum (solar maximum) for about seven years, when sunspots are recorded, and minimum (solar minimum), for the rest of the cycle. During solar maximum, the field fluctuates more and storms are generally more intense, though major storms can be recorded during solar minimum, as in February 1986. But the most intense storms are in solar maximum, and an example is given by the March 1989 storm, during which the Quebec electric blackout occurred. Very intense storms modify the magnetospheric shielding, as seen with low-altitude satellites.

We will review all those components of the environment and their space and time variations due to magnetic fields and their fluctuations (Anderson 1994; Barth 1997).

First of all, the origin of cosmic radiation is not fully understood. The major part comes from inside our galaxy (the Milky Way), so it is also called galactic cosmic rays (GCRs), but part of it originates in other galaxies. It is mainly composed of hydrogen and helium, though heavier ions are present. The ions present in GCRs are normally completely stripped and thus fully ionized. The spectrum is very hard, as TeV particles are usually observed. A second source of cosmic particles is anomalous cosmic rays. It is generally believed that their source is interstellar neutral gas diffusing in the heliosphere (the plasma cavity surrounding the Sun due to the interaction between the solar wind and the interstellar gas) and interacting with the solar wind, where it is singly ionized. There are cosmic rays everywhere in space, except near the magnetized planets, where the magnetospheric shielding acts to decrease the penetration of these particles. Outside magnetospheres, all cosmic rays are nearly isotropic, but near the penetration boundary related to the magnetospheric shielding the flux is anisotropic due to the field direction. There are rapid variations of cosmic rays, related to magnetic field variation. Nevertheless, as far as effects on satellites are concerned, they have no impact, and only the long-term variations are important. There is a solar cycle effect at low energies (lower than 10 GeV/nucleon), the flux of cosmic rays being lower at solar maximum than at solar minimum. This effect is due to the solar wind, which strongly diffuses the cosmic rays at solar maximum.

Another component of this energetic environment is due to the Sun itself. During solar flares, particles are ejected. These can be recorded near the Earth as solar energetic particle events when the ejection location is not too far from connected to the Earth by a field line. These particles are mainly protons and helium, with some heavier ions and electrons. The ions are weakly ionized. Because they are due to solar flares, they are recorded only during events, and the worst ones are

well known by the specialists (July 1959, November 1960, August 1972, September 1976, October 1989, for example). Intense events can be associated with coronal mass ejections. In that case, and if the flare is long enough, because the energetic particles reach the Earth within an hour while the solar wind perturbation (the magnetic cloud) reaches it in a few days, the flux of high-energy particles can be increased by nearly one order of magnitude during one day. This was the case, for example, in October 1989. The solar particles penetrate the magnetosphere like cosmic rays, with a limit given by the magnetospheric shielding. Apart from the beginning of the event, the flux is nearly isotropic outside the magnetospheres. Inside, there is an anisotropic effect near the penetration boundary, as in the cosmic ray case. The duration of events is variable, from a few hours for small events to some days for intense ones. Intense events are more likely to be recorded during solar maximum. A satellite can record half of all its single event effects in a single solar energetic particle event, depending on its orbital parameters.

Finally, radiation belt particles are trapped around magnetized planets. The ultimate source of these particles is in the solar wind for electrons and low-energy (less than 10 MeV) protons. Magnetospheres act as accelerators for solar wind particles, because they are accelerated crossing inward the dipole field lines. Another source of medium energy radiation belts (the 10 MeV range) is related to solar energetic particle events. When associated with an intense magnetic storm, particles from the interplanetary medium penetrate the stable trapping region and increases of the belt flux can be recorded. Higher-energy protons (up to hundreds of MeV) are also trapped near the Earth. This is related to the cosmic ray albedo neutron decay phenomenon. Because the particles are guided by the magnetic field, the radiation belt flux is highly anisotropic. Due to their gyration bounce and drift motion, particles near the planets are seen coming from a plane parallel to the surface, while near the (magnetic) equator particle flux is nearly isotropic. Radiation belt particles cover a large part of the inner magnetosphere, as they can bounce back and forth from mirror points located near the high-altitude atmosphere. Nevertheless, the flux is generally higher near the equator than near the footprints of the field line, one being in the north hemisphere and the other in the south. The radiation belts are not far from axisymmetric. Variations from axisymmetry are due to the magnetic field, which is composed of two parts:

1. The internal component, whose dipole is generally tilted and shifted from the rotation planet axis. On the Earth, the South Atlantic Anomaly is related to the tilt and the shift of the dipole. Multipole components add to this asymmetry.
2. The external component, due to the interaction between the solar wind and the internal field. This component creates also an asymmetry, which is seen at high altitudes as a day/night variation.

The Earth's proton radiation belt has one maximum (it is in fact a region located near the magnetic equator), the location of which depends on the particle energy. Table 11.6 gives the equatorial altitude of the maximum as a function of the proton kinetic energy. The minimum altitude for the belt is around 100 km, and the maximum for high energies is given by the magnetospheric shielding. This radiation belt is very stable, though very intense events modify it. Mullen et al. (1991) reported the creation of a second belt in the 30 MeV range at an altitude of 7,300 km following the March 1991 event. Other events can decrease the width of the belt, by a loss related to the modification of the magnetospheric shielding. The inward edge of the proton belt varies with the solar cycle. As solar activity in-

TABLE 11.6 Equatorial Altitude of the Proton Radiation Belt Maximum for Various Energies

Energy (MeV)	Altitude of the maximum (km)
0.1	14,000
0.3	13,400
1	12,100
3	7,000
10	4,500
30	3,200
100	2,500
300	2,500

creases, the high-altitude atmosphere extends and more trapped protons are lost. Thus, the proton flux is decreasing. This effect, which can be as high as a factor of 10, is effective up to an altitude of 2,000 km.

The electron radiation belt is composed of two different regions:

1. The inner electron belt, which is located at 5,000 km equatorial altitude for hundreds of keV, is very stable. Nevertheless, intense storms can increase it, as in the March 1991 case, or cause a loss at the outer edge of the belt.
2. The outer belt has a location which depends on energy. Table 11.7 gives an idea of the equatorial location for various energies. Nevertheless, this location, as well as the maximum flux, is extremely variable.

Between the two regions is a zone called the slot, where the flux is weak. The slot can be filled up with high-energy electrons in the case of a very intense event. The dynamics of the outer belt are related to strong magnetic storms (for Earth's magnetic index *Kp* greater than 4). The outer belt varies with the solar cycle. It has been shown that the maximum of the outer belt occurs in the declining phase of the cycle, some three years after the maximum, this being related to the extension of the coronal holes on the Sun. As for the protons, the low-altitude flux must vary with solar activity, being maximum when solar activity is low (solar minimum).

TABLE 11.7 Equatorial Altitude of the Outer Electron Radiation Belt Maximum for Various Energies

Energy (MeV)	Altitude of the outer belt maximum (km)
0.1	30,000
0.3	20,000
1	18,000
3	15,000
10	9,000

The outer electron belt is not as limited as for protons, the magnetospheric shielding being much less efficient for such light particles.

Other radiation belts (helium, carbon, oxygen, iron) exist in the Earth's magnetosphere, though they have not been so extensively studied. Due to the magnetospheric shielding, they are limited to low equatorial altitude regions. They are extremely variable with their sources (solar energetic particles) and their losses (magnetic storms).

The maximum energy of the radiation belt particles depends essentially on the size of the magnetosphere. This size depends on two parameters: the distance from the Sun, which gives the solar wind pressure in the vicinity of the planet, and the magnetic field moment, which by its magnetic pressure equilibrates the first parameter. Since Jupiter has the biggest magnetic field and is far from the Sun, its electron radiation belt contains electrons with 20 MeV kinetic energy and even more and thus emits in the radio frequency range (it is a big source of radio emission, one of the first discovered). Some of these high-energy electrons escape from the Jovian magnetosphere and are recorded in the Earth's vicinity when the two planets are connected by the Sun's magnetic field. Jupiter's radiation belts can cause problems for satellites sent to far space, as its flyby is usually used to accelerate these satellites. Some characteristics of the magnetized planets are given in Table 11.8. Less intense radiation belts can be found in Saturn's magnetosphere, as the magnetic field is much less intense than Jupiter's and the planet rings absorb energetic particles. Particular attention can be paid to Uranus. Because its dipole axis is nearly perpendicular to the Sun–planet line, it would be very interesting to study these radiation belts. Another interesting body is Mercury, with its small magnetospheric size.

The principal question to ask in order to predict the effects of energetic particles is how long the satellite encounters the different populations (cosmic rays, solar energetic particles, radiation belts). The answer depends on the orbital parameters of the satellite. For Earth-orbiting satellites:

- At synchronous orbit, the satellite is as if outside the magnetosphere for protons with energies greater than 1 MeV. It then records cosmic rays and solar energetic particle events. For lower-energy protons and electrons, it is always in the radiation belts. Effects can be due to these populations.
- At low Earth orbit, the satellite crosses the boundary of the magnetospheric shielding. This means that at high latitudes, cosmic rays and solar energetic par-

TABLE 11.8 Characteristics of the Solar System Magnetized Planets

Planet	Mercury	Earth	Jupiter	Saturn	Uranus	Neptune
Obliquity to orbit (°)	0.01	23.5	3.1	26.7	97.8	28.3
Magnetic moment ($G.Rp^3$)	0.0033	0.301[a]	4.28	0.21	0.228	0.142
Dipole tilt to rotation axis (°)	169	169.5[a]	9.6	t<1	58.6	46.9
Rp (km)	2,439	6,371	71,398	60,330	25,600	24,765

[a] The Earth's magnetic moment and dipole tilt vary with time and are given for year 2000.

ticles are recorded, whereas at low latitudes it measures radiation belt particles, especially in the South Atlantic Anomaly region.

These three populations are important to remind us to anticipate the effects they will have on the satellite system. Models are available for these populations, but in general these models give only orders of magnitude of the populations and thus their effects.

REFERENCES

Anderson, B. J. 1994. *Natural Orbital Environment Guidelines for Use in Aerospace Vehicle Environment*, NASA Technical Memorandum 4527, June.

Barth, J. 1997. "Modeling Space-Radiation Environments," IEEE Nuclear and Space Radiation Conference short course, Snowmass, CO.

Hargreaves, J. 1992. *The Solar-Terrestrial Environment*, Cambridge University Press, Cambridge.

Kivelson, M. G. and Russell, C. T. 1995. *Introduction to Space Physics*, Cambridge University Press, Cambridge.

Mullen, E. G., Gussenhoven, M. S., Ray, K., and Violet, M. 1991. "A Double-Peaked Inner Radiation Belt: Cause and Effect as Seen on CRRES," *IEEE Transactions on Nuclear Science*, vol. 38, pp. 1713–1717.

Smart, D. F. and Shea, M. A. 1985. "Galactic Cosmic Radiation and Solar Energetic Particles," in *Handbook of Geophysics and the Space Environment*, ed. A. S. Jura, Hanscomb AFB, MA.

Suess, S. and Tsurutani, B. 1998. *From the Sun: Auroras, Magnetic Storms, Solar Flares, Cosmic Rays*, American Geophysical Union, Washington, DC.

PART 3

AIRCRAFT ENVIRONMENT

Marc Pélegrin

11.8 TYPICAL FLIGHT PROFILE FOR COMMERCIAL AIRPLANES

Safety

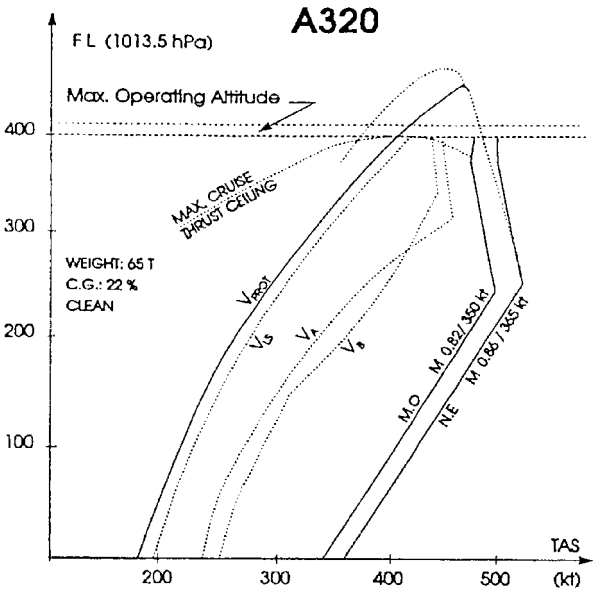
The three keywords for commercial air traffic are safety, efficiency, and environment. The local structure of the atmosphere in which the airplane flies is directly connected to safety. Accidents due to weather phenomena account for between 4 and 5% of the total number of accidents (Boeing source), and 5 to 7% domestic flight delays are due to meteorological causes, varying according to the season and the airport (Air France source).

For a flight of 1.5 to 2 hours' duration (gate to gate), accidents occur mainly during the takeoff and climb phases (more than 30%) and approach and landing phases (more than 50%). Some reasons are: at takeoff the aircraft weight could be at its maximum; *rotation,* the instant at which the pilot takes the initial climb attitude (angle of attack), corresponds to 1.3 V_{ST} (1.3 times the stall velocity); the landing phase implies a smooth junction between the airborne trajectory (altitude above terrain is related to *barometric pressure*) and the ground trajectory, which begins at the touchdown point; and atmospheric phenomena are more complex in the ground vicinity.

Until the 1950s, aircraft were considered as behaving like a rigid body; later, in flight, static deformations were included in the computation of the plane structure; the first two planes computed as deformable bodies were the B707 and the Caravelle, both introduced in 1958–1960. Nowadays, structural modes are taken into consideration at least up to the eighth first mode: two bending modes and two torsion modes of the wing, the symmetric and antisymmetrical mode of the jet engine masts, first bending and first torsion mode of the fuselage. Modes of rear empennage are also considered, namely when fuel can be stored in the rear for better balance (mainly during the cruise phase), leading to fuel savings. These modes are excited by atmosphere heterogeneities or pilot actions.

Parameters Available on Board

The only measurable parameters on board linked to the atmosphere are static pressure, p_s, dynamic pressure, p_d, and total temperature, T_a. From these data true airspeed (TAS) and Mach number are derived by the St. Venant or Rayleigh formula according to the Mach number. The local wind around the aircraft is derived from the true airspeed and the ground speed (if this is available on board). Precise localization systems associated with onboard inertial systems give the ground speed. The equations to be solved are:



SPEEDS DEFINITION

L_S : Lowest selectable speed. It corresponds to 1.13 V_S during take-off or following touch and go. It becomes 1.23 V_S as soon as any flaps/slat selection is made.

V_A : Design speed for maneuver

V_B : Design speed for maximum gust intensity and rough air speed (JAR 24.335 (d))

HIGH SPEED PROTECTION :

MMO/VMO : M.82/350 kt
M_{MO}+ 0.04/V_{MO}+15 kt : maximum steady speed with full nose down stick.

HIGH ANGLE ATTACK PROTECTION :

This protection has priority over all other protections.

V prot : min speed (corresponding angle of attack : alpha prot).
If alpha prot is exceeded, the angle of attack returns to and maintains alpha prot.

FIGURE 11.1 A320 flight envelope.

Subsonic Uncompressible Flow. Assuming that the local static temperature is available, the true airspeed is defined by:

$$V^2_{TAS} = \frac{2a^2}{\gamma - 1}\left[\left(\frac{p_d - p_s}{p_s} + 1\right)^{\gamma - 1/\gamma} - 1\right]$$

where a is the sound velocity corresponding to the local static temperature and γ is the ratio between the two heat coefficients (constant pressure, constant volume). But the static temperature is not measurable on board, and hence the value of a is not directly available.

Then a calibrated airspeed is defined by:

$$V^2_{CAS} = \frac{2a_0^2}{\gamma - 1}\left[\left(\frac{p_d - p_s}{p_0}\right)^{\gamma - 1/\gamma} - 1\right]$$

where a_0 and p_0 are the sound velocity and the static pressure at sea level for the standard atmosphere. The calibrated airspeed can easily be obtained on board—it is the ''speed'' that is shown on the panel instrument and is used by the crew to control the plane.

Compressible Flow. The Mach number is defined by:

$$\frac{p_d}{p_s} = \left[\frac{(\gamma + 1)^2 M^2}{4\gamma M^2 - 2(\gamma - 1)}\right]^{\gamma/\gamma - 1} \frac{1 - \gamma + 2\gamma M^2}{\gamma + 1}$$

It is important to know the TAS, or at least the CAS and the Mach number, in the event of flying in a wind shear zone, in order to avoid getting out of the flight envelope of the plane. Nowadays the TAS is computed or extracted from stored tables and presented on the instrument panel on the PFD (primary flight display).

The flight envelope for an A320 is represented in Figure 11.1; the envelope is graduated in V_{CAS} or M; the velocity used for piloting the aircraft is the V_{CAS}, which appears on the PFD. The V_{TAS} is computed and represented in the right corner of the PFD. In the near future, data will be automatically transmitted to ATC (air traffic control); the controllers will get ground velocities computed on board or derived from the radar tracking. In addition, atmospheric data collected and processed by planes will increase knowledge about the atmosphere.

11.9 THE ATMOSPHERE

The Standard Atmosphere

Perfect stability is assumed. The pressure is supposed to evolve according to the diagram presented in Figure 11.2. Such data are used to start the airplane computation, but the final computation should take into consideration the real atmosphere parameters, which can be extrapolated on a probability basis only.

The atmosphere is divided into (Figure 11.3):

- the *troposphere*, an 8,000–11,000 m high layer around the earth, according to the latitude and the period of the year. Horizontal and vertical movements (called

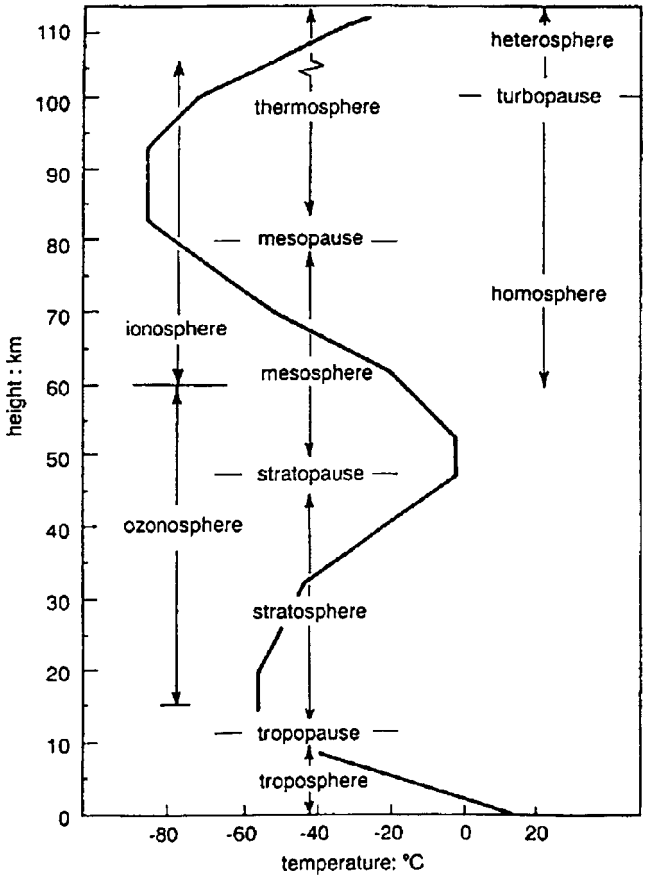


FIGURE 11.2 Standard atmosphere

turbulence) occur even in clear atmosphere. Inside clouds, namely in active cumulonimbus clouds, turbulence may reach values that can compromise the safety of the flight.

- The *stratosphere*, just above the troposphere, in which air movements are mainly horizontal. Clouds may be present only in the first 5 km of thickness.
- The *tropopause*, a transition layer between the troposphere and the stratosphere. The position and thickness of this layer vary with latitude and season.

The ATC works on ground velocities (not air velocities) to elaborate strategic positions of planes in a given airspace. The vertical separation is arbitrarily referenced to a barometric pressure (1013.25 hPa): 1,000 ft in the lower space and, from 2002, in the upper space. The horizontal separations should be reduced in the future to cope with higher densities of planes in a given airspace. Gradients of pressure, wind, and temperature should be taken into consideration in order to guarantee a

AVIONICS AND ASTRIONICS

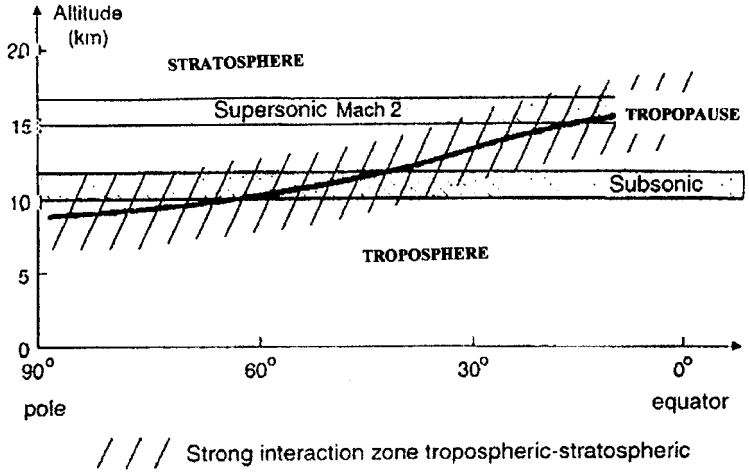


FIGURE 11.3 Troposphere, tropopause, stratosphere.

minimum separation distance in the isobaric surface on which the plane flies. Reference values are static temperature at altitude 0 on the geoid: 1013.25 hPa, temperature 288.15 K (15°C), density 1.2922 g/dm^3. The standard atmosphere is composed of nitrogen (75.5% in mass, 78.1% in volume), oxygen (23.1%, 21%), carbon dioxide (0.053%, 0.035%), argon (1.28%, 0.93%). The composition is constant from 0 to 50 km in altitude. The water content can vary from 30 g/m^3 in topical zones to 1 g/m^3 in polar zones; it is the main parameter for cloud formation. (See also Section 16.)

Thermal Equilibrium (Le Trent and Jancovici 2000)

The two main parameters that contribute to the Earth's temperature equilibrium and consequently to climate* are the energy received from the sun and the position and orientation of the earth in its orbit.

The atmosphere interacts with both incoming energy from the sun and radiated energy from the earth. The greenhouse effect is due to the reflection of this radiated energy; without this effect, the mean equilibrium temperature would be −18°C instead of +15°C. The energy reflected toward the earth is due primarly to the H_2O, CO_2, CH_4, and O_3 atmosphere content.

The energy emitted by the sun also varies; the frequency of fluctuation ranges between 11 and 12 years (solar activity) and millions of years. The Little Ice Age which occurred during the 17th and 18th centuries was due to a deficit in solar energy. The relative stability of the climate for several hundred thousand years is only partially explained. Presently, above the atmosphere the flux received is 1,365

* "Climate" comes from Greek and means "inclination" (of the rotation vector of the Earth, with regard to the ecliptic).

W/m^2, and that received on the surface of the earth, at a global mean value and day-night periodicity, is 345 W/m^2.

As to the variation of the rotation axis of the earth with regard to the ecliptic, the influence of planets such as Jupiter and Venus is dominant. Variation of the eccentricity of the annual cycle (main period 100,000 years), variation of the obliquity (angle between the rotation vector and the normal to the ecliptic plane; main period 40,000 years), equinoxial precession correlated with the mean distance to the sun (main period 20,000 years).

The three-atom gases H_2O, CO_2, and O_3 seem to play a dominant role in the energy balance of the planet, even though their concentration is very low. Excluding the two last centuries, the composition of the atmosphere seems to have been quite constant during the last 10,000 years: 270 ppmv (as measured from the composition of air bubles contained into ice samples taken from Greenland and Antarctica.

Energy Flow Distribution. For the incoming solar flux:

- 30% is diffused into space (6% corresponds to an interaction between the incoming photons from the Sun and the air molecules, mainly O_2; blue photons are emitted; 4% is directly reflected by the Earth surfaces, land, ocean, or ice/snow).
- 50% hits the Earth's surface and is absorbed, leading to a temperature increase; infrared radiation appears.
- 20% is directly absorbed by atmosphere; O_3 absorbs the UV radiation; the O_3 concentration is much more important above an altitude of 10–15 km, which is why the atmosphere concentration increases above the tropopause.

On the Earth's surface, the water evaporation produces a temperature drop and the condensation in the atmosphere produces a temperature rise. This is the main factor is the atmospheric heat. Heat transfer occurs from equatorial regions (from 30° S to 30° N) to polar regions, with a predominant transfer from tropical regions to subtropical regions (Hadley–Walker cells). The Earth's rotation (which induces the Coriolis force) makes transfers above the 30th latitudes unstable, giving rise to anticyclones and depressions with winds rotating around them.

Oceans interact with the atmosphere with a time lag of several hundred years. In the atmosphere, air movement is fast but carries little energy. In contrast, oceans carry a high level of energy but over a long time. However, due to its pattern, the Pacific Ocean has a dominant role on a short-term basis (some years); surface currents transfer energy from equatorial zones toward polar regions in both hemispheres (El Niño); in addition, every 2 to 4 years, warm water is carried from west to east. For the last few years, this phenomenon has been more active, with dramatic consequences—dryness in Australia, Indonesia, and northeastern Brazil and severe rains in California, Peru, and Argentina.

The Real Atmosphere

The real atmosphere differs from the standard atmosphere because its pressure and temperature vary. Winds are caused by these differences of pressure and temperature. The main cause of instability is the daily and annual periodicities of the (apparent) Sun motion; the direct consequence is the production of winds. The water content is not homogeneous even in the clear atmosphere (vapor is transparent)

Clouds. Clouds are generated by the ascending motion of moist air, the potential temperature, and the pressure inside a cloud decrease. Cloud formation depends upon the water vapor content of the atmosphere and the number and type of particles which act as centers of condensation and possibly icing.

Clouds are either droplets of water or ice crystals; liquid droplets may exist in negative temperature (from 0° to −35°C). The basic types of clouds are (Chaon 1999):

- *Cirrus* (Ci), high-altitude (7–15 km) isolated clouds in the form of delicate filaments or white or mostly white patches or narrow bands, silken or fibrous in appearance (altocirrus: ice crystals).
- *Cumulus* (Cu), low (below 2 km) and medium (2–7 km) altitude detached clouds, generally dense; they look like a cauliflower (several kilometers in diameter) with a side wall illuminated directly by the sun or not, as the case may be; their base is relatively dark and nearly horizontal.
- *Nimbus*, mainly *nimbostratus* (Ns), low-altitude, gloomy clouds, often dark, which generate rain or snow, very often continuously.
- *Stratus* (St), generally low-altitude clouds, looking like layers or extended flat patches at low, medium, or high altitudes; if the sun is discernible no halo is produced (except at very low temperature).

From the basic clouds mentioned above, the following clouds are derived:

- *Cirrocumulus* (Cc), thin, white patch, sheet or layer of cloud without shading, composed of very small elements.
- *Cirrostratus* (Cs), transparent, whitish cloud veil of fibrous or smooth appearance, totally or partly covering the sky; these produce halo phenomena.
- *Cumulonimbus* (Cb), accumulation of big, gloomy clouds, with considerable vertical extent: such thunderclouds may generate lightnings; the upper part often spreads out in the shape of an anvil.
- *Stratocumulus* (Sc), low-level layer having a dappled or wavy structure, with dark parts.
- *Altocumulus* (Ac), white or gray patch, sheet, or layer of cloud, generally with shading, composed of laminate rounded masses, rolls.
- *Altostratus* (As), grayish or bluish cloud sheet or layer of striated, fibrous or uniform appearance, totally or partly covering the sky.

Active cumulonimbus clouds are dangerous for aircraft due to pronounced turbulence and vertical velocities of 20 m/s or more in the updraft ascending core of the air/water, severe turbulence, lightning, and ice. They can be detected by radar aboard the aircraft.

Turbulence, Wind shear. The real atmosphere has no stability, though in many regions of the world, as in the temperate zone, the weather structure of the atmosphere evolves slowly. Weather is a consequence of air movements around the world; as for any system for which a good mathematical model exists, it should be predictable. Air movements are governed by partial differential equations and are known with reasonable certainty, but a set of accurate initial conditions (4-D) is not yet available, in spite of many meteorological satellites. Meteorologists proceed

by region (using ground grids with horizontal sizes varying from a few kilometers to hundreds of kilometers) and try to set coherent initial conditions for each grid. The computer then solves the equations and arrive at a correct (4/5) forecast for 48 hours.

Local random motion of air within the motion of a large mass of air (which covers an area of some tens or hundreds of square kilometers) is called *turbulence* and interacts directly with the aircraft structure. The size of the turbulence ranges from several meters to several kilometers.

To be certified, a plane must experience no damage (more precisely, it should stay in the elastic domain) when crossing a gust or flying in a turbulent area. Gusts are defined by specifications which vary slightly among the countries which certify planes. For example, in France, the two major conditions to be satisfied are:

1. A vertical gust of "1 + cos" type (Figure 11.4):

$$W(t) = 0 \qquad \text{for } t < 0 \quad \text{and} \quad t > \frac{d_m}{V}$$

$$W(t) = \frac{W_n}{2}\left(1 + \cos\frac{2\pi Vt)}{d_m}\right) \qquad \text{for} \quad 0 < t < \frac{d_m}{V}$$

where W_n = gust amplitude (m/s)—according to FAR 25.341, 20 m/s for height ≤20,000 ft
11 m/s at 50,000 ft
d_m = gust wavelength (m) (depends on the wingspan of the aircraft)
V = aircraft velocity
t = time

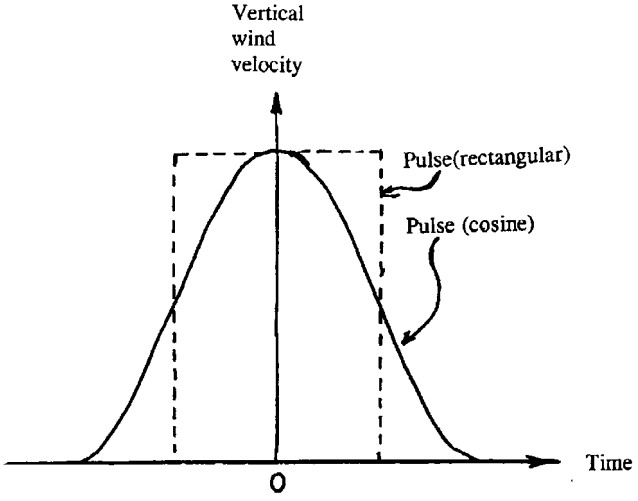


FIGURE 11.4 Vertical gust profiles.

2. A von Karman spectrum for the turbulence:

$$\Phi_W(\omega) = \sigma_W^2 \frac{L_W\{1 + (8/3)[1{,}339(L_W/V)\omega]\}^2}{[1 + 1{,}339(L_W/V)\omega]^{11/6}}$$

$$\sigma_W^2 = \frac{1}{2T}\int_{-T}^{+T} \phi_W(\omega)d\omega$$

where L_w = turbulence length scale (m)
σ_w^2 = mean square value of velocity (m^2s^{-2})
ω = pulsation (rd/s)

Note: numerical values must be coherent with the aircraft safety level reached at a given time. As this safety level increases, the amplitude of the gust or turbulence spectrum should be that which has a probability of occurrence of the same value as the safety level of the aircraft.

Meteorologists use four grades of turbulence, independently of the type of turbulence: *light, medium, severe,* and *extreme*.

Wind shear occurs when two layers of wind in the atmosphere have different velocities and/or directions. Due to friction between the two layers, a transition zone in between the two laminar layers is highly probable. Wind shear can exist at any altitude. Approach controllers are particularly interested in wind shear because it interacts with the safety of the landing.

Downburst. In the 1980s a new phenomenon was identified: *downburst*. A downburst is the collapse of a mass of cold air suspended at some thousand meters of altitude by active ascending movement of air. When the collapse occurs, a downstream of saturated air may reach velocity above 40 kts (Figure 11.5; note that *downburst* is called *microburst* in the figure). As the mass of air goes down, the local temperature increases. In the upper part of the downflow there may be droplets, which may disappear by evaporation below a certain altitude (depending on the local temperature) and consequently are difficult to detect. When the stream hits the ground, a giant vortex appears. This is a very dangerous phenomenon. Predetection is difficult, its detection requires permanent real-time analysis of the structure of the local atmosphere around the airport.

Downbursts have been clearly explained by Fujita (1985): ''Some aircraft accidents that occurred at low altitudes during convective activity were regarded as pilot error without blaming the weather systems as major contributing factors.''

Tornadoes and Microbursts. Tornadoes consist of ascending motion of saturated warm air in a column several hundred meters in diameter. They appear on hot seas or lakes (surface temperature higher than 27°C) mainly in the afternoon. Downbursts are frequently associated with tornadoes, an additional reason to avoid tornadoes. The most spectacular phenomenon in which tornadoes and downbursts were associated occurred in July 1987 at Teton-Yellowstone (U.S.A.) and was carefully studied by Fujita (1985). The U.S. Forest Service indicated that 1 million trees were uprooted in a 61 km^2 area over a period of 26 minutes. The analysis of the orientation of the fallen trees was a powerful tool of investigation. Over an area 2.5 km wide and 39 km long, four swirl marks of spin-up vortices (tornadoes) and 72 microburst outflows were identified.

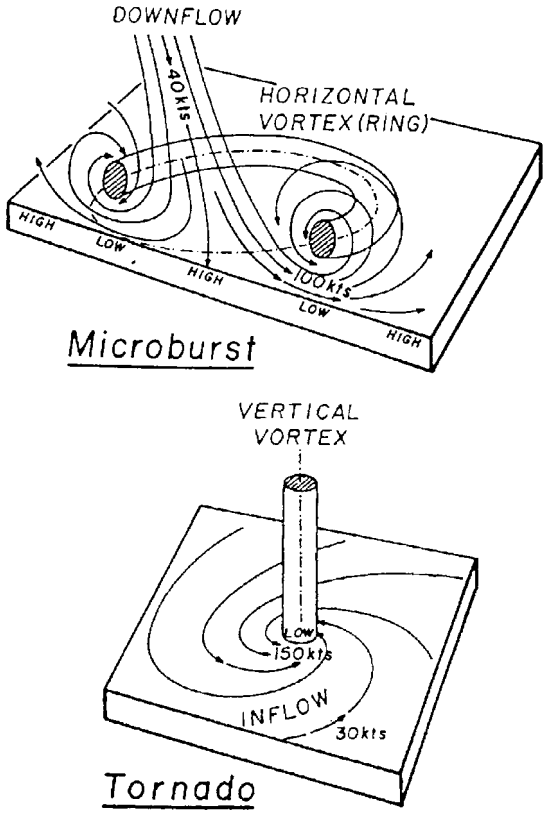


FIGURE 11.5 Downburst and tornado profiles (from Fujita 1985).

How could the tornado have maintained its fury against large frictional torque in the boundary layer over rugged terrain? Analysis of the damage caused along the trajectory of the tornado and on its sides suggests that the angular momentum of the tornado was supplied by microbursts as their outburst winds spiraled into the tornado center.

A tornado can be detected easily by radar or lidar or, most of the time, by direct observation because of the water content. However the side microbursts which accompany the tornado are often ''dry'' and not directly visible; it is recommended that flying be avoided at least 2.5 to 5 km away from the tornado.

Jet Streams. In the lower stratosphere, jet streams are frequent. These are ''tubes'' of air, roughly horizontal, several hundred meters or a few kilometers in diameter. Velocities may reach 200 m/s in the center. The flow is normally clear and laminar in the core, and the transition zone is highly turbulent. The direction is usually West to East.

Lightning. Upward convective motions in the troposphere may generate concentration of electrostatic charges. In the vicinity of a cumulonimbus cloud, strong electrostatic fields (500 kV/m) may be encountered by an aircraft flying in the cloud or in its vicinity, and electromagnetic perturbations may occur. On average, long-range aircraft receive a strike once every 3,000–4,000 hours, while short-range aircraft receive a strike every 2,000–3,000 hours; damage, if any, is rarely severe. Total destruction of the plane by lightning is very rare (less than three cases during the last 50 years). However, lightning is very often accompanied by strong adverse atmospheric conditions such as severe turbulence and icing (in an accident it is very difficult to decide which phenomenon was the real cause of the accident). Lightning is a discharge between zones in which the density of electrostatic charges is high and of opposite polarity. Lightning is composed of a short-duration impulse (about 200 ps to 2 μs) with an intensity of thousands of amperes and gradients reaching 100 kA/μs, followed by another, much longer pulse (several ms) but with an intensity much lower, about 100 A. Interference with airborne radio and electronic equipment is produced by the former, and damage can be caused to an aircraft by the latter since it contains more energy. Lower-power electronic chips and the increasing use of composite materials (though with conducting material incorporated) in aircraft mean that electronic equipment will have to be studied carefully.

Optical processors and an optical data bus will replace electronic equipment in the future (around 2005–2010).

Icing. In a cumulus cloud the water content is about 2.5 g/m^3 and the diameter of droplets is between 10 μm and 40 μm. Supercooled droplets can turn into ice when they collide with an aircraft structure, forming *rime ice* if the temperature is about -30°C or, if the temperature is close to 0°C, *glaze ice*. The accretion of ice may give rise to two horns. Supercooled droplets glide on the surfaces on which they have been deposited and may turn to ice somewhere on the wing, blades, or fuselage. The laminarity, if it was present before such a cloud was entered, is destroyed, the lift coefficient is reduced, and the drag coefficient is increased.

There are two types of icing clouds. Stratus clouds extend over a large area. Their content of water is low (0.1 to 0.9 g/m^3), and they produce continuous icing. Cumulus clouds have a water content of about 3 g/m^3. Their size rarely exceeds a few kilometers, and they can extend from 2 to 4 km up to 12 to 15 km in altitude.

Icing may also occur in clear air after an aircraft which has flown a long time through cool air moves into a clear, warm region. The water vapor in this warmer air condenses and freezes over the entire aircraft.

11.10 OTHER ATMOSPHERIC HAZARDS

Other hazards to aircraft due to interactions with the atmosphere are described below.

Turbulence Due to the Aircraft

The main parameter to be considered is the wake vortices which escape from the wings at their extremities. The wing-end vortex results from the difference of pres-

sures on the suction side (above) and pressure side (below) of the wing. From the right wing, the vortex rotates in the positive direction.

Behind a plane the structure of the atmosphere is modified in such a way that another plane crossing the wake vortices or penetrating into them can be exposed to a dangerous situation. According to Thomas Heintsch of the Institute for Flight Guidance and Control, Braunschweig University, the development of the wake vortices extends up to 200–250 wingspans behind the plane, which means approximately 10 km. The shape is quite constant (intensity and extension) during the first 50 wingspans. Then the intensity decreases and the lateral extension increases. Their intensity is proportional to the mass, balance, and load factor of the plane at a given time. Separation rules were established in 1970 (ICAO). Planes are classified into three classes: heavy (H), 130 T and up (250 passengers and up); medium (M), between 130 T and 10 T (50/250 passengers); and light (L), less than 10 T (less than 50 passengers). The minimum separation distances are (planes aligned on the ILS):

H/H = 3 nm
H/M = 5 nm
H/L = 6 nm
M/L = 4 nm

The separation rules are now obsolete, and real computation of the dangerous zone is possible.

The damping of the vortex is low (air viscosity); the local turbulence increases the expansion of a vortex motion. It is accepted that the decrease factor is $t^{-1/2}$ for a calm atmosphere and t^{-2} for a turbulent one.

Interference with the Ground (Puel and Saint Victor 2000). Due to friction with the local atmosphere, the vortices go down. If a vortex encounters the ground, reflections occur and interact with the initial vortex. Let's consider a wake vortex which is descending close to the ground. Local velocities are higher than those due to the mean wind, and additional decay of the vortex appears. However, the phenomenon is more complex and the descending vortex may generate a secondary vortex (opposite in rotation). The main vortex induces a lateral flow (with regard to the axis of the vortex) which then can suck the ground boundary layer and give birth to a bulb. Due to the pressure gradient, the vorticities of the two vortices are opposite and a vortex moving toward the initial vortex may be generated (Figure 11.6). The energy stored in the second vortex comes from the first one.

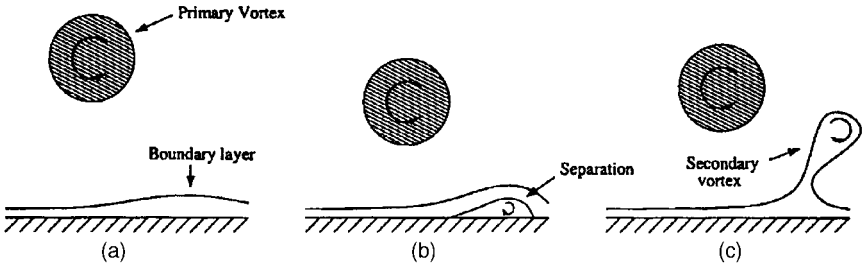


FIGURE 11.6 Interation between vortices and ground (ONERA).

The intensity of the vortex can be computed from the airplane parameters and the local characteristics of the atmosphere. A locally turbulent atmosphere damps the vortices more rapidly than a calm atmosphere. The local turbulence is slightly increased but has no structure, and the separation distance can be reduced.

Nowadays the trajectories of the vortices can be estimated (position and intensity) from the aircraft parameters and wind can be measured on the airfield. For airports equipped with lidar (or sonar), the intensity of the vortices may be roughly measured. Separation distances of planes can be computed in real time, including the lateral wind component with regard to the runway axis. The introduction of the A380 will impose real-time determination of separation between planes.

Birds

An aircraft is certified against collision with and ingestion of birds. Tests are performed on the ground: a (dead) bird of a specified mass is sent toward the cockpit windows, toward a propeller or inside a jet engine (if blades are destroyed, they should be self-contained inside the jet engine body).

The most dangerous (and frequent) case is ingestion when the plane is accelerating on the runway. If the ingestion happens before V_1 (the maximum speed after which the braking distance is higher than the length of the the runway in front of the plane) and is immediately detected, the takeoff should be abandoned. In addition to the risks attached to emergency braking, the risk of a fire resulting from the ingestion is slightly higher than that due to a failure of the jet engine without any ingestion. If the ingestion appears after, or is detected after V_1, the situation is much more critical; takeoff is mandatory even if V_2 has not been reached (V_2 is the recommended speed for rotation, i.e., the speed at which the longitudinal attitude of the plane corresponds to the initial climb). The ingestion is mainly detected by accelerometers set on the jet engine body which detect vibration of the pod.

The ingestion of birds is relatively frequent; Air France encounters some six to eight jet engine bird ingestions per year, leading to engine damages.

Meteorological Balloons

Some meteorological centers are qualified to send balloons equipped with instruments to measure atmospheric parameters across the atmosphere seven times a day (there are seven such centers in France). Balloon and nacelle are made so that collisions with aircraft are not hazardous, and the balloon is launched in accordance with present air traffic. However, some risk still exists.

Smoke from Volcanoes

There are presently 30 active volcanoes in the world, and during the last three decades about 200 volcanoes have been active. At least one fatal accident and one incident (four engines out but recovery after a while) occurred during the last 25 years due to ingestion of smoke ejected by a volcano. Even if there is no blowout, the smoke is composed of very hard dust which interacts with the engines and damages them. The largest particles of volcano smoke, most of which fall within

a few days, may constitute a danger for planes flying across the volcano's plume. The consequences may be the following:

- Jet engines being turned off because particles are deposited on hot parts (600°C to 800°C) and then form a solid state quite similar to glass
- Loss of aerodynamic data due to the pitot tube obstruction
- Erosion of front parts of the wings and opacification of glass windows
- Radio jamming due to electrical discharges encountered inside the plume
- Chemical corrosion due to acid droplets
- Fuel contamination by ash and soluble components such as Pb, Zn, and Cu

Between June 9 and 21, 1991, just after the Pinatubo eruption, nine incidents involving the replacement of 10 engines were registered.

Magnetic Storms

A magnetic storm is an ejection of charged particles coming from the Sun. There is a correlation with the Sun cycle (11 years). The energy involved may reach 10^{26} joules within a few minutes. When the particles reach the magnetosphere, the magnetic field is modified and electromagnetic inductions appear. This is a frequent phenomenon, and the consequences are well known: temporary degradation of the position precision of satellite positioning systems may occur, or no signals may be received for many minutes or hours; satellites, mainly geostationary ones, may be partially destroyed (internal flashes, destruction of solar pannels, etc.); prediction of occurrences and protection against such magnetic storms are difficult.

Traffic

The plane is not alone in the sky. The airspace is divided into six classes designated A to F. In each class a minimum of on-board equipment is mandatory. A flight can be operated under VFR (visual flying rules), where avoiding collision is the responsibility of the pilot, and IFR (instrument flying rules), where avoiding collision is the responsibility of the ground controller. In each class, IFR and VFR are possible under certain conditions, except for class A, a class in which only IFR flights are authorized. Before departure, the crew fills out a flight plan which describes the desired flight profile. According to the present traffic, the controller accepts it or modifies it. The plane is followed by ground controllers even when it flies above oceans; the crew reports the position of the plane at least every 20 minutes. Separation of planes is under the ground controller's responsibility. The situation is evolving rapidly thanks to automatic reporting systems such as ADS-B (Automatic Dependent System—Broadcast). The position of the plane, computed on board using GPS or ground systems (VOR, DME, ADF, LORAN,* etc.), is broadcast at a given frequency (which can be chosen from 1 second to 10 minutes). The data can be relayed by satellites and made available to ground control centers concerned with the flight.

*VOR: VHF omni-range; DME: distance measurement equipment; ADF: automatic direction finder; LORAN: long-range navigation.

Ingestion of Stones

During rolling on the runway or taxiways, stones can be ingested (or other objects: the Concorde accident on July 25, 2000, occurred due to ingestion of a piece of metal dropped on the runway by the plane which took off before). The structure around the pods (bottom part of the fuselage, wing, and pod itself) is sometimes modified by small ailerons in order to avoid the ingestion of stones thrown away by the wheel of the front landing gear. A small additional drag is induced.

11.11 THE IONOSPHERE

In the atmosphere an ionized layer is situated between 60 and 1,000 km, with a maximum concentration around 400 km. Electromagnetic waves crossing this layer are more pertubed if the frequency of the crossing wave is low; HF band and below are completely reflected.

For the frequencies used in GPS (1.2 and 1.5 GHz, plus 1.1 GHz in 2006), there is a slight energy attenuation and a slight increase of the traveling time, leading to an increasd distance of several meters between the satellite and the receiver. However, if two modulated frequencies are used, the perturbation may be corrected. This is why a third nonencrypted frequency will be transmitted by GPS satellites starting in 2006.

The delay Δt_1 which occurs on a frequency f_1 is related to the total electron content (TEC) encountered. If two frequencies f_1 and f_2 are used, there is a relation between the difference $\Delta t_1 - \Delta t_2$ and TEC. Then, the TEC being known, it is possible to compute the delay Δt_f which occurs on the frequency f used to determine the pseudorange (c = light velocity):

$$\Delta t_f = \frac{40.5}{c.f^2} \text{ TEC}$$

Magnetic Storms

Solar activity is not constant. Solar activity disrupts the ionosphere and affects, for instance, the pseudorange in GPS. Solar activity has a period of 11 years (high activity between 1999 and 2002). When a storm occurs, electrons and protons are ejected from the Sun's surface (for large storms, up to 10^{16} g).

REFERENCES

Chaon, J. P. 1999. *Cours de physique des nuages,* Meteo-France, Toulouse.

Fujita, T. 1985. *The Downburst,* University of Chicago, SMRP Research Paper 210.

Le Treut, H. and Jancovici, J.-M. 2000. *L'effet de serre,* Flammarion, Paris.

Puel, F. and Saint Victor, X. de. 2000. "Interaction of Wake Vortices with the Ground," Aerospace Science Technology, vol. 4.

PART 4

ELECTROMAGNETIC COMPATIBILITY

J. P. Parmantier

J. P. Catani, and

M. Crokaert

11.12 INTRODUCTION

Why compatibility in electromagnetics? The general answer is that electronic equipment has to operate in very different types of environments with which it has to remain compatible.

First, equipment does not have to be susceptible to the surrounding electromagnetic (EM) fields generated by the environment. The environment may be external to the entire system, as is the case with natural threats like lightning or electrostatic discharges, or human threats. Some are unintentional, but some systems, mainly military, are also concerned with intentional threats generated by EM weapons. The threat may also be internal to the system itself, generated by other pieces of equipment.

Second, a piece of equipment does not have to generate EM perturbation likely to interfere with another piece of equipment. This is why compatibility with its environment is required.

In aeronautics and aerospace, electromagnetic compatibility (EMC) has been known for a long time. The reason is that all the possible EMC problems are closely connected to reliability and safety. Indeed, in the air or in space, a failure of equipment may lead to serious casualties for the system's functions and, more serious, for people. Eventually, the entire transport domain has become concerned with EMC because of the increase of electronics in all its systems. Since the advent of EM weapons in the 1970s, military systems have also been involved in EMC, with the objective of being totally hardened to EM interference. EMC has become a discipline, thoroughly accounted for by any industry, from shipbuilders to household electrical manufacturers. Now, in addition to the functional aspect, EMC is also an economic challenge. To sell their products, manufacturers have to demonstrate their compliance with EMC standards. Therefore, they have to apply protections on their equipment and optimize them in terms of price, room and weight.

Subsection 11.13 deals with the physical process of EM coupling, which leads to the generation of EM interference. Subsection 11.14 presents the characteristics of the main EM threats and their associated standards. Subsection 11.15 introduces experimental and numerical tools commonly used in EMC design and analysis. In subsection 11.16, engineering methods for EMC are investigated, including a discussion of the control plan, the specifications, the conception, and the installation rules. Finally, the conclusion focuses on the future of EMC.

11.13 BACKGROUND OF EM COUPLING

Theory of EM Diffraction

The physical process that makes an interference act on a system is known as EM coupling. If an incident EM field is applied to an object, an induced current is generated on the surface of this object. This current is a potential interference for the object. Meanwhile, a scattered field is generated around the object. This field may also cause EM interference in its vicinity. On one hand, the theory of EM coupling has similarities to the theory of antennas because it is directly derived from the theory of diffraction (Stratton 1941). However, if antennas are mainly involved in far fields, EMC is mostly interested in near fields, which makes the usual approximations of antenna not applicable. On the other hand, EM coupling also has relations with circuit theory when the response of equipment connected with cable bundles is concerned or when electric protections as filters or limiters have to be considered. However, EM coupling involves additional information in terms of distributed equivalent sources induced by incident EM fields.

Considering the definition we gave of compatibility as being compatible with the ambient external EM field and as having the environment compatible with the EM-emitted field, two domains of analysis are commonly considered in EM coupling: EM susceptibility and EM emission.

EM susceptibility stands for an external stress applied onto the object under analysis. Two kinds of EM susceptibility problems are commonly distinguished: radiated problems, when an incident field is applied and conducted problems, when a current is forced on the object.

In radiated EM susceptibility, it is important to understand that the incident field E_t^{INC} is the field in the absence of the object. This condition is rigorous and comes directly from the theory of diffraction. When E_t^{INC} is applied, a current is induced on the surface in such a way that the total tangential field E_t^{TOT} on the surface verifies the following limit condition:

$$E_t^{TOT} = E_t^{SCA} + E_t^{INC} = Z_s \cdot J^{SCA} \tag{11.19}$$

where Z_s = the surface impedance of the lossy surface
J^{SCA} = the scattered current induced on the surface
E_T^{SCA} = the tangential scattered electric field

For example, if the surface is metallic, E_t^{TOT}, is zero and E_t^{SCA} verifies:

$$E_t^{SCA} = -E_t^{INC} \tag{11.20}$$

In conducted EM susceptibility, the current is directly applied on the system or forced with a generator to simulate the current induced by an incident radiated field or a perturbation generated by another part of the system.

EM emission is quite similar to conducted EM susceptibility in the sense that the source is also applied on the system or comes from a piece of equipment. But, in addition, we are mainly interested in the scattered fields radiated by the surface currents.

EM Coupling Phenomenon

After the general definition of scattering on a surface, the presentation of several physical processes will help the understanding of EM coupling.

Current Redistribution on a Surface. On an external surface, the first phenomenon to consider is redistribution of currents. At very low frequency, the current scatters with respect to the different resistance paths encountered, therefore following Ohm's law. On the whole frequency range, the current follows the paths of lower impedance. Particularly when frequency increases, the impedance due to the inductance of the structure becomes more important than the resistance. Because of the inductive effect, the current lines tend to separate from each other. At very high frequency, they follow the edges of the object. In the intermediate frequency range, there is a cutoff frequency where the resistive effect balances the inductive effect. For instance, this property explains why carbon materials progressively behave like metal.

EM Penetration Through a Surface. EM penetration through the surface may occur through two types of processes.

The first is EM diffusion and comes from the finite depth of the materials. It only concerns low frequencies when currents are able to penetrate into the materials. This phenomenon is also known as the skin effect: for an external excitation, the current will progressively concentrate on the external surface. The phenomenon may be understood as a generalization of the inductive effect on a finite surface applied here on the finite transverse dimension of the depth.

The second EM penetration type is due to scattering through apertures. ''Apertures'' is a generic name for a large variety of geometrical configurations: windows, holes, seams, junctions between panels, electromagnetic joints. The phenomenon is significantly dependent on manufacturing technology. Nevertheless, general rules may be established as a function of frequency. For instance, at low frequency, when the aperture is small compared to the wavelength, the scattered field is equivalent to the one radiated by an electric dipole and two magnetic dipoles (Degauque and Hamelin 1993; Boudenot and Laboune 1998). If the aperture is loaded with a resistive material, a cutoff frequency f_c appears on the magnetic field. Under f_c, the magnetic field penetrates, which is in agreement with a very general property in EM coupling; there is no perfect protection against the magnetic field for any kind of resistive material. Beyond f_c, the magnetic field is attenuated with a 20 dB per decade slope (see Figure 11.7). Also, the electric field is attenuated as soon as the very low frequency for almost all the resistive materials. The resistance of the junction area connecting the resistive material of the aperture to the metallic frame modifies the value of f_c (Boudenot and Labaune 1998).

Definition of EM Shields. The definition of currents induced on and through surfaces raises the important concept of EM shields. In EMC, a shield is a material that deviates the current, preventing it from running in undesired zones. Of course, the most efficient shield is metallic because all the current may be driven into it. But even if perfectly metallic, a shield is not efficient if the current is not able to flow on it. This means that the shield must be connected at both ends to ensure the derivation of the current. In addition, because currents circulate on shields, they are more likely to radiate inner scattered fields. This is why the geometry of an opti-

AVIONICS AND ASTRIONICS

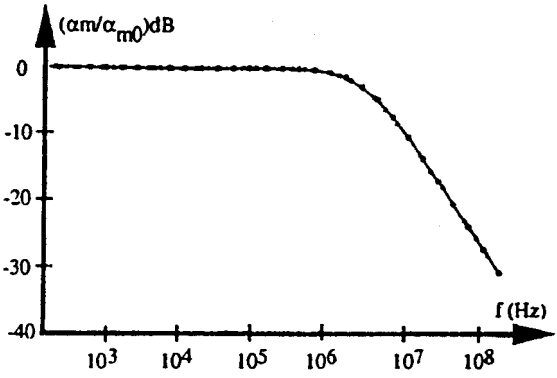


FIGURE 11.7 Ratio between the magnetic polarizabilities of a loaded aperture and a free aperture (from Bouderot and Labaune 1998). Magnetic/electric polarizabilities are linear coefficients defining magnetic/electric dipoles.

mized shield defined to protect a system must tend to keep the shape of a closed enclosure (generalization of the principle of Faraday cages).

EM Coupling on Cables. Because wiring is found everywhere in electrical systems, it is the most frequent cause of EMC problems. On the one hand, coupling on cables plays a particular role because cables are receiving antenna likely to transform the incident field in generators driving interference signals propagating at the equipment input. On the other hand, this interference propagating in clean zones may radiate undesired EM fields.

The other characteristic of wiring comes from its organization in bundles and branched harnesses. Therefore, there is electric and magnetic influence between wires in the same bundle. The so-called cross-coupling effect enables a perturbation on a wire to propagate from wire to wire.

To reduce radiation of cables or EM coupling of incident fields, bundles may be shielded, wrapping them in metallic screens. As seen before, because of the finite conductivity of the screen, the penetration of the magnetic field can never be totally stopped at low frequency. At higher frequency, a diffusion effect may appear in the depth of the screen. If the screen is made of a metallic coating, the scattering through small holes has also to be considered. Depending on the depth of the coating the diffusion effect may be hidden by the scattering through holes (Degauque and Hamelin 1993; Vance 1978). The penetration of the magnetic field is equivalent to applying a distributed voltage generator, linearly related to the external current I_{ext} circulating on the shield by the so-called transfer impedance Z_t. Associated with the penetration of the magnetic field is the penetration of electric field. The equivalent coupling model on the inner wires is a distributed current generator linearly related to the external common mode voltage V_{ext} developed on the shield by the so-called transfer capacitance C_t (see Figure 11.8). As seen above, the per-unit length circuit model supposes that the EM shield is correctly connected to the ground at both ends. In addition, the Z_t and C_t play a reciprocal role if one wants to determine the radiation of a shielded cable when the source of interference is on an inner wire.

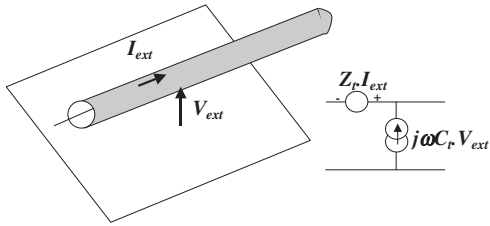


FIGURE 11.8 Equivalent model of EM coupling in a section of a shielded cable.

11.14 EM ENVIRONMENT AND EMC STANDARDS

In this subsection, we consider the external environment in which the system is likely to operate and the internal environment produced by the equipment of the system itself. Hereafter, we will present the most significant threats to account for in aeronautics and aerospace.

External EM Environment

Two types of external environments may be distinguished: natural and human-made environments.

Natural EM Environment. Natural EM environments such as lightning and electrostatic discharges (ESD) are threats which man has little power to avoid. The only possible action is to control their effects.

Lightning is a serious threat capable of leading to the destruction of the system. On aircraft, it is tolerated that the system is stressed by a lightning strike. The idea of protection is to maintain the evacuation of the injected current on the outer surface of the aircraft only. For instance, on radomes, lightning protection strips are installed onto the transparent material in such a way that the current does not flow into the antenna system. In the case of space launchers, the lightning strike is not tolerated in operation to avoid accidents. In addition, the launching pads themselves are secured with protections set all around the launchers. For example, ARIANE 4 and ARIANE 5 are protected by posts that deviate the possible lightning current outside the pad (see Figure 11.9).

The tests encountered in lightning standards are formulated in time domain. Different surge waveforms have been proposed describing the different phases of the propagating current in a lightning channel (harmonization documents from the Society of Automotive Engineers, SAE-4L). The frequency spectrum of all the waveforms is lower than 50 MHz, and the maximum threat has an amplitude equal to 200 kA (waveform A).

ESD is also a serious natural EMC threat. The general process is a local increase of static potential due to charge accumulations in different locations of the system creating difference of electric potentials likely to generate sparks. On aircraft, important problems occur on cockpit canopies where charges are deposited due to triboelectricity generating sparks on the canopy and in connectors of the heating circuits. In aerospace, ESD occurring on launchers when the stages made of dif-

AVIONICS AND ASTRIONICS

FIGURE 11.9 Lightning protection system on Kourou's ARIANE 5 launching pad (courtesy ESA/CNES).

ferent materials is always a relevant problem. In space, absolute potential voltages on satellites may generate typical current densities equal to 10 $\mu A/m^2$. Additionally, in the special case of geostationary satellites, the implantation of charges may lead to electric field surges up to 50 kV/m, with rise times lower than 10 ns. The waveform generated is variable and depends on the polarity of the discharge. Standardized tests to represent ESDs are not always available. Even if standards exist for human-origin ESD (CEI standards), there are no real identified standards for ESD on satellites or aircraft. However, on satellites the project of ISO (International Organization for Standardization) 14302 seems to be in the process of being accepted.

Human-Made Environment. The human-made EM environment is also likely to constitute a serious threat. Some threats are intentional and must be considered as generated by real weapons against which all military systems have to be protected (MIL-STD-461D, 462D, 464). Here we will mention two of the most important intentional threats: EMP and HPM.

EMP stands for the *EM pulse* generated after an atmospheric or extraatmospheric nuclear explosion (Lee 1980). The standardized threat is expressed in time domain with electric field pulses up to 100 kV/m and rise times lower than 10 ns. The frequency spectrum of this threat extends up to 100 MHz.

HPM stands for *high-power microwaves* and is associated with new types of weapons that appeared at the beginning of the 1990s. Here again, the threat is a pulse but the frequency content is larger, up to several GHz. The difference with

EMP is that the generation of power requires focusing antenna and the magnitude strongly depends on the distance of the source and the capability of available technology. This is why, up to now, this threat has not been totally standardized.

Other threats are unintentional and are mainly due to high-intensity fields created by communication systems. For the civil world, such threats are standardized under the name *high-intensity radiated fields* (HIRFs). They describe the environment created by high-intensity antennas and radar as the ones likely to be encountered in the vicinity of airports or on ships. Consequently, they concern aircraft and space launchers. In common civil EMC standards, the constraint imposed on the standardized external field is as large as 200 V/m for the electric field with a frequency spectrum ranging from DC to 18 GHz (RTCA DO 160 and documents from the SAE-4R subcommittee). The demonstration of those standards is not obvious because it is quite impossible to generate a plane wave illuminating a whole object with enough level in this frequency range. Illumination of parts of the object is generally the only available solution.

Internal EM Environment

First, systems must also comply with the internal environment generated by other systems. The first elements to consider are the on-board antennas, which can generate unwanted fields in the direction of their side lobes, for example. On satellites, sensors such as altitude sensors used to work with electric fields about 10 V/m may be perturbed by fields generated by on-board antenna up to 100 V/m.

Secondly, systems that are not made to transmit fields may also be significant interference generators. This is the case with power supplies producing coherent rays on both power and ground network. We think here of uninterruptible power supplies (UPS) and regulation of chopping, widely used for the efficiency and low weight they provide. On satellites, an EMC margin is imposed on modulators and emitters to tolerate a noise of about 1 V RMS on all power inputs. In addition, to reduce common mode in the power networks, a maximum limit of the resistance of the bars is specified and is obtained by increasing the section of the bars.

Finally, the problems generated by devices external to the system itself, but likely to interact with it, must be mentioned. This is the well-known case of portable electrical devices (PEDs), for which, so far, the only solution has been to forbid their use on aircraft.

11.15 EMC TOOLS

Experimental Tools

Experimental tools are necessary to demonstrate the standards. For radiated susceptibility, radiated antennas are required to generate the field levels imposed. At low frequency, the size of the antenna is large but high power is made available through the use of pulsed generators. For example, in EMP the size of simulators is much greater than the size of the systems under test.

For the special case of lightning on aircraft, the technique of coaxial injection is commonly used to simulate a uniform circulation of the currents on the surface or a part of the surface. The interest of the setup is that the current injected and

especially the return current can be controlled and considered as mainly symmetric. Nevertheless, it is obvious that this test modifies the circulation of the currents on the aircraft compared to a real lightning injection.

At higher frequencies, fields are generated by smaller antenna with more localized effects because the whole power cannot be applied on the whole structure. For small systems, radiated tests may be performed in anechoic chambers, but these are generally more appropriate for EM emission tests.

For conducted susceptibility, common mode currents are injected into the inputs of equipment. Generally the impedance matching is provided by an LISN (line impedance stabilization network). Current injectors are transformer-like devices enabling an equivalent voltage generator to be forced in wires or bundles of wires and therefore induce a current in the equipment.

Numerical Tools

Numerical tools have become unavoidable in aeronautics and aerospace whenever EMC is concerned. Indeed, the progress made in the last 10 years by computers now makes possible the running of large computer codes requiring large memory and calculation resources. Modeling is generally used at the design phase as an efficient method for analyzing the influence of different parameters on the system's response. In that sense it also helps to optimize the tests to be performed. However, EMC computer codes are also beginning to be used for their prediction capabilities. Therefore, they may eventually replace expensive and sometimes impossible-to-achieve tests at the qualification phase. This was the case for the Airbus A340, which was too large for full standardized lightning injection tests to be carried out and for which numerical demonstrations were made.

In the following, we will present the different types of calculation techniques available in EMC.

3D EMC Computer Codes. Three-dimensional (3D) codes describe the geometry of the system and solve Maxwell's equations. Necessarily, this description is approximated with a mesh sampling the geometry and simplified because accounting for the detailed geometry is quite impossible with the capabilities of present-day computers. These computer codes are now considered fully reliable for what is called the external problem, that is, the scattering of fields and currents by the outer surface of the system. However, up to now they have not really been applicable at very high frequencies (typically frequencies larger than 1 GHz on an aircraft and satellite). Of course, it will be impossible here to investigate all the available methods, and thus only the two families commonly used are mentioned.

First are volume methods, which enable dielectric and losses in materials to be described. They require the meshing of the entire calculation volume and the simulation of limit conditions or infinite medium with absorbing conditions. These methods are frequently developed in time domain, which means they offer a wide-frequency spectrum analysis with a single pulse. The most spread-out method is the finite difference time domain (FDTD) method, valued for its robustness and simplicity of implementation (Degauque and Hamelin 1990). The problem with this method is that the mesh made of cubic cells prevents a conformal description of the surfaces (see Figure 11.10).

The second type of 3D methods in EMC is surface methods based on the resolution of Maxwell's equations in their integral formulation (Degauque and

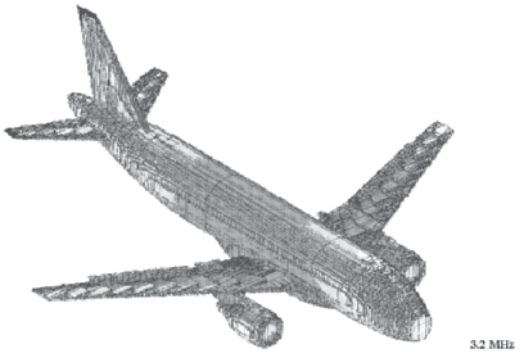


FIGURE 11.10 Current distribution on a Boeing 737 illuminated by an electromagnetic plane wave.

Hamelin 1993; Tesche et al. 1997) such as the method of moments in frequency domain. The interest of those techniques is that the shapes of objects on which scattered currents are calculated are described precisely and only the surfaces have to be meshed. Nevertheless, the drawback of those methods is that they require large amounts of memory that limit the calculations at high frequencies.

Cable Network and Circuit Codes. The drawback of 3D codes is that they are not able to handle the complexity of EMC cable bundle problems. Since the 1980s, several computer codes based on multiconductor-transmission-line-networks techniques allow this problem to be handled (Baum et al. 1986; Parmentier and Degauque 1996; Paul 1994). Both time domain and frequency domain techniques are available, but the latter offers the main advantages of accounting for the frequency dependence of transmission line parameters and provides models of cables independent on their length. Thanks to field-to-transmission line formalism (Lee 1980), it is possible to link those codes with 3D codes. The 3D codes calculate the distributed incident fields on the wiring path in the absence of the wiring; these wires are then used as generators for the cable network code. Compared to a 3D code, the calculation of a cable network code is fast because the network matrix is sparse. In addition, with the help of appropriate signal processing, the link between a 3D method in time domain and a cable network code in frequency domain is a very efficient technique (Parmantier and Degauque 1996).

If required, cable network codes may be linked to circuit codes to calculate the response at complex terminations. For this purpose, many SPICE-oriented computer codes, validated for a long time in other electrical domains, are available. The compaction with Thévenin equivalent (Parmantier and Degauque 1996) may be used to complete the effort to decompose the problem in subproblems as it is suggested in the theory of EM Topology (Baum 1990). Therefore, the decomposition is achieved in three layers: the incident field illumination with a 3D code, the propagation on the wiring with a cable network code, and the equipment response with a circuit code.

11.16 ENGINEERING METHOD

The EMC Control Plan

In industrial programs, the organization of EMC activities is under the responsibility of an architect and supervised by an EMC manager who is in charge of applying the EMC control plan. This plan covers all the activities in this domain, from the supply of equipment to the delivery of the system to the customer. The main purpose of this plan is to distribute the contract responsibilities to all the parties involved. EMC being the art of defining interfaces, it is very important to define the rules before conflicts happen. Therefore, the control plan also describes the way to manage nonconformities. Nevertheless, from the beginning of the program and the consultations for equipment purchase, it must specify the interface constraints in terms of EM emission limits, EM susceptibility, realization rules, and test methods.

EMC Specification

A General EMC Specifications document must contain all external EM constraints due to natural environment, to other systems, or to the system itself for autocompatibility.

In aeronautics, SAE-4 committees, RTCA Special Committee SC135, and EUROCAE working groups provide reference information on internal EMC on aircraft.

Up to now, there has been no real dedicated general standard for aerospace as in aeronautics. Nevertheless, in both domains EMC specifications have to be defined for each program to answer the technical clauses of the contract signed with the architect. Indeed, the architect generally has its own internal standard, coming from his experience and for which EMC specifications are mostly duplicated from one program to the other.

To maintain compatibility between systems, EMC margins are applied. By definition, this is the ratio between the susceptibility level to be demonstrated in a critical point of the system and the real perturbation level at this point. In the worst case, at least a 0 dB margin must be demonstrated, which means the perturbation level is equal to the susceptibility level. In aerospace, a 6 dB margin is commonly accepted on ground test results to account for the different conditions occurring in flight, but higher margins (20 dB) are required for sensitive equipment (pyrotechnics). The most reliable demonstration method consists of reproducing the perturbation signal in a critical point of the whole system and applying it once amplified by a factor equal to the margin required. This method has replaced old techniques consisting of the comparison of emission and susceptibility plots obtained on each piece of equipment separately. Indeed, because of the difficulty of accounting for nonlinearity and the combination of time and frequency characteristics, the incoherent summation of emission and susceptibility levels generally leads to incorrect conclusions.

EMC Test Plan

The methods for demonstrating the good behavior of the system are based on tests and analysis. They are described in the EMC test plan. Because of the market

competition, the reduction of development and manufacturing costs is the constant leitmotif of the architect. Therefore, the number of expensive tests, such as in-flight tests on aircraft, must be optimized or ground tests on ready-to-fly satellites performed only if strictly required. It is now common for no mock-up or prototype to be available to perform the design phase of a system. This is why, in aerospace, the minimum test plan is limited to the measurement of radiated fields in the launching configuration (EMC with the launcher) and the tests on electrical functions for telecommunications devices. Sometimes only a specific test for ESD susceptibility can be carried out.

Compared to aircraft, the difficulty for satellites and launchers is that it is very difficult to reproduce representative in-flight conditions. For example, the energy on a satellite is provided by solar panels. In tests, it is impossible to spread out this assembling of some 50 m^2 of photovoltaic cells and reproduce the illumination of a solar spectrum of 1,350 W/m^2. Consequently, a simpler laboratory EM source is used instead and the test is performed on an stand-alone satellite powered by its batteries only. The test is carried out in large anechoic class 100,000 clean rooms (see Figure 11.11). However, the connections between the satellite and the test bench are likely to modify the EM interface.

Additionally, for aircraft, some similar situations occur, such as for lighting. Even if, in this case, standards specify levels of current to inject in the structure, the generation of a real lightning channel connected to the structure is impossible.

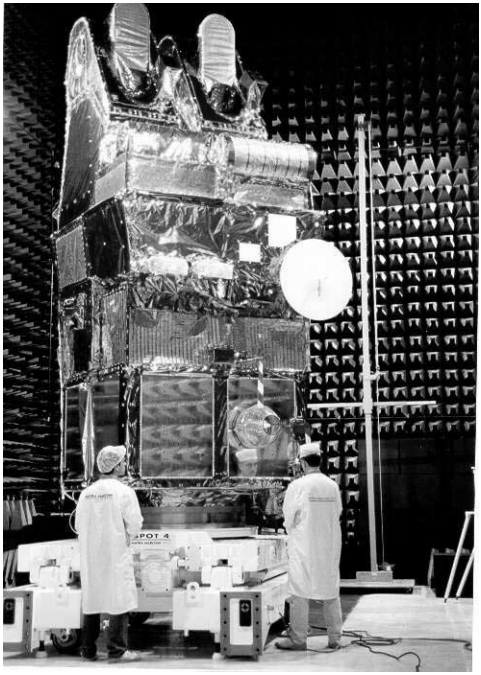


FIGURE 11.11 SPOT-4 satellite in an EMC test at INTESPACE (Toulouse, France, courtesy CNES/Renaut Claria).

Manufacturing Rules

Another document applicable to all the parties involved in a program defines the general rules for electrical manufacturing. The equipment constraints more than meet EMC needs. Some of these concern topics other than electricity, such as mechanical or thermal architecture or the integrity of on-board data transfer. Nevertheless, respect for EMC rules should theoretically avoid any redefinition. This document, written by the architect, collects all the conception and electrical manufacturing constraints. It covers very different topics, including electrical continuities, grounding of electric boxes or cable shields, installation of connectors, choice and position of cables, description of the interface circuits of power supplies, and the impedence and noise level of digital/analog data transfer devices.

On a satellite, it is frequent for the distributed power to reach 10 kW. For the near future, 20 to 30 kW are indicated. Primary currents overcome 200 A. The choice of the technology results from a deal between voltage losses to be maintained at the lowest level for both power saving and thermal dissipation, inductance reduction to minimize overvoltages at the switch-on and switch-off phases, and control of the magnetic moment up to a given level.

Even if aerodynamics constraints are not always relevant, all systems have to dominate their weight. Thus, whenever possible, the natural EM screen offered by the surface of the structure must be balanced with the distribution of cable shields and filters at equipment inputs.

11.17 CONCLUSION

In this Part, we have introduced the main concepts of EMC. We have first presented the theoretical concepts required to understand the basics of EM coupling phenomenon. We have then focused on aircraft and aerospace analyzing their EMC environment and associated standards. Experimental and numerical techniques used for EMC design of systems have been mentioned. Finally, the way to account for EMC in an industrial program has been presented.

Up to now, even though EMC is identified as a requirement impossible to miss in industrial projects, accounting for it at the very beginning of the project is not so frequent. Nevertheless, industry is at a crossroads where new progress in EMC modeling is likely to reduce dramatically experimental tests and offer a wider set of configurations. Manufacturers are already working on what they call ''electric mock-ups,'' on which EMC calculations should be systematically applied for the design of the actual system.

In this increasingly electronic world, manufacturers realize the importance of EMC for the quality of their products. In parallel, EMC should become a real electromagnetic discipline taught in engineering schools in the same way as classical disciplines such as antenna, microwaves, and electricity.

REFERENCES

Baum, C. E. ''The Theory of Electromagnetic Interference Control,'' in *Modern Radio Science 1990,* ed. J. B. Anderson, Oxford University Press, Oxford [also in *Interaction Notes,* Note 478, December 1989].

Baum, C. E., Liu, T. K., and Tesche, F. M. 1986. In *Fast Electrical and Optical Measurements,* ed. J. E. Thompson and L. H. Luessen, Nijhoff, Dordrecht, pp. 467–547.

Boudenot, J. C. and Labaune, G. 1998. *La compatabilité électromagnétique et nucléaire,* Ellipses, Paris.

Degauque, P. and Hamelin, J. 1993. *Electromagnetic Compatability,* Oxford University Press, Oxford [French edition: *Compatabilité électromagnétique,* Dunod, Paris, 1990].

EUROCAE ED-1 D/RTCA DO 160D, *Environmental Conditions and Test Procedures for Airborne Equipment.*

Lee, K. S. H. 1980. "A Complete Concatenation of Technology for the EMP Principle: Technique and Reference Data," *Interaction Notes,* December.

MIL-STD-461D, *Requirements for the Control of Electromagnetic Interference Emissions and Susceptibility*.

MIL-STD-462D, *Measurement of Electromagnetic Interference Characteristics*.

MIL-STD-464, *Electronic Environment Effects for Equipment.*

Parmentier, J. P. and Degauque, P. 1996. "Topology-Based Modeling of Very Large Systems," in *Modern Radio Science 1996,* ed. J. Hamelin, Oxford University Press, Oxford, pp. 151–177.

Paul, C. R. 1994. *Analysis of Multiconductor Transmission Lines,* John Wiley & Sons, New York.

Stratton, J. A. 1941. *Electromagnetic Theory,* McGraw-Hill, New York [French edition: *Théorie de l'électromagnétisme,* Dunod, Paris, 1961].

Tesche, F. M., Ianoz, M. V., and Karlsson, T. 1997. *EMC Analysis Methods and Computational Models*, John Wiley & Sons, New York.

Vance, E. 1978. *Coupling to Shielded Cables,* Wiley-Interscience, New York.

PART 5
INTRODUCTION TO RADAR

Florent Christophe

11.18 HISTORICAL BACKGROUND

The capability of detecting moving objects in the environment of antennas radiating electromagnetic waves was anticipated early after the experimental evidence of such waves, thanks to Heinrich Hertz in 1886. Practical applications of radar (radio detection and ranging), i.e., the detection of vehicles and estimation of their position were demonstrated as early as the 1920s in France and Germany. The development of radar was then made possible by the availability of a microwave pulsed power oscillator, the magnetron (a derivative of which was mass produced starting in the 1970s for microwave ovens).

The importance of radar was fully recognized following its role in defending Great Britain against air attacks during World War II. Much public and industrial research funding was then invested in the United States and various European countries for solving technological issues and bringing radar to its present wide range of applications. These applications, which include geophysics, meteorology, remote sensing, air traffic control, and acquisition of military targets, come from the long-range all-weather sensing capability of radar.

11.19 BASIC PRINCIPLES

If the beam radiated by a directive antenna such as a parabolic dish illuminates an object, some of the incident energy is backscattered and can be detected in a receiver connected to this same antenna. The direction of the beam and the round trip time delay between the transmitted and the received waveform—usually a repetitive pulse—are used to estimate the location of the scattering object.

Radar Equation

Let P_t be the power of the transmitter, connected to a directive antenna which has a gain G_t. The power density of the field incident on a target, at range R, is

$$D_i = \frac{P_t g_t}{4\pi R^2} \tag{11.21}$$

The interaction with the target is described through an equivalent radar cross-section (RCS), defined as a collecting surface σ that would isotropically reradiate all the received power. This collected power is then written as:

$$P_i = \sigma D_i \tag{11.22}$$

The incident power density D_i can be used in (11.9) for computing the power that has been reradiated to the receiving antenna:

$$P_r = P_i g_r (\lambda/4\pi R)^2 \tag{11.23}$$

where g_r is the gain of the receiving antenna. Combining (11.21)–(11.23) and assuming the same antenna is used for transmit and receive (i.e., $g_t = g_r = g$) result in a relationship known as the radar equation:

$$P_r = \frac{P_t \lambda^2 g^2 \sigma}{(4\pi)^3 LR^4} \tag{11.24}$$

where L is a loss factor accounting for various effects such as transmitter-to-antenna or antenna-to-receiver connection losses, propagation losses, etc.

For a radar pulse of duration τ, (11.24) can be used to derive the received energy $P_r \cdot \tau$, and the signal-to-noise ratio after dividing this received energy by the spectral power density of the receiver noise FkT_0 where F is the noise figure with respect to the reference noise temperature T_0 (usually 300 K), k being Boltzmann's constant of value 6.02 10^{-23} J/K:

$$S/N = \frac{P_t \tau \lambda^2 g^2 \sigma}{(4\pi)^3 LR^4 FkT_0} \tag{11.25}$$

Radar Cross-Section

Electromagnetic theory allows for a better understanding of the physics over the previous definition of radar cross-section, which was simply used for computing the power flow from transmitter to target and back to the receiver.

Being a linear process, the scattering of an incident electromagnetic wave by any object of dielectric properties deviating from the surrounding medium might be characterized by a scattering matrix S connecting incident field vector E_i to scattered field vector E_s through the equation

$$(E_s) = (S)(E_i) \tag{11.26}$$

The fields appear as vectors when projected on a polarization basis (usually vertical/horizontal, but also right circular/left circular or some other projection basis). First, diagonal terms of S correspond to copolarization scattering; second, diagonal terms to cross-polarization scattering.

In the far field of the scattering object, i.e., for range

$$R > 2D^2/\lambda \tag{11.27}$$

where D is a typical dimension of this object, E_s vanishes as $1/R$, and it can be shown that:

$$\sigma = 4\pi R^2 S_{rt}^2 \tag{11.28}$$

where S_{rt} is the term of the scattering matrix S corresponding to the transmit and receive polarizations.*

Depending on the shape and dielectric properties of the scattering object, various methods of solving Maxwell equations are available for computing the scattering

*A polarimetric radar would be sensitive to more than one term of the scattering matrix (S), adding characteristic information concerning the target.

matrix in the far field, hence the radar cross-section of targets. We will next indicate useful results for some canonical targets.

For a sphere of radius r large with respect to the wavelength, a high-frequency approximation, allows the derivation of

$$\sigma = \rho \pi r^2 \tag{11.29}$$

where ρ is the power reflection coefficient at the sphere surface: for a metal sphere, $\rho = 1$ and σ is then equal to the physical cross-section of the sphere.

For a metal plate of surface S, each dimension being large with respect to the wavelength:

$$\sigma = 4\pi \cdot S^2/\lambda^2 \tag{11.30}$$

when the incident wave is perpendicular to the plate (the so-called specular reflection situation) and vanishes rapidly for angles departing from perpendicular, as most of the energy is reradiated away from the incident direction. A smoother behavior, allowing for the design of passive radar calibrators, is obtained through corner reflectors. In such device, triple successive reflections combine into an outgoing ray parallel to the incident one whatever the orientation of the reflector, resulting in a quasi-isotropic behaviour like that of the sphere, but at higher RCS for the same physical cross-section. Table 11.9 indicates typical values.

Real targets behave quite differently from the above due to the combination of many scattering mechanisms, depending on the shape and material of the target, the wavelength, polarization, and the direction of the illuminating wave.

A simplified model for understanding complex target backscattering adds the contributions of canonical scatterers distributed across the skin of the target. The rapid change of relative phase of those scatterers with respect to frequency or angular presentation results in a noise-like appearance of respective diagrams, somehow similar to the measured ones.

More accurate modeling of the backscattering mechanisms of real targets makes use of multiple interactions, creeping waves, waveguide modes for air intakes, etc.

Clutter and Other Environmental Effects

The beam radiated by a radar antenna often illuminates the natural environment, which results in backscattering energy that may compete with possible targets that the radar is searching for at the same ranges. Main sources for such so-called clutter effects are surface scattering by irregularities of the soil—either bare or vegetated—and volume scattering, mostly coming from rain cells. As opposed to coherent scattering mechanisms responsible for target backscattering, natural environment backscattering is mostly built up by superposition of a large number of individual

TABLE 11.9 Radar Cross-section for Three Canonical Targets of 1 m^2 Physical Cross-section

	Sphere	Circular plate	Trihedral
Peak RCS @ 1 GHz	1 m^2	140 m^2	140 m^2
Peak RCS @ 10 GHz	1 m^2	14,000 m^2	14,000 m^2
Angular behavior	Isotropic	HPBW 10° @ 1 GHz, 1° @ 10 GHz	Quasiisotropic

scattering contributors with random position, i.e., making it an incoherent process in which the backscattered energy will be proportional to the surface or to the volume of the illuminated area, respectively. Since this backscattering energy can be associated with an equivalent radar cross-section through the radar equation (11.24), the ratio of such RCS to the illuminated surface or volume is defined respectively as the surface or volume reflectivity. Such reflectivity usually increases with various factors, mainly:

- For ground clutter: soil roughness, vegetation density, angle of incidence, frequency
- For atmospheric clutter: rain rate, frequency

Other Environmental Aspects. The interaction of radar waves with natural media, reported as clutter when negatively affecting target detection, may also be directly used in remote sensing techniques for earth resources or geophysical parameter characterization.

Other effects of wave propagation are to be expected on radar signals; for example, for long-range ground radar, atmospheric attenuation, tropospheric refraction, and interference of direct and ground surface-reflected rays need to be considered, or ionospheric effects for spaceborne radar observation of the earth. Part 1 of this section gave indications and references dealing with these effects.

Principles of Clutter Rejection. In many configurations, the clutter RCS is stronger than the RCS of expected targets, resulting in the risk of blinding the radar. Improvements come from exploitation of the Doppler effect,* which creates a frequency shift of the backscattered signal proportional to the radial velocity of the target (i.e., the projection v_r of the velocity on line of sight) according to the formula:

$$f_d = 2v_r/\lambda \tag{11.31}$$

Since the signals backscattered by clutter are almost stationary, they can be filtered out by zero Doppler rejection, which is achieved through the moving target indicator (MTI). Target signals remain unaffected to a certain extent as far as they do not approach zero Doppler, which could be the case for targets with tangential velocity, or blind velocities resulting from undersampling of Doppler frequencies. This effect will be illustrated below.

A further opportunity for reducing clutter effects is through exploiting the polarization properties of clutter backscattering; for example, quasispherical rain droplets produce backscattering almost cross-polarized with respect to a circularly polarized incident wave, and therefore most radars which are likely to operate in front of strong atmospheric clutter use circular polarization.

Detection Performances

When considering the received radar signal, one of two hypotheses will apply: either it is noise only, or, the addition of noise and signal backscattered by a target.

*The Doppler effect may be understood as the phase rotation induced by enlarging or shortening the round trip from radar to target and back when the radar and/or the target is moving.

For making a best guess, it can be shown that the optimum use of the received energy is matched filtering (i.e., processing by a filter tailored to the transmitted signal for minimal noise bandwidth) followed by thresholding. If the filter output is higher than the threshold, a target is detected. The performance of such processing is evaluated by the probability of correctly detecting a target (probability of detection) against the probability of erroneously declaring a target when in fact only noise is present (a false alarm).

The probability of a false alarm depends on the ratio between the threshold level and r.m.s noise level and the statistics of the noise, white Gaussian for thermal noise, or more complicated when clutter is added to Gaussian noise. Once the threshold is determined, the probability of detection might be computed starting from the signal-to-noise ratio and the statistics of the signal, derived from the fluctuation of RCS around its mean level.

For most situations, many successive pulses hit the targets and can contribute to improving the detection performances. The best case is that of coherent integration before thresholding, for which an improvement factor n in the signal-to-noise ratio is obtained, where n is the number of available pulses. When such coherent or Doppler processing is not feasible, noncoherent (or postprocessing) schemes are used, resulting in some degradation with respect to the optimal case.

Resolution, Accuracy, and Ambiguity of Radar Measurements

The resolution width of a radar is defined as the range of one of the parameters used for locating a target (distance, angular position, Doppler velocity when applicable) for which the output of the processing filter remains larger than or equal to half the peak power corresponding to the exact location of a target. With such a definition, two targets of equal RCS can be distinguished when they are separated by more than one resolution width. A resolution cell is the multidimensional patch having the resolution width along each axis. A high-resolution radar will have a resolution cell smaller than the targets (usually along the distance or Doppler velocity axis), thus allowing the analysis of target features such as its length. Information derived from such high-resolution radar is the basis for automatic target recognition.

Angular Resolution. Let $\Delta\theta$ and $\Delta\phi$ be the half-power beamwidths of the antenna in the horizontal and vertical planes. Due to the two-way effect, the half-power beamwidth of the radar—defining its angular resolution—is then $\Delta\theta/\sqrt{2}$ and $\Delta\phi/\sqrt{2}$.

Range Resolution Along the distance axis, when dealing with a radar pulse of duration τ, advancing or delaying the processing gate by $\tau/2$ from the exact round-trip time delay of the target will result in getting half the maximum power (only half of the signal is available), and the corresponding time resolution width is therefore τ. The related range resolution width taking into account the round trip is then

$$\Delta R = c\tau/2 \tag{11.32}$$

But we also have to consider the so-called pulse compression technique, for which the radar pulse of duration τ is modulated with a bandwidth Δf large with respect to $1/\tau$ ($k = \tau\Delta f$ is called the compression factor, being equal to 1 for nonmodulated

pulse). It can be demonstrated that matched filtering results in the resolution that would give a pulse of duration τ/k.* A formula which remains valid for both nonmodulated and pulse compression cases is therefore:

$$\Delta R = c/2\Delta f \tag{11.33}$$

Doppler Resolution. In Fourier analysis, when duration T is available for observing signals, a frequency resolution $1/T$ is achievable. Since the Doppler frequency is related to radial velocity through equation (11.31), the velocity resolution width can be derived:

$$\Delta v = \lambda/2T \tag{11.34}$$

where T is the duration available for coherent processing, smaller than or equal to the duration of the transmitted beam illuminating the target.

Accuracy of Radar Measurements. Some radars (e.g., tracking radars) are designed for achieving accurate location of the target; it can be shown that suitable processing, such as interpolating from overlapping resolution cells, allows for a standard deviation related to the resolution width by equation:

$$\varepsilon p = \Delta p/2\sqrt{S/N} \tag{11.35}$$

where p is one of the parameters to be measured (angle, distance, radial velocity). A signal-to-noise ratio of 13 dB, which is suitable for target detection, would then allow an improvement of a factor around 9 for estimating the parameter p when compared to the corresponding resolution width.†

The Ambiguities of Radar Measurements. A usual radar waveform is a pulse, the duration of which has to be short enough to overcome detrimental blinding of the receiver during transmission.§ Targets located closer to the radar than the blind range R_b are undetectable:

$$R_b = c\tau/2 \tag{11.36}$$

But a large integration time T is also needed for Doppler separation of targets from clutter and for improved signal-to-noise. The obvious solution is to repeat the initial short pulse, with repetition period T_r, resulting in n identical pulses received during $T = nT_r$. Such repetition introduces, on one hand, the risk of erroneously referring the received signal to the last transmitted pulse when actually it comes from a previous one; the error in distance would be then a multiple of the ambiguous distance R_a.

---

*This pulse compression technique allows for separating the power balance problem (a larger pulse duration gives a larger usable energy for a given available peak power from the transmitter) and the range resolution problem. Among others, classical modulations for pulse compression are swept frequency or pseudo-random biphase or polyphase modulations.

†In tracking radars, other improvements can be used when switching from a detection to a tracking mode: larger integration time and postprocessing such as Kalman filtering. A multiport antenna with associated receivers also allows for monopulse processing, which produces the angular accuracy indicated above without negative impact of the target RCS.

§Some radars make use of a continuous waveform (or CW), but these need separate transmit and receive antennas with specific attention paid to their decoupling.

$$R_a = cT_r/2 \tag{11.37}$$

On the other hand, spectral analysis techniques conclude that sampling a signal with period T_r introduces a repetition in its spectrum of period $1/T_r$; such repetition creates an ambiguity in radial velocity (or Doppler) analysis determined by:

$$v_a = \lambda/2T_r \tag{11.38}$$

Searching for long-range high-velocity targets implies ambiguities in either distances or velocities, and postprocessing will have to overcome this situation.*

11.20 TRENDS IN RADAR TECHNOLOGY

Improved performance and reduced life-cycle costs are driving factors for injecting new technologies in radar design. Next we will discuss some of the technological issues which are specific to radar.

Transmitters

The first generation of pulsed power microwave tube, the magnetron, was able to deliver megawatts of peak power for a few microseconds, the primary energy being stored in the modulator and delivered as a high-voltage pulse. Due to its self-oscillating behavior, pulse-to-pulse coherency for Doppler processing was difficult to achieve, and next generation of tubes, such as the klystron or the traveling wave tube, are amplifiers permitting longer pulses for pulse compression together with coherent integration. Such vacuum tubes rely on electron beams interacting with microwave cavities. But they suffer from the need to handle high voltages and use heated cathodes of limited lifetime. Furthermore, the handling of high peak powers may result in the need for pressurized transmission lines to the antenna for avoiding dielectric breakdown.

New-generation transmitters are therefore relying increasingly on solid state amplifiers through the association of single transistors able to deliver from a few watts to a few 100 watts of peak power, depending on frequency.

According to the scheme given below, specific components are to be associated with the radar transmitter for avoiding destruction of the low-noise receiver by part of the transmitted power: a circulator, which is a three-port device with a nonreciprocal ferrite core, allows transfer of more than 99% of the energy from the transmitter to the antenna, while energy coming from the antenna is transferred to the receiver. Depending on the transmitted power, further protection of the receiver might be brought by limiting diodes, possibly combined with a plasma switch. Insertion loss of those devices has to be kept as low as possible (less than 1 dB is currently achieved).

* The choice of a high pulse repetition frequency (or PRF) will make such a radar ambiguous in range, a low PRF will make it ambiguous in Doppler, whereas a medium PRF will make it ambiguous in both range and Doppler. The switching of one PRF to another within the same burst of pulses illuminating a target allows this ambiguity issue to be solved.

Antennas

The classical radar antenna is a parabolic dish fed by a horn at its focus, rotating around one axis for panoramic surveillance or two for target tracking. Improvements come from multiple reflectors, such as the Cassegrainian assembly for compactness (which is required for airborne radar) and sidelobe control. Close to the focal plane, multiple feeds are of interest: either for stacking beams at various elevation angles in a panoramic radar or for simultaneously receiving a single target echo in slightly separated beams for improving angular measurements in a tracking radar. Flat antennas built with slotted waveguides stacked together and suitably fed by a waveguide distribution of energy are now widely used for applications where weight or rotating inertia is constrained. In such an array antenna, each slot behaves as an individual radiating element whose contribution combined with that of the others builds a far field equivalent to what a parabolic dish of same aperture would produce.

But a breakthrough in radar antennas has been brought by electronic scanning, in which the mechanical rotation of the antenna is replaced by phase shifting of individual subsets of an array antenna. Electronic scanning might be achieved in one direction by a one-dimensional phase variation—each waveguide of the previous array antenna is fed through a phase shifter—or in two directions, which requires addressing each of the radiating elements individually. In such two-dimensional scanning arrays, the individual radiating elements can be dipoles, open waveguides, or metallic patches on a dielectric substrate. Phase shifter technology relies on either solid state diode switches or ferrite transmission lines, the insertion phase of which is modulated by an externally applied magnetic field. Insertion losses, power handling capacity, and accuracy are the main parameters for selecting these key components of an electronic scanning antenna.

An active array is an electronic scanning antenna combined with the splitting of the single high-power transmitter into many individual low-power solid state elements. These elements can be integrated into hybrid or even monolithic integrated circuits and brought together into a single transmit/receive module power amplification, phase shifting, switching, low noise amplification, and filtering for reception. Such transmit/receive modules are likely to be the core of most future radar systems as soon as low-cost production technology is available.

Receivers

The role of receivers is to transfer a very faint signal embedded into much higher parasitic signals (coming from clutter or even jammers) to a digital signal processor. Low-noise preamplification is now state of the art, with noise figures as low as 1 to 2 dB at frequencies ranging from 1 to 18 GHz. Frequency down-conversion can be performed with high dynamic range mixers, and bandpass filtering prepares for baseband down-conversion in both phase and quadrature for digitizing without loss of the phase. In an alternative scheme, a single channel is digitized with a residual carrier, and separation of real and imaginary parts of the complex signal is performed through numerical Hilbert filtering.

Due to the high dynamic range of signals resulting from the power -4 dependence with range—i.e., time delay referred to the transmitted pulse—compensation can be accomplished by applying an inversely varying gain prior to the final amplification or digital conversion. According to the highest clutter-to-noise or

jammer-to-noise ratio which can be expected, from 1 bit to 16 bits or even more analog-to-digital converters are required; clock rates are slightly higher than twice the radar bandwidth, the exact value depending on the roll-off of the antialiasing filter. In a radar using linear shift of the carrier frequency during the pulse (also called "chirp" radar), pulse compression can be performed by a surface acoustic wave or a bulk acoustic wave device having delay-versus-frequency characteristics inverse to the transmitted pulse. In such a device operating at intermediate frequencies of 50 to 150 MHz and time delays up to 100 microseconds, the design of the piezoelectric transducers allows tailoring of the required characteristics.

Signal Processors

The digital signal processor is in charge of performing tasks as various as pulse compression if not performed in the analogic part of the receiver, digital beam forming for a phased array antenna, Doppler filtering or MTI, thresholding or plot extraction, tracking, plot-to-track association, and display management. A popular radar display is the plan polar indicator (PPI), which creates a horizontal projection of the scene surrounding the sensor, according to the azimuth rotation of the antenna and the range of the echoes detected in the beam direction.

The computing load for the most demanding of functions may range up to many hundreds of floating point operations per nanosecond (or gigaflops). The availability of general-purpose digital signal processors approaching such figures makes it less and less necessary to develop costly on-purpose processors.

11.21 RADAR APPLICATIONS TO AERONAUTICS

Air Traffic Control

One of the main applications of radar to aeronautics is air traffic control (ATC). Long-range air surveillance is performed from selected national locations for controlling continental parts of major air routes. The associated radar sensors operate in the 1215–1400 MHz band, with large antennas (typical reflectors are 8×6 m, uniformly rotating around 6 rpm and powerful transmitters in the 100 kW range allows detection of general aviation and jetliners beyond 200 nm. Detection plots and the associated tracks built through consecutive illuminations of the target by the rotating antenna are made available to controllers by video or synthetic displays.

But beyond the cost of such a large radar (which makes it difficult to install in developing countries) there are also some major technical drawbacks:

- The accuracy of the elevation estimate at long range is much poorer than required for controlling vertical separation of routes of 2,000 or even 1,000 ft.
- The radioelectrical horizon at 200 nm is about 25,000 ft (see Part 1), which means that only jetliners close to their maximum flight level can be detected at such range.

Therefore, other sensors are required for filling those gaps, such as the so-called secondary radar. A secondary surveillance radar (SSR) can be understood as replacing the radio wave backscattering roundtrip with two single trips: on board

passenger aircraft or for flying under IFR conditions, a transponder is placed which detects the incident radar signal and retransmits a similar signal. This results in a power balance much easier to achieve in each single trip, through the R^{-2} telecommunications equation (11.12) instead of the R^{-4} radar equation (11.24) for round trip. The required signal-to-noise ratio is therefore reached for lower transmit power and antenna gain, resulting in reduced cost for the radar,* which can be introduced in a given territory with higher density than the previously described system, which we will now call *primary radar*. Due to technological constraints, the transponder has to receive and transmit at slightly different frequencies and with additional time delay. These parameters, which need to be accounted for in the secondary radar receiver and processor, are normalized according to ICAO regulations. In addition, the retransmitted pulse can be encoded with a message indicating the flight number (A mode) and on board measured flight level, (C mode) which allows accurate three-dimensional tracking from the ground.

In areas of high-density traffic and with possibly few secondary radars interrogating various transponders nearly at the same time, garbling situations may occur. Overloading either the transponders or the radar receivers with quasisimultaneous pulses results in track losses. To overcome such situation, S mode has been introduced, in which a given transponder is activated by a selective interrogation included in the encoded received pulse. In addition, S mode operation specifies a monopulse antenna for the radar, which allows for improved azimuth accuracy.

The availability of S mode transponders on board every aircraft operating under IFR conditions, which will occur as a result of ICAO recommendations, allows the implementation of a new airborne collision avoidance system. Under this concept, the interrogation, reception, and processing are performed on board, thus directly providing the pilot with the necessary information concerning the surroundings for en route collision avoidance.

Other Ground-Based Radars

Approach and Ground Surveillance Radars. For accurate guidance of aircraft at landing approach and for all-weather surveillance of taxiways and runways, short-range radars are used with high enough resolution—i.e. centimeter wavelengths and beyond with medium-sized antennas—for good accuracy and detection of the various possible obstacles.

Meteorological Radar. The adverse effects of rain clutter on air target detection has been pointed out. Specific radars have been designed for meteorological purposes and are operated for detecting strong rain cells associated with thunderstorms and contributing to avoiding lightning hazards to aircraft. Of further interest for improving safety at landing is from the ability of radar to make fine Doppler analysis of the rain echoes and therefore detect specific hazards due to wind shear at final approach. For the few cases of dry wind shears, detection through the specific Doppler signature of turbulent layers at UHF is considered.

*A linear antenna with azimuth-only directivity is used in secondary radar, since limited antenna gain satisfies the power balance and no direct elevation measurement is required. For increased azimuth accuracy, this linear antenna can be split into halves for monopulse processing.

Airborne Meteorological Radar

The need for aircraft to find a safe route far away from ground radars in case of thunderstorms has led to the equipment of airborne meteorological radars protected by a radome in the nose of the aircraft. Operating at centimeter wavelengths, such a radar provides the pilot with a display of heavy rain cells at ranges up to 15 nm, allowing the route to be adapted. A recently added option makes this nose radar also able to detect the specific Doppler signature of wind shears at the expense of a high-quality radome for avoiding detrimental ground clutter effects induced by sidelobes resulting from radome-to-antenna interactions.

Other Airborne Radars

In preparation for landing or for terrain awareness, direct measurement of altitude above the ground is available from a radioaltimeter, which is a low-power down-looking radar.

11.22 OVERVIEW OF MILITARY REQUIREMENTS AND SPECIFIC DEVELOPMENTS

In addition to the need of detecting, identifying, and, if required, directing actions for shooting down targets flying beyond the altitude/velocity domain of civilian aircraft, the major specificity of air defense radars is robustness against adverse countermeasures. Therefore, ground systems have to face many challenges: modern targets of military interest may have reduced RCS (they are stealthy) and can be associated with active countermeasures (jammers) which can be carried by the target itself for self-protection, or come from standoff; passive decoys may also be encountered.

Against jammers, beam and waveform agility is required, which can be brought to the radar by the technology of active arrays. Against decoys, improved resolution may help and is also of interest against stealth targets when associated with increased radiated power.

But a drastic revision of radar design principles may be necessary for restoring detection performances in the modern countermeasures environment: low-frequency sensors (at UHF and below) and bistatic configurations (with widely separated transmitter and receiver) are among the promising techniques.

Other applications for radars on board military aircraft include:

- Air-to-air detection, which is performed, for example, with a large rotating antenna protected by a lenticular radome mounted piggyback on board a large standoff platform, but is also associated with fire control in the nose radar of combat aircraft.
- Air-to-ground surveillance and reconnaissance, where synthetic aperture (which will be presented below and illustrated in Part 11 for space applications) allows for high enough resolution at long range.

Those missions where radar is the key sensor are also to be performed despite adverse countermeasures.

11.23 OVERVIEW OF RADAR APPLICATIONS TO SPACE

Earth Observation Missions

LEO satellites are suitable platforms for global observation of the earth with revisit time matched to natural phenomena. Radar observation of the sea surface at or near vertical incidence has proven its value for oceanography through roughness analysis (inducing wind strength and direction), as well as sea-level measurements with centimetric accuracy (inducing salinity or temperature gradients, currents, etc.). In addition to its wide bandwidth, the major technical challenge with such radar is to remove both instrumental and radiowave propagation offsets and drifts, which would reduce the accuracy of the altitude measurements.

Radar has also demonstrated its capability for imaging the earth surface, offering all-weather operations as opposed to passive visible or infrared observation, and, surprisingly, with a ground resolution not as degraded as one could expect from the wavelength ratio. This comes from the exploitation of the Doppler spreading of ground clutter, known as synthetic aperture radar (or SAR) techniques. Since each scatterer on the ground has an apparent radial velocity:

$$v_r = v_p \cos(\alpha) \tag{11.39}$$

where v_p is the platform velocity with respect to the ground and α is the angle between line of sight and velocity vector, it appears from equation (11.31) that a Doppler analysis of the scattered signals results in the angular analysis of the ground scatterers in the antenna footprint. It can be combined with the usual time delay analysis into a two-dimensional mapping of the ground. The range resolution of 10 m can be easily obtained with a 15 MHz bandwidth waveform (see equation (11.30)), which projects on the ground for an incident angle of 60° as 20 m cross-track. Along-track resolution can be derived from combining equations (11.39), (11.31), and (11.36) for α close to 90°:

$$\Delta x = \lambda R / 2Tv_p \tag{11.40}$$

which is similar to the far-field resolution of a synthetic antenna of aperture length $2Tv_p$, where T is the coherent processing time for the Doppler analysis.

The resolution of 20 m for R = 1,000 km at a wavelength of 6 cm, v_p = 6,000 m/s corresponds to T = 0.25 s, which looks more practical to implement than the real antenna of length 3 km that would directly give the same along-track resolution!

Further implications of spaceborne SAR and related technologies are given in Part 11.

Other Applications of Radar to Spacecraft

Radar has many other applications to spacecraft operations; for example, surface-based radars for trajectography of launch vehicles or for in-orbit satellite or even debris observation.

The synthetic aperture radar presented in the previous section is also of interest for planetology studies from dedicated orbiters,* the limited mass, energy, and data stream available for the payload resulting in reduced performance with respect to Earth observation.

On-board radar also has applications in rendezvous or docking applications.

* Radar echoes from the Moon or planets have also been observed from Earth-based experiments, making use of very large dish antennas such as at the Arecibo Observatory.

PART 6

OPTICAL FIBERS AND LASERS

Jean-Claude Mollier

11.24 OPTICAL FIBER THEORY AND APPLICATIONS

Basic Characteristics

***Geometry* (*Figure 11.12*).** An optical fiber is a cylindrical dielectric structure that can guide a light beam over distances ranging from tens of meters to tens of kilometers. It mainly consists of a cylindrical core (radius a, refractive index n_1) surrounded by a cladding of slightly lower refractive index n_2.

Refractive Index Profiles. The refractive index distribution in the transverse direction is given by:

$$n(r) = n_1[1 - 2\Delta(r/a)^g]^{1/2} \quad 0 \le r \le a \tag{11.41a}$$

$$= n_1[1 - 2\Delta]^{1/2} = n_2 \quad r \ge a \tag{11.41b}$$

where the index contrast Δ is defined as:

$$\Delta = (n_1^2 - n_2^2)/2n_1^2 \tag{11.42}$$

Optical fibers are classified according to the value of the parameter g, which can be varied to get different profiles:

- Step index fiber ($g \infty$)
- Graded index fiber: $g = 1$ (triangular profile), $g = 2$ (parabolic profile), etc.

Spectral Loss. Depending on the application, two kinds of dielectric material are currently used to manufacture optical fibers: silica (with dopants) and plastic. The fibers with core and cladding made from silica exhibit low loss and are systematically used for long-distance applications. Plastic-clad fibers and all-plastic fibers are less expensive and have higher mechanical strength than silica fibers. But they

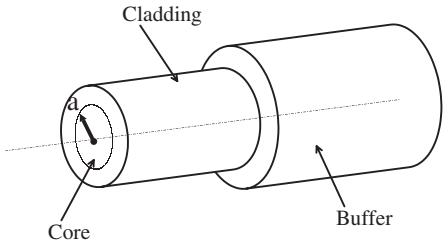


FIGURE 11.12 Optical fiber geometry.

AVIONICS AND ASTRIONICS

are presently used for short-distance applications (less than 100 m) due to their higher losses. Two typical loss spectra are represented in Figures 11.13 and 11.14 for silica fiber and plastic fiber, respectively.

Table 11.10 summarizes the basic characteristics of several optical fibers.

Ray Theory of Optical Fibers

This simplified theory is valid when the core radius a is much larger than the operating wavelength λ (multimode fibers). In this geometrical optics approach, light consists of a number of rays being reflected or refracted at the interface between the core and the cladding.

***Meridional and Skew Rays* (*Figures 11.15 and 11.16*).** The rays injected into the core of the fiber are either confined to the meridian planes (meridional rays) or not confined to any plane (skew rays), depending on the injection conditions: launch angle and distance to the fiber axis.

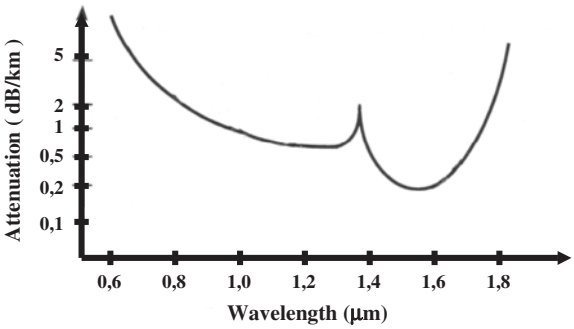


FIGURE 11.13 Spectral attenuation in silica fiber.

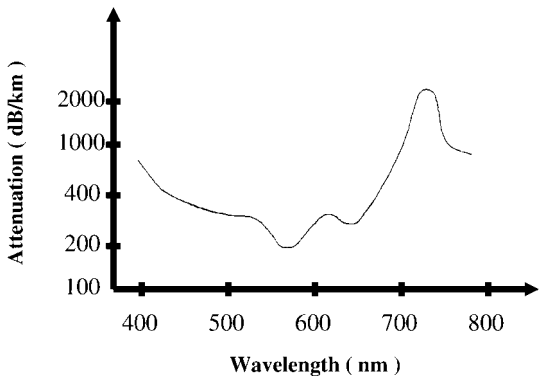


FIGURE 11.14 Spectral attenuation in plastic fiber.

TABLE 11.10 Optical Fiber Properties

Dielectrics material	Plastic	Silica plastic	Silica Ge O_2 doped core			
Index profile	Step index	Step index	Step index		Graded index	
Core/cladding diameters (μm)	980/1000	400/480	100/140	8/125	50/125	62/125
Attentuation (dB/km)	200	10	2–5	0.2 0.5	1–3	0.9–3
Wavelength range (μm)	0.45–0.7	0.7–1	0.8–1.6	1.3–1.6	0.8–1.6	0.8–1.6
Applications	Lighting	Optical power transfer	Short-distance telecoms	Long-distance telecoms	Information networks Medium-distance telecoms	

***Ray Paths* (*Figure 11.17*).** The rays excited in the fiber can be classified into three different types: bound rays, refracting leaky rays and tunneling leaky rays.

Bound rays remain guided in the core of the fiber as they undergo total reflection at the core–cladding interface.

Refracting leaky rays leak out from the fiber a very short distance.

Tunneling leaky rays leak out gradually from the core of the fiber. The attenuation distance of the corresponding optical power can vary from a few millimeters to hundreds of meters.

Numerical Aperture. Bound rays exist when the angle of incidence at the core–cladding interface is sufficiently large. It can be shown, by applying Snell's law, that the total internal reflection takes place only if the angle of incidence at the external air–core interface θ_0 is sufficiently small.

For a step index fiber, the maximum acceptance angle is given by:

$$n_0 \sin \theta_0 = \sqrt{n_1^2 - n_2^2} \tag{11.43}$$

Equation (11.44) defines an important parameter of the optical fiber called the numerical aperture (NA) as:

$$\text{NA} = \sqrt{n_1^2 - n_2^2} = n_1\sqrt{2\Delta} \tag{11.44}$$

where Δ was defined in (11.42).

For typical telecommunication optical fibers, NA varies from 0.1 to 0.25. For plastic fibers, NA varies from 0.4 to 0.5.

Wave Theory of Optical Fibers

To get an accurate representation for light propagation in fibers, electromagnetic analysis based on Maxwell's equations is essential. If z denotes both the fiber axis

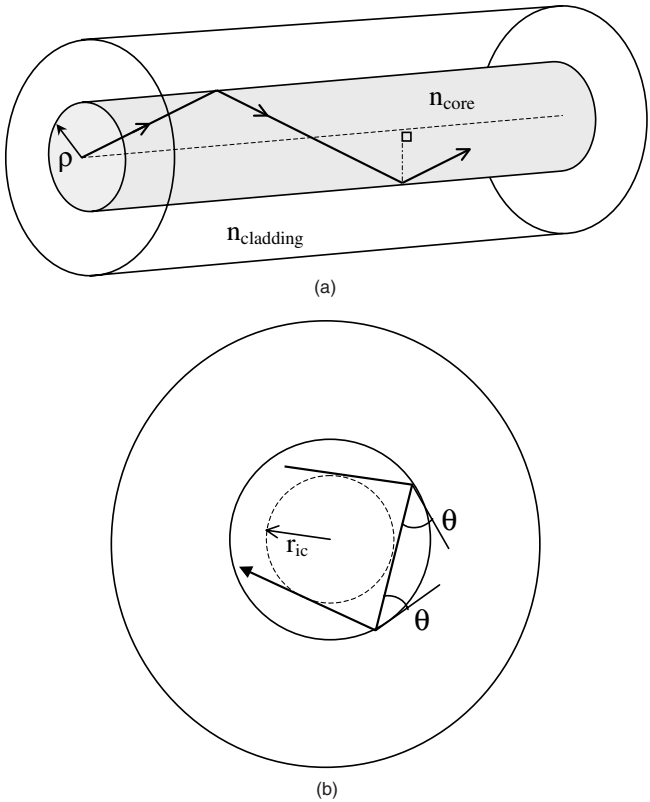


FIGURE 11.15 (a) Meridional ray in a step index fiber; (b) skew ray in a step index fiber.

and the direction of wave propagation, the angular (Φ) and radial (r) distributions of the wave can be represented by:

$$\vec{E} = \vec{E}_0(r,\Phi) \exp i(\omega t - kz) \tag{11.45a}$$

$$\vec{H} = \vec{H}_0(r,\Phi) \exp i(\omega t - kz) \tag{11.45b}$$

where $\vec{E}$ and $\vec{H}$ are the electric and magnetic fields, which must satisfy boundary conditions at the core–cladding interface. In addition, $\vec{E}$ and $\vec{H}$ must remain finite on the fiber axis ($r \rightarrow 0$) and decay to a negligible value outside the cladding ($r \rightarrow \infty$).

Using these boundary conditions allows the derivation of an eigenvalue equation for the propagation constant $k = \omega/v$, with v denoting the phase velocity of the wave.

Weakly Guiding Approximation. For most optical fibers, the core and cladding indices are nearly the same ($\Delta << 1$) and the eigenvalue equation can be simplified, leading to the so-called weakly guided modes: all the guided modes are linearly

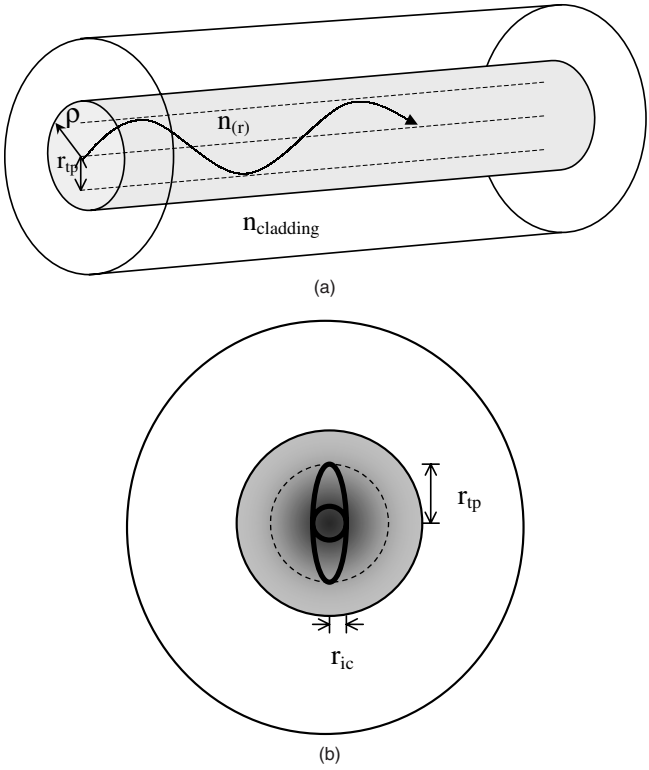


FIGURE 11.16 (a) Meridional ray in a graded index fiber; (b) skew ray in a graded index fiber.

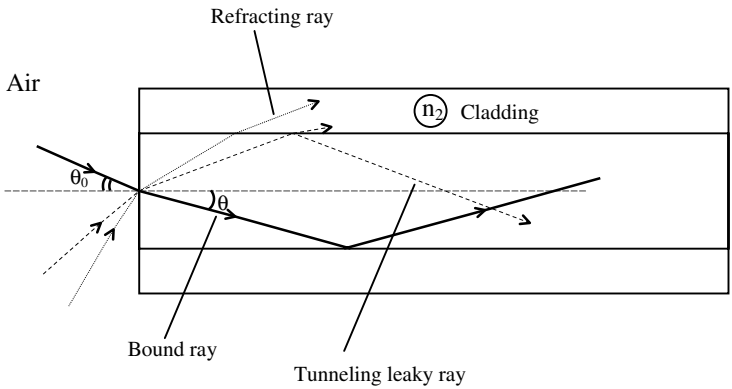


FIGURE 11.17 Various types of rays in a step index fiber.

polarized (fixed direction of $\vec{E}$ in the fiber) and are commonly called LP_{mn} modes. The integers m and n specify the angular and radial variations of the fields (number of extrema).

In order to represent conveniently the dependence of k upon the light wavelength and the parameters of the fiber, two dimensionless parameters are introduced:

1. The normalized frequency as:

$$V = ak_0\sqrt{n_1^2 - n_2^2} = ak_0n_1\sqrt{2\Delta} \tag{11.46}$$

where $k_0 = \omega/c = 2\pi/\lambda$ is the free space propagation constant (or wave number).

2. The normalized propagation constant b as:

$$b = \frac{(k/k_0)^2 - n_2^2}{n_1^2 - n_2^2} \tag{11.47}$$

The variation of b with V forms a set of universal curves which are plotted in Figure 11.18 for a step index fiber. For a given fiber and a given operating wavelength, V is readily derived from (11.46). Then b is obtained from the curves which represent the guided modes LP_{mn}. Finally, k is derived from (11.47) for each mode of interest.

The total number of guided modes is roughly approximated by $V^2/2$ for a step index fiber and $V^2/4$ for a parabolic index profile fiber.

Cutoff Wavelengths. The condition $b = 0$, or alternatively $k = k_0n_2$, corresponds to the cutoff of the mode. The electromagnetic wave is no longer guided in the core of the fiber and is known as an evanescent wave.

Single-Mode Fiber. As can be seen from Figure 11.18, when the condition $0 < V < 2.405$ is verified, we will have only one guided mode (LP_{01} mode). Such a fiber is referred to as a single-mode fiber and is of tremendous importance in optical communication systems.

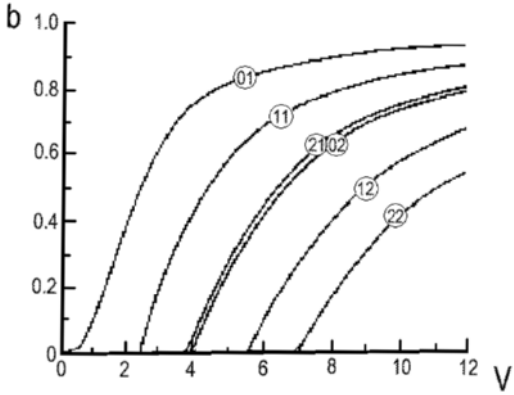


FIGURE 11.18 Universal curves of normalized propagation constant b for a step index fiber corresponding to guided modes LP_{mn}.

The above condition allows us to determine either the lowest operating wavelength λ_l for a given fiber or the highest value of core radius a at a fixed wavelength:

$$\lambda_l = \frac{2\pi a n_1 \sqrt{2\Delta}}{2{,}405} \tag{11.48}$$

$$a < \frac{2{,}405\ \lambda}{2\ \pi n_1 \sqrt{2\Delta}} \tag{11.49}$$

The core radius for single-mode silica fibers is typically in the range 3–5 μm.

Confining Factor. The confining factor Γ (<1) is defined as the ratio of the power carried in the core to the power injected in the fiber. For a given *LP* mode, Γ increases markedly with V. It is essential to maintain Γ as high as possible in order to minimize the wave attenuation related to the loss mechanisms in the cladding. For a single-mode silica fiber, typical values for Γ are in the range 0.7–0.8, corresponding to $1.8 < V < 2.2$.

Dispersion in Optical Fibers

Different light components (different modes and/or different wavelengths) propagate through the fiber at different velocities. This phenomenon, called dispersion, determines the information-carrying capacity of a fiber optic communication system. When the information is coded in the form of pulses (bits), each pulse becomes larger, due to fiber dispersion, by an amount strongly depending on the refractive index profile.

Modal Dispersion. In a multimode fiber, the dominant dispersion mechanism is the intermodal dispersion: at a given operating wavelength (or given V), different modes travel with different k, i.e., different velocities (see Figure 11.18).

For a step index fiber, the full width at half maximum (FWHM) τ of the broadened pulse can be estimated as:

$$\tau \simeq L \cdot n_1\ \Delta / c \tag{11.50}$$

where all symbols have been defined in previous subsections.

For a parabolic graded index fiber, the FWHM τ is roughly given by:

$$\tau \simeq L \cdot n_1\ \Delta^2 / 8c \tag{11.51}$$

As expected, τ is proportional to the length L of the fiber. Relations (11.50) and (11.51) clearly show that Δ must be made as small as possible to minimize the pulse spreading.

Chromatic Dispersion. In a single-mode fiber, the dominant dispersion mechanism is the chromatic dispersion, which results from two independent phenomena:

1. *Waveguide dispersion,* determined by the difference in propagation velocity between the core and the cladding
2. *Material dispersion,* determined by the variation of propagation velocity with wavelength

In that case the FWHM of the broadened pulse is given by:

$$\tau \simeq L \cdot |D_c| \cdot \Delta\lambda \tag{11.52}$$

where $\Delta\lambda$ is the spectral width of the optical signal launched into the fiber and D_c is the chromatic dispersion coefficient.

A plot showing D_c along with its two components—waveguide dispersion D_w and material dispersion D_m—is shown in Figure 11.19. Chromatic dispersion disappears at a certain wavelength. In other words, pulses can propagate without spreading and/or distortion whatever the length of the fiber is.

That zero-dispersion wavelength depends on fiber geometry and refractive index values in the core and the cladding. For a conventional fiber, it is approximately equal to 1.3 μm. The information transmission rate R (expressed in bits/s) is often defined as

$$R \simeq \frac{1}{2\tau} \tag{11.53}$$

and the information-carrying capacity of the fiber optic link is measured in terms of the $R \cdot L$ product, expressed in Gbps · km (symbol G means giga).

Typical Values

Single-mode silica fibers: 5–50 Gbps · km

Multimode silica fibers: 0.05–0.5 Gbps · km

Plastic fiber: 5–10 Mps km

Dispersion—Shifted and Dispersion-Flattened Fibers. In a silica fiber, the losses attain a minimum value of about 0.2 dB/km at 1550 nm. The shift of the zero dispersion wavelength to a region around 1,550 nm can be accomplished by slightly decreasing the core diameter and modifying the refractive index profile (triangular profile).

To achieve low dispersion in a large range of wavelengths (1.3–1.6 μm), more sophisticated index profiles are used (Figure 11.20). The resulting fiber is called dispersion-flattened fiber.

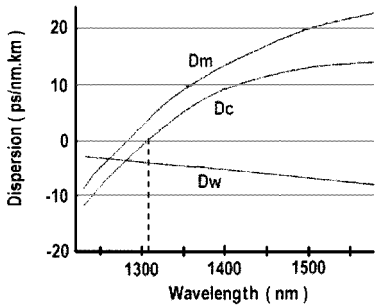


FIGURE 11.19 Chromatic dispersion D_c, material dispersion D_m, and waveguide dispersion D_w for conventional single-mode fiber.

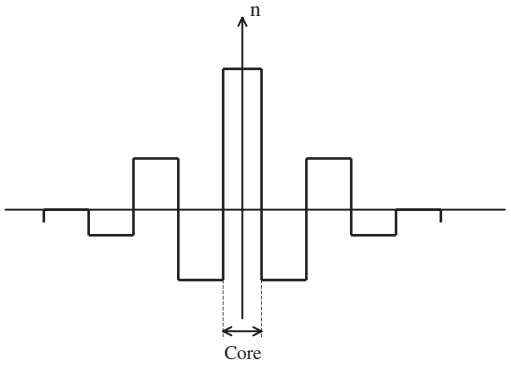


FIGURE 11.20 Index profile for dispersion-flattened fiber.

Figure 11.21 compares the plot of the chromatic dispersion coefficient D_c versus λ for the three different types of fibers mentioned above.

Fabrication Technique for Optical Fibers

Two main types of material are currently used: plastic and glass. Plastic fibers offer some advantages in terms of cost and ease of manufacture, but their high transmission losses limit their use to short-haul applications (less than a few hundred meters). Glass fibers exhibit the best characteristics suitable for telecommunications applications.

The refractive index of different kinds of glass may be varied over a relatively wide range, but control of the impurity content is very difficult except for silica: the index of pure silica is 1.45 at 1 μm. Dopants are added to lower (B_2O_3) or raise

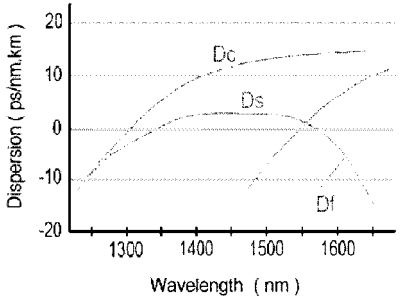


FIGURE 11.21 Dispersion for conventional single-mode fiber (D_c), dispersion-shifted fiber (D_s), and dispersion-flattened fiber (D_f).

(GeO_2) the refractive index. Thus, a typical fiber might consist of an SiO_2:GeO_2 core with a pure silica cladding.

The preparation of silica or glass fibers is a two-stage process in which initially the pure glass is produced and converted into a rod (or preform) suitable for making the fiber. A drawing technique is then used to form the fiber. In a third stage, the fiber is coated (for protection) with a layer of plastic by passing it through a bath of molten plastic and a curing oven, before being wound on a drum (Figure 11.22). Several vapor phase deposition techniques have been developed to produce preforms having minimal impunity content. One of the most popular is the modified chemical vapor deposition (MCVD) method, illustrated schematically in Figure 11.23.

In a first stage (a), a doped silica layer is deposited onto the inner surface of a pure silica tube. The deposition occurs as a result of a chemical reaction between the vapor constituents that are introduced into the tube heated to about 1,600°C:

$$SiCl_4 + O_2 + 2H_2 \rightarrow SiO_2 + 4HCl$$

and

$$GeCl_4 + O_2 + 2H_2 \rightarrow GeO_2 + 4HCl$$

Good control of the index profile is obtained by repeating this process with different concentrations of the dopants in order to built up thin layers of doped silica sequentially.

In a second stage (b), the tube is heated to its softening temperature (~2,000°C). The tube is thus collapsed into a solid rod called the preform.

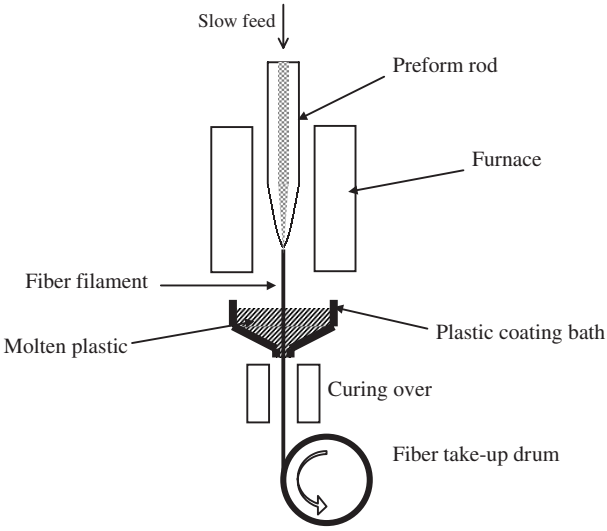


FIGURE 11.22 Schematic diagram of fiber-drawing apparatus.

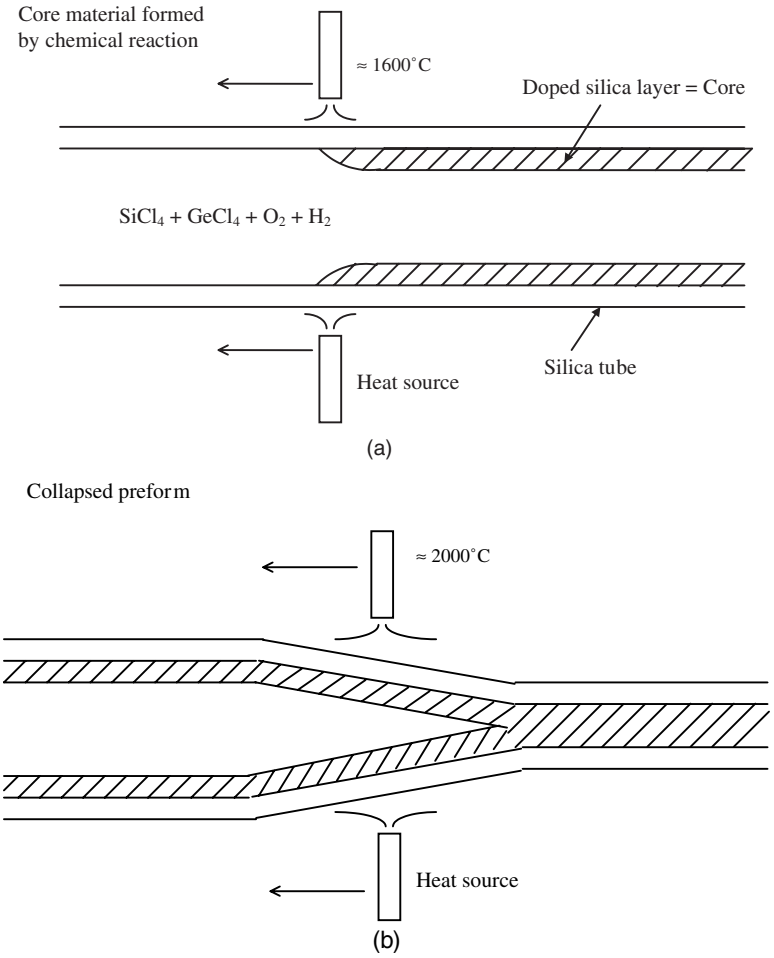


FIGURE 11.23 (a) Production of fiber preform by MCVD. (b) Collapsed preform.

11.25 LASERS

Introduction

Laser (light amplification by stimulated emission of radiation) was originally used to describe the processes that amplify an electromagnetic wave of appropriate wavelength, or enhance a collection of photons of suitable frequency. But it has rapidly come to mean an optical oscillator, and so *laser* now denotes an optical device that can deliver a coherent and highly directional light beam in the ultraviolet, visible, or infrared spectrum. Laser wavelengths range from about 0.2 μm to 10 μm and beyond.

AVIONICS AND ASTRIONICS

Analogy between Lasers and Electronic Oscillators

A laser, just like an electronic oscillator contains three basic components (see Figure 11.24): an amplifier, a power supply, and an optical cavity or resonator producing the desired positive feedback. The amplifying medium can be a gas, a liquid, a solid, or a semiconductor crystal. The optical feedback is usually provided by two reflecting mirrors, but lasers have been built with three or four mirrors.

In an electronic oscillator, the signal at the amplifier output is partly reinjected into the amplifier through the feedback element. In a laser, an electromagnetic wave travels back and forth between the mirrors through the amplifying medium.

The oscillation condition can be expressed as:

$$\text{Loop-gain equal to one} \Rightarrow \quad g = \alpha + \frac{1}{2L}\ln(1/r_1 r_2) \tag{11.54}$$

$$\text{Round-trip phase-shift multiple of } 2\pi \Rightarrow f_o = m\frac{v}{2L} \tag{11.55}$$

where $r_{1,2}$ = reflectivity of each mirror
g and α = gain and loss per unit length respectively
v = light velocity in the amplifying medium which is supposed to fill in the whole cavity
m = integer of very large value
f_o = oscillation frequency

Equation (11.54) expresses how much gain is needed to compensate for material and mirror losses. Equation (11.55) shows that oscillation occurs at an infinite number of equidistant frequencies. In practice, optical oscillation is possible only in a limited frequency range related to the amplifier bandwidth ΔF (for lasers in free oscillation, ΔF ranges from 10 MHz up to 4 THz, where M and T denote 10^6 and 10^{12} respectively).

Specific Properties of Laser Beams

Compared to conventional light sources, lasers have three unique properties: a narrow beam, a limited spectrum, and coherent emission.

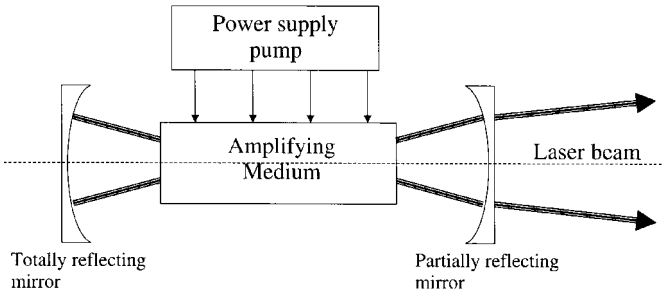


FIGURE 11.24 Schematic representation of a laser.

Narrow Beam. Most gas or solid-state lasers emit beams with a diameter limited to a few millimiters at the device output and with a divergence angle less than 1 milliradian (the actual beam divergence depends on the type of Laser and the optics used with it). In contrast, a conventional light source radiates light in all directions.

Limited Spectrum. Most laser beams contain only a narrow range $\Delta\lambda$ of wavelengths and thus can be considered quasi-monochromatic for all practical purposes. Typical values for $\Delta\lambda$ can be readily derived from ΔF and the oscillation wavelengths λ_0 through the relationship:

$$\Delta\lambda = (\lambda_0^2/c) \cdot \Delta F \tag{11.56}$$

where c denotes the light velocity in vacuum.

$$\Delta\lambda \sim 10^{-3} \text{ nm} \rightarrow 10 \text{ nm} \qquad (1 \text{ nm} = 10^{-9} \text{ m})$$

Conventional light sources emit light over a very broad spectrum (10^3–10^4 nm).

Coherent Emission. The electromagnetic waves radiated from the laser output are in phase with one another over a distance l_c, called the coherence length. This (longitudinal) coherence is related to the coherence time τ_c by $l_c = c \cdot \tau_c$. The coherence time itself is a direct measure of the monochromaticity of the laser:

$$\tau_c \sim 1/\Delta F \tag{11.57}$$

Some stabilized gas lasers can exhibit a coherence length over 1 km. Spatial (or transverse) coherence is a measure of the transverse distance y_c across the beam over which a constant phase relationship exists. y_c is related to the divergence angle $\Delta\theta$ of the beam by the approximate expression:

$$y_c \sim 0.6\lambda_0/\Delta\theta \tag{11.58}$$

In contrast, conventional light sources are called incoherent because their coherences are negligible.

Basic Aspects of Laser Physics

Transitions between Energy Levels—Gain Coefficient. Light may behave either as if it were an electromagnetic wave of frequency ν or as if it were a stream of particles called photons, each of them carrying a discrete amount (or quantum) of energy $h\nu$ (Planck's constant h equals $6.6\ 10^{-34}$ J · s). Matter is composed of discrete particles: atoms, molecules, or ions. The electrons bounded within these particles can occupy only certain discrete energy levels. They can make jumps between any two levels E_i and E_j in three ways involving photons of energy $h\nu_{ij} = E_i - E_j$ (Figure 11.25): spontaneous emission, stimulated emission, or absorption.

Spontaneous Emission. An electron spontaneously falls from E_i to E_j, emitting a photon with random phase and propagation direction. The probability A_{ij} of such a jump is about $10^6 - 10^8$ s^{-1} and the corresponding lifetime $\tau_i = 1/A_{ij}$ of the level E_i is shorter than 1 μs.

In contrast, some energy levels between which jumps are unlikely to occur have lifetimes as large as 0.1 to 1 ms. These levels are said to be metastable.

AVIONICS AND ASTRIONICS

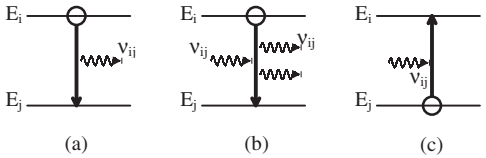


FIGURE 11.25 Schematic representation of spontaneous emission (a), stimulated emission (b), and absorption (c) between two levels of energy E_i and E_j.

Stimulated Emission. An electron makes a jump from E_i to E_j by the action of an incident photon and an identical photon is created. Both photons are traveling in the same direction, carrying the same energy and in phase (or coherent) with one another.

The level E_i is depleted at a rate given by the product population density $N_i \cdot$ probability B for such jumps to occur.

Absorption. An electron absorbs a photon and jumps from E_j to E_i. Level E_j is depleted at a rate given by the product $N_j \cdot B$ (probabilities of stimulated emission and absorption are identical and denoted by the Einstein B coefficient).

Gain Coefficient per Unit Length. This is defined as:

$$g(\nu_{ij}) = B(N_i - N_j)\frac{h\nu_{ij}}{c} \tag{11.59}$$

At the thermodynamic equilibrium, the population density N_j in the lower energy level is always larger than the population density N_i in the upper one, yielding a negative gain or attenuation.

For the gain to be positive, in order to amplify a stream of incident photons, a population inversion would be needed. Techniques to create such a situation are known as pumping techniques.

Transition Lineshapes—Broadening Mechanisms. In practice, the energy levels are not so sharp as previously described and transitions may occur in a range of frequencies rather than at a unique frequency ν_{ij}. Besides the finite lifetime of the upper energy level E_i, which gives rise to the natural transition broadening, two types of broadening mechanisms increase the transition linewidth:

1. *Homogenous broadening* such as lattice vibrations in crystals and atomic collisions in gas and liquids. The corresponding lineshape, related to the gain coefficient $g(\nu)$, is a Lorentzian written as:

$$g(\nu) = \frac{c^{te}}{(\nu - \nu_0)^2 + (\Delta\nu/2)^2} \tag{11.60}$$

   where ν_0 and $\Delta\nu$ denote central frequency and linewidth of the transition $(E_i \rightleftarrows E_j)$ respectively.

2. Inhomogeneous broadening that occurs when the environment or properties of particles are nonidentical. For example, the presence of imperfections and impurities in a crystal alters the physical environment of atoms from one site to

another. In a gas, the random distribution of particle velocities leads to a distribution in the emission frequencies, producing a Doppler broadening. This gives rise to a Gaussian lineshape expressed as:

$$g(\nu) = c^{te} \cdot \exp[-(2.77)(\nu - \nu_0)^2/(\Delta\nu_D)^2] \tag{11.61}$$

with the full width at half maximum (FWHM) given by:

$$\Delta\nu_D = 2\nu_0 \left(\frac{1.38kT}{Mc^2}\right)^{1/2} \tag{11.62}$$

where M = the mass of particles involved in the transition
T = the absolute temperature,
k = Boltzmann's constant
c = light velocity.

Typical values of linewidths for various lasers are given in Table 11.11.

Pumping Techniques. It is impossible to create a population inversion $N_i > N_j$ with just two energy levels E_i and E_j interacting with an electromagnetic wave at the frequency ν_{ij}. Most coherent amplifiers and lasers work with four or three energy levels.

Four- and Three-Level Pumping Scheme. In a four-level system (Figure 11.26), atoms are raised from the lowest (or ground) energy level O to the short-lived level 3 by using an external source of power (see below).

Atoms decay rapidly from level 3 to level 2, which is long-lived, so that it accumulates population.

TABLE 11.11

Laser medium	Transition linewidth	Transition wavelength (μm)
He–Ne	1.5 GHz I	0.633
Nd^{3+}: YAG	120 GHz H	1.06
CO_2	60 MHz I	10.6
Er^{3+}: silica fiber	4 THz H/I	1.55
InGaAsP	1 THz H/I	1.55

H and *I* indicate the type of broadening mechanism.

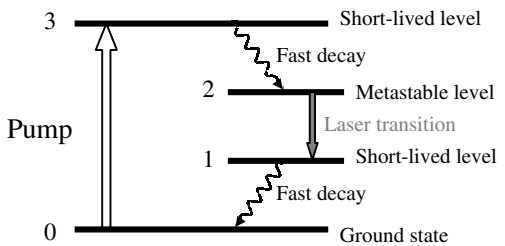


FIGURE 11.26 Four-level pumping scheme.

Level 1 is short-lived, i.e., atoms decay rapidly from level 1 to level 0, and thus there is little population accumulation in level 1.

As a result, population density N_2 is much larger than population density N_1. The stimulated emission predominates over absorption, providing an amplification at the frequency:

$$\nu_{21} = (E_2 - E_1)/h \tag{11.63}$$

In a three-level system (Figure 11.27), level 1 has become the ground level. Again atoms decay rapidly from level 3 to level 2 which is long-lived.

The laser transition occurs between levels 2 and 1.

Those requirements for the different level lifetimes lead to a limited number of laser transitions from a given medium. Nevertheless, the very large variety of amplifying media allows the whole electromagnetic spectrum to be covered from the ultraviolet to the infrared, the radiation wavelengths ranging from about 0.2 μm to 10 μm and beyond.

Pumping Techniques. Two main types of pumping are used to create population inversion: optical and electrical pumping.

Optical pumping is done with either an incoherent light emitted by a powerful lump or a laser beam. The broadband light from an incoherent source is well suited to solid-state and liquid lasers which exhibit very large transitions. The availability of powerful CW or pulsed lasers at many wavelengths makes laser pumping attractive and practical for nearly all kinds of lasers.

Electrical pumping is accomplished by means of an intense electrical discharge and is particularly suited to gas lasers. Electrons produced and accelerated by the high-voltage collide with atoms or molecules in the ground state and involve raising these atoms or molecules to an excited energy level.

Laser Resonators

Most resonators have either plane or spherical mirrors of circular (or sometimes rectangular) shape, separated by a distance L. Mirror dimensions are smaller than a few centimeters, and resonator length L ranges from a few centimeters (or even less) to one meter.

Resonator Stability. From a geometrical-optics point of view, a light ray slightly out of alignment with resonator axis (paraxial ray) must be reflected back and forth between the mirrors without escaping from the resonator. The confinement condi-

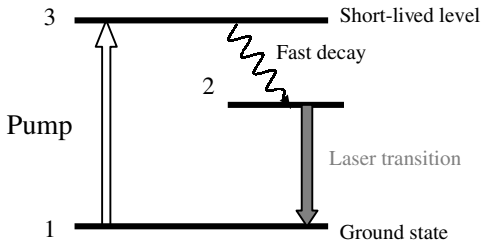


FIGURE 11.27 Three-level pumping scheme.

tion for a paraxial ray in a resonator composed of two mirrors of curvature radii R_1 and R_2 separated by a distance L is written as:

$$0 \leq \left(1 - \frac{L}{R_1}\right)\left(1 - \frac{L}{R_2}\right) \leq 1 \tag{11.64}$$

This condition is expressed graphically in Figure 11.28 for various resonator configurations.

Cavity Modes. From a wave-optics point of view, waves that develop inside the laser cavity must obey the boundary conditions imposed by the finite size and the curvature of the mirrors (as well as the finite lateral size of the amplifying medium).

This leads to the development of standing wave patterns called TEM_{pqm} modes inside the resonator. These modes are classified according to the number p (or q) of nulls that appear in the cross-sectional Profile (or intensity pattern) of the beam and according to m, which denotes the number of half-wavelengths along the distance L.

A sampling of these mode patterns is shown in Figure 11.29.

Gaussian Beam. The transverse intensity distribution of the lowest-order mode TEM_{oom} is a circularly symmetric Gaussian function centered about the resonator axis. For that reason, the wave TEM_{oo} is called a Gaussian beam.

The properties of a Gaussian beam are summarized below (Figure 11.30).

- Intensity:

$$I(\rho,z) \text{ proportional to } \exp - [2\rho^2/w^2(z)] \tag{11.65}$$

where ρ and z denote the radial and axial distances respectively.

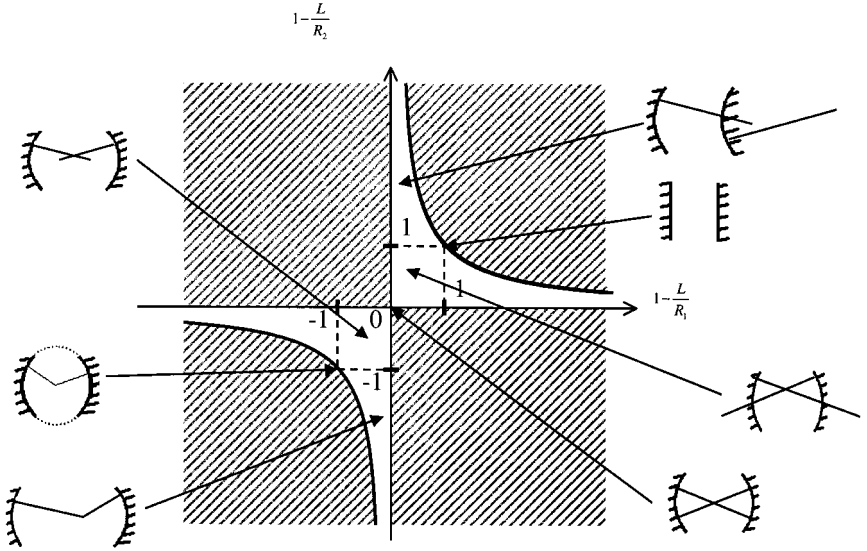
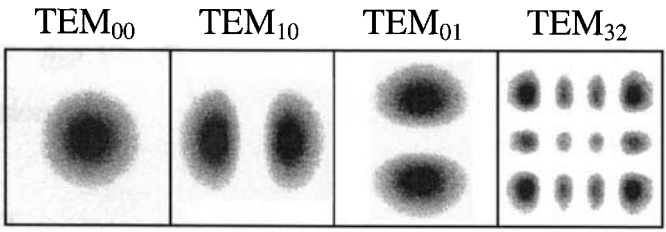
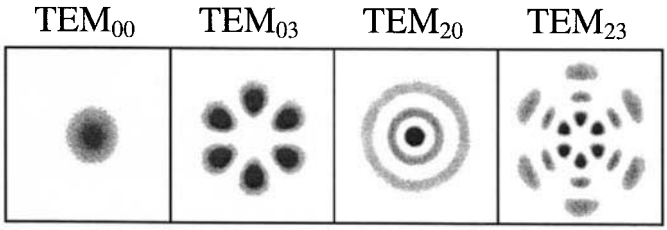


FIGURE 11.28 Resonator stability for various resonator configurations.

AVIONICS AND ASTRIONICS



(a)



(b)

FIGURE 11.29 Mode pattern for various TEM laser modes: (a) Cartesian symmetry (x,y); (b) circular symmetry.

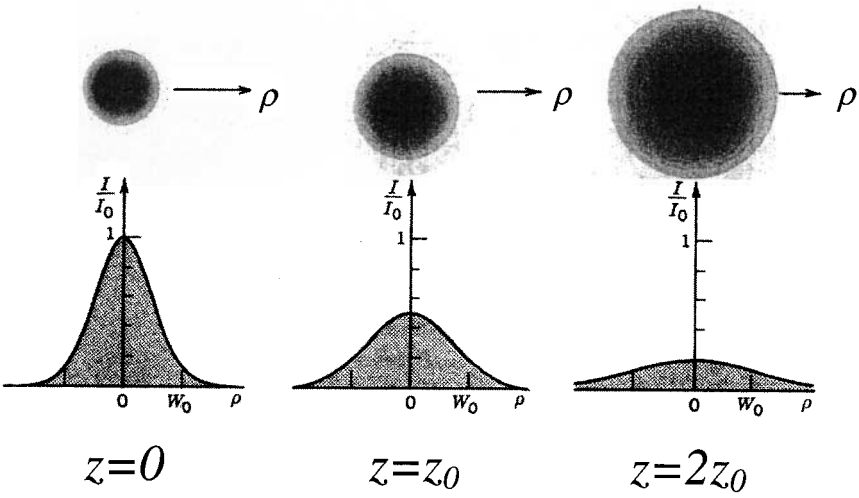


FIGURE 11.30 Gaussian beam intensity versus radial distance ρ at different axial distance z.

- Beam radius: Radial distance w at which I drops by the factor $1/e^2$.

$$w = w_0[1 + (z/z_0)^2]^{1/2} \tag{11.66}$$

  where the Rayleigh range z_0 is related to the smallest radius or beam waist w_0 by:

$$z_0\lambda = \pi w_0^2 \tag{11.67}$$

- Beam Divergence: Far from the waist, the angular divergence θ_0 of the beam increases approximately linearly with z according to the relation:

$$\theta_0 = \frac{\lambda}{\pi w_0} \tag{11.68}$$

Multimode Oscillation. The frequencies of Gaussian beams self-sustained in the resonator are equidistant from one another and the frequency spacing $v/2L$ (see Eq. 11.55)) is usually smaller than the gain bandwidth of the amplifying medium. Consequently, most lasers simultaneously oscillate on several (N) cavity modes and are said to be longitudinal multimode.

Typically, $N \simeq 5$ for an He-Ne laser and $N > 500$ for an Nd:YAG laser.

The selection of a single mode is possible by using intracavity frequency selective elements to modify the frequency spacing of the resonator modes: for example a tilted etalon (Fabry–Perot resonator) whose mirror separation d is much shorter than L has a large spacing $v/2d >> v/2L$, so that only one etalon mode can fit within the amplifier bandwidth.

Pulsed Lasers

Three techniques are commonly used to produce short pulses with very high peak powers: Q-switching, cavity dumping, and mode locking. Cavity dumping is sometimes combined with Q-switching and mode locking to produce special pulse characteristics.

Cavity Dumping. Both cavity mirrors are totally reflective. By inserting an acoustooptic deflector in the cavity, a light beam can be switched out of the resonator during a very short period of time. Pulse duration $\simeq$10–50 ns.

Q-switching. If the quality factor Q of the resonator is kept artificially low, the energy produced by stimulated emission will gradually accumulate. When the losses are suddenly removed, a large population inversion occurs in a high-Q cavity, producing a high power burst of light during a short time.

Three types of Q-switching are in use:

1. A rotating prism as the rear cavity mirror. Periodically the rotating mirror is aligned with the output mirror, causing cavity Q to increase suddenly and during a short duration of about 1 μs.
2. Insertion of an electrooptic modulator inside the cavity. The laser beam is linearly polarized by means of a polarizing element within the cavity (such as Brewster angle windows on the amplifier). When the electrooptic modulator is turned on, the beam polarization is rotated, preventing the beam from being

amplified and the oscillation from occurring. Pulse duration of about 10 ns is obtained with this type of *Q*-switching.

3. Insertion of a saturable absorber into the cavity. When the amplifying medium is pumped with sufficient energy, beam intensity builds up from the spontaneous emission and an intense emission is produced that saturates the absorber, causing it to change from high loss to low loss. Pulse duration cannot be shorter than 1 μs.

Mode Locking. Pulsed laser action is achieved by coupling in phase a great number of longitudinal modes. This powerful method is particularly suited to lasers whose gain bandwidth is very large, such as Nd:YAG or Ti:Al_2O_3 lasers. Mode locking allows the generation of ultrashort pulses: 30 fs − 10 ps (1 fs = 10^{-15} s).

11.26 SPECIFIC LASER SYSTEMS

Carbon Dioxide Laser

Operating Principle. Transitions occur between energy levels related to rotational and vibrational motions of CO_2 molecules. In fact, a gas mixture of CO_2, nitrogen, and helium is used, in the approximate ratio CO_2:N_2:He = 1:1:8. An electrical discharge involves raising nitrogen molecules to a long-lived excited energy level. It thus accumulates population, which transfers its energy by collision to CO_2 molecules, thereby populating the upper laser level. The radiative decay rate of this level is about 20 times slower than the decay rate from the lower level, thus allowing for a population inversion to be established.

In practice, a very large number (~100) of laser transitions between closely spaced rotational levels are possible in the 9–11 μm wavelength range, the strongest transition being centered at 10.6 μm.

Helium plays a double part: its large thermal conductivity makes heat transfer easier and it helps the CO_2 molecules to drop from the lower laser level to the ground state.

Laser Structures

1. Sealed-tube lasers: Conventional gas discharge lasers which have lengths of up to 2 m and can be either pulsed or CW (continuous wave). Maximum output power ~50 W per meter of cavity length.
2. Waveguide lasers: The tube is constructed in the form of a waveguide excited by intense radio frequency (RF) fields. The result is a compact CW (or pulsed) laser with an output power ranging from several watts to 50 W.
3. Longitudinal-flowing gas lasers: Slow axial-flow lasers produce CW output limited to 80 W per meter of cavity length. Fast axial-flow lasers (where a pump is used to move the gas quickly) produce CW output ~ 800 W/m due to better cooling of gas mixture.
4. Transverse-flow lasers: This structure makes pumping easier and more efficient and produces very high power, ranging from 5 to 25 kW.

Applications. CO_2 lasers are used in a great number of application areas: materials working (cutting, welding), heat treatment of metal surfaces, lidar (light detection and ranging) for analyzing the atmosphere or measuring wind velocity, surgery, etc.

Neodymium Laser

Operating Principle. The neodymium may be incorporated (in the form of impurity ions Nd^{3+}) into various host materials, either single crystals—yttrium aluminum garnet (YAG), yttrium lithium fluoride (YLF)—or glasses of different compositions. Neodymium is a four-level system pumped at wavelength bands near 0.73 and 0.80 μm. The strongest laser transition is near 1.06 μm, with precise wavelength dependent on the host.

This transition is either homogeneously broadened (Nd:YAG) with a linewidth $\Delta\nu \simeq 120$ GHz, or inhomogeneously broadened (Nd:glass) with a linewidth $\Delta\nu \simeq 3$ THz.

Laser Structures. Two types of optical pumping are currently in use: lamp pumping and diode laser pumping.

For lamp pumping, hollow reflective cavities transfer pump light from linear lamp to the amplifier rod (Figure 11.31).

Diode laser pumping is usually done from the end of the rod (Figure 11.32). A coating deposited on the input section of the rod allows for reflecting light at 1.06

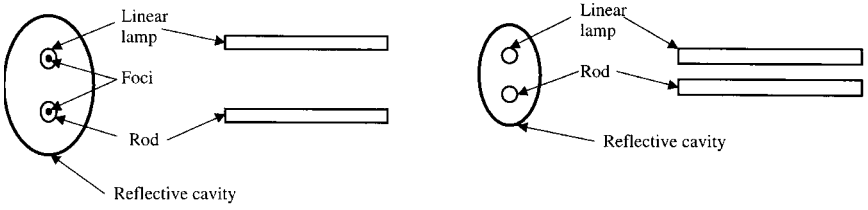


FIGURE 11.31 Cavity configurations for flashlamp pumping.

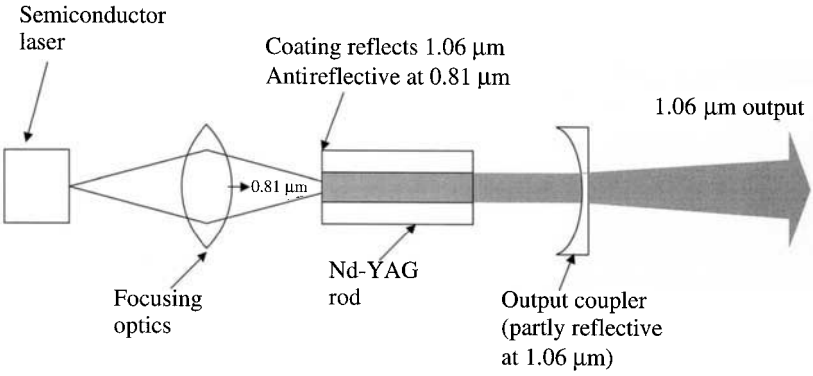


FIGURE 11.32 Diode laser pumping of a neodymium laser.

μm (mirror of the laser resonator) and for transmitting pump light at 0.81 μm into the rod.

A wide range of output powers is available, depending on configuration, road size, pumping source, and wavelength (harmonic generation can shift the wavelength into the green and ultraviolet). CW or average powers (in pulsed mode) range from a few milliwatts to several kilowatts. *Q*-switched lasers provide pulses lasting from a few nonoseconds to hundreds of nanoseconds, depending on pumping technique.

Applications. Like CO_2 lasers, these solid state lasers have found many applications in science, industry, medicine, and military equipment, including materials working (cutting, drilling), electronics fabrication, cataract surgery, measurements of deformation under stress, and lidars (wind shear detection, target identification, etc.).

Semiconductor Diode Laser

Operating Principle. Semiconductor lasers are based on coherent emission of photons due to the stimulated recombination of electrons and holes in a forward biased pn junction.

Population inversion occurs if the injected current is high enough to ensure that electronic density of the conduction band is greater than that of the valence band. The minimum current, called the threshold current, is of the order of a few tens of milliamps. The resonant cavity is formed by the reflecting facets at each end of the semiconductor structure.

The amplifying media are made from the compounds of group III and group V elements of periodic table, resulting in AlGaAs and InGaAsP materials. The emission wavelength ranges either in the band 650–870 nm (ternary compound) or the band 1.1–1.65 μm (quaternary compound).

Laser Structures. The basic structure of a semiconductor laser is shown in Figure 11.33. The pn junction is part of double heterojunction formed by the addition of semiconductor materials grown in layers of various thicknesses (10–1,000 nm) in the *y*-direction. These layers are composed of materials having a wider energy bandgap and a lower index of refraction than the junction (or active) region. As a result, both optical beam and charge carriers are confined in that active region. One

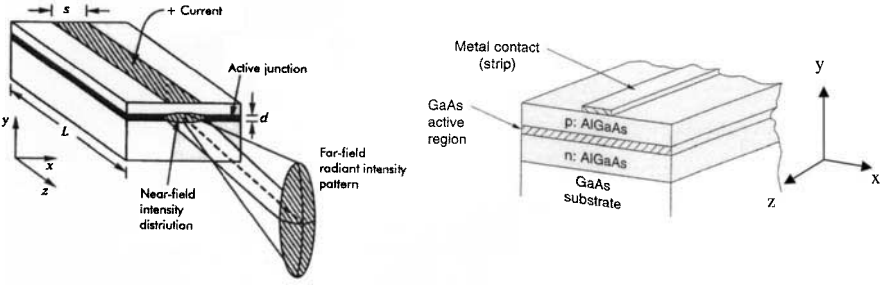


FIGURE 11.33 Schematic representation of a basic semiconductor laser.

simple way of providing optical and electron confinement in the x-direction is to made the metal contact to the top of the laser through a stripe, etched through an insulating oxide.

A better lateral confinement is provided by the buried heterostructure shown in Figure 11.34. The active region (about 2 μm wide) is surrounded on all four sides by material of lower refractive index and larger bandgap, and the laser is called an index-guided laser.

The gain bandwidth is typically on the order of 10^{12} Hz (or 10 nm) and thus several longitudinal modes can oscillate. In addition, the laser beam experiences a great angular divergence due to the small size of the active region ($\sim 2 \times 05$ μm).

Simple frequency operation is achieved by using Bragg reflectors or gratings, which provide optical feedback for laser action. These structures are known as distributed feedback (KFB) or distributed Bragg reflector (DBR) lasers (Figure 11.35).

Applications. Currently the largest market for semiconductor diode lasers is compact disk players, but two fields of application have been growing for the last few years.

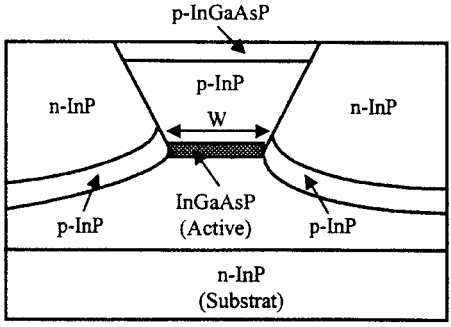


FIGURE 11.34 Simplified cross-sectional view of a buried hetero-structure laser.

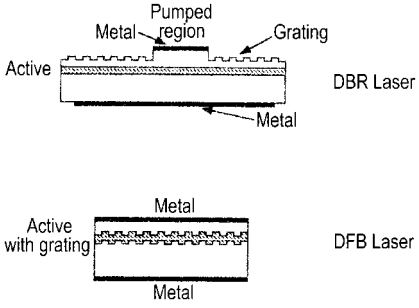


FIGURE 11.35 Simplified cross-sectional view of single-frequency laser structures.

1. Long-distance light-wave communication systems, which need diode lasers as sources in the second (1.3 μm) and third (1.55 μm) wavelength window (minimum dispersion and loss of optical fibers respectively).
2. Local area networks, which include computer networks, high-definition TV, avionic systems, and satellite networks. Other uses of semiconductor lasers are pump sources for solid state lasers, printers, and laser radars.

REFERENCES

Buck, J. A., *Fundamentals of Optical Fibers,* John Wiley & Sons, New York (1995).

Ghatak, A. and Thyagarajan, K., *Introduction to Fiber Optics,* Cambridge University Press, Cambridge (1998).

Hecht, J. *The Laser Guidebook,* McGraw-Hill, New York (1992).

Iga, K., *Fundamentals of Laser Optics,* Plenum Press, New York (1994).

Lecoy, P., *Téléoms sur fibres optiques,* Hermès, Paris (1997).

Morthier, G. and Vankwikelberge, P. *Handbook of Distributed Feedback Laser Diodes,* Artech House, Boston (1997).

Papannareddy, R., *Introduction to Lightwave Communication Systems,* Artech House, Boston (1997).

PART 7

PHOTOVOLTAICS

Etienne Rapp

11.27 SOLAR RADIATION

The average solar power per unit area received in the Earth's orbit is 1,368 W/m^2. This power decreases like the reciprocal of the square of the distance to the Sun (Figure 11.36 gives the Sun's spectrum). For distant missions it is insufficient, which compels use of RTGs. However, the European Space Agency has optimized cells developed for LILT (low intensity, low temperature) conditions to be used for the Rosetta mission for an insolation representing only 5% of the flux received on Earth.

11.28 PHOTOVOLTAIC CELL

A photovoltaic cell (see Figure 11.37) is made from a *p-n* junction obtained by doping a crystalline or amorphous semiconductor material (space cells are all made on single crystal substrates). A metal layer forming the electric contact is applied to the rear face. On the front face, this contact has the form of a grid to let the light through. In addition, a low-reflecting coating is applied. This bare cell will be equipped with metal interconnectors (designed to assemble cells electrically with each other) and a coverglass to protect it from UV rays and ionizing radiation.

Photons are absorbed in the semiconductor material and convey their energy to the minority carriers of the semiconductor (electrons or holes). This energy is then used to cross the potential barrier and be collected by the metal contacts of the

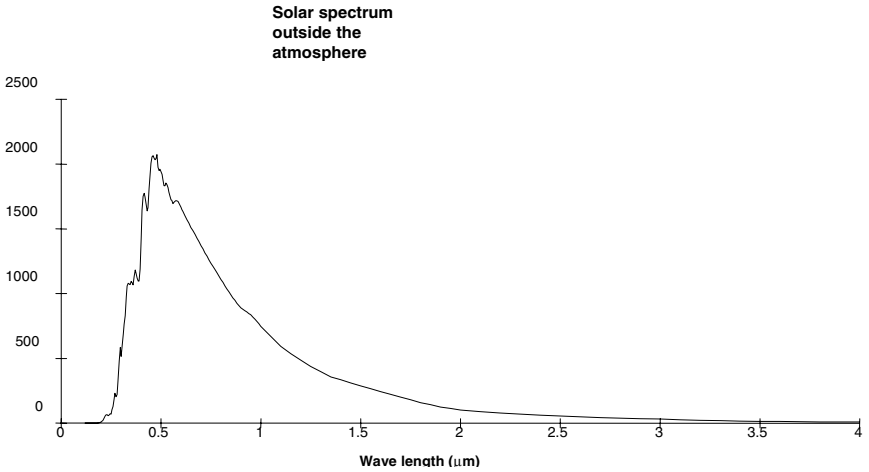


FIGURE 11.36 Spectral distribution of solar radiation.

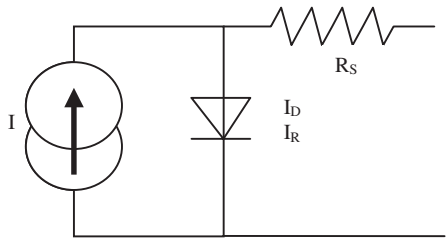


FIGURE 11.37 Equivalent diagram of a photovoltaic cell.

cell. When the cell structure is sufficiently ordered, and for usual illumination, the current collected is proportional to the luminous intensity

The equivalent electrical diagram of a photovoltaic cell is a current source in parallel with a diode. A typical electrical characteristic is presented in Figure 11.38. It is that of a 17.4 cm^2 gallium arsenide cell on board the STENTOR satellite.

The electrical characteristic of the photovoltaic cell can be simply modeled with an analytical formula of a diode coupled with an electrical source.

$$I = I_L - I_{D0} \cdot \exp(q(U - E_g/kT)) - I_{R0} \cdot \exp(q(U - E_g/2kT))$$

$$V = U - R_s \cdot I$$

where I_L = is the line current
U = the internal voltage of the cell
E_g = the material energy gap
R_s = the internal resistance of the cell
I and V = the currents and voltages across cell terminals

The product $I \cdot V$ is the power supplied by the cell. It is conveyed via a maximum $P_{\max}$ at a point $I_{p\max} \cdot V_{p\max}$ of the characteristic. Photovoltaic efficiency is the ratio

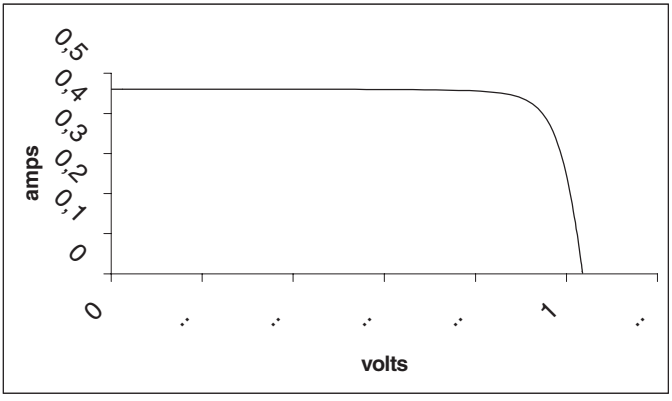


FIGURE 11.38 Typical electrical characteristic of a gallium arsenide cell (STENTOR).

between this power and incident solar power. It is generally provided in laboratory conditions at 25°C (ISO standard).

The photovoltaic cell must be of a material adapted to the type of solar radiation, whose useful wavelength ranges from 0.35 to 1.1 μm, corresponding to energies of from 1.1 to 3.5 eV. If the energy gap is too great, the energy transmitted to carriers is too low to allow them to generate current. If the gap is too slight, current is high but the voltage across the cell terminals is too low and efficiency drops. The optimum gap is 1.4 eV, i.e., that of gallium arsenide.

Direct gap materials (GaAs, InP), in which energy is transferred directly from the photon to the electron can be differentiated from indirect gap materials (silicon) in which this transfer requires a phonon intermediary. Efficiency is greater with direct gap materials.

The *p-n* junction may be made by the diffusion of dopants in the semiconductor material or through epitaxy (crystal growth) from a substrate.

Cells with the best performance are gallium arsenide cells with up to 22% efficiency. These, however, are heavier and more fragile than silicon cells, but above all, material and epitaxy costs are not yet competitive with those of conventional silicon cells made by diffusion. Efficiency of the latter does not exceed 17%.

We now use epitaxied gallium arsenide cells on a single-crystal germanium substrate. This substrate is more robust and less costly than gallium arsenide. The energy utilization rate is approximately 20%.

Two- or even three-junction cells are now marketed. A two-junction cell is a GaAs cell on a germanium substrate on which a GaInP cell is epitaxied, serialized via a tunnel diode. The GaInP cell senses the most energetic photons and converts them into current under a 1.4 V voltage. It lets the weaker photons through, which are retrieved by the GaAs cell under a 1 V voltage. The correct pairing of the two enables 23% efficiency.

Three-junction cells use an extra tunnel diode between the GaAs junction and the germanium substrate. The latter is doped to make a low gap junction, collecting residual infrared photons. Cell voltage is increased by 300 mV. Efficiency reaches 26 to 27%.

American industrialists and laboratories are now working on high-performance, improved three-junction and four-junction photovoltaic structures (InGaP/GaAs/InGaAsN/Ge) whose efficiency may exceed 35%.

11.29 THE SOLAR ARRAY

The solar array is composed of thousands of cells. The association of serial cells is used to obtain the required voltage of 50 or 100 V. The current is obtained by assembling these strings in parallel. In practice, for reliability and modularity reasons, the array is composed of separate electrical sections. For example, with a rigid solar array with deployable panels, we will use two electrical sections per flap.

Dimensioning the solar array for satellite life calls upon many environmental factors.

Maximum efficiency requires the solar array to be kept continuously facing the sun. This condition is rarely put into practice. Depending on the mission, the orbit,

and general optimization of the satellite, the solar array will be fixed or oriented towards the Sun by steering devices with one or two degrees of freedom.

Cell efficiency drops as temperature increases. We try to keep the cells at a low temperature.

The balance between incoming and outgoing radiant fluxes (first thermodynamic principle) determines cell temperature.

Incoming fluxes are:

- Direct solar light, light reflected by the Earth
- Infrared Earth radiation

Outgoing fluxes are:

- Solar radiation reflected by the panel
- Electrical power supplied to the satellite
- Infrared fluxes radiated by the forward and rear faces of the panel

The emissivity of their surfaces must be increased to reduce the balance temperature of cells.

The operational temperature is approximately 50°C to 70°C.

Array performances are affected by loss factors linked to the design losses, manufacture, and aging in the space environment, such as:

- The ratio between active and total area of arrays, which can range from 80 to 95%
- The margins required for guaranteeing the reliability of performances
- Calibration uncertainties of measurements on cells
- The dispersion of performances of cells assembled in series
- The degradation of cells upon the soldering of interconnectors
- Optical losses (or gains) upon bonding of the coverglass
- Irradiation by ionizing particles trapped in the magnetosphere (protons and electrons)
- Irradiation by ionizing particles from solar flares (high-energy protons)
- The darkening of transparent material by the ultraviolet part of the solar spectrum
- The impact of dust particles and micrometeorites which erode optical surfaces

The last may have sufficient weight and speed to cut through the panel and create short circuits in the solar array. Major power loss results if this risk has not been integrated in the design.

The solar array located outside the satellite is subjected to severe constraints. From a mechanical viewpoint, major panels are highly sensitive to acoustic noise under the fairing at satellite launch.

From a thermal viewpoint, due to its great thermooptical emissivity and low heat capacity, a solar array cools considerably when the Sun disappears. At eclipses, the temperature of a solar array of a geostationary satellite drops from 50°C to −180°C. At the end of an eclipse, it returns to its balance temperature in a few minutes. This cycling will be repeated a thousand times throughout satellite life. In low orbit, cycling range is narrower (from −100°C to +90°C) but the number of cycles

is several tens of thousands. This specific feature leads to the absolute necessity for rigorous design and long and costly tests necessary for space qualification. The choice of materials used and assembly technologies is always critical, even for the smallest elements.

Dielectric materials are used for coverglasses. They charge at a different temperature than that of the satellite. They are frequently at a negative potential, but less negative than the satellite. This inverted potential gradient situation may be dangerous. An electrostatic discharge between a metal element (interconnector) and the coverglass may form an electrical arc between two arrays of cells. This arc is transformed into a short circuit and reduces array power. These accidents occur with high-voltage arrays when the available current is sufficient to maintain the arc.

Current solar arrays are made on carbon skin honeycomb plates, insulated by a Kapton sheet. Cells and coverglasses are bonded (and even cold bonded) by very flexible silicon adhesives to absorb the thermoelastic distortion of components. Cell thickness is reduced to 100 μm for silicon and to 140 μm for cells on germanium substrate to reduce the weight.

The power required may now reach 18 kW. It has doubled every four years since the 1970s.

Silicon solar arrays can be of considerable size, over 100 m^2. Their performance is approximately 130 W/m^2 (end of life) and 50 W/kg. They have not yet been replaced by multijunction cell arrays due to the cost of the latter. The choice of the type of array depends on the balance between launch costs per kg and the watt cost. Some arrays use both types of cells, optimizing the cost and power for each mission.

Multijunction arrays are sometimes equipped with solar concentrators in the form of plane mirrors on either side of the array of cells. This makes for a gain in weight and recurrent costs, as the same power is obtained with fewer cells. This gain is obtained through increasing complexity, thermal constraints, and uncertainty about aging. Their performance would reach 70 to 90 W/kg.

Other solar arrays use higher concentration factors, using Fresnel lenses. They are used on interplanetary probes such as Deep Space 1.

PART 8

AIRCRAFT FLIGHT CONTROL SYSTEMS

Dominique Brière

Sylvain Prudhomme

11.30 FOREWORD

Flight control is a very complex engineering challenge that encompasses many disciplines. Flight control systems (FCS) design includes understanding and modeling of the aircraft dynamics, deriving control requirements, hardware and software development and implementation, and finally in-flight validation and operation. It goes far beyond flight control laws (FCL) design, although this particular discipline is generally supported by valuable and intensive theoretical and applied research.

Historically, the technologies available for the first FCS generations and the tools used for their development were very simple. The continuous demand for enhancement in aircraft flight performance has led to progress in technology, more sophisticated FCS, and higher integration in the overall aircraft architecture. Particularly, fly-by-wire (FBW) technology revolutionized aircraft FCS in the 1980s. Several books have been published since that time that will be of interest for students or engineers interested in FCS and willing to learn more (McLean 1990; Blakelock 1991; Stevens and Lewis 1992; Pratt 2000). This section obviously cannot go into detail for all FCS design issues. It should be considered as an introduction that aims only to help the understanding of key issues. Basic information will be found on:

- Flight control objectives and requirements
- Actuation principles and aircraft flight dynamics
- Control laws design process
- Technological issues for flight control implementation

When necessary, information will be particularized for civil and military FCS design. The Airbus FBW experience will be detailed and some open and challenging control problems that pertain to meeting high-level performance requirements and developing flight controls for nonconventional aircraft configurations will be presented.

11.31 FLIGHT CONTROL OBJECTIVES AND PRINCIPLES

Flight Control Objectives

In very general terms, flight control first aims to ensure safe and suitable aircraft maneuvering in the overall flight envelope. The adjectives *safe* and *suitable* have slightly different meanings for civil and military applications.

In civil aviation, *safe* must be understood in terms of airworthiness and certifiability by civil authorities (FAR 25 in the United States, JAR 25 in Europe). Aircraft manufacturers must prove the correct functioning of the FCS in all possible flight conditions, including cases of failures. Such proof can be obtained by extensive analysis and simulation and by flight testing and may be very costly. Because of these stringent safety objectives, the FCS is one of the most critical aircraft systems.

In practice, *suitable* means that the FCS must provide good handling qualities (HQ), i.e., predictable dynamic response to inputs applied to the aircraft. Different requirements are set for two kinds of inputs:

- The aircraft must have a good response to commanded maneuvers (e.g., pilot demand on the stick or rudder).
- Any perturbation should be alleviated (wind shear, failure, etc.) in order to desensitize the aircraft behavior and maintain both desired trajectory and flight comfort.

The distinction between the two kinds of inputs already introduces a first consideration for the architecture of the FCL and different levels of specifications for their design. Feedback control is required to maintain stability and alleviate perturbations. Feedforward control is required to shape demands from the pilot into control signals sent to the actuators. For civil aircraft HQ, each manufacturer develops its own experience in addition to regulation rules and has specific requirements and criteria. Since the civil aircraft mission is unique, the best uniformity in dynamic behavior is generally sought over the entire flight envelope in order to provide the pilot with safe and comfortable handling. Landing and takeoff are exceptions, with specific requirements in terms of climbing, rotation, descent, approach, and flare. Higher-level requirements are set for autopiloting and flight envelope protection and monitoring.

For military aircraft, safety does not lead to the same constraints and must generally be compromised with performance and flexibility for various operational needs. However, a key safety issue is to ensure the stability of aircraft that generally have relaxed or negative natural stability margins, traits that are favorable for high-performance maneuvering. Flight control must provide the best actuation efficiency for any of the various flight configurations encountered in a mission, including weapons carriage, terrain following, ground attack, combat, and refueling. Military HQ requirements can be found in Ministry of Defence (1983) and MIL-1797-A.

Flight Control Principles

The physical principle of flight control is the generation of external forces and moments, either for maintaining aircraft equilibrium or for breaking it in order to induce dynamical motion. The only way to generate such forces is to exploit onboard available power such as thrust and hydraulic/electrical power. The latter is used for moving conventional flight control surfaces through dedicated actuators and for modifying aerodynamic forces and moments as desired on the aircraft (McCormick 1995). At high speed, aerodynamic efficiency is high and only small surface deflection is necessary to generate large forces. Low-speed flight conditions are more critical because of the need for higher aerodynamic coefficients. They generally drive the sizing of conventional moving surfaces. Some additional sophisticated active control technologies (ACT) now are included in FCS with ambitious objectives much beyond the aforementioned basic ones, mainly for military

applications such as described below. Performance requirements have pushed the technology.

Development of hardware actuator technology is a specific discipline of FCS design in which tremendous progress has been made. Some advances from the Airbus experience are presented below. Other descriptions of technology can be found in Mair and Seabridge (2001). We will only retain here the following trends, which tend towards best efficiency at the lowest cost and lowest weight with the highest operational reliability:

- Transmission of power to the control surfaces by mechanical means belongs to the past.
- Hydraulic power is progressively being replaced by electrical power (all-electrical aircraft concept)
- When necessary, generation of hydraulic power is localized rather than centralized (electro-hydraulic actuator technology).

Actuators are controlled using local feedback loops and designed separately before being integrated in the aircraft as specific stability-augmented, high-bandwidth stand-alone subsystems. For this, the assumption is made that couplings with the aircraft dynamics can be neglected. Couplings may appear for specific military aircraft applications, which necessitates the integration of actuator design in overall aircraft design.

Aircraft Modeling

For sizing flight control surfaces and designing the FCL, a good understanding of the aircraft flight dynamics is necessary. As illustrated in several books dedicated to that area (Stevens and Lewis 1992; McCormick 1995; Boiffier 1998; Pratt 2000), most issues related to flight control can be studied using the general equations that govern aircraft dynamics. This section is not the place for going into details. We only discuss the most pertinent issues, using very few and very simple equations.

The general equations of aircraft dynamics have the following form:

$$M(\dot{X}) = F(X,U,W) \tag{11.69}$$

where X = the vector of generalized aircraft internal states (from simple rigid aircraft states such as incidence, pitch angle, and rate to more sophisticated ones that describe the structural dynamics, for example)
U = the vector of inputs to the actuators (thrust, moving surface deflection, etc.)
W = the vector of generalized unmeasured disturbances applied to the aircraft
M = the operator of generalized masses and inertia
F = the operator of generalized forces and moments applied to the aircraft, some being very simple (weight is assumed concentrated at the CG), others more complex (thrust, aerodynamic forces, and moments are nonlinear functions of states X and controls U)

Actually, the construction of such equations relies on the availability of data from propulsion, structures, and aerodynamics. Permanent model updating throughout the aircraft development process positions FCS design as one of the ultimate steps, particularly concerning trimming, HQ assessment, and FCL design.

Trimming consists of finding the controls U_0 that lead to equilibrium $\dot{X} = 0$ for some flight conditions X_0 and steady disturbance W_0 by solving the following equation for U_0

$$F(X_0,U_0,W_0) = 0$$

Trimming values X_0, U_0 are acceptable only if they lead to a stable static equilibrium, i.e., if any small instantaneous disturbance is naturally damped so that X always returns to X_0.

HQ assessment consists of studying the dynamic behavior in the neighborhood of trimmed conditions

$$X = X_0 + x; \quad U = U_0 + u; \quad W = W_0 + w$$

and more precisely of analyzing the locally tangent model obtained after linearization and M inversion:

$$\dot{x} = A(X_0,U_0,W_0)x + B(X_0,U_0,W_0)u + E(X_0,U_0,W_0)w \tag{11.70}$$

Its eigenvalues $\lambda(A)$ are called modes and have physical meanings. Rigid modes such as short period, phugoid, and Dutch roll characterize the low-frequency aircraft dynamic behavior. They should be stable and fall within an acceptable frequency band (neither too fast nor too slow). They can be manipulated by the FCL in order to provide the pilot with suitable HQ.

FCL design consists of computing the gains which will be used to generate signals sent to the actuators $U = U_0 + K(X - X_0) + \bar{K}\bar{X}$ for stability augmentation and HQ enhancement via feedback gain K and for tracking pilot demand $\bar{X}$ via feedforward gain $\bar{K}$, such as illustrated in Figure 11.39.

These three issues are developed in the sequel.

Control Allocation and Aircraft Trimming

Control allocation consists of finding the best correlation between available control surfaces and aircraft motions to be controlled. In the most general case, considering that all moving surfaces (and all additional control technologies) contribute to the generation of forces and moments in all directions would lead to a very complex multivariable control problem whose complexity goes beyond this Part. For conventional aircraft such as illustrated in Figure 11.40, the architecture of wing and

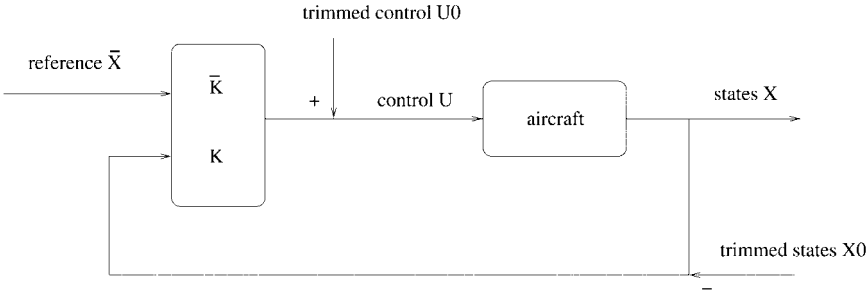


FIGURE 11.39 Generic FCL architecture.

AVIONICS AND ASTRIONICS

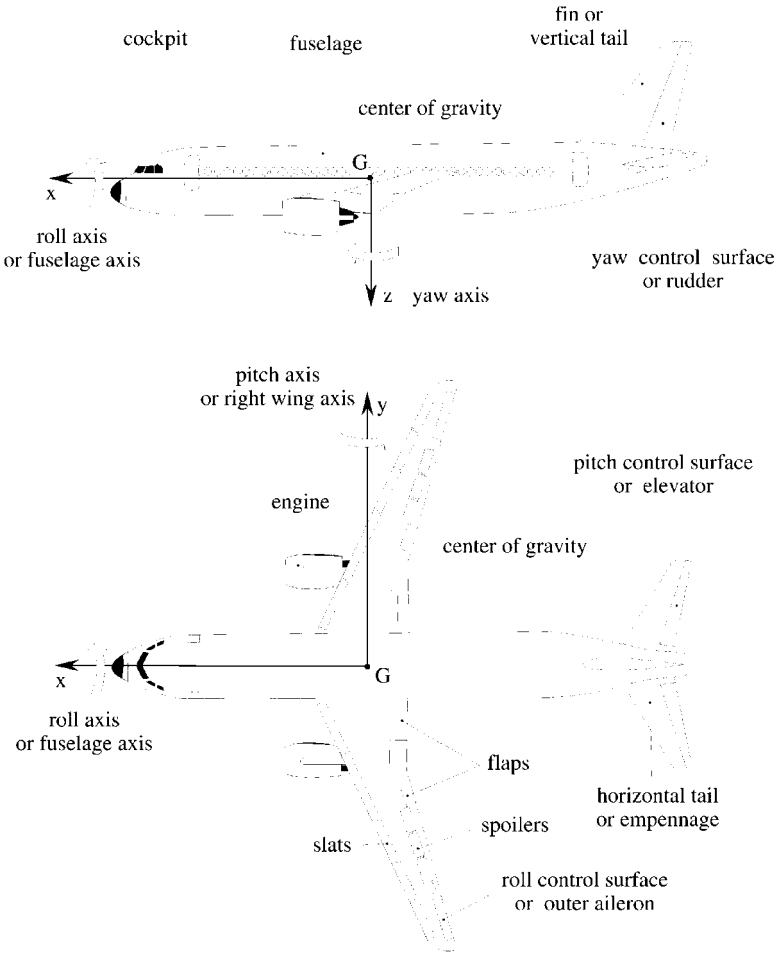


FIGURE 11.40 Conventional actuation architecture.

tail allows natural control allocation by decoupling longitudinal and lateral actuation. For longitudinal motion, thrust generates forces against drag and allows low-frequency longitudinal axis motion control, high-lift devices (flaps and slats) generate extra lift coefficient and vertical forces which provide some low-frequency control on the vertical axis, and the horizontal tail (when trimmable) and the elevator generate moments for low and higher frequency pitch control, respectively. For lateral motion, the ailerons and rudder generate moments in roll and yaw, respectively.

For conventional aircraft, the longitudinal and lateral dynamics are decoupled because of important items:

- Mass and inertial distributions lead to the diagonal dominance of M and a natural splitting of states X into decoupled X_{lon} and X_{lat} on the left-hand side of equation (11.69).

- Conventional tail configurations allow the generation of decoupled forces and moments into F_{lon} and F_{lat} on the right-hand side.
- For most steady trimmed flight conditions in either X_{0lon} or X_{0lat}, aerodynamic coupling effects involved in cross-terms $F_{lon}(X_{lat})$ and $F_{lat}(X_{lon})$ are expressed very simply, so that equation (11.69) splits into two decoupled ones:

$$M_{lon}(\dot{X}_{lon}) = F_{lon}(X_{lon},U_{lon},W) + F_{lon\text{-}lat}(X_{0lat},W)$$
$$M_{lat}(\dot{X}_{lat}) = F_{lat}(X_{lat},U_{lat},W) + F_{lat\text{-}lon}(X_{0lon},W)$$

Trimming then appears as a decoupled problem, first in longitudinal for given flight conditions X_{0lon}, X_{0lat},W_0 by solving the following equation in U_{0lon}:

$$F_{lon}(X_{0lon},U_{0lon},W_0) + F_{lon\text{-}lat}(X_{0lat},W_0) = 0 \qquad (11.71)$$

then in lateral by solving the dual equation in U_{0lat}. Another simplification is possible, but only for conventional two lifting surface configurations, which makes longitudinal trimming easier. The aerodynamic forces on the wing are not dependent on the controls on the horizontal tail. This has two major advantages:

- Equation for longitudinal trimming (see eq. (11.71)) further simplifies into:

$$F_{wing}(X_{0lon}) + F_{tail}(U_{0lon}) + F_{lon\text{-}lat}(X_{0lat}) = 0$$

and can be solved explicitly in U_{0lon}. Actually, the longitudinal trimming problem is itself decoupled in lift (wing), drag (thrust), and pitch (tail). This allows simple expressions of static constraints for actuator sizing, including margins to ensure static stability over a given range of mass distribution and CG location.

- Actuator sizing and wing performance are decoupled (at first order). The tail can be sized without penalizing the drag of the wing, the total drag being computed as $C_{Dwing} + C_{Dtail}$, and minimized using remaining degrees of freedom. High-speed and low speed trimming can be solved independently without compromising performance.

Control surfaces are sized with different kinds of criteria. Among the most stringent constraints are those related to the requirement to maintain aircraft equilibrium in the presence of a maximum disturbance W_0 expected to be encountered in flight (wind, engine failure, as imposed by certification rules), mainly at landing. Some additional criteria coming from dynamical constraints (see below) may be taken into account.

The general architecture of military aircraft introduces coupling effects in both mass/inertia distribution because of compactness and force/moment generation because of actuator geometry. Control allocation is more complex than for civil aircraft. It must be considered that some control surfaces can generate forces in several directions and that in-flight dynamic allocation may be necessary either for reconfiguration in case of failure/damage or for permanent optimization of total efficiency by exploitation of redundancies.

HQ and Dynamic Stability

HQ can be assessed such as in Hodgkinson (1999) after trimming using equations (see eq. (11.72)). The coefficients in the time-invariant matrices A,B,E are aerodynamic derivatives $\partial F/\partial X(X_0,U_0)$ and $\partial F/\partial U(X_0,U_0)$, such as C_α, C_{mq}. For con-

ventional aircraft configurations, the simplifications already introduced in the previous sections apply:

- Mass and inertia distribution allow separation of x into x_{lon} and x_{lat}
- Control allocation allows separation into longitudinal and lateral contributions
- Terms related to aerodynamic coupling effects generally vanish when trimmed so that decoupled equations for both motions yield:

$$\dot{x}_{lon} = A_{lon}\, x_{lon} + B_{lon}\, u_{lon} + E_{lon}\, w \qquad (11.72)$$
$$\dot{x}_{lat} = A_{lat}\, x_{lat} + B_{lat}\, u_{lat} + E_{lat}\, w$$

or in the Laplace domain:

$$x_{lon} = (sI - A_{lon})^{-1}\,(B_{lon}\, u_{lon}(s) + E_{lon}\, w(s))$$
$$x_{lat} = (sI - A_{lat})^{-1}\,(B_{lat}\, u_{lat}(s) + E_{lat}\, w(s))$$

Each equation determines specific modes, the number of which depends on the complexity of the retained modeling. In longitudinal, $\lambda(A_{lon})$ includes at least the short period and phugoid modes. In lateral, $\lambda(A_{lat})$ includes Dutch roll, spiral, and roll. Removing negligible terms leads to understandable formulas for frequencies and damping ratios, which have physical interpretation. Such analytical analysis forms the basis of HQ assessment. A good knowledge of the masses and inertias (input from structures) as well as the main aerodynamic coefficients (for example C_α, C_α for short period) is necessary. Those related to steady aerodynamics are well predicted by CFD, at least at high speed. Wind tunnel testing (WTT) or flight testing may provide updated estimates. For coefficients related to unsteady aerodynamics (C_{mq}), CFD prediction accuracy around 30% is unacceptable. Dynamic WTT can provide a better 10% accuracy.

An additional requirement is to avoid aircraft pilot coupling (APC). This phenomenon (formerly called pilot induced oscillations, PIO) is due to an implicit feedback loop introduced by the pilot. He may inject demands correlated to his feeling of the aircraft motion into the actuators, potentially destabilizing the overall closed loop if not properly phased. Such requirements are more stringent for military aircraft. Description of APC categories, criteria, and requirements can be found in Pratt (2000).

In the case where the natural behavior of the aircraft would not be good enough from a HQ point of view, different options are possible:

- Update aircraft configuration and design to make it better.
- Enhance the HQ with an automatic flight control system (AFCS), such as developed in below.

11.32 FLIGHT CONTROL SYSTEMS DESIGN

The FCS Design Process

The FCS design and development processes are long-term programs. There are a variety of approaches used among manufacturers based on their own experience and particular needs. Pratt (2000) and Moir and Seabridge (2001) give an overview

of technical activities undertaken in the development programs and their scheduling in the standard *V*-diagram (requirement-specification-development-integration-testing process) of Figure 11.41. The main concern is to meet requirements while guaranteeing certifiability at the lowest cost. This implies very structured development methodologies in order to keep traceability and control, including ground and flight testing (Bourcier 2001). The following subsections will be devoted to a more detailed presentation of the FCL design. In the first steps, models used for such design only include prediction values from structures and aerodynamics. They are refined throughout the ongoing aircraft design process. Final knowledge of the aircraft dynamics is only available after ground and flight tests, just before operation. The FCL design process must also include tools and procedures for model identification (Ljung 1987; Söderstrom and Stoica 1989) and for FCL tuning (Prudhomme and Gimonet 2001).

FCL Design Principles

The most sophisticated control design techniques may be very powerful but useless if basic aircraft stability is not understood. Before moving to modern design, this subsection gives basic physical principles of control law design (McLean 1990; Stevens and Lewis 1992). Most civil or military aircraft do not meet all the HQ requirements in open loop, generally because modes (short period or Dutch roll) are not adequately damped. Some military aircraft may even be unstable. Stability augmentation systems (SAS) provide additional artificial damping by connecting a feedback loop between measurements (generally angular rates) and the commands sent to the actuators in a single input, single output (SISO) scheme. Taking the example of longitudinal motion and assuming perfect lateral decoupling, the second order short period approximation of equation (11.72) gives the following open loop transfer function between elevator deflection $u_{\text{lon}} = \delta_q$, vertical wind w, and pitch angle $x_{\text{lon}} = \theta$ in the Laplace domain:

$$\theta(s) = \frac{1}{\omega^2 + 2\zeta\omega s + s^2}\left(k\delta_q(s) + l\ \omega(s)\right) \tag{11.73}$$

The natural damping ratio ζ is proportional to C_{mq} and C_α and may be small or

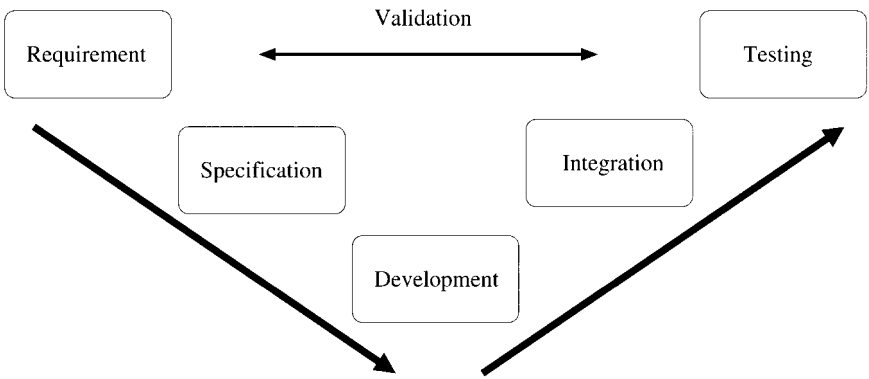


FIGURE 11.41 Standard *V*-diagram development process.

negative, i.e., unacceptable from an HQ point of view. It can be modified by a feedback loop using measurement of pitch rate $q = \dot{\theta}$ and computing:

$$\delta_q = -Kq + \bar{\delta}_q$$

This leads to the closed loop transfer function

$$\theta(s) = \frac{1}{\omega^2 + (2\zeta\,\omega + kK)s + s^2}\,(k\,\bar{\delta}_q(s) + \omega(s))$$

If K is correctly tuned, the closed loop damping ratio is augmented from ζ to $\zeta^x = \zeta + kK/2\omega$. It allows artificial stabilization of a naturally unstable aircraft. Similarly, using a measurement of θ enables the tuning of the disturbance rejection bandwidth by manipulation of the natural frequency ω and the θ/w transfer function. For lateral motion, the principle of SAS is similar: it consists of feeding yaw rate measurement r on the rudder actuator $u_{\text{lat}} = \delta_r$ for augmenting natural damping of the Dutch roll mode, with possible difficulties (not detailed here) due to couplings with roll and spiral modes.

In manual control, control augmentation systems (CAS) aim to provide acceptable aircraft time responses to pilot inputs. As illustrated in Figure 11.42, the laws are more sophisticated than for SAS. An additional feedforward gain $\bar{K}$ for filtering the pilot inputs is designed simultaneously with the SAS feedback. For the same example of longitudinal control, using the following control signal:

$$\delta_q = -Kq + \bar{K}(\bar{x}_{\text{lon}} - x_{\text{lon}})$$

in equation (11.72) leads to a modified closed loop transfer function

$$x_{\text{lon}}(s) = \frac{T}{\omega^2 + k\bar{K}T + (2\zeta\omega + kK)\,s + s^2}\,(k\bar{K}\bar{x}_{\text{lon}}(s) + l\,w(s))$$

where $T(s)$ is the x_{lon}/θ transfer function. Both gains contribute to

- The augmentation of the closed loop damping and the manipulation of the natural frequency

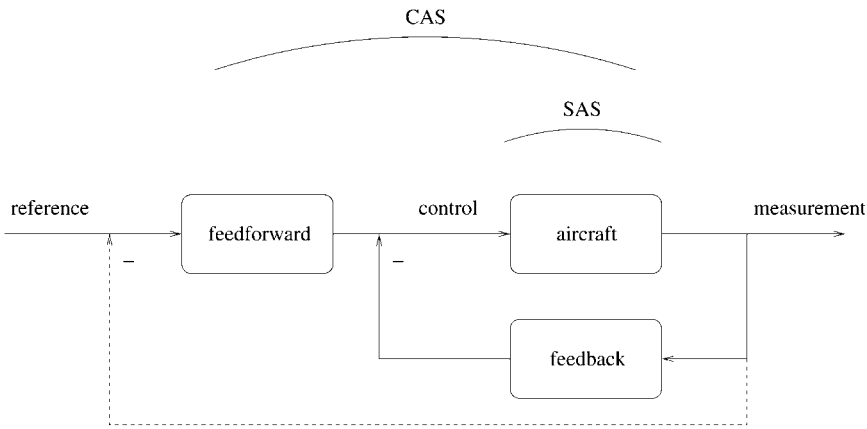


FIGURE 11.42 Illustration of SAS and CAS architecture.

- Disturbance rejection in a specified bandwidth (small x_{lon}/w for high $K,\overline{K}$)
- Additional features for the pilot, depending on the selection of reference $\overline{x}_{lon}$ (vertical acceleration N_z or pitch angle θ) and on gain structure: proportional gain ($\overline{K}$ constant) is for short-term response, proportional integral (PI) gain also offers long-term reference tracking ($x_{lon} = \overline{x}_{lon}$).

Autopiloting is the ultimate specification level. It consists of automatic tracking of target variables without any pilot input. Methods used for CAS may apply with specific selection of reference variables (altitude and speed in cruise, pitch angle or slope in climb, glide angle for landing, etc.). Gain $\overline{K}$ always include some integral terms (PI control with high gains at low frequency) for meeting the main requirement of zero steady state error. These ultimate levels in requirements do not induce additional needs in terms of technology (actuators or measurements). Requirements generally can be met by specific FCL design. The interrelationship among flight control, guidance and flight management is illustrated below and developed in Moir and Seabridge (2001).

For military aircraft, APC prevention consists of designing gains with values high enough for performance, but not too high, in order to avoid possible coupling between the stability-augmented aircraft and the pilot's own gain in the controlled bandwidth.

Modern FCL Design

SAS and CAS show limitations for multi-input, multi-output (MIMO) or high-order dynamic systems. Modern control design is an extension of these basic methods and opens perspectives for high-performance control of complex aircraft configurations showing high-order dynamics such as flexible aircraft or strong couplings between axes, or having redundant movable control surfaces. It provides unified tools for formulating all control objectives, including stability augmentation, tracking, disturbance rejection, etc. It has been used in the past for the integration of flight and propulsion control (FADEC) into a single MIMO integrated flight propulsion control system (IFPCS).

For MIMO control design, transfer functions are only formally used for the purpose of notation simplicity in representing generic aircraft model

$$y(s) = [G_u(s)G_w(s)]\begin{bmatrix} u(s) \\ w(s) \end{bmatrix}$$

and two-degree-of-freedom dynamical control laws simultaneously using multidimensional measurements y and reference signals $\overline{y}$

$$u(s) = [-K(s)\overline{K}(s)]\begin{bmatrix} y(s) \\ \overline{y}(s) \end{bmatrix}$$

Feedback loops have the same structure as in Figure 11.42, but in MIMO:

$$y(s) = (I + G_{uK})^{-1}[G_u\overline{K}G_w]\begin{bmatrix} \overline{y}(s) \\ w(s) \end{bmatrix}$$

It is clear from these equations that feedback MIMO transfer function $K(s)$

mainly deals with stability through $(I + G_{uK})$ and disturbance rejection through y/w and that feedforward gain $\overline{K}$ mainly deals with tracking through $y/\overline{y}$.

State space representation is more convenient for practical control law computation in the time domain. In its generic form for either longitudinal, lateral, or coupled motion, the aircraft model writes:

$$\dot{x} = Ax + Bu + Ew$$

$$y = Cx + Du + Jw$$

Using matrices A,B,C,D,E,J, different techniques are available for computing a generic control law under its own state space realization. For static output feedback (K constant) and assuming the aircraft is proper ($D = 0$), the obtained closed loop writes:

$$\dot{x} = (A + BKC)X + [B\overline{K}\ E + BKJ]\begin{bmatrix}\overline{y}\\ w\end{bmatrix}$$

$$\begin{bmatrix}y\\ u\end{bmatrix} = \begin{bmatrix}C\\ KC\end{bmatrix}X + \begin{bmatrix}0 & J\\ \overline{K} & KJ\end{bmatrix}\begin{bmatrix}\overline{y}\\ w\end{bmatrix}$$

The order of the dynamics is unchanged. Gain K directly fixes the closed loop modes and HQ. However, the number of degrees of freedom in K (number of inputs · number of outputs) may be too small to meet all requirements. Dynamic feedback design offers more versatility by augmenting the closed loop dynamics. It is more challenging, however, since it requires more sophisticated tools.

For detailed information on theory and application of modern MIMO design, readers should refer to one of the many recent books. Some available techniques are based on algebraic manipulation of state space equations. One of the most powerful ones for linear systems is eigenstructure assignment (Liu and Patton 1998), which consists of assigning selected closed loop eigenvalues. It is very convenient for controlling the HQ of civil aircraft since frequencies and damping ratios are closely related to transient time responses and amplitudes of overshoots for standard step pilot demand and since it allows modal decoupling. The Airbus experience is based on this technique (Farineau 1989), as detailed below. For military applications, dynamic inversion also exploits algebraic manipulation of nonlinear control equations (Slotine and Li 1991) for constructing nonlinear control signals leading to linear closed loop behavior. Because they rely on a full description of the systems to be controlled, enhanced versions of these basic algebraic methods must be used for improving robustness. Other linear control design techniques are based on minimization of some control performance index (Maciejowski 1989), with various levels of sophistication: linear quadratic (LQ), Gaussian (LQG), loop transfer recovery (LQG/LTR), H_2 and H_∞ optimal control. Most control design problems can be formalised under the standard form of Figure 11.43. The augmented system $M(G)$ includes the aircraft model G and additional shaping functions to characterize and specify the control problem (tracking, disturbance rejection, autopiloting, etc.). Resolution consists of computing the control laws K which minimize the closed loop transfer between hexogeneous inputs w and controlled outputs ε in either H_2 or H_∞ norm. For instance, if w characterises initial values of aircraft states, states are measured ($y = x$) and ε represents weighted states $\sqrt{Q}x$ and control effort $\sqrt{R}u$, we have a standard LQ problem. If w are stochastic noises and ε the state estimation error, we have a LQG problem. If w is a standard pilot demand

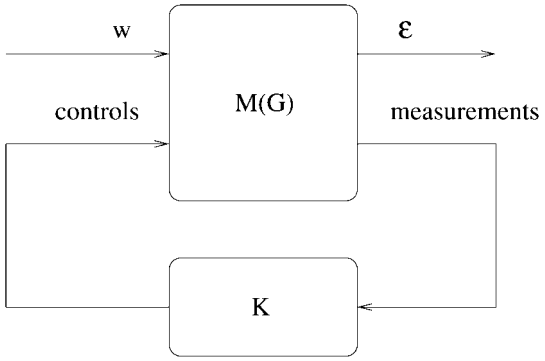


FIGURE 11.43 Standard form for control design.

and ε the distance to the outputs of a target model, we have model following control. Individual applications can be found in Magni et al. 1997. A comparison of best design practices in modern control is available in Moorhouse (2000).

Modern design techniques have recently been extended to robust control by inclusion of robustness analysis using the μ theory. The basic idea is to put aircraft model under the linear fractional transformation (LFT, also called M-Δ form) of Figure 11.44. Such modeling is a particular challenge (Ackermann 1993), but it

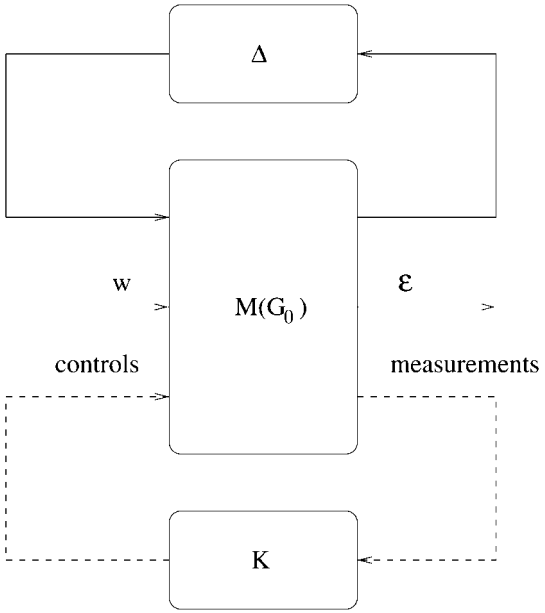


FIGURE 11.44 LFT representation for robustness analysis or robust design.

offers capabilities for representing a continuum of models using a nominal model G_0 with an artificial feedback containing different kinds of uncertainties (parameter variations, unmodeled dynamics, etc). Tools are available for computing robustness margins (μ-analysis), i.e. amplitude of parameter variations which lead to unstability when applied to the nominal model (Ferreres 1999) and for robust control design (μ-synthesis, linear parameter variant LPV synthesis).

For nonlinear military aircraft, some nonconventional control design techniques also are available (such as QFT and neural network control) but are beyond the scope of this Part.

Automatic FCL Implementation

For the sake of simplicity, for physical interpretation of models and because of tool capability, automatic FCLs generally are designed in the continuous time domain. However, on-board implementation tends to be performed on computers working in the discrete time domain. Transforming FCL from continuous to discrete is a specific automatic control problem for which solutions are given in Fargeon (1989) and McLean (1990). The principle is to have sampling periods small enough so that digital control is perceived as quasicontinuous by the actuators and the aircraft. Several sampling periods may be necessary for digital control of subsystems having different bandwidths. Actuator control must be very fast (few ms sampling period) because of their high bandwidth (several Hz). In civil applications, flight control may be slower, with a piloting bandwidth of about 1 Hz and gain scheduling may be slower still for aircraft that maneuver smoothly. However, requirements are more stringent and sampling periods shorter for military aircraft maneuvering faster.

FCS design teams have developed their own skills and technologies in terms of architecture, interface and redundancy. Figure 11.45 shows a typical FCS interface with other systems. Other information on general system architecture is available in Pratt (2000) and Moir and Seabridge (2001). Figure 11.46 illustrates the very

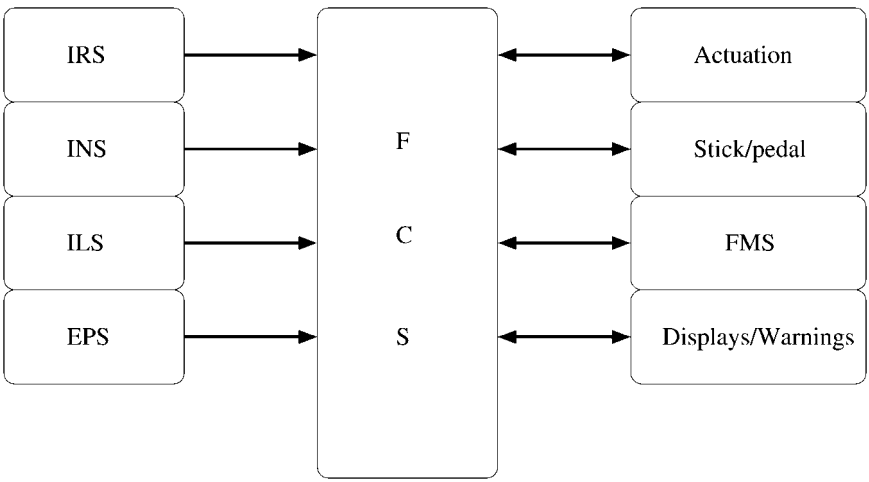


FIGURE 11.45 Typical FCS interface with other systems.

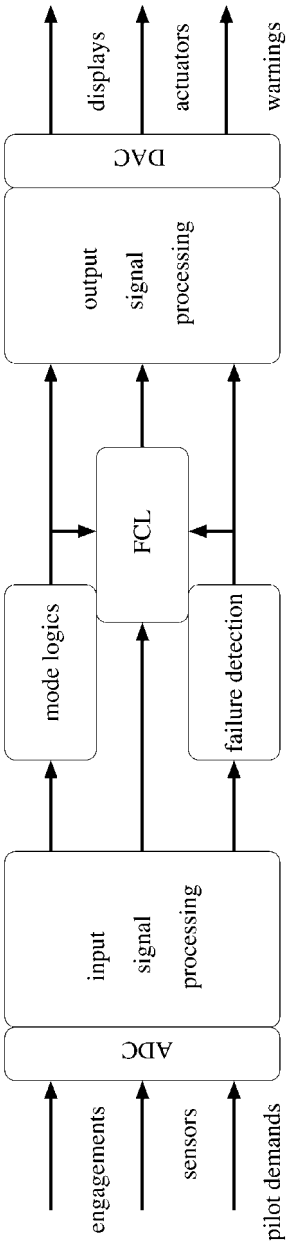


FIGURE 11.46 General FCC architecture.

generic architecture of a flight control computer (FCC). The Airbus experience is presented below. Details on avionics technology can be found in Part 10 and in Spitzer (1987) and Collinson (1996).

11.33 AIRBUS FLY-BY-WIRE: AN EXAMPLE OF MODERN FLIGHT CONTROL

Since the A320's entry into service in 1988, the Airbus fly-by-wire (FBW) aircraft fleet has become a diverse and rapidly growing family of civil aircraft. At the end of 2001, about 2000 Airbus FBW aircraft had accumulated more than 20 million flying hours and were operated worldwide by a large variety of airlines with various cultures and backgrounds. Consequently, Airbus FBW may be considered as a legitimate reference in the domain of flight control systems. The purpose of the following subsections is to present the main flight control system features shared by this wide family, from the shorter A318 to the larger A380. The high-lift system that controls flap and slat surface extension at low speed is often designed as a secondary flight control system and is outside the scope of this presentation.

Background

Until the 1950s, flight control surfaces were mechanically linked to the cockpit controls and consequently actuated by the pilot's muscles. On such aircraft, flight control surfaces were balanced with an additional mass and horn and their rear part fitted with tabs to aerodynamically amplify or compensate pilot inputs. This technology is still successfully applied on turboprop aircraft such as the ATR 42 and ATR 72. However, aircraft size and speed could not have been increased without installing servo-controls (Figure 11.47) either to amplify the efforts generated by the mechanical linkages or to replace the linkages altogether. Direct mechanical actuation of control surfaces still exists as a complement to servo-controls or as a

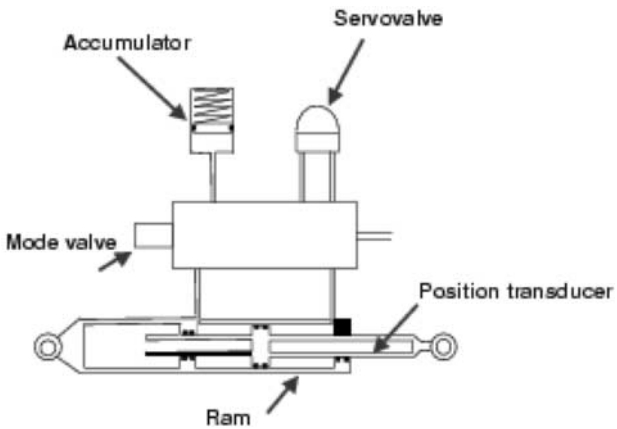


FIGURE 11.47 Servo control principle.

backup after hydraulic failure on many transport aircraft designed in the 1960s and the 1970s such as the DC9, Boeing 737, and Fokker 100. We will now discuss the main historical stages of the Airbus flight control systems development (first flight years are indicated in parentheses).

Caravelle (1955). This 100-seater turbofan aircraft represents the first example of a transport aircraft with permanently servo-controlled flight control surfaces. This means that on the Caravelle, mechanical linkages are no longer used to drive the surfaces directly but are instead used to transmit pilot commands to the servo-controls. The Caravelle hydraulically powered flight control system is probably the first example of a full-time critical active control system on a transport aircraft. All-weather automatic landing is another important innovation of this aircraft that opened the way for the use of electricity in critical applications. The Caravelle can be considered an early precursor of the Airbus family.

Concorde (1969). For the first time on a transport aircraft, pilot orders are transmitted to surface servo-controls by electrical wires. FBW is chosen on the Concorde to minimize the risk of HQ degradation in supersonic and transonic flight that could result from the nonlinearity (backlash, hysteresis) inherent in a mechanical transmission. A mechanical flight control system nevertheless remains installed, but only as a passive backup that will become active only after failure of the dual duplex electrical flight control systems. All the flight augmentation and flight envelope protection functions are achieved by analog computers (pitch and yaw damping, artificial feel, automatic trim, angle-of-attack protections).

A300 (1971). This aircraft launches the European Airbus story. All the flight control surfaces are permanently actuated by hydraulic servo-controls, as on the Caravelle and the Concorde, and mechanically signalled, as on the Caravelle. The flight augmentation function and envelope protection are performed by analog computers, as on the Concorde.

In the beginning of the 1980s, digital technology becomes attractive for civil aircraft application, thanks particularly to microprocessor industrialization, and quickly supersedes analog technologies. The first applications on Airbus aircraft are the digital autopilot and flight-augmentation computers; the same basic flight control architecture is maintained.

A310 (1982). FBW with digital computers appears on Airbus A310 and A300-600 with spoiler control. These upper wing surfaces provide roll control in conjunction with the ailerons when antisymmetrically deflected on one wing and in-flight speed braking or on-ground lift dumping when symmetrically deflected on both wings. The main features of the A310 spoiler control will be applied later on the Airbus FBW family. They are summarized here:

- All the mechanical linkage and servomotors previously needed on the basic A300 spoiler control system are eliminated and all spoiler actuators electrically signalled. System weight is notably reduced and its installation simplified.
- The command/monitoring computer is the basic architecture building block. It allows detecting and neutralising of command channel failure. Spoiler actuators are controlled by four pairs of command/monitoring units installed in two boxes.
- Each command or monitoring unit includes all the electronic components it needs to perform its function, without sharing resources with another: input and output circuits (analog, discrete, digital), processing component, electrical supply.

- All functions, including spoiler actuator servo-loops, are software based and fully under control of digital units. Actuator servo-valve command current and feedback position acquisitions are multiplexed in order to reduce drastically the electronic volume dedicated to each actuator.
- Roll control is optimized according to airspeed and flap position. Thus, spoiler contribution to roll control is significantly increased in comparison to the previous A300 mechanically signalled system. This makes it possible to eliminate the outboard aileron on A310 and A300-600 and maintain only the inboard aileron.

The EFCU also controls the automatic decoupling of the two pitch control channels at low speed and, on pilot demand, the automatic return to zero of rudder trim.

But the main step is the *A320* (1987), prepared by full-scale experimentation on Concorde (1978), then on the Airbus A300 (1983). This experimentation validated the essential evolutions further applied on the Airbus FBW families:

- Enhanced flight control laws and flight envelope protection
- Side-stick pilot controls

The A320 launched the short/medium-range Airbus single-aisle family, with the stretched *A321* (1993), then the shrunk *A319* (1995) and *A318* (2002) versions. The same flight control philosophy is maintained on the four-engine *A340* (1993) and twin-engine *A330* (1994) long-range family, their latest derivative version *A340-600* (2001) and *A340-500* (2002), the very large *A380* (target: 2004), and the military transport aircraft *A400M* (target: 2005).

Architecture Principle

In a mechanically signalled flight control system, pilot commands are sent to servo-controls by mechanical cables and rods. However, handling qualities would not be acceptable without an electrical system with computers being added to the basic mechanical system in order to perform flight-augmentation functions and flight envelope protection such as yaw damping, pilot control artificial feel, and stall warning (stick buzzer, stick pusher). The computer output is incorporated into the mechanical system by hydraulic and electrical motors. Failure of these functions is generally not hazardous as long as the basic mechanical system remains operative but handling qualities are downgraded.

The FBW term is used when the pilot (or autopilot) orders are no longer sent by mechanical cables and rods to servo-controls, but instead by computers and electrical wiring (see Figure 11.48). Consequently, computers become a basic and essential part of the flight control system.

In comparison to the previous mechanically signalled architecture, FBW enables weight savings and simplification of system installation. Evolution can be limited to this change. In addition, FBW technology offers the possibility of enhancing HQ, improving safety in emergency situations, improving cockpit ergonomy, and facilitating cross-crew qualification. This is the way selected by Airbus.

All Airbus FBW aircraft share some main features:

- Side-sticks without motorization
- Enhanced flight control laws and flight envelope protection

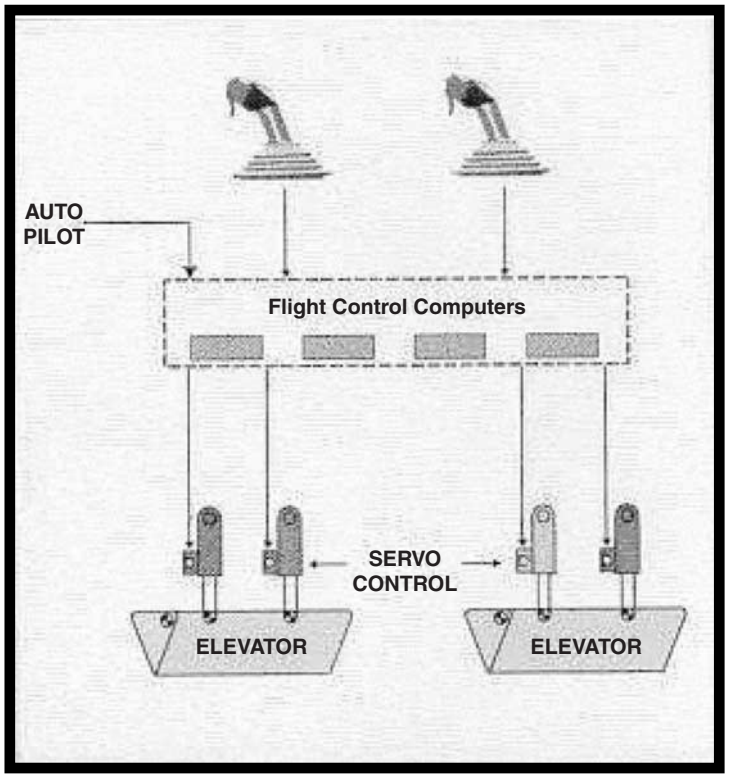


FIGURE 11.48 Airbus elevator control FBW principle.

- Architecture based on a set of highly dependable and dissimilar FBW digital computers
- Ultimate backup independent of the basic FBW system

A total flight control loss or uncontained surface runaway may lead to severe consequences. The probability of an occurrence of such failures must be extremely improbable. Some basic design principles are applied to fulfill these stringent safety objectives:

- *Redundancy:* Control channel, power supplies, surface actuators, sensors, etc. are at least duplicated so that the consequence for an aircraft of any failure or combination of failures remains acceptable according to its effect on safety. As a general rule, a single failure should have minor consequences
- *Monitoring:* Computers must permanently monitor not only the computation they perform but also the signals they receive and the actuators they command. In addition, periodic tests may be necessary to detect hidden failures that cannot be detected by computer on-line monitoring.

- *Segregation:* Components, electrical routes, and hydraulic circuits must be installed to reduce the risk that a common cause, such as projections of fragments resulting from an engine burst, will lead to a total system failure.
- *Robustness:* The flight control system must be demonstrated to be immune from external aggressions such as lightning strike, electromagnetic perturbations, high and low temperature, loss of cooling air, vibrations, supply source perturbations
- *Dissimilarity:* this additional precaution is selected by Airbus to reduce the risk that a common hardware or software defect may lead to total flight control loss (e.g.: a common defect on a batch of electronic components).

Pilot Controls

Large displacement central control columns are necessary to drive mechanical systems with enough accuracy and actuators to introduce flight augmentation, flight envelope protection, and autopilot system. Hence, linked and back-driven central pilot controls are a consequence of the mechanical technology.

On a fly-by-wire aircraft, pilot control no longer drives mechanical linkages, but transducers. Large-displacement central columns are no longer suitable to sense pilot orders; lateral well-in-hand side-sticks may be used (see Figure 11.49). In addition, side-sticks improve cockpit ergonomy and comfort with better visibility of instrument panels, easier accessibility, and the possibility to install a pull-out table in front of each pilot.

FIGURE 11.49 Airbus cockpit, with lateral side-sticks.

Controlling the aircraft with the rudder only on lateral axis and the tail trimmable horizontal stabilizer (THS) on the pitch axis was previously demonstrated with A300 flight tests. The A320 FBW takes advantage of this result and employs a mechanically signalled servo-control on the rudder and an actuator on the THS (THSA), controls activated from the rudder pedestal and pitch trim hand-wheel, respectively. This architecture allows side-stick and FBW implementation while maintaining a mechanically signalled backup.

Jamming of the cockpit controls is a rare but very severe failure case. It may be caused by any jamming of the pilot control itself or the mechanical linkage it drives, or by pilot incapacitation. If the two sticks are linked, the only way to avoid losing aircraft control if such jamming occurs is to install a decoupling device between the two sticks. A large force must be applied on the unfailed stick to declutch it from the affected one, which leads to a significant transient on the aircraft trajectory. Thus, a simple failure is capable of causing a severe situation on a mechanically signalled flight control aircraft or on a FBW aircraft with linked pilot sticks.

On Airbus FBW, side-sticks are no longer physically coupled, roll and pitch orders from each side stick being algebraically added together by the flight control computers. Therefore, the jamming of one stick will have no effect on the other. If such a failure occurs, the opposite pilot can immediately take control of the aircraft with his unfailed stick and cancel orders of the failed stick. Furthermore, he can quickly and easily neutralize orders from the failed stick by depressing a switch located on its handle.

Dual input awareness is adapted to this architecture without altering its basic advantage. Several ways are implemented to alert the crew: visual warning by flashing lights in front of each pilot (basic solution on Airbus FBW), audio warning, and now tactile warning by vibrating the handles (but not the sticks).

Autopilot orders are sent from autopilot computers to flight control computers by ARINC 429 digital buses. Servomotors are no longer useful, as the mechanical linkage has been eliminated. The active failures they were capable of introducing on flight controls are consequently eliminated. A solenoid will freeze each side-stick at neutral as soon as the autopilot mode is engaged by the crew and gives nonambiguous tactile information. It may be overridden by depressing the instinctive disconnect switch on one of the sticks or by applying a firm, but not excessive, force on each side-stick. Either of these actions will immediately disconnect the autopilot and restore manual control.

Control Laws

The flight control law objective is to achieve an accurate, stable and safe aircraft response to pilot input at any point in the flight domain, whatever the aircraft configuration and flight phase. It is obtained on Airbus aircraft by closed loop control laws that incorporate flight envelope protection (Figure 11.50):

- Inside the normal flight domain, aircraft static stability is neutral, cross-effects between axes are automatically compensated (turn coordination, longitudinal compensation), elevators are automatically trimmed at neutral by the THSA. Aircraft pitch and roll attitudes are held as soon as the stick is at neutral and small stick corrections are sufficient to compensate for a configuration change, thrust variation, speed variation, wind gust effect, etc. The natural behavior of the lateral

AVIONICS AND ASTRIONICS

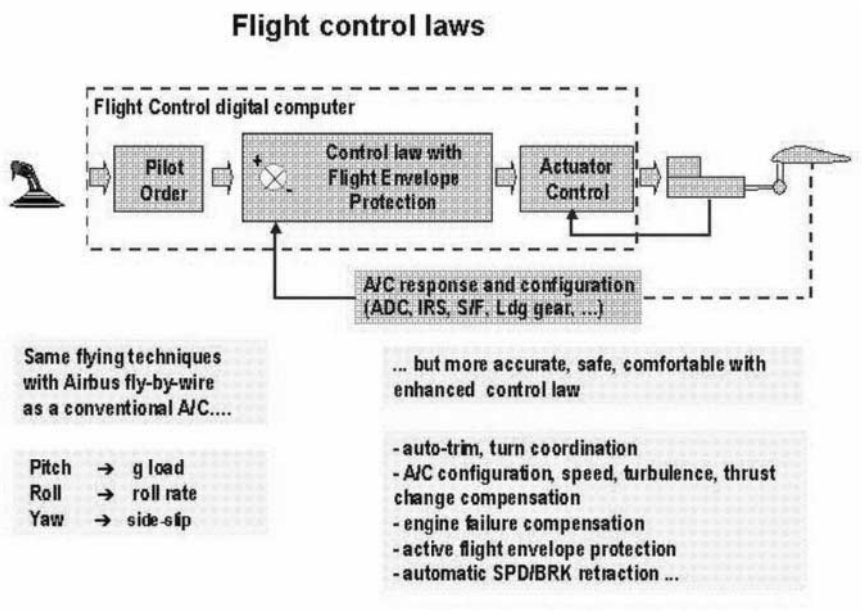


FIGURE 11.50 Airbus flight control principle.

control law automatically compensates for a lateral dissymmetry. Hence, in case of engine flameout, the aircraft will reach a limited bank angle and sideslip while staying in the normal flight domain.

- Outside the normal flight domain, positive stability is restored: the pilot must permanently apply a stick deflection to maintain the aircraft inside the peripheral domain. With the stick at neutral, the aircraft will come back to and stay at the limit of the normal flight domain.
- By applying full stick, the pilot will command the aircraft to fly at the limit of the peripheral domain.

The aircraft being thus protected against stalling and overstressing, the pilot can quickly and safely obtain maximum aircraft potential by a full stick deflection. Consequently, aircraft maneuverability is significantly improved.

A control law must not rely on a single resource if its failure may have severe consequences. Therefore, at least two independent sources are usually needed for each type of parameter to perform the normal control law described above (Figure 11.51). For example, three independent sources provide inertial or air data parameters on Airbus. If one fails, the normal law remains engaged. If a second one fails and is undoubtedly isolated, the control law will automatically switch to alternate mode, in which enhanced flight envelope protection is replaced by a conventional warning. If the three sources have failed, the control law will revert to direct law without auto-trim. An appropriate warning and resolution procedure are

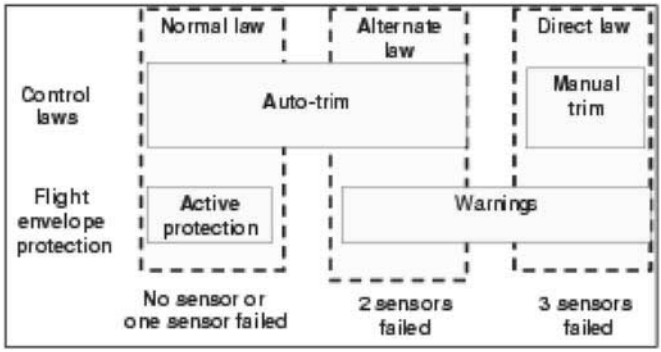


FIGURE 11.51 Control law reconfiguration (example of IRS failure).

simultaneously provided to the crew. Hence, the best possible handling qualities and safety level are provided, depending on the available resources.

In addition, FBW may also alleviate structural loads and consequently reduce the aircraft structural weight. This is obtained both by the basic control law and by specific load-alleviation functions:

- The basic control laws are designed with the target of guaranteeing that the pilot is able to command the maximum full-stick order without overstressing the structure. It is achieved on the pitch axis by limiting the load factor demand to 2.5 g and by avoiding exceeding the speed of the above design case (VMO/MMO protection)
- Complementary control laws may also be implemented in order to alleviate loads in specific cases. Foe example, on the A330/A340, the aileron and outboard spoilers are symmetrically deflected upwards when the normal acceleration command exceeds 1.3 g.

Flight Control Computers

All the flight control functions are software based (control laws, actuator control and monitoring, failure detection, reconfiguration logic, etc.). This gives the designer greater flexibility to adapt them to aircraft characteristics and fulfill new or specific requirements.

Flight control functions are implemented in a set of redundant, dissimilar, and fail-safe digital computers. Each FCC computer is capable of detecting and isolating any failure affecting the function it performs and the actuators it controls thanks to its dual type architecture. Each computer is composed of two fully separate command and monitoring units (Figure 11.52) without hard-wired synchronization. Each one owns all the hardware resources it needs to perform all its software-based functions (analog, discrete and digital input and output circuits, processors, electrical supplies) and to perform its self-monitoring (check-sum, watchdog, on-ground automatic self-test on electrical power-up).

The FCC automatically reconfigures the FBW system in the case of simple or multiple failures that affect themselves, the signals they receive, or the actuators

AVIONICS AND ASTRIONICS

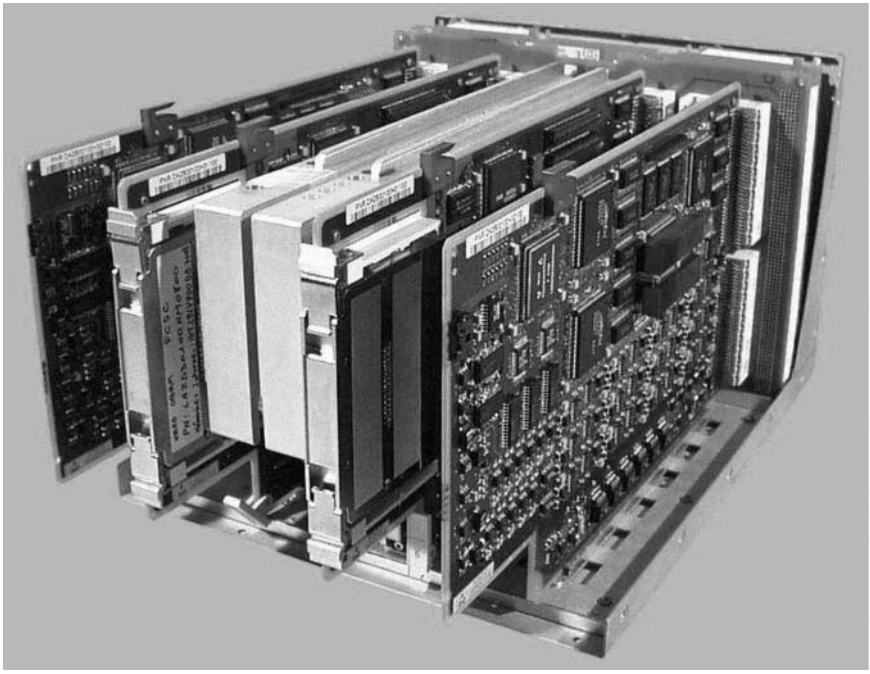


FIGURE 11.52 Airbus Flight control computer (A340-600 FCSC).

they control. The reconfiguration logic is distributed between all flight control computers and follows a predefined sequence. The principle of this logic is as follows:

- In normal operation, one predefined computer is in charge of computing the control order. Digital buses to the other flight control computers dispatch this control order, and all servo-controls are slaved to the master control orders.
- If the master computer detects an internal failure, it will fall into a fail passive state by switching off all its output. Immediately, the second computer, following the predefined logic, will take over in order to compute the control order in place of the previous one.

One computer alone is sufficient to control the aircraft, but with reduced maneuverability.

For actuator control, the same principle is applied, but servo-control failure will not lead to the total loss of the computer that controls it. The actuator reconfiguration logic is also fully automatic and adapted to each type of surface. Two dissimilar types of flight control computers are used in order to prevent a single common hardware or software defect from leading to total FBW loss.

All flight control computer function, control laws, input and output monitoring, actuator control and monitoring, and reconfiguration logic are specified using a graphic language based on a well-defined library of symbols. Functional specifications are consequently accurate, complete, and unambiguous and may be auto-

matically coded, first in order to validate them on simulation tools, then to implement them with a qualified coding tool in the flight control computers.

Flight Control Surface Actuation

On the A320 and A330/A340 family, hydraulically powered actuators control all surfaces. With the A380, a new type of electrically powered actuator will appear.

Each surface is controlled by one, two, or three actuators depending on the consequence of total surface actuation loss, taking into account that it may be combined with other failures.

Two servo-controls are installed on each aileron and elevator. In normal operation, one servo-control is active and the other one on standby. This last one will automatically become active if the first one fails. Mode change is obtained by positioning a valve controlled with a solenoid. In standby mode, both ram chambers are put into communication by the mode valve via a restrictor, and the servo-control becomes a damper. The damping coefficient is defined to avoid a detrimental surface flutter if only one servo-control remains connected to the surface in damping mode.

Each surface actuator receives its order from the flight control computer by a dedicated electrical channel. Thus, a failure of one servo-control or its control channel will not be propagated on other servo-controls of the same axis, as could be the case with mechanically signalled flight controls.

The rudder surface is powered by three synchronized servo-controls. These are mechanically signalled on the A320 and A330/A340 families and electrically signalled on A340 500/600 and later Airbus aircraft.

Each spoiler surface is controlled by a single electrically signalled servo-control. In case of control channel electrical failure, the hydraulic jack is commanded to the retracted position and the spoiler surface fully applied on the wing.

The THS is controlled by a fail-safe screw jack powered by two hydraulic servomotors. These are electrically signalled in normal operation. If all electrical channels have failed, the THSA is mechanically signalled from hand-wheels located on both sides of the center pedestal in the cockpit, one for the captain, the other for the first officer. After touchdown, the THSA is automatically positioned at a neutral position acceptable for a further takeoff. Before takeoff, the THS will be mechanically positioned at its optimum position, depending on the aircraft's center of gravity location.

System Installation and Protection

Physical segregation between the various channels of a redundant system is the basic precaution to reduce its vulnerability to a common external aggression. It is obtained on an FBW system by sending signals in separate and distant electrical routes, by installing computers in different areas in such a manner that the aircraft remains controllable in case of an external failure such as an engine burst or a local fire.

Lightning strikes and electromagnetic aggression are a concern on any aircraft. For aircraft certification, it shall be demonstrated by test that FBW may suffer the very high level required for certification with margin, not only without any damage but also without any disconnection or upset. In-service experience confirms these results.

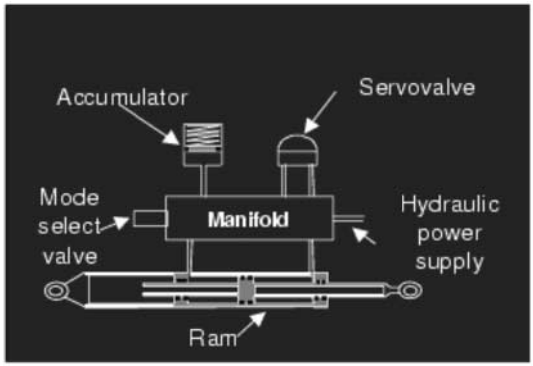


FIGURE 11.53 Airbus servo-control principle.

Maintenance Assistance

The failure isolation logic is performed by two dedicated flight control data concentrator (FCDC) units that permanently analyze all the monitoring maintenance messages that they receive from all the FCC and deduce from this analysis which line replaceable unit (LRU) has most probably failed. Maintenance assistance is provided by automatically reporting any failure to a centralized maintenance system. This system gives information on the failure environment (flight phase, time, etc.).

The FCC permanently monitors the whole FBW system, including standby functions. Consequently, the number of potential dormant failures is drastically reduced and scheduled maintenance task intervals significantly increased in comparison with mechanically signalled systems. For example, the ground spoiler function needs a weekly check on mechanically signalled A300 B2/B4 and is completely eliminated on the A320 and A330/A340. All safety checks below check A (600H on A320) have been eliminated on Airbus FBW aircraft.

In addition, all flight control equipment is specified and designed without TBO (time before overhaul).

Airbus Fly-by-Wire Advances since the A320's Entry into Service

Airbus FBW evolves to cope with new aircraft characteristics. New and mature technologies are progressively implemented on Airbus aircraft in order to reduce weight, production cost, and maintenance cost and improve passenger comfort and safety. All this is managed with the goal of maintaining the basic Airbus principle.

A330/A340. Some significant architecture changes have been introduced.

- Yaw damper availability has been increased in order to damp Dutch roll in any situation, including ultimate mechanical control. The vertical fin is accordingly sized to reduce its area and weight and reduce aerodynamic drag. For that reason, an alternate yaw damper has been added on the rudder axis that will automatically operate in case of a total loss of the nominal yaw damping function. The backup

yaw damper unit (BYDU) is electrically signalled by electronic modules fully independent from the FBW computer package, including electrical supplies.

- The two A320 flight augmentation computers have been eliminated and their flight control functions incorporated into FBW computers (control and monitoring of all yaw axis servo-motors: yaw damper, rudder trim, and rudder travel limitation actuators).
- A specific control law has been added to improve comfort in turbulence by damping low-frequency vertical and lateral fuselage structural modes. Structural modes and handling modes are sufficiently distant that they can be decoupled by conventional filtering techniques.

A340-500/600. These versions are derived from the current A340, but with a longer fuselage and larger wing area in order to extend aircraft capacity and range. The first structural mode frequencies have consequently been decreased, both on longitudinal and lateral axes, and are close to HQ modes. The following significant advances of the FBW system have been developed in order to maintain good HQ and passenger comfort.

- Control laws are designed to control HQ modes and low-frequency structural mode simultaneously.
- On the rudder axis, FBW orders are sent directly to electrically signalled servo-controls. The mechanical rudder control is fully eliminated and an ultimate electrical backup is added. This backup is highly segregated from the FBW computers and aircraft electrical network.
- The inner autopilot loop is located in the FBW computers to reduce transmission delay from the AP order to the servo-control.

Perspectives for the A380

The very large A380 is presently under development. Several challenging innovations have been introduced on the A380 flight control system and presented below. They should also be applied to the A400-M transport aircraft.

Electrically Powered Actuators. Safety objectives on a modern transport aircraft require three independent power sources for primary flight control surface actuation. Up to now, hydraulics has been the only practicable solution to supply servo-controls, taking into account all the aspects: weight, performance, and cost. In other respects, the total loss of electrical sources is safety-critical on FBW aircraft obviously, but also on other transport aircraft due to the growing place of avionics. On Airbus FBW aircraft, total electrical failure is prevented by the multiplicity and diversity of supplies (engine-driven generators in normal operation, hydraulically driven generator and batteries in emergency situations) and by stringent segregation rules applied in the electrical network. Consequently, hydraulics may be at least partially eliminated, provided that electrical actuators are used in addition to hydraulic servo-controls for surface actuation. Ten years ago, a new and attractive technology of improved weight-to-power ratio, brushless electrical motors, appeared on the market. Airbus launched a trade study to evaluate their impact on FBW architecture and to select the most appropriate type of flight control actuators.

Benefits of power by wire are:

- Aircraft weight, reliability, production cost and maintenance cost are reduced by eliminating the hydraulic circuit and all of its components (pumps, reservoirs, filters, pipes, selectors, etc.). This gain has been confirmed whatever the aircraft size, from a 100-seater to the largest A380.
- Aircraft safety level is improved at different levels. First, slight damage on a hydraulic circuit may lead to leakage and to its complete loss, even if hydraulic fuses are installed. With an electrical circuit, only downstream equipment is lost in the case of a circuit rupture. An electrical circuit may also be installed in an area not suitable for an hydraulic circuit and more segregated from the remaining hydraulic circuits than the hydraulic circuit it replaces. The aircraft remains controllable if a common failure mode affects all the hydraulic circuits.

New types of electrical actuators will be introduced for Airbus FBW:

- Three types of actuators have been envisaged: IAP (integrated actuator package), EMA (electromechanical actuator), EHA (electo-hydrostatic actuator). For the first application, trade studies lead to selecting EHA for the following reason: EHA and conventional servo-control jacks are identical or very similar. Therefore, the experience gained with conventional servo-controls on important safety items is directly applicable to EHA: damping function integrity, lack of jamming, fatigue justification. This is not the case with EMA. Furthermore, EMA airborne experience is limited to low-speed and/or low-power actuators.
- Several EHA demonstrators have been developed by several suppliers and successfully tested in the laboratory and in flight on A320 and A340. This new technology is mature enough for new programs.

Permanent Active Stabilization. The first step has been implemented on the A330/A340 rudder control. The next will be the permanent active stabilization on the 380 elevator control. Tail horizontal stabilizer is accordingly downsized.

Full Electrical Ultimate Backup. THSA control is the last remaining mechanical control on the A340 500/600. It has been eliminated on the A380 and replaced by an electrical control. Consequently, ultimate backup becomes fully electrically signalled and will become automatically active if all flight control computers have failed, as on previous Airbus aircraft. It will be similar to the A340-500/600 rudder ultimate backup and employed on the elevator, rudder, and ailerons.

Hydraulic Pressure Increase. On all present transport aircraft, hydraulic circuits are pressurized at 3,000 psi (210 Bars). This pressure will be increased to 5,000 psi (350 Bars) on A380. An important weight savings will be obtained.

Autopilot Integration. The autopilot function will be integrated into the primary flight control computers. The autopilot function is being installed in 3 PRIM computers in place of two, as in the previous A320 and A330/A340 programs, and its availability will consequently be increased. In addition, the total electronics volume will be reduced, autopilot and flight control sharing the same basic resources.

AFDX Bus. The new AFDX multiplexed bus has been introduced in the A380 avionics. It will be used by FCC to communicate with other systems. Flight control critical information will continue to be transmitted with ARINC 429 networks (in-

ertial and anemometric parameters, information exchanged between fly-by-wire computers, etc.)

Conclusion on AIRBUS Fly-by-Wire

Airbus has set a new standard by introducing FBW on the A320, the target being not only to reduce weight and cost, but also to improve handling qualities and safety, improve aircraft survivability in extreme situations, reduce pilot workload, and facilitate cross-crew qualification and to reduce maintenance tasks.

Nevertheless, new and mature technologies conttinue to be progressively implemented in new versions to improve performance and safety while minimizing costs and maintaining the basic Airbus systems philosophy.

11.34 SOME CONTROL CHALLENGES

For illustration, we introduce three kinds of challenges in this subsection:

1. Design of FCL for aircraft having complex dynamics
2. Integration of advanced active control technologies (ACT)
3. Flight control of nonconventional configurations

FCS design engineering will also have to face many more challenges not detailed here.

Flexible Aircraft Control

Flexible aircraft control is a challenge for both military and civil aircraft. Trends toward design lighter military aircraft with new materials and larger transport aircraft lead to a more flexible structure. Difficulties in FCS design come from:

- The significant reduction of the distance between the conventional rigid aircraft flight control bandwidth and the flexible modes (sometimes overlapping)
- The high order of the total aircraft dynamics (typically 50 to 100 states), which now include structural modes coming from finite element models (FEM)
- The need to improve the accuracy of aerodynamics prediction
- The necessity to use sensors to extract information on the structural dynamics.

Two strategies are possible for FCL design (RTO 2000). The first strategy is to try to recover (on the flexible aircraft) the rigid control performance that could be designed on the rigid aircraft by deleting the flexible contribution in the measurements used in the feedback loop. The most frequently employed technique is to introduce notch filters having zeros at flexible mode frequencies. This method, however, leads to softening of the control and degradation of the rigid performance. An alternative is to use geometrical filtering, i.e., a combination of measurements for minimizing flexible residuals in the feedback. Robustness may be poor against sensor failure, however, and it is not efficient if rigid and flexible bandwidths over-

lap. A more ambitious strategy is to perform simultaneous control of the rigid aircraft and of its structural dynamics. This is a fully multivariable, multiobjective control problem (Alazard et al. 1999) which can be solved after being formalized under the standard form of Figure 11.43 if models of the flexible aircraft are available with reasonable accuracy, if actuators have a high-frequency bandwidth compatible with structural modes to be controlled, and if sensors showing the contribution of the structural dynamics are available at different locations on the aircraft fuselage and wing.

Active Control Technologies

Active control technologies (ACT) generally mean nonconventional control for enhancement of military aircraft performance and maneuverability (RTO 2001). It mainly includes the popular active flow control (Gad-el-Hak 2000). However, other technologies could be considered, such as noise control, compressor instability control, (Gravdahl and Egeland 1999), and combustion control, since they will all involve the same kinds of methodological issues. Most are mainly considered as isolated monodisciplinary technical challenges. However, ACT design is clearly multidisciplinary and could lead aeronautical engineering much beyond performance enhancement on given aircraft configurations. It may reopen the whole process of aircraft design. For instance, active aeroelastic wing technology opens a new design approach for aircraft wing structure (Flick et al. 1999). Integration of advanced fluidic actuation technologies may revolutionize FCS architectures (Moir and Seabridge 2001; Pratt 2001). Consequently, two main problems must be solved in terms of control design for ACT: developing control strategies for the technology itself and integrating it in aircraft FCS.

Design can be considered conventional if the two problems can be solved separately: technology is controlled with a local feedback loop and optimized separately before being integrated in the aircraft as a specific locally controlled, stability-augmented, stand-alone subsystem. Design becomes nonconventional either when the technology has a nonconventional dynamic behavior or when strong couplings appear between the dynamics involved in the technology and the aircraft's own dynamics, particularly when some attempts are undertaken to make actuators play a more ambitious role than just controlling the rigid-body flight. The two problems may need to be solved at once. Depending on complexity, it may still be possible in practice to distinguish two steps. The first step remains local control of the nonconventional technology, considering aircraft dynamics as an external input. This is a very difficult control design problem for two reasons. The first is that in most ACT, subsystems to be locally controlled have a distributed nature (vortices or shocks or separations in active flow control, wing distortion in structural control, pressure distribution in compressor instability control, etc.). Their dynamics are governed by partial differential equations (PDE), for which standard control design tools do not apply. *Local* and *distributed* refer to the aircraft scale and to the dynamics of the subsystem to be controlled, respectively. They should not be confused. The second reason is that such local control is very demanding since it should lead to a closed loop controlled subsystem having predictable and uniform dynamic behavior for any input from aircraft dynamics. As an extension of conventional design, the nonconventional, locally controlled subsystem could then be used for aircraft flight control, with the additional difficulty of dealing with couplings in a fully multivariable control design problem.

Flight Control of Nonconventional Aircraft Configurations

Some assumptions made for modeling conventional aircraft and allowing simplified approaches for control surface sizing, HQ assessment, and FCL design are not valid for nonconventional configurations such as delta wing, tailless, three-surface, flying wing, STOL, etc., for different reasons, such as:

- Couplings between the longitudinal and lateral dynamics may be significant because of different mass and inertia distributions and different aerodynamics (more compact aircraft, for instance).
- In the absence of a horizontal tail, the wing actuators have to perform adverse horizontal trimming and high-lift generation simultaneously, depending on CG location. Trimming also affects wing performance.
- Controlling lateral motion without a tail requires alternative technologies (thrust vectoring, for instance, or nonconventional actuators such as crocodile flaps).

Control surfaces sizing, HQ assessment, and FCL design for such nonconventional aircraft require the invention of new FCS design processes, including specific modeling, nonconventional requirements, and new types of compromises.

11.35 CONCLUSION

For the first aircraft generations, technological and methodological issues of FCS design were considered independently (McRuer and Graham 1981). Through development of modern military and civil aircraft programs, technology has been pushed by needs in performance and reliability. Airbus fly-by-wire technology is a key illustration. Sophisticated methods have been invented simultaneously which now are part of global industrial processes. By supporting the understanding of the requirements, the specification, development, and integration of efficient and validated solutions, these tools allow the design of highly integrated FCS with high maturity, integrity, and reliability. The trend is to integrate FCS design into a global and optimized systems design process (integrated modular avionics technology, more electrical aircraft and power optimized aircraft concepts), then to reach final integration into the overall aircraft design.

REFERENCES

Ackermann, J. 1993. *Robust Control: Systems with Uncertain Physical Parameters,* Springer Verlag, Berlin.

Alazard, D., Buckarles, A., Ferreres, G. Magni, J. F., and Prudhomme, S. 1999. "Towards a Global Methodology for Flexible Aircraft Control," in *Proceedings of RTO Structural Aspects of Flexible Aircraft Control,* Ottawa, pp. 18–21.

Babister, A. W. 1980. *Aircraft Dynamic Stability and Response,* Pergamon Press, Oxford.

Blakelock, J. H. 1991. *Automatic Control of Aircraft and Missiles,* John Wiley & Sons, New York.

Boiffier, J. L. 1998. *Dynamics of Flight: The Equations,* John Wiley & Sons, New York.

Bourcier, M. A. 2001. *Logistics Test and Evaluation in Flight Test,* NATO-RTO-SCI Agardograph 10.

Collinson, R. P. G. 1996. *Introduction to Avionics,* Chapman & Hall, London.

Cook, M. V. 1997. *Flight Dynamics Principles,* Arnold, London, and John Wiley & Sons, New York.

Etkin, B. and Reid, L. D. 1996. *Dynamics of Flight,* John Wiley & Sons, New York.

Fargeon, C., ed. 1989. *The Digital Control of Systems,* North Oxford Academic, Oxford.

Farineau, J. 1989. "Lateral Electric Flight Control Laws of a Civil Aircraft Based upon Eigenstructure Assignment Techniques," in *Proceedings of AIAA GNC Conference,* Boston, August.

Federal Aviation Administration (FAA). Federal Aviation Regulation 25 (FAR 25), *Airworthiness Standards: Transport Category Airplanes*.

Ferreres, G. 1999. *A Practical Approach to Robustness Analysis with Aeronautical Applications,* Kluwer Academic, Dordrecht.

Flick, P. M., Love, M. H., and Zink, P. S. 1999. "The Impact of Active Aerolastic Wing Technology on Conceptual Aircraft Design," in *Proceedings of RTO Structural Aspects of Flexible Aircraft Control,* Ottawa.

Gad-el-Hak, M. 2000. *Flow Control: Passive, Active and Reactive Flow Management,* Cambridge University Press, Cambridge.

Gravdahl, J. T. and Egeland, O. 1999. *Compressor Surge and Rotating Stall Modelling and Control,* Advances in Industrial Control, London, Springer.

Hodgkinson, J. 1999. *Aircraft Handling Qualities,* AIAA Educational Series, American Institute of Aeronautics and Astronautics, Reston, VA.

Joint Aviation Authorities. JAR-25, *Joint Aviation Requirements for Large Aeroplanes*.

Liu, G. P. and Patton, R. J. 1998. *Eigenstructure Assignment for Control System Design,* John Wiley & Sons, New York.

Ljung, L. 1987. *System Identification: Theory for the User,* Prentice Hall PTR, Upple Saddle River, NJ.

Maciejowski, J. M. 1989. *Multivariable Feedback Design,* Electronic Systems Engineering Series, Addison-Wesley, Reading, MA.

Magni, J.-F., Bennani, S., and Terlouw, J. 1997. *Robust Flight Control: A Design Challenge,* Lecture Notes in Control and Information Sciences 224, Springer, London.

McCormick, B. W. 1995. *Aerodynamics, Aeronautics, and Flight Mechanics,* John Wiley & Sons, New York.

McLean, D. 1990. *Automatic Flight Control Systems,* Prentice Hall, Englewood Cliffs, NJ.

McRuer, D. and Graham, D. 1981. "Eighty Years of Flight Control: Triumphs and Pitfalls of the System Approach," *Journal of Guidance and Control,* vol. 4, pp. 353–362.

MIL-1797-A, *Flying Qualities of Piloted Vehicles*.

Minister of Defence (MoD). 1983. Design and Airworthiness Requirements for Military Aircraft, UK MoD Definition Standards 970.

Moir, I. and Seabridge, A. 2001. *Aircraft Systems: Mechanical, Electrical, and Avionics Subsystems Integration,* American Institute of Aeronautics and Astronautics, Reston, VA

Moorhouse, D. J., ed. 2000. *Flight Control Design: Best Practices,* NATO-RTO-SCI Technical Report 29.

Nelson, R. C. 1998. *Flight Stability and Automatic Control,* McGraw-Hill, New York.

Normand-Cyrot, D., ed. 1998. *Perspectives in Control,* Springer, London.

Pratt, R. W., ed. 2000. *Flight Control Systems,* Progress in Astronautics and Aeronautics 184, American Institute of Aeronautics and Astronautics, Reston, Va, and Institution of Electrical Engineers, Stevenage, Herts, UK.

RTO-MP36. 2000. *Proceedings of Structural Aspects of Flexible Aircraft Control*.

Prudhomme, S. and Gimonet, B. 2001. "Flight Control Performance Enhancement Avoiding Aircraft Identification," paper delivered at AIAA-GNC Conference.

Slottine, J. J. E. and Li, J. W. 1991. *Applied Nonlinear Control,* Prentice Hall, Englewood Cliffs, NJ.

Söderstrom, T. and Stoica, P. 1989. *System Identification,* Prentice Hall, New York.

Spitzer, C. R. 1987. *Digital Avionics Systems: Principles and Practice,* McGraw-Hill, New York.

Stevens, B. L. and Lewis, F. L. 1992. *Aircraft Control and Simulation,* John Wiley & Sons, New York.

Tischler, M. B., ed. 1996. *Advances in Aircraft Flight Control,* Taylor & Francis, London.

PART 9

SPACE BORNE INSTRUMENTS

Jean-Paul Agultes

11.36 SPACE BORNE SYNTHETIC APERTURE RADAR (SAR)

Introduction

Synthetic aperture radar is a category of imaging side looking radar (SLR) for which the resolution along the speed axis is determined by using the Doppler shift instead of the antenna beam footprint, which cannot be narrow enough in the spaceborne case because of the huge distance to ground. Indeed, standard SLR (non-SAR) is never used for imaging from space. Principles of SLR and SAR are defined in Part 5. This subsection focuses on the design specificity of a spaceborne SAR.

The progression of the original idea of SAR (Wiley 1951) from airborne to space was slow, since the adaptation of technology to the space environment was far from simple. Indeed, the introduction of satellite altitude and speed in the SAR equations as well as the operational requirements conventionally attached to space remote sensing greatly increase the SAR design constraints.

- As shown below, the maximum available swath over resolution ratio, called merit factor, is drastically reduced (by a factor of at least 25) because the ambiguity and timing constraints are tightened. The minimum required antenna area is boosted (by a factor of at least 1,250). It ranges from 5 m^2 in the X band to 50 m^2 in the L band and drives the overall vehicle dimension, which is far from the case in airborne. Moreover, in the spaceborne case the swath (or the resolution) depends strictly on the antenna aspect ratio (length over width)
- For spaceborne, unlike airborne, the swath is much smaller than the ground field of view and contains a very limited incidence range. Electronic elevation beam steering was required early on in order to obtain incidence variability and reduce the time delay to get an image of a given ground target.
- Besides the conventional strip SAR mode, new imaging modes have been necessary to implement in order to circumvent the strict dependency on the antenna aspect ratio and give some image versatility. The scansar mode and spotlight modes allow the swath (or the coverage) to be instantaneously traded against resolution.
- Active antenna technology that is not necessary for airborne SAR (at least civilian SAR) is favored in space for many reasons: beam steering agility, high RF peak power required for a long-distance trip, high reliability (inherent redundancy of transmit and receive electronics).

Elachi (1988) describes design and applications of a spaceborne SAR instrument. Aguttes (2001) shows that beside inherent constraints in comparison to airborne systems the space context can enable significant simplifications provided that both SAR and space environment constraints are considered together from scratch for the design of the whole satellite leading to integration of the instrument and the carrier.

SAR geometry and Imaging Modes

The velocity vector V and the line of sight T, defined by an incidence i, are supposed to be orthogonal. The antenna size that counts is the effective one obtained by projection of the actual one along the line of sight into the plane containing V and orthogonal to T. In this plane, the dimension along V is the effective antenna length L and the dimension normal to V is the effective antenna height H. Two orthogonal axes define the image frame, the range axis (also called elevation axis) in a plane normal to V, and the Doppler axis (also called azimuth axis) along V (see Figure 11.54).

Strip image mode is the standard imaging mode that drives the sizing of the antenna according to the relationships that will be introduced next.

Scansar mode enables swath widening by the use of electronic beam steering in elevation. The swath enlarged by a factor N is formed by N sub-swaths obtained by N sequential positions of the beam in elevation. The illumination time of a given target, delimited by the fly-over of the azimuth beam footprint in strip mode, is here reduced by N because of the scansar time multiplexing process. Consequently, the Doppler excursion and the SAR resolution are degraded by the same factor (more precisely, by an $N + 1$ factor).

Spotlight mode provides the opposite effect to scansar, that is to say an improvement of the resolution at price of reduction of coverage. Backward scanning of the beam enables an increase of illumination time and of the Doppler excursion by a factor N and thereby an improvement of the resolution by a factor $1/N$. Spothlight mode cannot be continuous and introduces coverage gaps along the orbit (see Figure 11.55).

Maximum Merit Factor and Minimum Antenna Area

***Merit Factor* (*Swath/Azimuth Resolution*).** Merit factor and minimum antenna area are driven by the ambiguity (see Part 5 for introduction of radar ambiguity). Figure 11.56 gives the geometry on a range/elevation axis. It represents the positions on the ground of the swath, classically delimited by the 1 dB beam aperture and the first outer (the one with a positive range deviation of $C/2$ PRF) ambiguity of the inner swath edge (the edge at nearer range). The geometry and radar timing in the spaceborne case ensure that this first ambiguity drives the design. Indeed, it

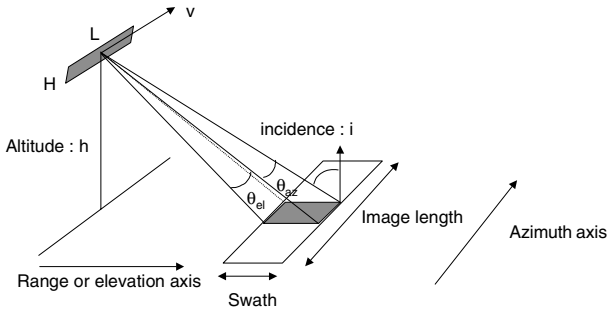


FIGURE 11.54 SAR geometry.

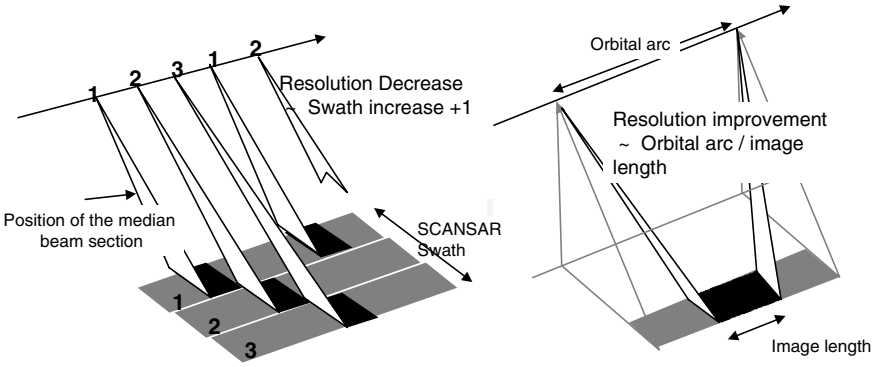


FIGURE 11.55 Scansar and spotlight modes.

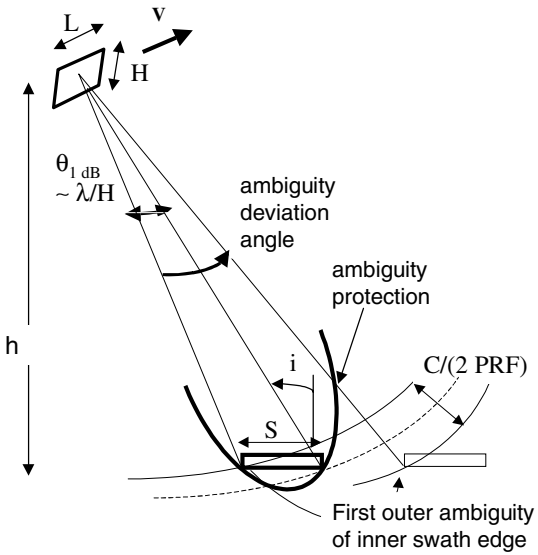


FIGURE 11.56 Range axis geometry.

is generally located on the edge of the main beam and is not deviated enough in range to be on the sidelobes. Even the cumulative contribution of all the other ambiguities of greater rank (range deviation of value $+/-\ k(C/2$ PRF) which fall on the sidelobes is smaller than the first ambiguity. The ambiguity protection is the worst on the two swath edges. For any other target in the swath the protection given by the main beam fall-off is greater. The other configuration equivalent to those given in Figure 11.56 occurs with the outer swath edge and the first inner ambiguity. For a typical ambiguity protection in the range -22dB, -25 dB, the protection given on first ambiguity must be in the range -24 dB, -27 dB and corresponds to a main beam fall-off in the range -12 dB, -13.5 dB. Because of the pattern of

a fine (SINC shape) beam, such fall-off is achieved when the ambiguity deviation angle is about twice the swath aperture.

We have the following relations:

$$\text{Ambiguity deviation angle} > 2\ \theta_{1\text{dB}} \tag{11.74}$$

$$C/(2\ \text{PRF}) > 2S\sin(i)$$

$$S < C/(4\ \text{PRF}\sin(i)) \tag{11.75}$$

where C = velocity of light, S = Swath

On the Doppler axis we have to consider the Shannon criteria (to sample the phase history correctly) and to avoid aliasing by Doppler ambiguous targets (targets having a history spectrum shifted by k PRF).

$$\text{PRF} > 1.2B_{\text{dop}} \tag{11.76}$$

where B_{dop} = Doppler bandwith classically delimited by the 3 dB azimuth beam aperture and 1.2 is a typical margin factor on account of the unperfected protection against ambiguous targets given by azimuth beam shape.

$$B_{\text{dop}} = 2(V/\lambda)\theta^{3\text{dB}} \tag{11.77}$$

$$\theta^{3\text{dB}} = \lambda/L \text{ (for a fully efficient antenna area, no directivity loss)} \tag{11.78}$$

The Doppler effect is equivalent to a chirp modulation of the successive radar echoes that gives a timing resolution $1/B_{\text{dop}}$. The azimuth resolution is the product of the speed by this timing resolution (see Part 5).

$$r_{\text{az}} = V/B_{\text{dop}} = L/2 \tag{11.79}$$

The combination of these constraints (relations (11.75), (11.76), and (11.79)) leads to an inherent limitation of the merit factor:

$$M_f < C/(4.8V\sin(i)) \tag{11.80}$$

We see that this limitation is much tougher for space than for airborne since V is typically 25 times greater. For modern SAR able to provide high incidence (greater than 50°), M_f is about 10,000.

Minimum Antenna Area (for an Antenna without Directivity Loss). From Figure 11.56, we get the following relations:

$$S = \theta_{1\text{dB}}R/\cos(i) < M_f r_{az} \tag{11.81}$$

$$\theta_{1\text{dB}} = (\lambda/2H) \text{ (for a fully efficient antenna area, no directivity loss)} \tag{11.82}$$

From (11.79)–(11.82), we derive:

$$A = L\,H > 4.8\lambda R(V/C)\tan(i) \tag{11.83}$$

The minimum antenna area constraint imposed by the ambiguity is much tougher in space than in airborne systems since R can be at least 20 times greater and V/C 25 times greater. In space, ambiguity is the key sizing factor of the antenna. We note also the great influence of the maximum operating incidence with a factor 3 on A between 40° and 60°. Since, on the other hand, the maximum incidence and

the altitude drive the ground field of view and therefore the revisit performance, we see the significance of these two geometric parameters in the sizing of a spaceborne SAR mission.

Beam Shaping: Interest and Limits. All the previous relations concern the general case where the swath is covered by an unshaped beam (no illumination weighting, no directivity loss). Beam shaping enables steeper beam fall-off, reduction of the ratio of ambiguity deviation angle over swath aperture below the value of 2, and therefore greater merit factor. However this performance improvement is very costly in term of antenna area. Indeed, an increase of 50% on merit factor requires a beam fall twice steeper and a double antenna height (see Figure 11.56).

Mean Transmit Power

The received power P_r from a target of area dS is given by:

$$P_r = P_t(4\pi A/\lambda^2)(1/(4\pi R^2))dS(1/(4\pi R^2)A/L$$

where P_t = mean transmit power
$(4\pi A/\lambda^2)$ = transmit gain
$(1/(4\pi R^2))$ = propagation loss
L = total loss (including the deviations between the physical antenna area A and the efficient one after beam shaping, widening, or steering)

The signal-to-noise ratio after range and azimuth matched filters is given by:

$$S/N = P_r T_{ill}/N_o$$

where T_{ill} = illumination time = $R\theta^{3\text{dB}}/V = R\lambda/(2Vr_{az})$
N_o = noise density = FKT_o = noise factor × Boltzmann constant × 300 K

Then the minimum required power for $(S/N > 1)$ is for a point target of cross-section dS:

$$P = 8\pi LFKT_o VR^3 \lambda r_{az}/dSA^2 \tag{11.84}$$

and for a scattered target with backscatter coefficient σ_0,

$$P = 8\pi LFKT_o VR^3\ \lambda/\sigma_0 r_{rd} A^2 \tag{11.85}$$

When A is the minimum antenna area, we have:

$$P = 8\pi LFKT_o RC^2/23\sigma_0 r_{rd}\lambda V \tan(i)^2 \tag{11.86}$$

For LF ranges between 10 and 15 dB, a typical breakdown is as follows (typical values in X band):

- Noise factor F (1.2 dB)
- Atmospheric loss, e.g., rain (typically 1 dB for 2 paths)
- RF: antenna radiator losses, path between radiators and amplifiers (typically 5 dB for two paths)
- Processing loss: nonperfect matched filter (typically 0.5 dB)
- Sampling and coding losses (typically 0.5 dB)

- Azimuth and elevation antenna beam fall-off (typically 1.7 dB + 2 dB for two paths)
- Antenna directivity loss, not to consider in (11.86) since 11.86 is established with a fully efficient minimum antenna area
  - Unwanted directivity loss (2 × 0.2 dB for active antenna, 2 × 1.5 dB for parabolic)
  - And/or counterpart of desired beam shaping/widening, desired electronic off-steering

Radar Bandwidth and Data Rate

The radar bandwidth is directly related to the range resolution and the incidence (see Part 5):

$$B = C/(2r_{rd}\sin(i))$$

SAR data rate results from the sampling of the radar echo received with a repetition rate PRF.

The echo duration is:

$$D_{\text{echo}} = \frac{2S \cdot \sin(i)}{c} + \frac{D_c}{PRF}$$

where D_c/PRF = transmit pulse length = transmit duty cycle/PRF.

The sampling frequency must be greater than B. A typical margin factor of 1.1 can be considered. From (11.79) we have $PRF = 1.2\ V/r_{az}$.

Each data sample is coded on I and Q on N bits (typically 4 or 5 bits).

Finally, we have:

$$\text{Data rate} = 2N \cdot \left(\frac{1.3 \cdot V \cdot S}{r_{rd} \cdot r_{az}} + \frac{1.1 \cdot D_c \cdot C}{2 + r_{rd} \cdot \sin(i)}\right)$$

Note that in an ideal case where

- The range chirp compression (remove the effect of the second component of the echo duration and the 1.1 margin factor)
- The filtering of the Doppler bandwidth (from PRF to $B_{3\text{dB}}$) or the azimuth compression (remove the 1.2 margin factor)

are done on board, data rate is smaller and directly derived from the swath registration:

$$2N(V/r_{az})\ (S/r_{rd})$$

SAR Information: Amplitude, Polarization, Interferometry

Each pixel of an SAR image contains two terms, the amplitude and the phase. Most SAR applications are based only on the amplitude since the phase has two additive components with one fully random between pixels (coherent backscattering mechanism), the other one resulting from the distance to the satellite ($4\ \pi\lambda$/distance).

The phase can provide information only in differential mode for the same pixel. Phase difference between polarization channels together with ratio of amplitudes helps in discrimination of different types of backscatter. Phase difference $\Delta\varphi$ between two images with small incidence change is called SAR interferometry and is based on the difference of the pixel to satellite distance. This image of the distance difference modulo λ is called an interferogram. It provides a measurement of the relief whose accuracy is function of the baselength B_{orth} formed by the positions of the SAR antenna as indicated by Figure 11.57.

Another application of interferometry is the fine measurement of terrain displacement between two fly-overs of the SAR. In this case the relief is seen as a measurement error and its contribution to $\Delta\varphi$ can be removed if the relief is known or reduced by selection of a small B_{orth}.

One key limitation of multipasse interferometry results from the variation of the phase between the images due to non-geometrical factors (atmospheric propagation, backscatter). Single-pass interferometry based on two antennas separated by a boom on a single spacecraft or two spacecraft flying in formation solves the problem at the price of technical complexity. However, it cannot apply to the displacement measurement.

11.37 OTHER MICROWAVE INSTRUMENTS

Other Active Microwave Instruments

Radar Altimeter. The radar altimeter is used in space for establishing altitude maps with a very high accuracy, mainly over oceans (1 cm) and ice caps (10 cm). The use is more limited over land because of the relief and the steep slopes.

The altimeter uses a vertically oriented antenna and a pulse waveform. For ocean applications the range profile of the echo collected within the antenna footprint is sampled, averaged with other several echoes, and filtered (maximum likelihood estimation) around a range profile model (Brown) function of the altitude and sea-wave height. Such processing enables an accuracy of few cm of altitude, although each range gate width is about 40 cm (radar bandwidth about 300 MHz).

Correction of the ionosphere propagation delay is achieved by duplicating the measurement on two frequencies, typically 13.5 GHz and 5.3 GHz.

A spaceborne altimeter is generally coupled with an accurate orbit tracking device (DORIS or differential GPS) so as to get the true sea topography, that is, the sea-level height with respect to the earth geoid.

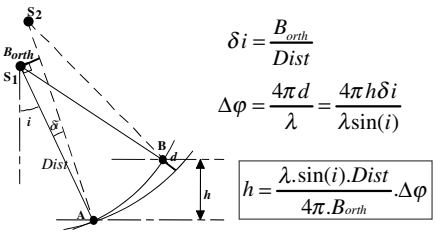


FIGURE 11.57 SAR interferometry.

The altimeter provides measurement along the satellite projected track on ground. Topography maps are obtained by selecting orbit parameters which give a grid of very closed tracks.

Scatterrometer. A scatterometer is a spaceborne radar designed specifically to measure ocean near-surface wind speed and direction. As the wind blows over the ocean, the surface is roughened by the generation of cat's paws (centimeter-scale capillary waves). These, in turn, modify the surface backscatter (reflected signal or echo) properties. Because the backscatter is function of angle between the wind speed vector and the antenna beam axis, several measurements over the same pixel with different azimuth angles give both intensity and direction of wind speed. As shown by Figure 11.58, this can be obtained with antennas, each crossing being seen by the three antennas thanks to the motion of the satellite. Unlike the altimeter and like SAR, the scatterometer provides a strip image. The pixel resolution is obtained by the range for the direction along the antenna footprint and by the beamwidth for the direction across.

Passive Microwave Instrument: Radiometer

Unlike radar, a radiometer is a passive instrument because it does not transmit a signal. It measures the natural radiation, called *brightness,* produced by the media

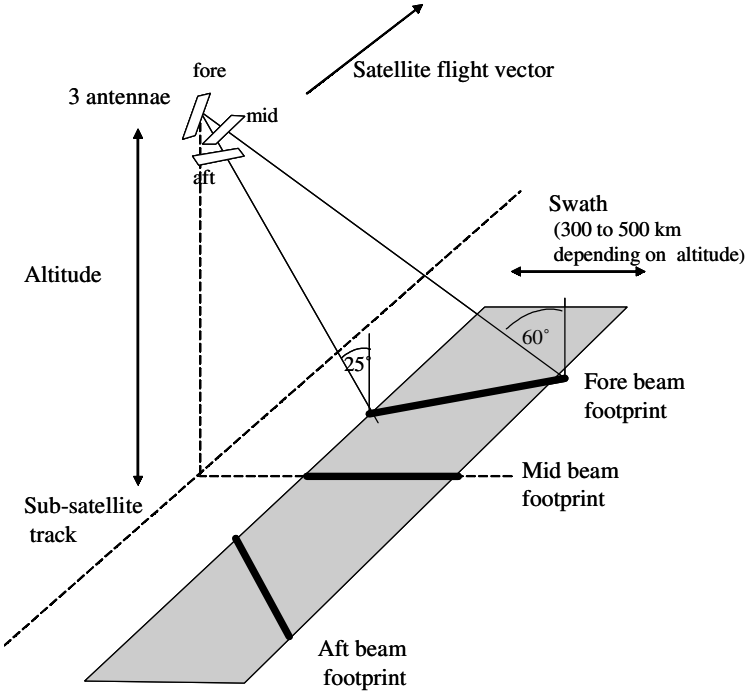


FIGURE 11.58 Scatterometer geometry.

seen by the antenna beam. The power P_A received by the antenna is expressed as follows:

$$P_A = kT_B\Delta f = \varepsilon kT\Delta f$$

where T_B = brightness temperature, the quantity observed by the radiometer
Δf = measurement bandwidth
k = Boltzmann constant
T = physical temperature of the media
ε = emissivity of the media seen by the antenna

The emissivity ($0 < \varepsilon < 1$) is a function of the properties of the media, the frequency, the polarization, and the incidence of observation.

The spatial resolution of a radiometer is given by the antenna footprint and is therefore poor as compared with SAR, even if the frequencies used can be very high (from 1 GHz to 300 GHz). Figure 11.59 shows the architecture of a radiometer.

The sensitivity of a radiometer is defined as the smallest detectable change in temperature and given by:

$$\Delta T_{\min} = \frac{T_{\text{sys}}}{\sqrt{\Delta f \cdot \tau}}$$

where T_{sys} = system temperature
Δf = measurement bandwidth
τ = integration time

The system temperature is given by:

$$T_{\text{sys}} = \frac{T_B}{L_a} + \left(1 - \frac{1}{L_a}\right) T_{pa} + (L - 1)T_{pl} + L(F - 1)T_0$$

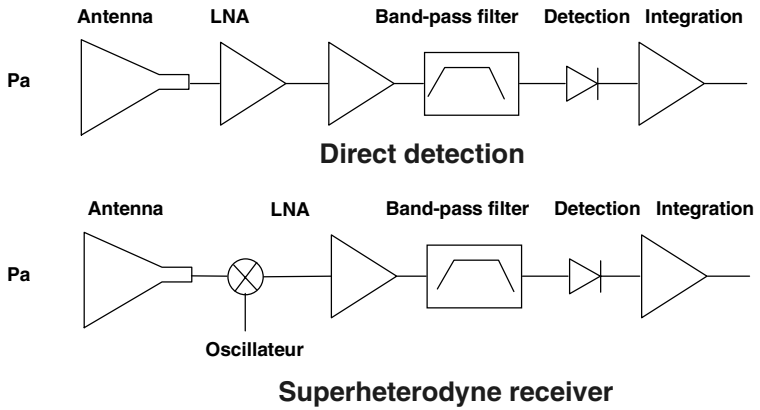


FIGURE 11.59 Radiometer principle.

where L_a = antenna loss
L = loss provided by the line between the antenna and the amplifier
T_{pa} = physical temperature of antenna
T_{pl} = physical temperature of the line between the antenna and the amplifier
F = noise factor of the amplifier
T_0 = 293 K

REFERENCES

Elachi, C. 1988. *Spaceborne Radar Remote Sensing: Applications and Techniques,* IEEE Press, New York.

Aguttes, J. P. 2001. "Radically New Design of SAR Satellite: Short Vertical Antenna Approach," *IEEE Transactions on Aerospace and Electronics Systems,* vol. 37, no. 1, pp. 50–64.

11.38 SPACECRAFT SENSORS AND INSTRUMENTATION

Introduction to Space Optical Instrumentation

Earth or planet observation, exploration of the universe (star formation, life outside the Earth), or satellite equipment: all these missions need optical instrumentation.

We can divide optical instrumentation into two main parts:

- Passive instruments measure optical radiation emitted or diffused by space objects. They represent the main part of optical instrumentation: cameras for Earth control, infrared radiometers for temperature measurement, spectroimagers for physical and chemical discrimination and remote sensing.
- Active instruments: optical radiation is emitted by instruments and interacts with observed objects, and the resulting radiation is detected by an active instrument. Lidars are used for telemetry or sounding, space lasers for optical communications between two satellites or inside one satellite, instrumentation for metrology or handling (manipulation of atoms).

These instruments can be classed by their field of application. See Table 11.12 (the given values are only an order of magnitude and are not given as accurate values).

Imaging Cameras

Image Formation. For Earth observation, these instruments can be distinguished mainly by their spatial resolution.

Implemented on low-orbiting satellites (Sun-synchronous orbits), typical spatial resolution is below 10 m, or less. Ikonos satellite spatial resolution is 1 m. For military applications, resolution is much better. Techniques for imaging use CCD arrays. Association of optical telescopes and CCD arrays (linear or bidimensional) can be chosen.

TABLE 11.12

Mission/instrument	Field of view L(km)/ resolution δL (m)	Spectral domain (μm)	Radiometry
Observation and imagery	$10 \leq L \leq 100$ $0.1 \leq \delta L \leq 100$	Diffused: $0.4 \leq \delta\lambda \leq 2.5$ Emission: $3 \leq \delta\lambda \leq 12.5$	Good SNR, needed for image treatment
Spectroimagery	$50 \leq L \leq 200$ km (low spectral resolution) $5 \leq L \leq 30$ (moderate spectral resolution)	3 to 10 spectral bands: $0.05 \leq \delta\lambda \leq 0.5$ 10 to 50 spectral bands: $0.01 \leq \delta\lambda \leq 0.1$	Accurate intercalibration (between channels) Absolute radiometrics measurement $\leq .5$ %
Radiometry: temperature measurement	$200 \leq L \leq 1500$ $100 \leq \delta L \leq 1000$ (low and geostationary orbits)	Atmospheric windows (3–5 μm and 8–12 μm) absorption bands (H_2O, CO_2, etc.)	Accurate absolute measurement
Radiometry: global radiative measurement	$500 \leq L \leq 1{,}500$ $1{,}000 \leq \delta L \leq 10{,}000$	From UV to far infrared global band and <5 spectral bands in atmospheric and absorption windows	Repetitive and accurate measurement
Spectroscopy	$500 \leq L \leq 1{,}000$ $5 \leq \delta L \leq 50$	Infrared absorption lines $0.01 \leq \delta\sigma \leq 1$ cm $-$ 1	Moisture and vertical temperature measurement
Lidars: altimetry, winds, chemistry	Frequency of laser pulses $\leq$100 Hz Vertical accuracy $\leq$100 m	UV laser (Nd:YAG and frequency $\times$ 3)	Windspeed $\leq$1ms $-$ 1 Moisture concentration $\leq$10%

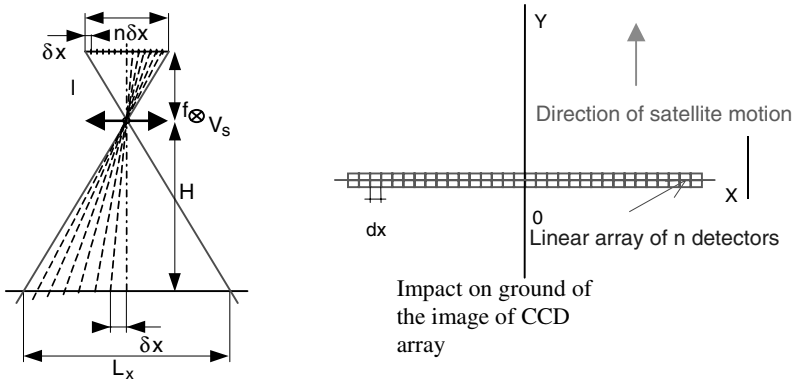


FIGURE 11.60 Principle of imaging by push broom analysis.

Image formation principle is given in Figure 11.60.

$$\text{Field of view} = L_x = n \cdot \delta x \cdot \frac{H}{f}$$

where n = number of pixels in an array
δx = pixel size
f = focal length of the instrument
H = altitude of the spacecraft

Numerical and typical application:

$$n = 24.000, \quad \delta x = 6.5\ \mu\text{m}, \quad H = 700\ \text{km}, \quad f = 6\ \text{m}, \quad \text{then } L_x = 18.2\ \text{km}$$

Sampling δL, and spatial resolution on ground δX: usually there is a confusion between these two definitions. Properly speaking, we must introduce a difference and give an accurate definition of spatial resolution:

$$\delta L = \frac{H}{f} \cdot \delta x$$

with the preceding numerical values: $\delta L = 0.76$ m.

What about spatial resolution δX? We should say that it is a specification given by the user, and we will examine the consequences of this definition. Spatial resolution is linked with the transfer function in frequency domain. We must now introduce some definitions in the frequency domain:

$$\text{Cutoff frequency of optics:} \quad \nu_c = \frac{D}{\lambda \cdot f}$$

where ν_c is expressed in cycles/mm, λ (wavelength, in μm) and D = optical aperture of optics (in mm).

$$\text{Useful band (max frequency):} \quad \nu_M = \frac{H}{2 \cdot \delta X \cdot f} \text{ (in cycles/mm)}$$

Now we can give an accurate definition of resolution. It is the value of global transfer function at the maximum frequency. We give, first, an example of numerical values (instrument as above) and $D = 500$ mm and $\lambda = 0.65$ μm:

$$\nu_c = \frac{500}{0.65 \cdot 10^{-3} \cdot 6000} = 128 \text{ cycles/mm}$$

If the specified resolution at ground is $\delta X = 1$ m:

$$\nu_M = \frac{700 \cdot 10^{+3}}{2 \cdot 1 \cdot 6000} = 58.3 \text{ cycles/mm}$$

The global transfer function takes into account other degradations of optics (aberrations, diffraction, defects: polishing, integration, etc.), detector (window effect, charge diffusion, transfer inefficiency), structure deformations (thermal effects, moisture, etc.), motion of satellite, distortion effects which affect FTM (desynchronization effects with mosaic detectors, etc.), integration, launching, and space effects. Figure 11.61 below gives a numerical example of typical degradations of FTM.

One line shows the theoretical curve (no aberrations, no other defect) and a second line is a typical curve. It is not possible to describe these effects with values. Typical values are on the second curve and can be expressed as follows: at a maximum frequency, 58.3 C/mm, or 0.7 m^{-1}, the maximum value of FTM is 0.3.

Signal-to-Noise Considerations. The second aspect of optical cameras to be considered is signal and noise. Figure 11.62 describes the main parameters used to compute signal and noise. Signal, expressed in volts or amperes or number of electrical charges (electrons), corresponds to the following equation:

$$F(W) = L_C \cdot \frac{S_{\text{pup}} \cdot S_{\text{det}}}{f^2} \cdot \tau_{\text{atm}}(\theta_S, \theta_V, \lambda) \cdot \tau_{\text{opt}}(\lambda) \cdot \cos^4 \varphi$$

$$= L_C \cdot \pi \cdot \frac{D^2}{4 \cdot f^2} S_{\text{det}} \cdot \tau_{\text{atm}}(\theta_S, \theta_V, \lambda) \cdot \tau_{\text{opt}}(\lambda) \cdot \cos^4 \varphi$$

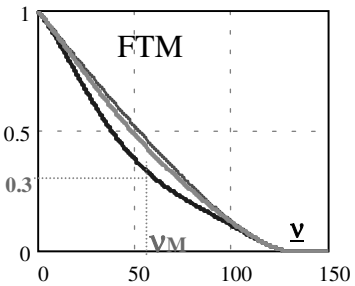


FIGURE 11.61 Optical transfer function and useful band.

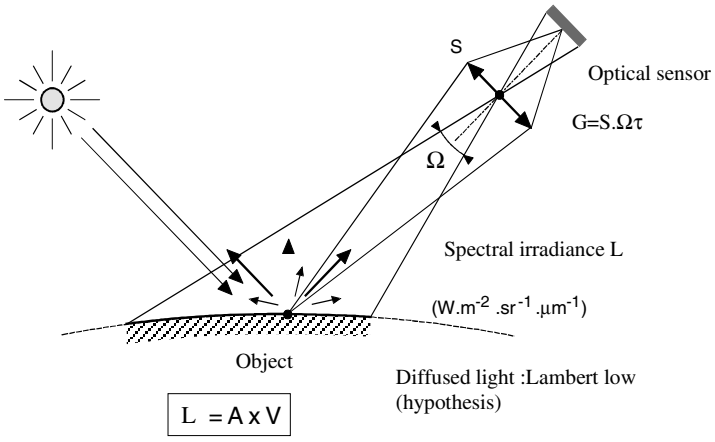
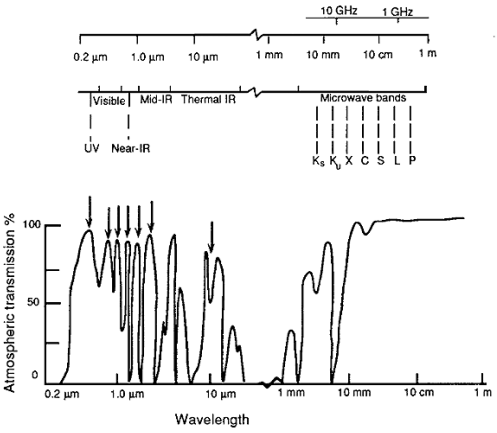


FIGURE 11.62 Sketch of light from the sun, diffused by ground and detected by instrument.

where S_{pup} = the area of entrance pupil of telescope
S_{det} = effective area of detector
τ_{opt} = the transmission of instrument
$\tau_{atm}(\theta_s,\theta_v,\lambda)$ = the transmission of atmosphere taken from Sun to ground and ground to spacecraft (see Figure 11.63).
L_c = spectral irradiance, $\delta X(\mathrm{W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}})$

Angles of incidence of the incoming rays: θ_s and angle of emerging rays outside atmosphere: θ_v are different. θ_v takes into account latitude, seasonal, and daily effects. These angles can be very large (>60°). φ is the field angle (dependence of



$$\tau(\theta_s,\theta_v,\lambda) = \exp\left(-\sigma(\lambda).h.\left(\frac{1}{\cos(\theta_s)}+\frac{1}{\cos(\theta_v)}\right)\right)$$

FIGURE 11.63 Optical windows in atmosphere and transmission.

the power 4, when Lambertian low of irradiance). $\sigma(\lambda)$ is optical thickness and function of wavelength.

The signal can then be expressed by a formula including the system parameters δX and H:

$$F(W) = L_c \cdot \pi \cdot \frac{D^2}{4 \cdot f^2} S_{\text{det}} \cdot \tau_{\text{atm}}(\theta_s,\theta_v,\lambda) \cdot \tau_{\text{opt}}(\lambda) \cdot \cos^4\varphi$$

$$= L_c \cdot \pi \cdot \frac{D^2 \cdot \delta X^2}{4 \cdot H^2} \cdot \text{atm}(\theta_s,\theta_v,\lambda) \cdot \tau_{\text{opt}}(\lambda) \cdot \cos^4\varphi$$

L_c can be expressed in terms of radiance $E(\lambda,\theta)$ at ground and spectral relectivity $\rho(\lambda,\theta)$:

$$L_c(\lambda,\theta) = \frac{E(\lambda,\theta) \cdot \cos(\theta)}{\pi} \cdot \rho(\lambda,\theta)$$

Now let us give some numerical and typical values:

$$\tau_{\text{atm}}(\lambda = 0.6\ \mu\text{m},\ \theta_s = 67°) = 0.46; \quad \tau_{\text{atm}}(\lambda = 0.6\ \mu\text{m},\ \theta_v = 55°) = 0.73$$

then:

$$\tau_{\text{atm}} = 0.73 \times 0.46 = 0.33$$

and (the Sun is the light source):

$$L_s(\theta_s = 67°,\ \theta_v = 55°,\ \lambda = 0.6\ \mu\text{m}) = 260\ \text{W} \cdot \text{m}^{-2} \cdot \text{sr}^{-1} \cdot \mu\text{m}^{-1} \cdot \rho$$

The main problem in Earth observation is to detect small variations $N_e\Delta\rho$ of ground albedo ρ. Let us take a typical value: $N_e\Delta\rho = 0.5$ % and $\delta\lambda = 115$ nm. With these values, a variation of $\rho = 0.5\%$, give a variation: $N_e\Delta L_s = 0.15\ \text{W} \cdot \text{m}^{-2} \cdot \text{sr}^{-1}$.

$N_e\Delta F$ of incoming flux or minimum signal expressed in amperes, detected during an observation time Δt, can be computed with the help of following equations. Satellite time at ground for one pixel is Δt.

$$\Delta t = \frac{\delta L}{V_{\text{sat}}} \cdot V_{\text{sat}}$$

is the motion velocity at ground of satellite ($H = 700$ km, $V_{\text{sat}} \approx 6.7\ \text{km} \cdot \text{s}^{-1}$).

The number n_s of photoelectrons detected during Δt is:

$$n_s = \frac{\delta L^3}{H^2} \cdot \frac{\pi \cdot D^2}{4} \cdot \frac{(N_e\Delta L) \cdot \eta \cdot \lambda \cdot \tau_{\text{atm}} \cdot \tau_{\text{opt}}}{q \cdot h \cdot c} \cdot \frac{1}{V_s}$$

where η = is the quantum efficiency of photodetector (in e^-/photon)
q = electrical charge of electron
h = the Planck constant
c = velocity of light
τ_{atm} = global transmission of atmosphere
τ_{opt} = transmission of optics
δL = the spatial sampling at ground

We notice that n_s is proportional to the third power of pixel sampling (≈ resolution).

The curves (Figure 11.64) are examples of realistic values obtained with the same mission specifications as above (spatial resolution, altitude, $N_e\Delta L$, λ, etc.).

It can be shown that with an aperture of 700 mm and a realistic CCD ($\eta = 0.5$ e^-/photon), the incoming photons number is very low. We must now estimate the signal-to-noise ratio. In a first approximation, and considering a Poisson statistic distribution, the noise can be written as following:

$$n_b = \sqrt{(n_s + n_{\text{obs}} + n_{\text{par}}) \cdot (1 + 2 \cdot \varepsilon) + \sigma_e^2 + \sigma_{\text{ech}}^2}$$

where n_b = the variance of global noise
n_{obs} and n_{par} = number of dark signal and spurious photons during integration time
ε = the inefficiency transfer of electrical charges in CCD register
σ_e and σ_{ech} = the variances of electronics and sampling

On the left curve above, we can see that signal-to-noise ratio is below 1 for 0.7 m spatial resolution and with very good electronics (<10 electrons), a typical transfert inefficency ($\varepsilon = 0.1$), low dark current and spurious flux (20 photons), and a very good electronic quantization (10 bits over a low dynamic).

In that case, the minimum resolution is for $\delta L > 1$ m ($S/B > 1$). In another case, we must use another technique: time delay integration detectors (TDI), which operate an integration of signal using a mosaic of detectors.

When the integration time is multiplied by $p = 20$ (using p detectors on the same arrow and transferring the electrical charges in synchronism with satellite motion), signal-to-noise ratio can be improved by a factor p:

$$\sqrt{\frac{p_o}{2}} \le p \le \sqrt{p_o}$$

The factor 2 is dependent on technology used for TDI (quantum efficiency of an individual detector can be affected by absorption in optical and electrical layers in the pixel area).

For 0.7 m spatial resolution, $S/B = 0.32$ (above example).

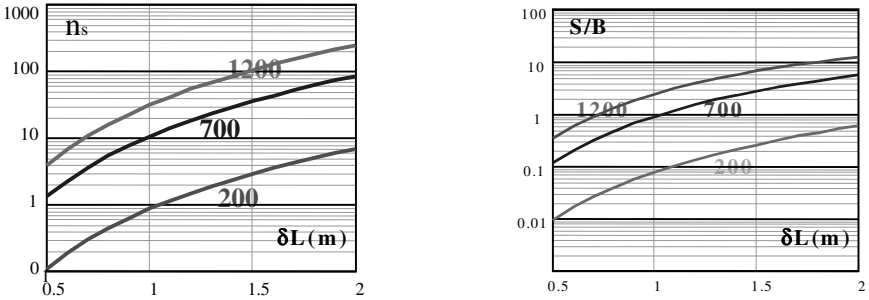


FIGURE 11.64 Number of photoelectrons detected versus spatial resolution and for different optical apertures (left); signal-to-noise ratio versus spatial resolution and for different optical apertures (right). Satellite is in a heliosynchronous orbit.

In order to obtain $S/B = 1$, and using the best technology, we need $p = 10$ integration stages, or 20 with more realistic technology.

Radiometers

Radiometers are used for measuring radiative flux (w) from the Earth or from the atmosphere. These measurements are of great importance in studying climatic variations, to monitor the flux transfer from the equator to the polar regions and incoming values for atmospheric models.

Figure 11.65 shows how incoming flux, visible and infrared (from the Sun and in arbitrary units), is absorbed, diffused, and emitted by atmosphere, clouds, land, and ocean. As we can see, the major part is absorbed by oceans and then radiated in the infrared domain. But scattering and radiation from clouds are also very important.

Of course, it is necessary to measure this flux in all electromagnetic channels, from UV to far infrared, and visible measurements are of great importance. We will now restrict our description to infrared radiation and give an example of temperature measurements from infrared radiation.

The background for this operation is blackbody radiation measurement. It is a direct relation between temperature (K), wavelength (μm) and radiation ($w/m^2/sr/\mu m$). Figure 11.66 shows the well-known Planck function for realistic temperatures from the Earth. The central problem is to detect this radiation from space and establish the link between electrical signals measured by instrument (radiometer) and incoming radiation.

A radiometer with an aperture D, transmission τ, working at wavelength λ, in spectral band $\delta\lambda$, during an observation time δt, detects optical radiation coming from a given area on the Earth δL, at altitude H, at temperature T (K), and with emittance $\varepsilon(\lambda)$.

Observation from space exists only in atmospheric windows, as shown above. These windows are mainly between 3 and 5 μm and 8 and 12 μm. It is also sometimes useful to measure radiative flux in absorption domains: H_2O (about 5 to 7 μm), O_3 (near 9.6 μm), and CO_2 (14 to 16 μm).

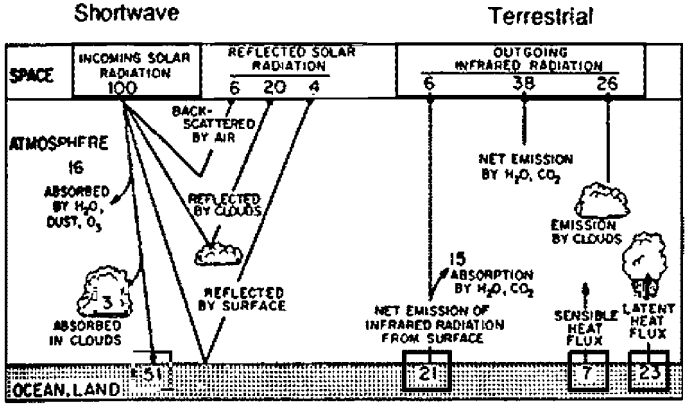


FIGURE 11.65 Balance of incoming, scattered, radiated, and emitted flux.

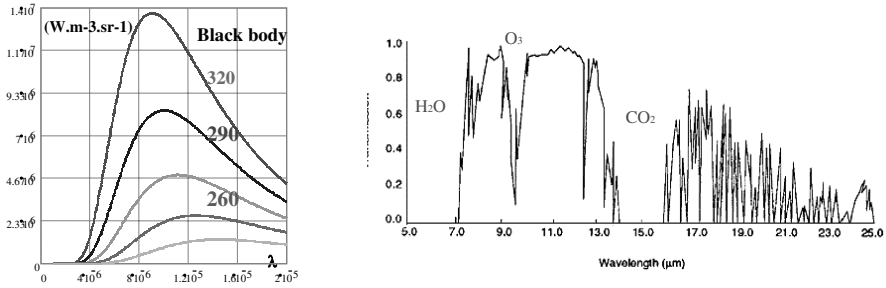


FIGURE 11.66 Blackbody curves for typical Earth and atmospheric radiation (left). Spectral transmission of atmsphere in infrared domain. Strong absorption bands are evident from CO_2 (15.5 to 17 μm) and H_2O (5 to 7 μm).

This optical signal is then detected by a cooled infrared detector and converted into an electrical signal. This signal is then compared to reference signals: blackbodies at low temperature and a zero optical signal giving an internal and external calibration.

The main problem with infrared instrumentation is the determination of the minimum signal to be detected: what is the signal and what is the noise? In infrared detection it is usual to introduce noise equivalent power (NEP):

$$\text{NEP} = \frac{\delta x \cdot \sqrt{B}}{D^*} = \sqrt{2h\nu B \cdot (\Phi_u + \Phi_p)}$$

where δx is pixel, or detector size (cm), B is electrical bandwith (Hz), and D^* is specific detection (expressed in $W^{-1} \cdot Hz^{1/2} \cdot cm$). This represents the detection performance. The best performance is obtained for cooled photovoltaic detectors. This value depends on the wavelength. Typical values of D^* are between 10^{+10} to 10^{+12}, Φ_u and Φ_p are signal and spurious flux (W) emitted by the background.

Signal corresponds to this approximation:

$$\Phi_u = B(\lambda,T) \cdot \delta\lambda \cdot \left(\frac{\delta x}{f}\right)^2 \cdot \pi \cdot \frac{D^2}{4} \cdot \tau_{\text{atm}} \cdot \tau_{\text{opt}}$$

where $B(\lambda,T)$ is Planck's function depending on λ and temperature T, $\delta\lambda$ is spectral band, D is the diameter of the entrance pupil, and τ_{atm} and τ_{opt} are transmission of atmosphere and instrument.

The important parameter is the minimum variation of temperature to be detected, $N_e\Delta T$ (noise equivalent temperature), corresponding to instrument noise. It can be shown that:

$$Ne\Delta T = \frac{4 \cdot N^2 \cdot \sqrt{B} \cdot}{\pi \cdot D^* \cdot \delta\lambda \cdot \{[\partial B(\lambda,T)]/\partial T\} \cdot \delta x \cdot \tau_{\text{atm}} \cdot \tau_{\text{opt}}}$$

Infrared space radiometers in a geostationary orbit (Meteosat, Meteosat second generation, GOES, GEOS Next, etc.) have, in a first approach, similar characteristics, such as:

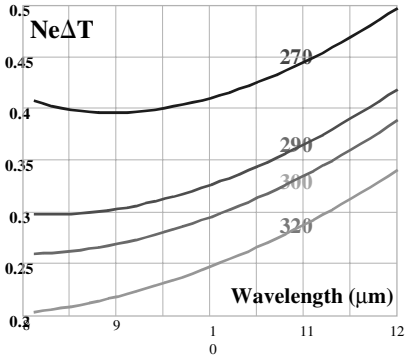


FIGURE 11.67 Minimum temperature detection $Ne\Delta T$ versus wavelength for various atmosphere temperatures ($D^* = 1.5.10^{+10}$, $N = 1.5$, $B = 30$ KHz, $\delta\lambda = 2$ μm, $\lambda = 10$ μm, $\delta x = 70$ μm, $\tau_{atm} = 0.7$).

- $N_e\Delta T$ lies in about the order of 0.1 to 0.3 K at 290 K.
- The whole Earth disk (at 36,000 km altitude) is observed with a delay of 30 minutes (first generation) and 15 minutes (second generation).
- The optical window is 8 to 12 μm for temperature measurements and narrow absorption bands (located at 5 to 7 μm, and in other domains).
- Typical entrance pupil diameters are 400 to 600 mm.
- Instantaneous field of view at ground is on the order of a few kilometers. Typical: 2 to 5 km.

Figure 11.67 give typical temperature resolutions of these instruments.

11.39 SPECTRO-IMAGERS

The aim of spectro-imagers is to carry out remote sensing of different classes of vegetation on the ground. Vegetation species can be distinguished by the spectral properties of their reflectance. Figure 11.68 shows the differences between different species. A main difference is in visible bands near 0.7 μm, where a strong variation of reflectance for green vegetation exists.

These properties are used in order to distinguish classes of vegetation, water pollution, crops, deforestation, etc. Observation is made mainly in visible and near infrared, but mineral objects are better detected in infrared (1.5–1.7 and 2–2.3 μm). Numerical classification methods are used to help this selection.

Optical spectro-imaging instruments are needed for these operations. These are cameras which form an image of the landscape in each spectra band (see Table 11.13).

A short description of the optical principles of these instruments follows. At the focal plane, a camera provides an image of the landscape. This image is then

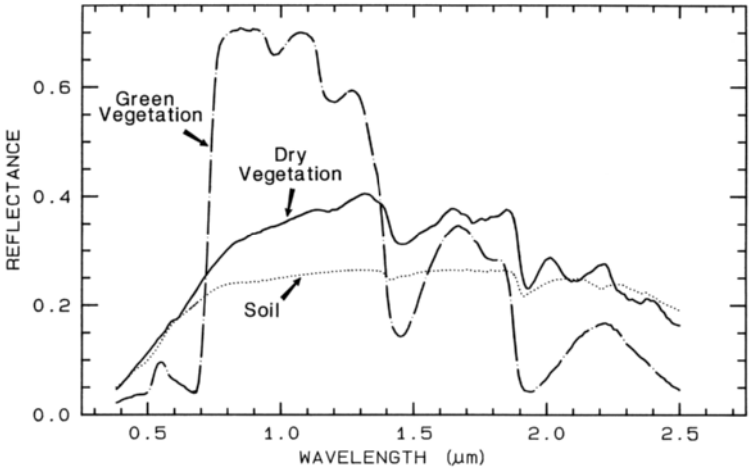


FIGURE 11.68 An example of spectral reflectance curves from various species versus wavelength.

dispersed by a spectrograph and the final image is focused on a CCD or other detector.

There are many different optical techniques. An optomechanical scanner forms an image of one or a few pixels at the entrance of a spectrometer slit. The dispersed image is then focused on an array of detectors. In order to form an image, we have to add a scanning mirror at the entrance pupil of the instrument. Scanning is done by an oscillating mirror (Thematic Mapper on Landsat) or rotating mirror. Spatial resolution is very poor with these instruments. Figure 11.69 shows the optical scheme of this type of instrument. Spectral bands are given by a dichroic beam splitter. Dichroics and narrow optical filters are sufficient because in most cases there is no need for good spectral resolution (typically 5 to 10 bands of 10 to 50 nm width, corresponding to the spectral shape of multilayer optical filters). A drawback of this design is the polarization effects that can occur with tilted dichroic plates.

The simplest solution is a continous filter (wedge filters). Figure 11.70 shows a diagram of this type of instrument.

The second sort of spectro-imager is a camera followed by a spectrometer. Figure 11.71 shows this optical principle.

When spatial and spectral resolutions are difficult to achieve simultaneously, the optical concept is more complex. Figure 11.72 gives an example of a dispersing system using prisms.

There is a compromise between spatial and spectral resolution. Spatial resolution is linked to signal-to-noise ratio. Let G be the product of transmission, τ the effective area of entrance pupil S_{pup} and solid angle of ground pixel δx_g. G is an invariant and is the same at the focal plane (see equation (11.87)). δx is pixel size, f'_2 is focal length of spectrometer, and S_2 is area of exit pupil (spectrometer).

$$G = \tau \cdot S_{\text{pup}} \cdot \frac{\delta x_g^2}{H^2} = \tau \cdot \frac{\delta x'^2}{f_2'^2} = \tau \cdot \frac{S_2}{f_2'^2} \cdot \delta x'^2 \tag{11.87}$$

Dispersive properties must also be taken in account. R is dispersive power of the spectrometer:

TABLE 11.13 Some Space Applications of Spectro-imagers[a]

Project	Resolution (m)	FOV (km)	Spectral bands	Radiometric resolution	Applications
SPOT 4 HRVIR Pbroom	10 and 20	60	4 VIS + MIR	$NE\Delta\rho = 0{,}5\%$	Remote sensing teledetection
CHRIS (G.B.)	25 and 50	19	37 to 118	$S/N = 200$ 12 bits	0.415 to 1.01 μm
MSTI (USAF)	30	7.1	15	$Ne\delta\rho \leq .001$	VIS + MIR
ADEOS OCTS Scanner	680	1,500	40 to 60	$NedL = 10^{-4}$	VIS + MIR
PRISM (ESA)	50	50	4 VIS	$S/B > 200$	Remote sensing
MERIS ENVISAT Pbroom	300	1,150 km (68°5)	Programming among 151 bands $390 < \lambda < 1{,}040$ nm	$0.5 \times 10^{-3} = \text{NE}\Delta\rho$	Ocean color monitoring
EOS-A ASTER Pbroom	15 (VIS) 30 (MIR) 90 (IRT)	60	3 VIS + 6 MIR + 5 IRT	$0.5 \times 10^{-2} = NE\Delta\rho$ $1.3 \times 10^{-2} = NE\Delta T$ 0.3 K = $NE\Delta T$	Agricultural resources

[a] The satellites listed have not all yet been launched.

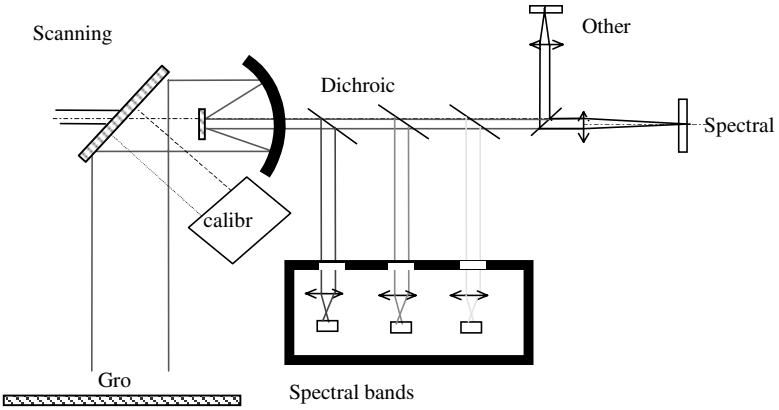


FIGURE 11.69 Optical scheme of scanning spectro-imager.

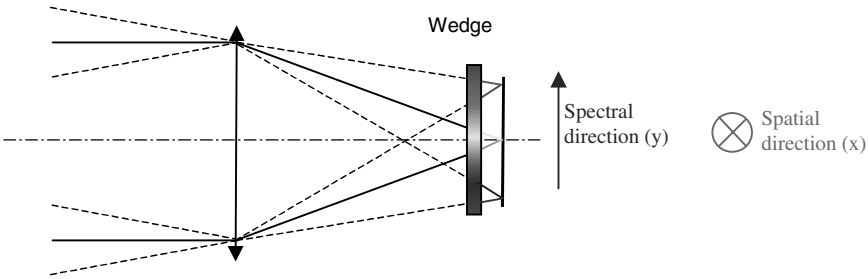


FIGURE 11.70 In the focal plane, the CCD array is in front with a wedge filter. The spectral and spatial directions are crossed.

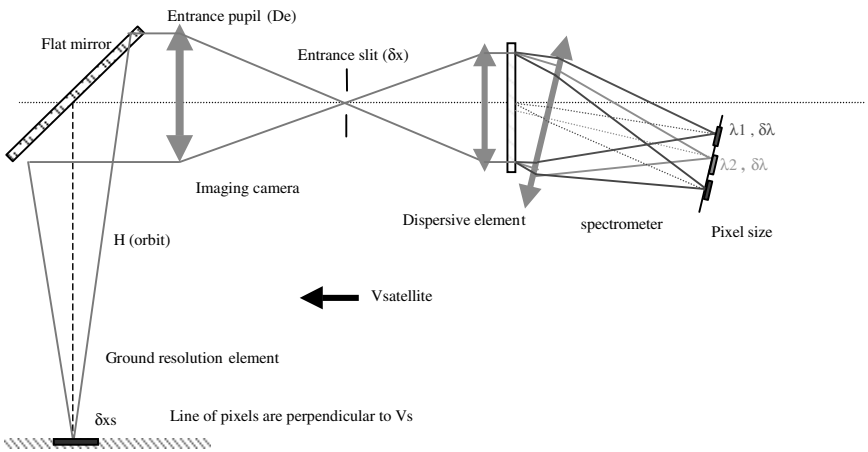


FIGURE 11.71 Optical sketch of space spectro-imager camera.

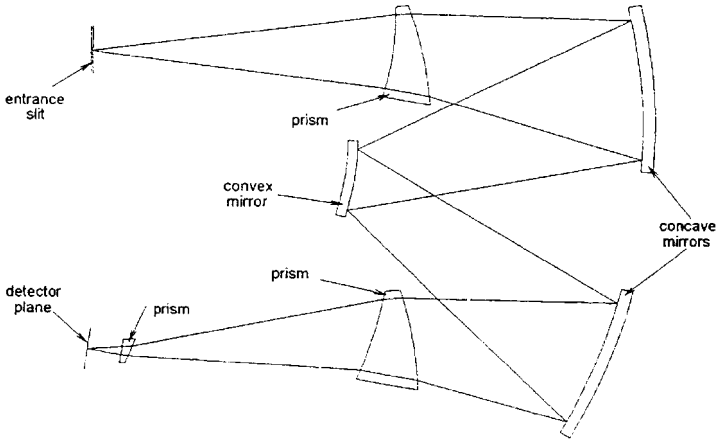


FIGURE 11.72 Aberration compensation when dispersion is made by prism (Offner relay).

$$\Delta\lambda = \left(\frac{\partial\lambda}{\partial i_2}\right)\frac{\partial x'}{f'_2} = R\,\frac{\delta x'}{f'_2} \tag{11.88}$$

In the case of a grating,

$$R = \frac{\cos i'_2}{n \cdot k}$$

where n is the number of rules/mm, i'_2 is exit angle from the grating, and k is the order of grating. For a prism,

$$R = L \cdot \frac{dn}{d\lambda}$$

where L = is the length of the prism base and $dn/d\lambda$ is the dispersion of the glass.

Optimization of optical characteristics of instrument consists of associating equation (11.87), conservation of flux, and equation (11.88), need for spectral resolution. In order to have feasible lenses, we must add a third equation:

$$\frac{f'_2}{D'_2} \geq N_2 \tag{11.89}$$

where N_2 is f/number and must be sufficiently large (typically $N_2 > 1$). The designer must choose this number before starting optimization.

In conclusion, for a given spatial resolution at ground δx_s and for a given spectral resolution $\Delta\lambda$, G is determined (for geometric and radiometric needs). Then we can rewrite the preceding equations:

Condition for imaging and flux conservation:

$$\delta x' > D_e \cdot \delta x_g \cdot N_2 \tag{11.90}$$

Feasibility of lenses:

$$\frac{f'_2}{D'_2} \geq N_2 \tag{11.91}$$

Spectral resolution need (for a grating spectrometer):

$$f'_2 \leq \cos(i'_2) \cdot \frac{\delta x'}{n \cdot k \cdot \Delta\lambda} \tag{11.92}$$

These three equations, (with the condition that we have earlier determined D_e) allow us to determine D'_2, f'_2 and pixel size $\delta x'$. Special care must also be taken for optical transmission τ, as has been done.

11.40 ACTIVE INSTRUMENTATION: SPACE LIDARS

Active instruments use an optical source (laser) which interacts with the atmosphere or the ground. Backscattered flux gives the needed information: Doppler shift (velocity measurement), time delay (distance to the target), spectral width (temperature), intensity absorption (concentration, etc.) It is then possible to make remote measurements from space with instruments equipped with very sharp and stable optical sources, such as frequency-stabilized lasers, and with appropriate detection systems.

We can distingish three sorts of space lidars: backscatter lidars, dial lidars, and wind lidars.

Backscatter Lidars

A short impulse is emitted by a laser, Nd:YAG (1.06 μm), or at a doubled or tripled frequency towards the ground. The returning impulse is scattered and absorbed by aerosols and molecules in the atmosphere. The returning signal is as in Figure 11.73, where we can see a large signal at a given altitude which corresponds to a cloud (large cross-section and absorption).

Direct detection is made with the help of a photodiode with a large-bandwith preamplifier.

The basic lidar equation is:

$$P_{\text{recpt}}(x) = \frac{E}{\tau} \cdot \beta(x) \cdot \frac{C \cdot \tau}{2} \cdot \frac{S_{\text{pup}}}{x^2} \cdot \left(\exp\left(-2 \int_0^x \alpha(y) \cdot dy \right) \right) \cdot \tau_{\text{opt}}$$

where $\beta(x)$ = atmospheric backscatter coefficient ($m^{-1} \cdot sr^{-1}$)
E = laser energy per pulse (J)
$C \cdot \tau/2$ = vertical spatial resolution of lidar (in relation to electrical bandwith)
S_{pup}/x^2 = solid angle of backscatter
α = extinction coefficient and the integral of this coefficient gives the transmission of atmosphere
τ_{opt} = transmission of instrument
τ = laser pulse duration (typically 30 to 50 ns)

Good detection is strongly related to atmospheric backscatter coefficient. Typical

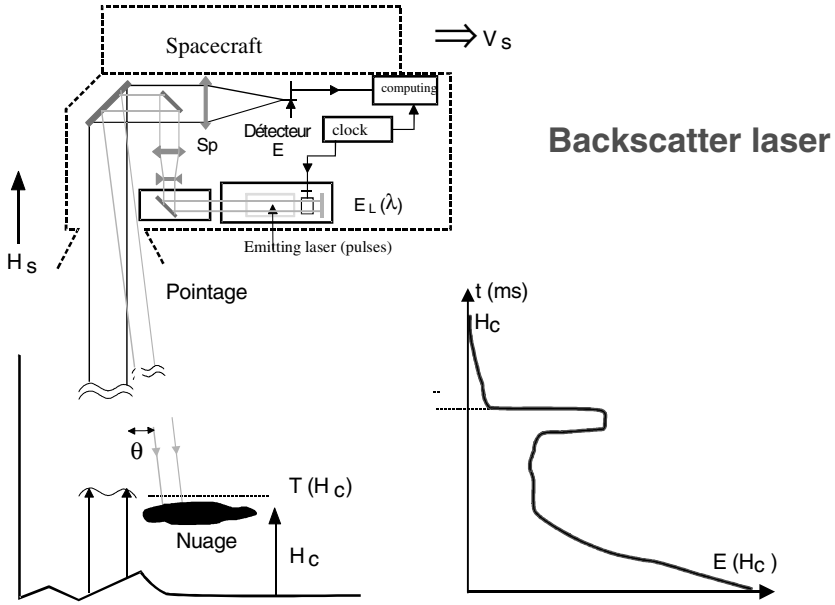


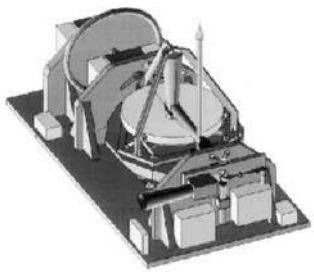
FIGURE 11.73 Schematic of backscatter space lidar.

values are 10^{-3} to 10^{-10}; such large dynamics do not allow good measurements everywhere. We must therefore consider only a small vertical distance of atmosphere (see Figure 11.74).

Another application is to monitor altitudes in order to permit computation of surfaces and shapes of polar ice sheets (MOLA on the Mars orbiter, GLAS).

Dial Lidars

Dial lidars measure atmospheric concentrations of water vapor, minor species, pollutants, etc. Two different wavelengths very close each to each other are emitted (simultaneously or with a short delay) from space to atmosphere. They correspond



Typical values of actuals lidars (Picasso, ATLID, etc)
Boundary layer: optical depth of fine clouds.
Vertical resolution 100–200m
Pulsed energy per pulse: 0.1 J; frequency: 100 Hz
Diameter of reception telescope: 800 mm to 1 m
Mass: 200–300 kg power consumption: 300 to 500 W
Photo: Atlid (image from MMS)

FIGURE 11.74 ATLID, a backscatter lidar designed by MMS for ESA.

TABLE 11.14 Choice of Lasers

Type of laser	Wavelength	Mission
Nd:YLF → SHG → OPO (BBO 832) and BBO 747	305 and 308 nm on-line	O_3-Dial (ORACLE, LaRC): 0.5 J
Nd:YLF → 2SHG → Sapphire:Ti (or OPA) → mixing	315 and 320 nm off-line	
HO:Tm:YLF diode pumped Q-switch power oscillator + injection sealed by 2 cw Ho:Tm:YLF lasers (on and off-line)	2.05–2.06 μm	Dial for water vapor and carbon dioxide
OPO based on $KNbO_3$ by tuning a Ti:Sapphire (pump)	mid-infrared <4 μm	Aerosol concentration
Other crystals:KTP, KTA		Lab experiment only (LSIOM, Lyon)
Cr:LiSAF Q-switched diode pumped and injection sealed by DFB diode laser	815.3 nm	Water vapor (airborne application)

to the center and edge of the spectral absorption line of species (typical values 40 to 100 pm). Return signals are then compared (the ratio of these return signals is proportional to concentration).

$$n_c(x) = \frac{1}{\sigma_{on}(x) - \sigma_{off}(x)} \cdot \frac{d(\text{Log } \{[V(\lambda_{off},x)]/[V(\lambda_{on},x)]\})}{dx}$$

where σ_{on} is backscatter cross-section at wavelength λ_{on} and σ_{off} is the same at wavelength λ_{off}, and $V(\lambda_{on},x)$ is the detected signal corresponding to scattering in atmosphere in layer dx at altitude x and for λ_{on}.

These lidars are very difficult to realize due to high optical power and narrow spectral lines at given frequencies. The choice of lasers depends strongly on the available technology, as shown in Table 11.14. It cannot be said that an appropriate solution exists, but solutions with OPO and solid state laser seem to offer a new opportunuity.

Figure 11.75 gives the principle of dial lidar: two pulses on- and off-line are emitted and compared at return.

Wind Lidars

Wind lidars represent the most promising area of instruments for meteorological and climatological applications (see Figure 11.76). The general mission consists of measuring windspeed vector fields over the global area of the Earth in three dimensions (mainly horizontal speeds and wind shear).

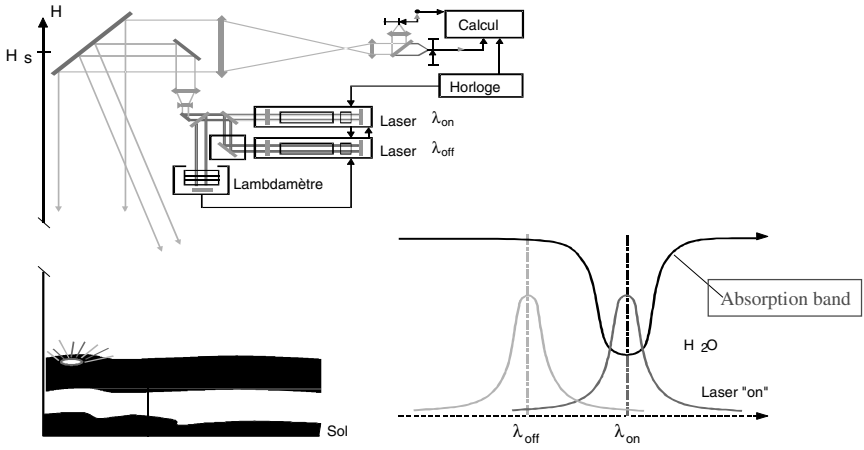


FIGURE 11.75 Principle of dial lidar.

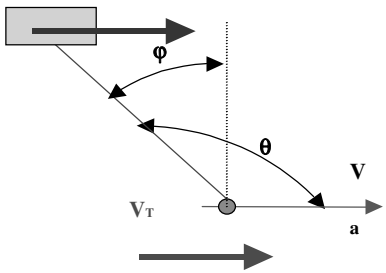


FIGURE 11.76 Principle of measurement (v is wind velocity and Doppler measurement is the projection of v into V_s).

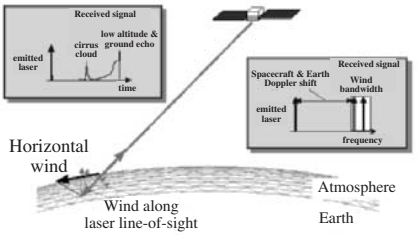
For the horizontal dimension, sampling distance on the ground must be less than the coherence length of winds at about 100 to 200 km^2. For the vertical dimension, the useful altitudes vary from the boundary limit ($\approx$ 1 to 2 km) to 8 to 12 km. It is opportune to have a vertical sampling of less than 500 m to 1 km. The repetition rate for valuable weather forecasting is $<$ 5 days.

Why are optical lidars a good solution for wind measurements? The laser measurements afford maximum accuracy because laser beams interact best with aerosols (size $\approx$ 1 to 10 μm).

Heterodyne Lidars

The basic principle of Doppler shift measurement is as follows (see Figure 11.77).

The frequency shift is given by the equation below. V_T and V_S are Earth and satellite velocity; these effects induce a Doppler shift that cannot be ignored.



Principle of wind lidar from space : lidar emits pulses towards atmosphere and retrieves frequency of backscatter. Received signal is Doppler shifted from emitted laser , due to spacecraft, Earth rotation and wind velocities.
(figure from MMS and Aladin project)

FIGURE 11.77 Frequency measurement from space: wind, satellite, and Earth motion.

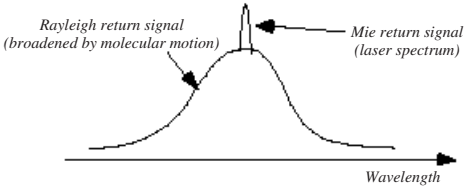
$$\frac{\Delta f}{f} = 2 \cdot \frac{(\vec{V}_a + \vec{V}_T + \vec{V}_S) \cdot \vec{N}}{C} = 2 \cdot \frac{(\vec{V}_a + \vec{V}_T + \vec{V}_S)}{C} \cdot \sin \varphi$$

The values below give an idea of these different Doppler shifts. Satellite Doppler is the most important.

$f(\lambda = 10.6\ \mu\text{m}) = 2.83.10^{+13}$ Hz
$\Delta f_a(V_a = 1\ \text{ms}^{-1},\ \varphi = 30°) = 95$ kHz
$\Delta f_T(V_T = 460\ \text{ms}^{-1}) = 43$ MHz
$\Delta f_s(V_S = 6.7\ \text{kms}^{-1}) = 632$ MHz

At the same wavelength (10 μm), we can estimate the wind dynamic: -9.5 MHz $< f_a <$ + 9.5 MHz and for 1 ms^{-1} resolution ($\approx$ 100 kHz), the total dynamic (Earth + satellite + wind motion) is $\approx$ 700 mHz ($f/\delta f$ = 7,000).

Figure 11.78 describes the instrument when heterodyne detection is used. The master laser (frequency stabilized by injection seal) emits short energetic pulses (more than 1 J). This signal, after diffusion by particles in motion (Doppler effect), returns to the instrument 3 or 4 ms later and is mixed by a local oscillator. The detected signal is at the difference between these frequencies and is the total Doppler signal. After detection, some conditioning is done in order to increase the signal-to-noise ratio. Many individual pulses are added and an optimal filtering, such as between spectral filtering, is applied.



Scattered signal by aerosols and molecules

FIGURE 11.78 Rayleigh and Mie spectrum: it is easier to measure Mie signal (narrow) than Rayleigh signal, but the probability of a Mie signal is much lower than of a Rayleigh signal.

Direct Detection

Direct detection is used for molecular (Rayleigh) and particular (Mie) backscattered signals. This recently developed method is used to measure windspeed. The basic principle is that it is easier to design a good detection system than to develop a high-power, robust laser.

The signal (Nd:YAG laser at 0.35 μm) is scattered by aerosols (Mie) and performs wind velocity measurement (less than 1 m/s) well at low altitudes, where there is a good concentration of these particles. But at higher altitudes the situation is quite different. Concentration of large particles (size on the order of wavelength or more) become very low and the equivalent cross-section of these particles (β) is too small, even in the UV region. Good detection is not possible. It is possible to detect molecular diffusion (larger signals), but Doppler scattered signals with molecules are broadened by molecular motion (temperature effects). Figure 11.78 shows these phenomena.

In the absence of wind, no Doppler shift occurs (Figure 11.79). Wind creates a Doppler displacement of these figures. This displacement can be detected thanks to different measures in two spectral bands.

The instrument uses a triple Nd:YAG laser (0.355 μm) for two main reasons: strong backscatter cross-section of molecules and particles (Mie and Rayleigh effects) and eye safety.

Typical spacing for Rayleigh measurements: Separation between A and B is 5 GHz, linewidth is $\approx$2 GHz. For a Mie receiver, the spectral line is much narrower, typically 100 MHz. Note: spectral laser width is about 30 MHz.

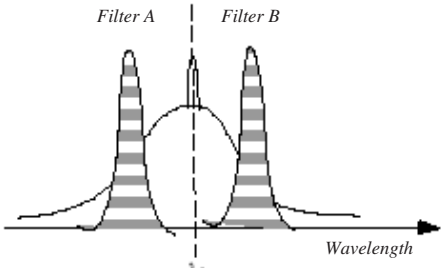
Detection is always critical: a large receiver telescope (typically 1 m) is needed, and laser energy must be high (typically 0.1 J). Pulse rate must also be high (100 Hz).

Optical design consists of two separate channels, one measuring spectral shift of the Rayleigh signal (Figure 11.80).

Theoretical accuracy of detection is given by the following formula:

$$\delta V = \frac{C}{2 \cdot \lambda} \cdot \frac{\delta \lambda_{\text{rms}}}{\text{SNR}}$$

where C is velocity of light in a vacuum, $\delta\lambda_{\text{rms}}$ is the width detected (Mie or Rayleigh), and SNR is signal-to-noise ratio.



Principle of detection using two optical filters on either side of a Doppler curve. Displacement of these curves creates a dissymmetry in detection at detectors A and B.

FIGURE 11.79 There is no Doppler shift if there is no wind. Measurement consists of detecting two identical signals at two detectors filtered at A and B.

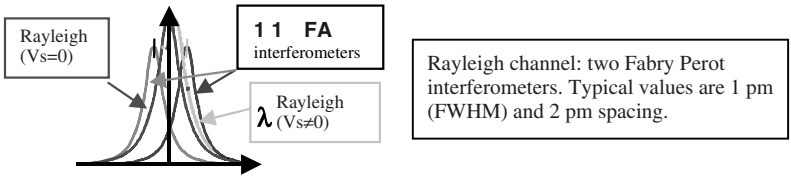


FIGURE 11.80 Optical measurement with filters.

***Numerical Example* (*ALADIN*)**

Wind Doppler bandwidth (± 100 m/s): 700 MHz (at 0.355 μm)
Spectrometer bandwidth: 210 MHz
Altitude range: 20 km
Estimated accuracy of wind measurement: 1.3 m/s
Vertical resolution: 0.5 to 1 km

These values correspond, in a row approximation, to actual design specifications of some actual space projects.

11.41 PASSIVE SOUNDING FROM SPACE

Atmospheric measurements using this technique from space consist of determining temperature and humidity from vertical profiles. The major advantage of space measurement is regular, long-term observation on a global scale.

The method consists of conducting radiometric measurements at the nadir or at the limb in very narrow spectral absorption bands. These measures determine brightness temperatures from absorption lines of a gas (CO_2, etc.).

Absorption lines are strongly correlated with vertical profiles of atmospheric temperature: radiant energy $I(v)$ is the sum of thermal emission from different absorbing layers of the atmosphere. A layer emits infrared flux in the direction of the upper atmosphere. This light is absorbed by upper layers, which also emit infrared radiation. Global radiation outside the atmosphere corresponds to the weighted sum of all these individual emissions.

The following equations give the principle of this operation:

$$I(v) = \tau(0) + \int_0^{\infty} B_v(T(\log(P)) \cdot F_v(\log(P)) \cdot d(\log(P)) \tag{11.93}$$

where F_v is the equivalent weighting function: as

$$F_v(\text{Log } P) = \frac{\partial \tau(v)}{\partial(\log(P))} \tag{11.94}$$

The problem now is to measure optical radiation $I(v)$ in narrow spectral bands (CO_2, O_3, etc.) in order to determine $B_v(\log(P))$, assuming the weighting function is known.

Figure 11.81 shows that by selecting wavelengths so that emittance is maximum in a given layer and weak in the other layers, it is possible to use measurements at each wavelength to derive the temperature of corresponding layer.

This is an inversion method and will be performed on the ground.

For measurement of $I(v)$ an optical spectrometer with high spectral resolution (typically 0.01 $cm^{-1} \leq \delta\sigma \leq 1\ cm^{-1}$) is needed.

Most instruments in space use spectroscopy by a Fourier transform. From an interferogram taken by instrument (Michelson interferometer), we perform a Fourier transform, which gives an even spectrum of radiated emission in atmosphere (see Figure 11.82).

A Fourier transform interferometer is shown in Figure 11.83. Pixel element ΔL at ground is focused in the entrance aperture δx of the interferometer. When moving mirror M2 from path difference $-\Delta$ cm to $+\Delta$ cm, it can be shown, using spectroscopic Fourier transform properties, that:

$$\Delta\sigma_{th} = \frac{1}{2 \cdot \Delta}$$

where $\Delta\sigma_{th}$ is theoretical resolution (expressed in cm^{-1} if Δ is path difference expressed in cm.

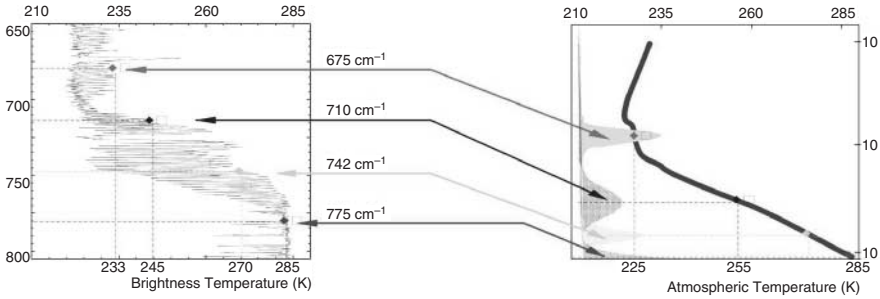


FIGURE 11.81 Correlation between CO_2 absorption spectrum and atmospheric temperature profile (from CNES/Eumetsat documentation: IASI mission).

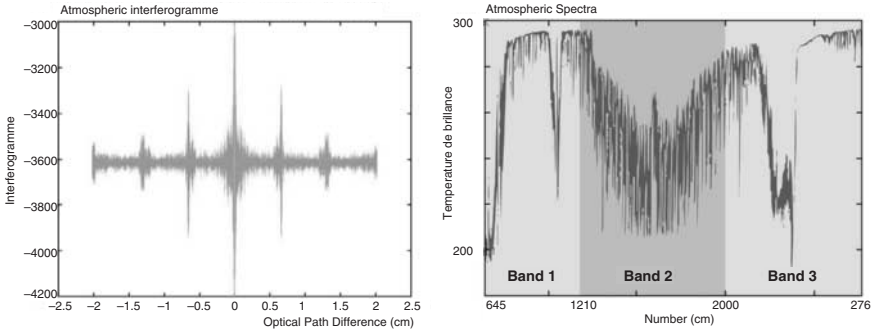


FIGURE 11.82 The figure on the left shows simulated interferogram (over ± 2 cm optical path difference) and right figure is Fourier transform a (atmospheric spectrum).

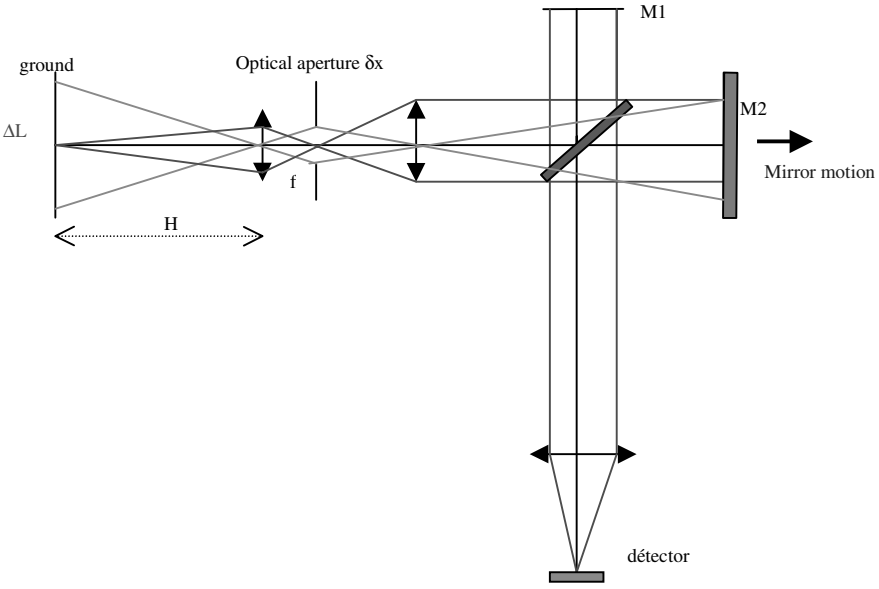


FIGURE 11.83 Diagram of a Fourier transform spectrometer for atmospheric sounding.

The relation between the interferogram and the spectral shape of light is given by the equation:

$$\Phi(\delta) = G \cdot \int_{-\infty}^{+\infty} B_p(\sigma) \cdot d\sigma + G \cdot \int_{-\infty}^{+\infty} B_p(\sigma) \cdot \exp(j2\pi\sigma\delta) \cdot d\sigma$$

where $B_p(\sigma)$ is the even part of the spectrum, $\Phi(\delta)$ is the interferogram for a path difference δ (in cm, if σ is expressed in cm^{-1}), and G is a radiometric constant, taken into account aperture size, field of view, transmission, and irradiance of incoming light

$$\delta = 2 \cdot V \cdot T$$

corresponding to displacement of mirror, where V is the motion of plane mirror and T is time of displacement.

Spatial resolution ΔL and spectral resolution $\delta\sigma$ are linked. When we must increase ΔL, because signal is too low, for example, then spectral resolution decreases. The following equations show this problem:

$$R = \frac{\sigma}{\delta\sigma_f} = \frac{2\pi}{\delta\Omega}$$

where R is resolution power limited by the field of interferometer, and

$$\delta\Omega = \frac{\pi}{4}\left(\frac{\delta L}{H}\right)^2$$

representing solid angle of aperture of interferometer.

Global resolution is given by approximate equation:

$$\delta\sigma_g = \delta\sigma_{th} \otimes \delta\sigma_f$$

Sounding Spectrometer: Example

Infrared Atmospheric Sounding Interferometer (IASI) mounted on Metop satellite (ESA). Orbit is heliosynchronous.

Main Characteristics

Spectral range: 645–2,760 cm^{-1} (3.6–15.5 μm).

Spectral resolution: $\delta\sigma_g = 0.35–0.5\ cm^{-1}$.

Radiometric resolution: 0.25–0.5 K at 290 K.

Ground resolution: 25 km (cloud-free).

Vertical resolution: 1 km in low troposphere ($\delta T = 1$ K). This instrument also measures humidity profiles (10%) and detects other gases such as: CO, CO_2, N_2O, and CFCs.

The Fourier transform is computed on board, but inversion and data processing are done on the ground. Total power is about 200 W and mass is 210 kg.

PART 10

IN-FLIGHT COMPUTING

Frederic Boniol

Jack Foisseau

11.42 INTRODUCTION TO AVIONICS

In this Part, computer-based technologies used for supporting more and more aircraft functions are presented. The term *avionics* (*avi*ation electron*ics*) refers to aircraft equipment serving aircraft functions such as communications, navigation, and automatic flight control. In fact between aircraft functions and embedded equipment there is an intermediate concept, called *system*, including many pieces of equipment, some of them being computers with a significant amount of software. Figure 11.84 is a class diagram introducing the basic notions and their relationships.

The main top-level aircraft functions are:

1. Provide and distribute communications (with ATMS, airlines, crew, etc.)
2. Plan, generate, and control aircraft movement
3. Provide crew, passenger, and cargo environment and services
4. Detect and analyze aircraft conditions for flight
5. Provide airframe movement and attachment capability
6. Provide containment and internal support

Each top-level function may be decomposed into subfunctions. For instance, function 2 above includes navigational guidance, aircraft performance monitoring, flight planning, thrust control, ground braking, aerodynamic configuration control, auto-managed and controlled flight, fuel management, etc.

Many aircraft functions performed in the past by mechanical, hydraulic, or electrical systems are now implemented as computer-based systems. A *system* is an aggregate of equipments. For instance, the flight control system of an aircraft includes equipment (like the rudder pedal, the ailerons, elevators, spoilers, hydraulic actuators, electrical and hydraulic connections, sensors, and computers inputting/outputting many signals) enabling the pilot (human or autopilot) to control the

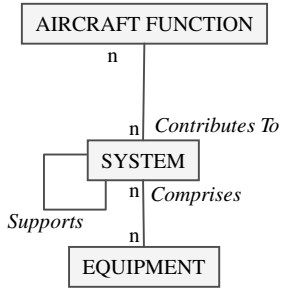


FIGURE 11.84 High-level conceptual model for avionics.

aircraft. For other systems in interface with pilot and/or crew, additional equipment types are displays, input devices, etc. Sensors and actuators are connected to computers communicating by means of data buses.

ATA Specification 100

Aircraft systems obey a standard classification for identifying aircraft equipment: ATA (Air Transport Association of America) classification 100, used by all airframers and aircraft equipment manufacturers for the preparation and distribution of technical data about systems. It divides an aircraft into major systems (see Table 11.15).

ATA 100 is a numbering system providing a unique identifier for each aircraft hardware component (the first two digits identify a major aircraft system). Table 11.16 shows the subdivisions specified by ATA 100 for two major systems.

With the use of digital information, a companion specification (ATA 2100) was designed to standardize the generation and distribution of ATA 100 data in electronic form.

11.43 REQUIREMENTS FOR AVIONICS

In the case of avionics, three requirement types are important: real-time constraints, safety-related regulatory requirements, and economic requirements.

TABLE 11.15 ATA 100 Computer-Based Systems

ATA chapter	System
21	Air conditioning
22	Auto flight
23	Communications
24	Electrical power
25	Equipment/furnishings
27	Flight controls
28	Fuel
29	Hydraulic power
30	Ice and rain protection
31	Indicating/recording systems
32	Landing gear
33	Lights
34	Navigation
35	Oxigen
36	Pneumatic
38	Water/waste
42	Integrated modular avionics (IMA) and avionics data communication
49	Airborne auxiliary power
54	Nacelles/pylons
71	Power plant
76	Engine controls

TABLE 11.16 Two Examples of ATA 100 Classification

ATA chapter	ATA section	System/subsystem
22		**Auto flight**
	22–10	Autopilot
	22–20	Speed attitude correction
	22–30	Auto throttle
	22–40	System monitor
24		**Electrical power**
	24–10	Generator drive
	24–20	AC generation
	24–30	DC generation
	24–40	External power
	24–50	Electrical load distribution

Real-time Constraints

Real-time constraints are of great importance when dealing with the control loop of dynamic systems. The total time delay is important for stability and other performance aspects of closed-loop control of the flight control system, the automatic pilot system, the flight management system, etc. The time lag between data input and the corresponding output (orders to actuators, for instance) is due to (a) computing, (b) communication latency, and (c) storage time. Typically, the input is preprocessed before being used in the main computational cycle, and the output will be postprocessed after being computed. The pre- and postprocessing may involve unit conversion, reasonableness checks, comparison with inputs from other computers (voting, etc.), and integrity checks on the communication links. The time required for these activities may vary but could increase the effective lag between input and output orders above the value corresponding to the specified iteration rate. This could impact the performance of the closed-loop control system, and consequently the stability of the aircraft. Guaranteeing and verifying real-time constraints are major tasks when designing and validating avionics systems.

Regulatory Requirements

Regulatory requirements are passed to aircraft manufacturers through FARs (Federal Aviation Regulations) established by the FAA (Federal Aviation Authority). European countries have the JAA (Joint Aviation Authority), editing JARs (Joint Aviation Regulations). Compliance with these requirements must be shown before the commercial aircraft is allowed to be operated. Such requirements pertain mainly to the safety of the aircraft and its occupants. The goal of system safety is to optimize safety by identifying safety-related risks, eliminating or controlling them by design and/or procedures.

Economic Requirements

Economic requirements are often associated with direct operating costs (navigation fees, crew, maintenance, fuel and oil, etc.) and also with indirect costs driven by

various factors such as use of automated system engineering tools, dispatch reliability, use of standardized computer resources, etc. Such economic requirements come from airlines and address the costs of purchasing, operating, maintaining, and servicing the aircraft.

11.44 PHYSICAL ARCHITECTURE

Fundamentals of Architecture

There are three types of fundamental avionics architecture: centralized, federated, and distributed.

Centralized Architecture. A centralized architecture is characterized by signal conditioning and computations taking place in one or more computers in a line replaceable unit (LRU) located in the avionics bay with sensor and command signals transmitted over data buses. The main advantages are the computers' accessibility within the bay; a safe computer environment, simplifying qualification; and easily written and validated software. However, there are two main drawbacks: many long buses to collect and transmit data/commands, and system vulnerability to hazardous events occurring in or near the avionics bay. Centralized architectures are commonly found in rockets, missiles, and satellites.

Federated Architecture. A federated architecture is characterized by each major system sharing input and sensor data from a common set of hardware and consequently sharing their computed results over data buses. *Partitioning* (see ATA classification) is then an intrinsic feature of this architecture. One of the main advantages is that such an architecture permits independent design, configuration, and optimization of the different systems. Consequently, changes in system hardware/software are relatively easy to make. Federated architectures are commonly found in modern military aircraft (F-16, EFA, Rafale, etc.) and commercial transport aircraft (Airbus 320/330/340 Boeing B777, etc.).

Distributed Architecture. A distributed architecture is composed of multiple processors throughout the aircraft that are assigned computing and control tasks in real time. Limited signal processing may be performed in or near the sensors and the actuators. Such an architecture presents several advantages: fewer and shorter buses, faster program execution due to many computer resources, and reduced vulnerability thanks to distribution. However, such an architecture needs more processor types and then more complex software generation and validation procedures. Furthermore, some of the processors, located near sensors and actuators, may be in a more severe and less accessible environment. This could be an important drawback for maintenance costs. Distributed architecture are not yet implemented but are being examined for future avionics systems.

MIL-STD-1553B and ARINC-429 Federated Architectures

Modern military and commercial aircraft include federated avionics architectures, but based on different digital data buses.

MIL-STD-1553B-Based Architecture. This is a serial bidirectional two-wire bus based on a time-division command/response multiplex mode. Communication between different units can take place at different moments in time, and all communications between units share the same bus. Figure 11.85 shows a hypothetical architecture based on three redundant 1553 data buses which operate at 1 megabit per second (Mbit/s) in a half-duplex mode using Manchester II data encoding.

Hardware is the tangible portion of a 1553 bus. The major hardware elements are the bus controller, remote terminals, and optional bus monitors.

The bus controller (BC) is in charge of all data flows on the bus and initiates all information transfers. It transmits and receives data in command/response mode. This sole control of information transmission on the bus shall reside within the BC. The BC also monitors the status of the system. BCs are generally separate LRUs, but they can be part of one of the other LRUs on the bus.

A bus monitor (BM) receives and stores selected bus traffic. A BM will not react to any traffic received unless the traffic is specifically addressed to it. BMs are generally used to receive and extract data for off-line purposes (flight test, maintenance, etc.). They are strictly passive and are used only as a test device.

Remote terminals (RTs) are the largest fraction of units in a 1553 bus system. No RT can speak on the bus unless spoken to first by the BC and specifically commanded to transmit. RTs respond only to valid commands specifically addressed to them or to valid broadcast (all RTs simultaneously addressed) commands. Because of the address field size, there can be only up to 31 RTs on a given bus. The address of an RT is established by external wiring. This approach allows any LRU of a given type to be used in any location without internal changes to it being made. The RT can be separate from any subsystem it serves, or it can be embedded in it.

The avionics architecture of the F-16 C/D fighter aircraft, shown in Figure 11.86 is a typical 1553B-based architecture with three dual redundant 1553 data buses.

The primary bus BC function of these three buses is provided by the fire control system. Each bus supports a top-level function: navigation and engine control, defensive avionics, etc.

The EuroFighter Aircraft (EFA) avionics is also based on 1553 buses. It has a flight-critical fly-by-wire control system and a utilities control system (UCS).

The UCS, shown in Figure 11.87 is composed of five computers (MC 68020 microprocessors) integrated by a 1553 data bus, and it implements centralized control for the secondary electrical power, hydraulic management, fuel management, fuel gauging, environmental, etc. These systems are integrated in the same federated architecture based on the same redundant 1553 bus.

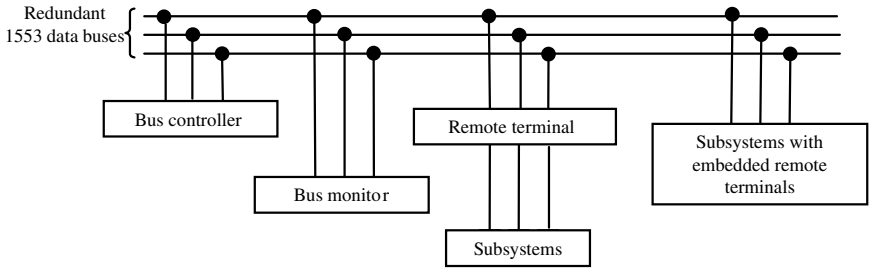


FIGURE 11.85 Typical MIL-STD-1553B-based architecture.

AVIONICS AND ASTRIONICS

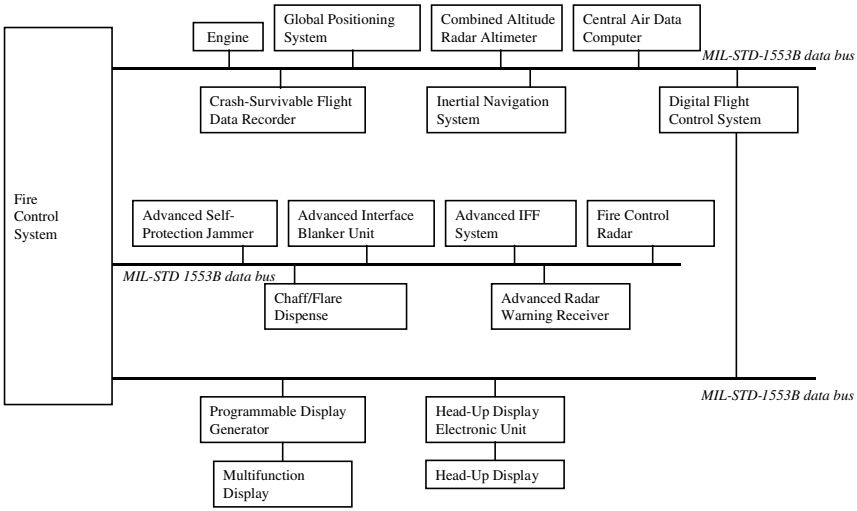


FIGURE 11.86 Example of a MIL-STD-1553B-based architecture: F-16 avionics system.

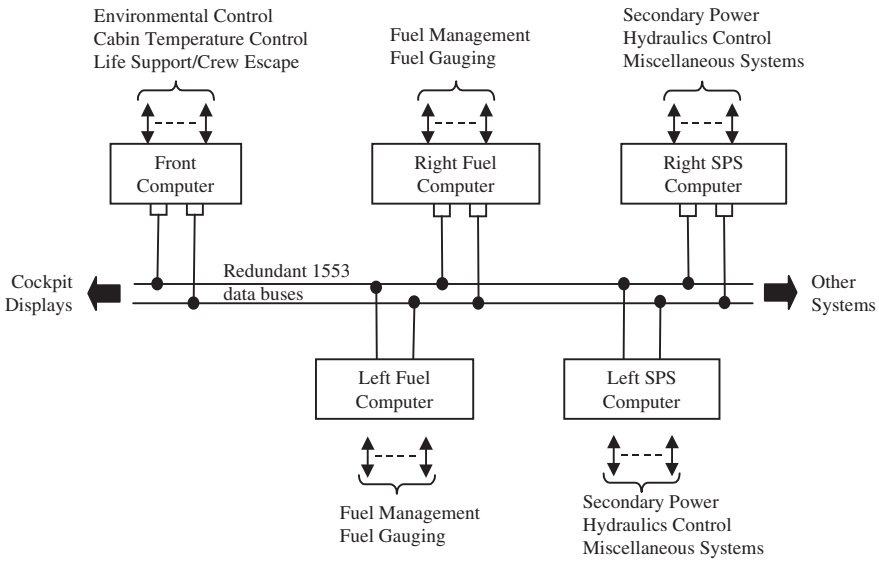


FIGURE 11.87 Example of a MIL-STD-1553B-based architecture: EFA avionics system.

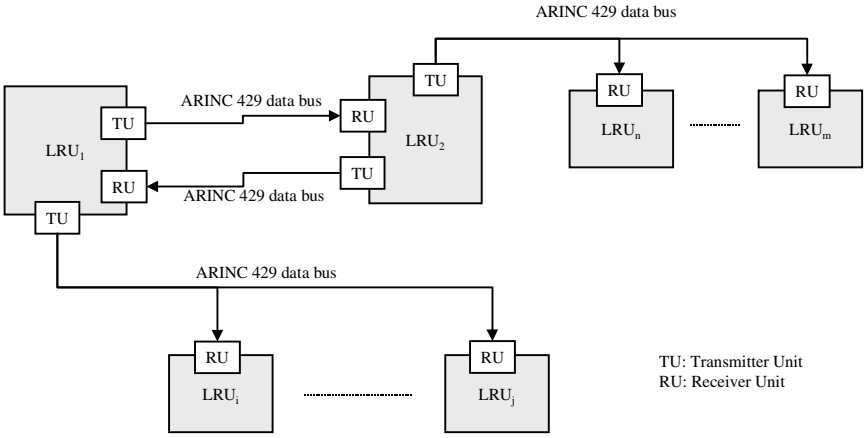


FIGURE 11.88 A typical ARINC 429-based architecture.

ARINC-429-Based Architecture. Just as MIL-STD-1553B is the basis for digital buses in modern military aircraft, ARINC-429 is the basis for digital buses in modern civil aircraft. These two standards establish vastly different approaches to data buses that reflect the differences in design philosophy and performance priorities for military and civil aircraft. Requirements for minimum weight and maximum flexibility drove 1553 to operate at 1 Mbit on a bidirectional bus, while certification requirements drove 429 to operate at either 12 to 14.5 or 100 kbit/s on a monodirectional bus.

ARINC 429 is a monodirectional bus, i.e., with only one transmitter T and multiple receivers $R_0 \ldots R_n$ ($n \leq 20$). If R_i needs to communicate with T, a separate bus with R_i as the source and T as the sink is required (see Figure 11.88).

This can lead to complex architecture with a great number of ARINC 429 buses. However, due to the absence of multiplex traffic, there is no need for bus controllers and remote terminals. Consequently, such an architecture is more easily certified. This is a major advantage for civil aircraft.

Communications on 429 buses use 32-bit words with odd parity. ARINC specification 429 identifies two transmission rates: 100 kbit/s (high-speed bus) for transmitting large quantities of data or flight-critical information, and 12 or 14.5 kbit/s (low-speed bus) for general-purpose, low-criticality information.

ARINC 429-based architectures are commonly found in commercial aircraft designed and built in the 1980s. For instance, the avionics system of the Airbus family A320/A330/A340 is based on such digital data buses (A340 avionics architecture includes more than 300 ARINC 429 buses).

Integrated Modular Avionics (IMA)

Integrated modular avionics will make the current method of packaging line-replaceable units in electronic equipment obsolete. Functions now found in relatively large stand-alone units will be performed within a large set of hardware modules, forming a subsystem with a common chassis, common fault-tolerant proc-

essing, centralized power supplies, and flexible aircraft interfaces. From a computing point of view, IMA is basically a standard avionics architectural model containing removable components called line-replaceable modules (LRMs) communicating through multiplex digital data buses (ARINC 629, Ethernet-based, etc.). This replaces many black boxes and enables a large variety of avionics functions to share processing, memory, I/O functions, and common software such as a real-time operating system that independently processes application software while maintaining robust partitioning between software modules for the most critical flight functions. An IMA architecture example is the ARINC 629-based architecture, as in the Boeing B777. Other protocols are being studied, such as Ethernet-based deterministic protocols.

ARINC-629-Based Architecture. ARINC 629 is a multiple-transmitter digital data bus which operates at 2 Mbit/s. It uses word formats very similar to those in 1553, but without a bus controller. Each terminal has an autonomous access to the bus through a fair distributed protocol. There are two protocol schemes for ARINC 629: basic and combined. In the basic protocol, the bus operates in either the periodic or aperiodic mode. In the combined protocol, the periodic messages have priority and are guaranteed to be transmitted on a set schedule, while the aperiodic data are transmitted only when the bus is free.

The Boeing B777 Airplane Information Management System (AIMS) is a typical example of ARINC 629-based IMA architecture.

The central features of AIMS, shown in Figure 11.89, are the two cabinets containing LRMs. AIMS functions performed in both cabinets include flight management and engine data interface, electronic flight instrument system (EFIS), engine indicating and crew alerting system (EICAS), central maintenance, communications management, etc.

Within each cabinet (see Figure 11.90), the LRMs are interconnected by dual ARINC 659 backplane data buses. The cabinets are connected to the triplex redundant ARINC 629 system bus and to the triplex redundant fly-by-wire bus, and are also connected via ARINC 429 buses to the EFIS and EICAS displays.

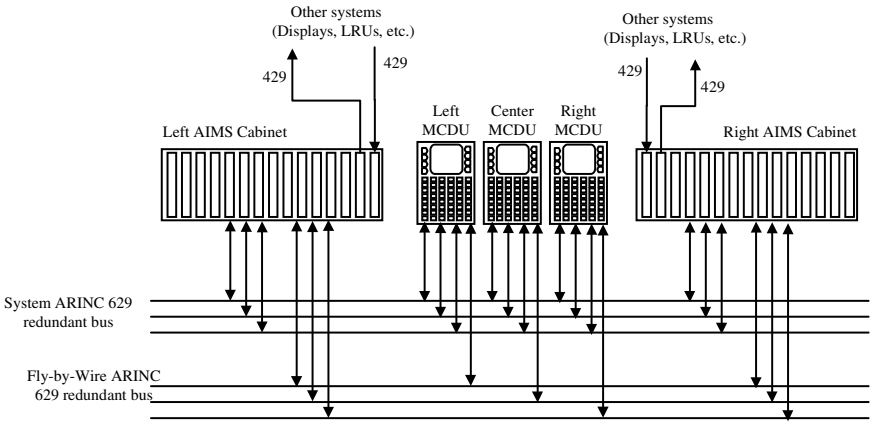


FIGURE 11.89 Example of an ARINC 629-based IMA architecture: B777 avionics system.

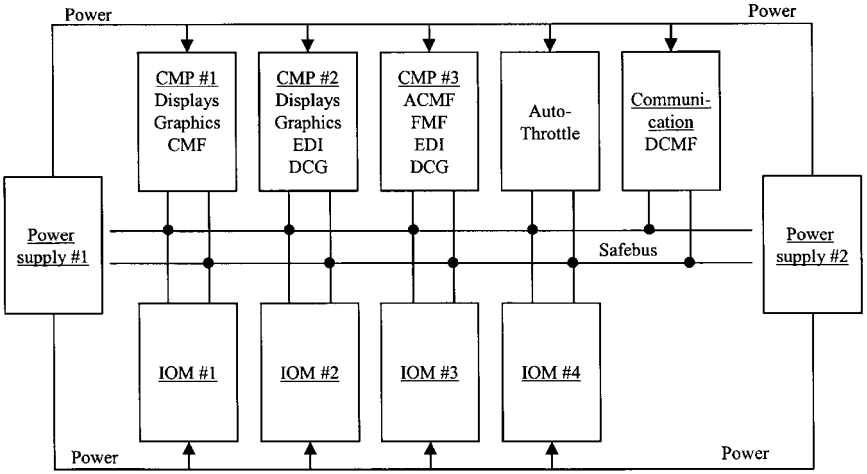


FIGURE 11.90 B777 AIMS cabinet architecture.

11.45 AVIONICS LOGICAL ARCHITECTURE

Fault tolerance and maintenance are two major points of view. To meet the fault-tolerance requirements (safety and availability), avionics fault-tolerant systems generally use several types of fault tolerance, such as redundancy and dissimilarity, and have the capability for automatic, dynamic reconfiguration of the system. To meet the maintenance requirements, avionics designers have developed specific logical architecture and built-in test functions.

The Fault-Tolerance Point of View

The possible system configurations range from the simple single microcomputer with its own internal computer bus to configurations using 100 or more microcomputers and over 250 ARINC 429 data buses (e.g., the Airbus A340). Critical functions are usually implemented as *fail operational* (i.e., the function continues to operate in a normal manner after any single system component fails). Essential functions are usually implemented as *fail passive* or *fail soft* (i.e., a single failure will either cause loss of the function with no aircraft perturbation or cause a disturbance which is limited to a safe and acceptable level).

Avionics architectures are generally based on redundant channels configured as either independent channels (no interconnection or sharing of control signals between the parallel channels) or cross-strapped (interconnections do exist). Cross-strapping may be accomplished by either analog cross-feed or intercommunication between processors. Redundant channels can operate in either active or active-standby mode. In the active mode, all redundant channels are operating simultaneously. In the active-standby mode, some of the channels are controlling the system while the others are standing by, ready to assume control in the event one or more of the active controlling elements is declared faulty by the fault-detection logic. Some of the possible architectures are discussed in the following paragraphs.

Simplex Architecture. The simplex configuration is a single dedicated computer that may have one or more sensors/subsystems connected to input/output ports as well as actuators connected to the output ports. The principal fault-detection technique consists primarily of the data validity checks with built-in test equipment for self-monitoring.

Dual Architecture. A dual architecture is composed of two identical parallel channels performing the same functions. Outputs of the two channels can be compared.

Figure 11.91 shows a typical dual flight control system. In normal operation, channel A is active and channel B is in standby mode. Built-in test equipment (BITE) and software tests monitor the components of each channel. A faulty component in channel A is automatically replaced by the corresponding one in the B channel (and vice versa).

Such an architecture assumes that elementary faults can be detected by embedded functions of equipment. However, certain types of faults, e.g., latent (failure to zero), bias, slow drift, and scale factor shift, are not readily detectable. Consequently, single–dual configurations have not been widely used for critical functions.

Command/Monitor Computer. This is a particular dual configuration with cross-strapping only for output comparison. This concept has been developed and proven by Aerospatiale for Concorde and more recently for the Airbus family A320/A330/A340. Such a configuration, depicted in Figure 11.92, is composed of a pair of dissimilar channels: a command channel ensuring the function allocated to the computer (for example, control of a control surface) and a monitoring channel ensuring that the command channel operates correctly.

Each channel includes one or more processors, their associated memories, input/output circuits, a power supply unit, and specific software (hardware and software of the two channels are dissimilar). Both channels are simultaneously active. The difference between the results computed by each channel is compared. If this difference is above an allowable threshold for a sufficiently long period of time, the channel which detects this failure cuts the links between the whole equipment (i.e., the two channels) and the outside world. System tolerances (most notably sensor inaccuracy, rigging tolerances, and computer asynchronism) must be taken into account to prevent undue failure detection, and errors which are not detectable (within the signal and timing thresholds) are assessed with respect to their handling quality and structural loads effect.

Command/monitor computers are fail-passive (internal failures are not propagated outside the computer). This allows the design of a fault-tolerant system composed of one active equipment and N-1 standby equipment. Reconfiguration in such a system is then very simple and natural. The second equipment becomes active when the first produces no more data.

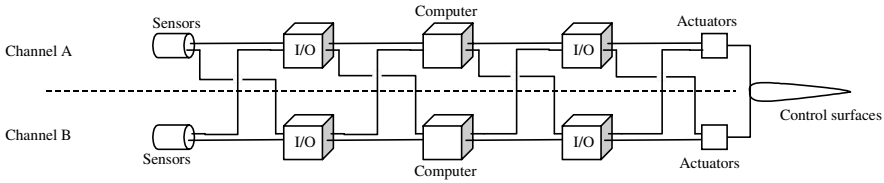


FIGURE 11.91 Typical dual logical architecture.

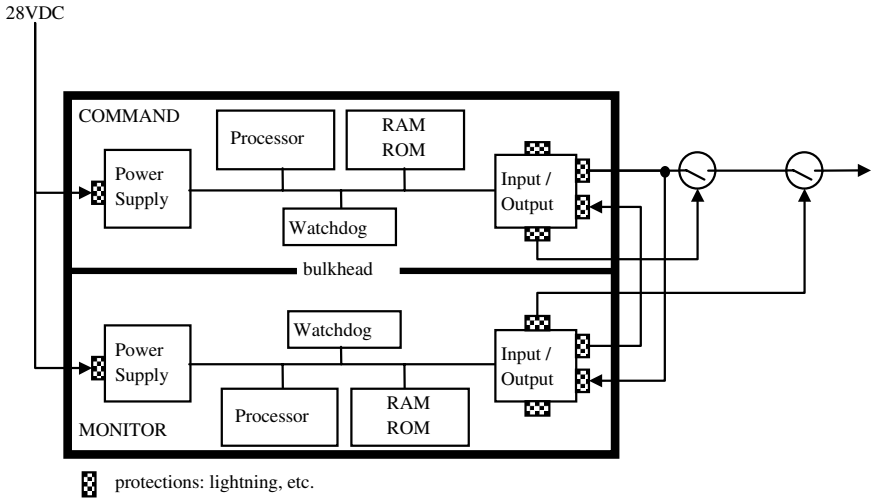


FIGURE 11.92 Command/monitor architecture.

Control/monitor digital equipments are commonly found in the essential and critical digital systems (automatic pilot, flight control system, break system, fuel management system, etc.) of Airbus built in the 1980s (see example of A320 flight control system presented below).

Dual Command/Monitor Architecture. Figure 11.93 depicts a typical dual–dual configuration.

While this may appear to be a four-computer system, the system is typically operated as two similar command/monitor equipments E1 and E2, each composed of two subchannels which exchange data. Data exchange also occurs between E1 and E2.

In normal operation, E1 is active and E2 is in standby mode. In case of channel 11 or channel 12 failure, the first equipment is disconnected and no data are sent

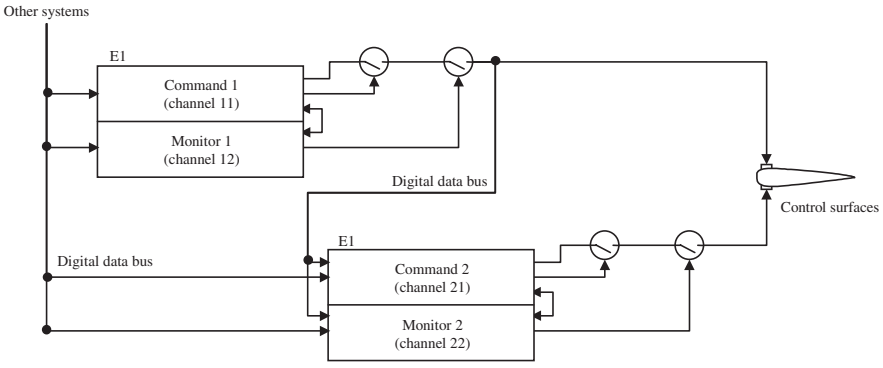


FIGURE 11.93 Dual command/monitor architecture.

to the actuators nor to the other equipment E2. After a confirmation delay, E2 then becomes active. Thus, despite one failure, the whole system remains operational.

Obviously, additional equipments similar to E1 and E2 could be added to achieve any level of reliability. Such an architecture can be found, for instance, in the Airbus A320 flight control system.

N-uplex Architecture. A last common redundant architecture that can be found in a great number of civil aircrafts is a triplex, quadruple, and more generally *N*-uplex configuration. Figure 11.94 depicts the triplex configuration of the Airbus A340 Air Data and Inertial Reference System (ADIRS) computing the state data of the aircraft.

Because of its high critical level, this system is composed of three similar computers (ADIRU) simultaneously active. Consequently, the whole ADIRS system produces three altitudes, three Mach numbers, etc. These data are then compared and selected by each user system, the vote being performed in each system (FMGEC in the considered example). Two comparison monitorings may be used: bit-by-bit comparison and differential comparison.

One of the main differential comparison techniques is median selection requiring a minimum of three channels to tolerate one failure. Another method that is commonly used to develop the control signal in cross-strapped controllers is an averaging cross-feed. In this method, each channel accepts input from all other channels and computes their average which is then used as a signal. Only two channels are required to tolerate a single failure if autonomous false-detection techniques are used. A third channel would allow two failures.

Another way commonly found in military aircraft is bit-by-bit comparison. Contrary to differential comparisons, which do not require perfect agreement between the channels, bit-by-bit is based on the assumption that the outputs are in total agreement except in the case of a fault. In this method, the resulting signal of an

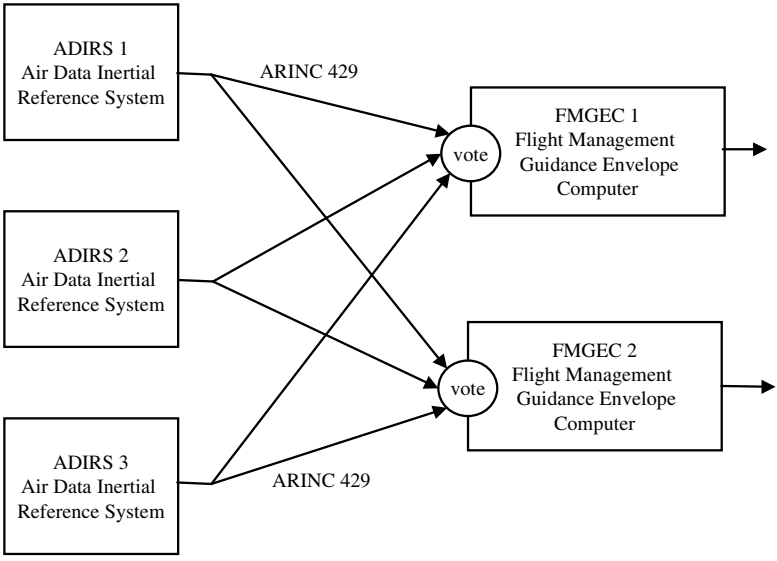


FIGURE 11.94 Typical triplex architecture: the A340 ADIRS architecture.

N-uplex architecture is derived from a majority vote of all channel signals on a bit-by-bit basis. However, to use this method, certain criteria must be satisfied, such as time synchronism. This last constraint prevents the use of bit-by-bit comparison for most civil aircraft. Unlike military aircraft, civil avionics subsystems compute separately in a asynchronous way; parallel equipments do not share a common clock or synchronization mechanism.

Fault-Tolerant Software

Flight-critical systems require fault-tolerant software to complement the fault-tolerant hardware. Many of the concepts for fault-tolerant hardware have parallels in fault-tolerant software. The software designer must use software design concepts as adeptly as hardware concepts. Fault-tolerant software falls into three categories: multiversion programming, recovery blocks, and exception handlers (however, note that all of these techniques are subject to error if the software specification is incorrect).

Multiversion, or *N-version,* programming requires the development of two or more versions of a program that performs a specific function described in the software specification. These versions should be developed by separate software teams and may even be designed to operate on different processors. They accept a common input from an executive-level program, which in turn also compares the results of the different versions to detect faults. For instance, the two-version programming method is used for the Airbus A320 command/monitor equipment development. Channel command and channel monitor are separately developed by separate teams.

Recovery blocks are another concept in fault-tolerant software. Acceptability checks are made on the results from a primary version of a program. If the results fail the acceptability checks, an alternative version of the program different from the primary version is invoked and the process of computation and acceptability checks is repeated. If no version produces an acceptable result, the software block is judged to have failed.

Now that a fault has occurred and been detected, it now must be recovered from. Recovery can be either backward or forward. Backward recovery is exemplified by recovery blocks, where, in the case of a fault, the executive software reinitializes the program using the same input values as used in the previous cycle and attempts to execute the program again. Forward recovery is demonstrated in *N*-version programming, where the outputs are compared and erroneous values generated by faulty software are ignored and only the correct value is passed to the user. Thus, even though a fault may have occurred, recovery has been made without any loss of time.

Another form of fault tolerance in software is *run-time assertions*. Watchdog timers check the time for a block of code to be executed, and if the code is not completed within the prescribed time, an error is assumed to have occurred.

An emerging concept that uses software in lieu of hardware replication to achieve fault tolerance is *analytical redundancy*. In the case of a faulty sensor, analytical redundancy combines data from the remaining functioning sensors with data from other sources in the aircraft in algorithms that compute the most probable value from the failed sensor. This computed value is then used in the same ways as a value from a functioning sensor. An equivalent concept can be applied to flight control actuators and surfaces where, if an element is lost, the remaining functioning elements can be combined in a way to offset the loss. Analytical redundancy and

its companion concept for actuators are two of the cornerstones of reconfigurable flight control systems. Such analytical redundancy techniques have not yet been implemented in civil aircraft but are being examined for future military avionics systems.

Alarm and Maintenance System

Maintenance is a major factor in operating costs and therefore life-cycle costs. For this reason, built-in test equipment (BITE) and centralized maintenance function (CMF) are two main systems which impact the global architecture of the whole avionics. CMF and BITE form a distributed system composed of one or two computers CMCs (centralized maintenance computer) and a BITE function integrated in the LRUs, sensors, and actuators of the aircraft. The aim is first to recognize and correctly identify faulty elements to be changed (at least 90% of the possible faults should be recognized by the BITE components), and secondly to display and store fault data for all BITE-equipped LRUs on the aircraft. For that purpose, each BITE function integrated in each LRU tests its host LRU. In case of internal fault, or if some failure is detected on input data from other systems, the LRU's BITE component generates a message to each CMC indicating the failure. However, in order to circumvent intermittent faults, multiple successive failures over an appropriate time period are required before the BITE declares a unit failed.

In the military context, MIL-STD-2165, *Testability Program for Electronic Systems and Equipment*, requires for the BITE that 95% of all possible critical failures should be detected in less than 1 second and that 100% should be detected in less than 1 minute. Eighty-five percent of all other failures must be detected in less than one minute.

The CMF builds on the capabilities of the BITE to consolidate information on failures from the whole avionics system into a single database along with auxiliary information such as airspeed, altitude, time, and date when the failure occurred.

The A320 CMF, shown in Figure 11.95 is a representative example of BITE and CMF architecture in civil aircraft. It can monitor up to 70 systems, but not all of them directly, and can store the 200 most recent faults or all faults for the last 63 flight legs. The CMF also records universal time coordinated (UTC), flight number and phase, and any cautions and/or warnings that were in effect at the time of the fault. The interface with the CMF is through a multifunction control/display unit (MCDU) in the cockpit. All fault messages are displayed on the MCDU screen in standard formats using plain English, or can be sent to the cockpit printer and to the aircraft communications addressing and reporting system (ACARS) VHF data link. This latter opportunity allows advance notification to maintenance personnel at the flight destination, who can begin preparation to work on the aircraft when it arrives.

11.46 AVIONICS EXAMPLE: THE AIRBUS A320 FLIGHT CONTROL SYSTEM

The first electric flight control system for a civil aircraft was designed by EADS-Airbus and installed on Concorde. This is an analog, full-authority system for all control surfaces. It copies the stick commands onto the control surfaces. A me-

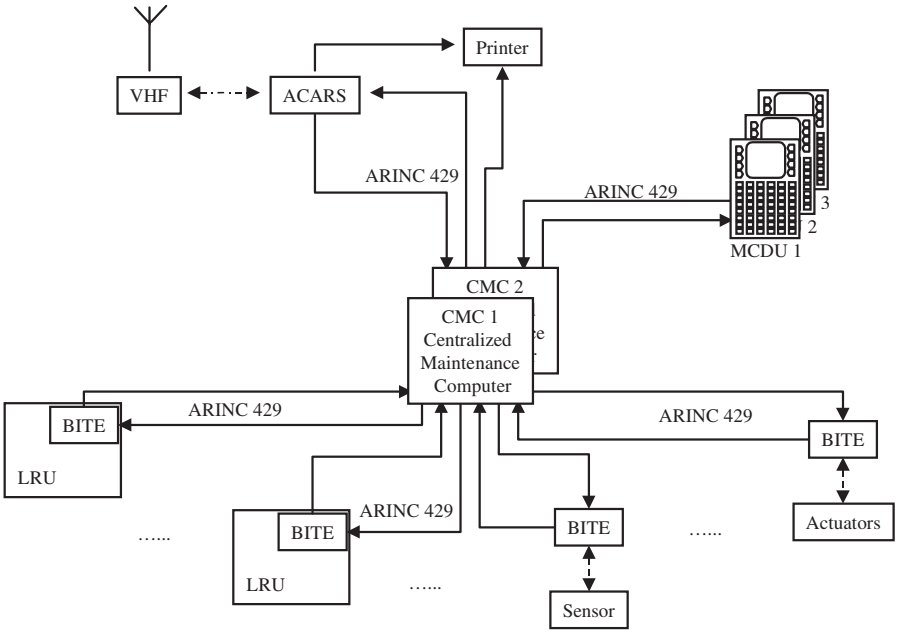


FIGURE 11.95 Airbus A320 centralized maintenance function architecture.

chanical backup system is provided on the three axes. All control surfaces (slats, flaps, and spoilers) are controlled electrically by high-level control laws under all circumstances. The global architecture of the whole system is highly redundant and is based on the fail-passive command/monitoring computers.

Airbus A320 Flight Control System Architecture

Seven computers form the heart of the flight control system: two elevator aileron computers (ELACS), three spoiler elevator computers (SECS), and two flight-augmentation computers (FACS). The allocation of these computers to the flight control surfaces is shown in Figure 11.96.

The redundancy aspect is handled at the system level. The functions of the system are divided between all the computers so that each one is permanently active at least on one subassembly of its functions. For any given function, one computer is active and the others are on standby (“hot spares”). As soon as the active computer interrupts its operation, one of the standby computers almost instantly changes to active mode without a jerk or with a limited jerk on the control surfaces. This architecture ensures that the failure of any one ELAC, SEC, or FAC will not substantially affect the performance of the system.

The ELAC contains a pair (command and monitor part) of MC 68000 processors, each of which has approximately 40,000 words of software.

The SEC contains a pair of Intel 80186 processors, each of which has approximately 40,000 words of software. Dissimilar software is used. The functions of

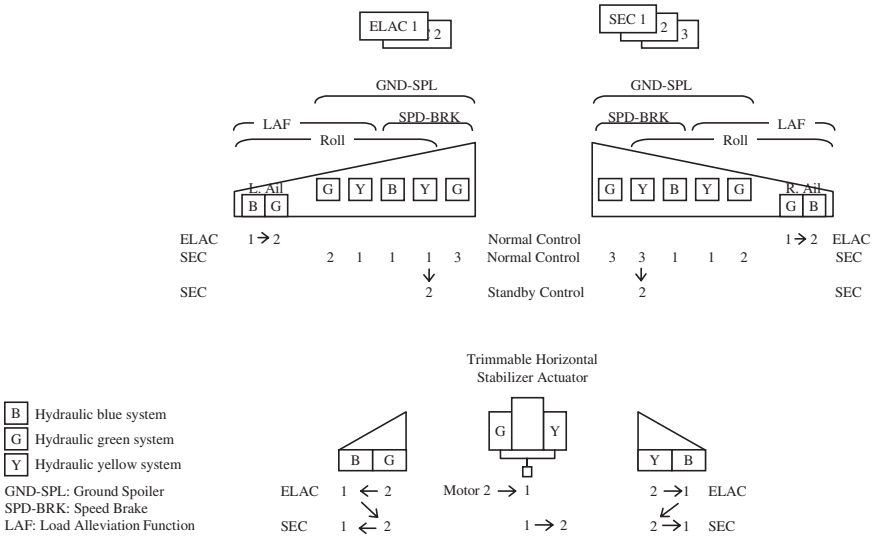


FIGURE 11.96 Airbus A320 flight control system (partial) architecture.

ELACs and SECs are identical except that ELAC control of the spoilers is via the SECs. The basic control law and all protection systems are in both ELACs and SECs. In the case of ELAC failure, the ailerons are no longer active and the aircraft is controlled by the SECs. The ELACs and the SECs are designed and manufactured by separate companies.

In addition to the ELACS and SECs of the A320, two computers are used for rudder control (FAC). However, they are not redundant with the ELACs and SECs. The FACs have the same dual structure as the ELACs and SECs. Each contains 120,000 words of software.

Redundancy Management. The Airbus A320 flight control system is highly critical and must be sufficiently dependable. This requires redundancy, dissimilarity, and reconfiguration management.

The electrical power is normally supplied by two generators, each driven by a different engine. Moreover, an auxiliary generator (APU), batteries, and a ram air turbine (RAT) are available. If the two engines shut down, the RAT is automatically extended. It then pressurizes a hydraulic system which drives a third electrical generator. The FCS computers are connected to at least two power sources.

The aircraft has three hydraulic systems (identified by G [Green], B [Blue], and Y [yellow] in Figure 11.96), one of which is sufficient to control the aircraft. Two systems are pressurized by each engine, the third being pressurized either by an electric pump or by the RAT.

The computers and actuators are also redundant. This is illustrated by the A320 pitch control (left and right elevator, plus trimmable horizontal stabilizer, THS) (see Figure 11.96). Four control and monitoring computers are used, but one is sufficient to control the aircraft. In normal operation, one of the computers (ELAC2) controls the pitch, with one servocontrol pressurized by the green hydraulic for the left elevator, one pressurized by the yellow hydraulic for the right elevator, and by

electric motor 2 for the THS. The other computers control the other control surfaces. If ELAC2 or one of the actuators that it controls fails, ELAC1 takes over (with the servocontrols pressurized by the blue hydraulic for elevators, and with THS motor 1). Following the same failure method, ELAC1 can hand over control to SEC2. Likewise, pitch control can be passed from one SEC to the other depending on the number of control surfaces that one of these computers can handle. These priority orders are pictured by arrows in Figure 11.96. Note that three computers would be sufficient to meet the safety objectives. The additional computer is fully justified by the availability constraints: it is desirable to be able to tolerate a takeoff with one computer having failed.

Functional Reconfiguration. The high redundancy of the system permits hardware or software failures to be tolerated. However, other failures must be considered, such as air data and inertial reference unit (ADIRU) failures.

In the normal case, i.e., if the three ADIRUs are available, the pilot has full authority within a safe flight envelope provided by protections included in the control laws and by addition of protection orders to the pilot orders. Flight control is in g-load factor mode.

If only one ADIRU is available, it is partially monitored by comparison with other independent information sources (in particular, an accelerometer). The safe flight envelope is now provided by warnings, as on a conventional aircraft. Flight control is still in g-load factor mode. If all ADIRUs are lost, the flight envelope protections are also lost and the flight control law is in a degraded mode: direct mode. This law has gains which are a function of the aircraft configuration (the position of the slats and the flaps) and here again allow flight control similar to that of a conventional aircraft.

Airbus A320 Flight Control System Experience

The A320 has accrued around 3.5 million flight hours. During these revenue flights, fault-tolerance mechanisms have been activated, due, for instance, to occurrence of some hardware failures (in a consistent manner with predicted reliability), loss of both ELAC during one flight following an air conditioning failure and the subsequent abnormal temperature. In all cases, failure detection and reconfiguration were successful, including the automatic takeover by the SECs in the last case, which justifies the use of dissimilarity as a fault-tolerant mechanism.

11.47 ENGINEERING OF AVIONICS SYSTEM

The avionics system is a complex process merging engineering activities at the aircraft level, system level, and equipment level. Major activity types are requirements development, solution development, verification, validation, and certification.

Safety Engineering

The development of more and more complex and integrated avionics requires a sound methodology to ensure failure conditions have been adequately addressed.

When dealing with computer-based systems, applicable regulatory requirements are FAR 25.1309/JAR 25, stating:

> The aircraft systems and associated components, considered separately and in relation to other systems, must be designed so that—(1) The occurrence of any failure condition which would prevent the continued safe flight and landing is extremely improbable, and (2) the occurrence of any other failure condition which would reduce the capability of the aircraft or the ability of the crew to cope with adverse operating conditions is improbable.

Risks are characterized by *hazard severity* and *likelihood of occurrence*. An hazard is defined (FAA Order 8040.4) as a "condition, event or circumstance that could lead to or contribute to an unplanned or undesired event." Table 11.17 shows typical severity definitions

For the likelihood of occurrence, qualitative and quantitative definitions do exist. Table 11.18 gives the quantitative definition as a probability of occurrence (PoO) per operational flight hour.

Assessment of risk is made by combining the severity of consequence with the likelihood of occurrence in a table such as Table 11.19.

Failure conditions refer to aircraft/system functions. They denote conditions with effects on the aircraft and its occupants, both direct and consequential, caused or contributed by one or more failures. They carry a severity burden which can be classified and lead to safety objectives in terms of probability requirements. Failure conditions are technology-independent.

TABLE 11.17 Definition of Hazard Severity

Hazard severity	Definition
Catastrophic	Results in multiple fatalities and/or loss of the system
Hazardous	Large reduction in safety margin or functional capacility Crew physical distress/excessive workload Serious or fatal injury to small number of aircraft occupants
Major	Significant reduction in safety margin or functional capacity Significant increase of operator workload Physical distress to occupants of aircraft
Minor	Slight reduction in safety margin or functional capability Slight increase workload Some physical discomfort to aircraft occupants
No safety effect	No effect on safety

TABLE 11.18 Quantitative Definition for Likelihood

Likelihood of occurrence	Quantitative definition
Probable	$PoO > 1.10^{-5}$
Remote	$1.10^{-7} < PoO < 1.10^{-5}$
Extremely remote	$1.10^{-9} < PoO < 1.10^{-7}$
Extremely improbable	$PoO < 1.10^{-9}$

TABLE 11.19 Risk Assessment

Severity likelihood	No safety effect	Minor	Major	Hazardous	Catastrophic
Probable					
Remote					
Extremely remote					
Extremely improbable					

High risk Medium risk Low risk

Failures refer to systems, equipment, and installations. They contribute to occurrences of failure conditions and denote the inability of an item to perform the intended function (loss of function, malfunction of a part of system, etc.).

Faults refer to systems, equipment, or installation components. They may cause or contribute to occurrences of failures. A fault is an undesired anomaly in an item.

Errors refer to systems, equipment, or installation components as affected by, or affecting, human activity. They may cause or contribute to occurrences of faults. With respect to software, mistakes in requirements, design, or code are errors.

A *failure mode* is the way in which the failure of an item occurs.

Safety regulatory requirements are verified in a *certification plan* submitted by the aircraft manufacturer. A major part of it is a functional hazard assessment (FHA) which identifies hazard categories for specific components so that the proper designs and redundancies can be implemented.

The FAA, in cooperation with the Society of Automotive Engineers (SAE), issued a recommended process (see SAE ARP 4754) incorporating the principles of system engineering into the certification process. Conducting a functional analysis is suggested for all levels of aircraft systems. The three major safety analyses are FHA, PSSA, and SSA:

- FHA (functional hazard analysis) allows identification of failure conditions to set safety objectives.
- PSSA (preliminary system safety analysis) and SSA (system safety analysis) are used to identify failure, faults, and errors in the design and realization of the systems.
- While PSSA evaluates a proposed system architecture to determine how failures can lead to the identified failure conditions, SSA evaluates the implemented system.

Many methods are available to perform the safety analyses: fault tree analysis (FTA), dependence diagram (DD), Markov analysis (MA), failure modes and effects analysis (FMEA), failure modes and effects summary (FMES), common cause analysis (CCA), zonal safety analysis (ZFA), particular risk analysis (PRA), and com-

mon mode analysis (CMA). An overview of these methods can be found in a second SAE guidelines document, ARP 4761.

Avionics Software Engineering

Software development should also follow an organized way to make sure it meets the objectives related to its contribution to fault, failures, and finally to failure conditions. Safety analyses are used to discover safety requirements for software.

The DO-178B, issued by RTCA (this document is also referred as EUROCAE/ED-12B), provides guidance for software planning, development, verification, configuration management, quality assurance, certification, and maintenance. It is used to guide the development of flight-critical software for commercial aircrafts. It defines five standardized levels of development assurance for software, as shown in Table 11.20.

Note that software does not fail in a probabilistic way. A system containing a level A software running on a hardware with a 10^{-3} probability of failure per flight hour does not mean that the system has a 10^{-9} probability of failure per flight hour. The assurance levels give information about the development of software, showing how the failure condition classification has been taken into account.

Avionics Systems Engineering Tools

Software tools are used for automating much of the system/software development process: tools for requirement management, system modeling, performance analysis, behavioral simulation, testing, documentation, etc. For software item development, tools exist for coding (like automatic code generation) and debugging tasks.

11.48 FUTURE AVIONICS

The cost of avionics systems is now around 40% of the cost of an aircraft. More and more computers have full authority control. Mechanical backup mechanisms tend to disappear. Avionics systems are intrinsically real-time, and a failure to meet a deadline will impair the operation of the system. Therefore, avionics systems should be robust and safe.

TABLE 11.20 Software Development Assurance Level

Failure condition severity	Software level
Catastrophic	A
Hazardous	B
Major	C
Minor	D
No safety effect	E

Avionics will soon integrate new functions for improving safety, efficiency, availability (air traffic management, airline operations management, cockpit crew operations, cabin crew services, etc.) and also passenger services (such as connection with the external world). In-flight computing will be more and more coupled with ground computing through many operational and passenger networks. The aircraft is becoming a flying complex information system. Information technology is a strategic asset.

ACRONYMS

ACARS	aircraft communications addressing and reporting system
ADIRS	air data and inertial reference system
ADIRU	air data and inertial reference unit
AIMS	airplane information management system
APU	auxiliary power unit
ARINC	Aeronautical Radio Inc.
ARP	aerospace recommended practice
ATA	air transport association
ATMS	air traffic management system
BC	bus controller
BITE	built-in test equipment
BM	bus monitor
CMC	centralized maintenance computer
CMF	centralized maintenance function
EFA	Euro Fighter Aircraft
EFIS	electronic flight instrument system
EIA	Electronics Industry Association
EICAS	engine indicating and crew alerting system
ELAC	elevator aileron computer
FAA	Federal Aviation Administration
FAC	flight augmentation computer
FAR	Federal Aviation Regulation
FHA	functional hazard analysis
FMGEC	flight management guidance envelope computer
IEEE	Institute of Electrical and Electronics Engineering
IMA	integrated modular avionics
JAA	Joint Aviation Authority
JAR	Joint Aviation Requirements
LRM	line replaceable module
LRU	line replaceable unit
MCDU	multiple purpose control display unit

PSSA	preliminary system safety assessment
RAT	ram air turbine
RT	remote terminal
RTCA	Radio Technical Commission for Aeronautics
SAE	Society of Automotive Engineers
SEC	spoiler elevator computer
SSA	system safety assessment
THS	Trimmable horizontal stabilizer
UCS	utilities control system

FURTHER READING

Avionics Library, *Performing a Safety Certification for Avionics Components and Systems,* Avionics Communications Inc., Leesburg, VA (1995).

Avionics Library, *Principles of Avionics Data Buses,* Avionics Communications Inc., Leesburg, VA (1995).

Avionics Library, *Validating Digital Systems in Avionics and Flight Control,* Avionics Communications Inc., Leesburg, VA (1993).

Brière, D. and Traverse, P., ''AIRBUS A320/A330/A340 Electrical Flight Controls, a Family of Fault-Tolerant Systems,'' paper delivered at Fault-Tolerant Computing Symposium, 1993.

Electronic Industries Association, *System Engineering,* EIA-632 (1998).

Fishbein, S. B., *Flight Management Systems—The Evolution of Avionics and Navigation Technology,* Praeger, Westport, CT (1995).

Federal Aviation Administration (FAA), *System Safety Handbook,* FAA, Washington, DC (2000).

Scott J., *Systems Engineering for Commercial Aircraft,* Ashgate, Aldershot, UIT (1997).

Paasch, S., ''Software and the Safety Assessment Process'' FAA, Roston (1999).

RTCA DO-178B, *Software Considerations in Airborne Systems and Equipment Certification,* RTCA, Washington, DC (1992).

Society of Automotive Engineers (SAE), in cooperation with the FAA, Guidelines for the Certification of Highly Integrated and Complex Aircraft Systems, SAE ARP 4754, SAE, Dearborn, MI (1996).

Society of Automotive Engineers (SAE) *Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment,* SAE ARP 4761, SAE, Dearborn, MI (1996).

Spitzer, C., *Digital Avionics Systems,* McGraw-Hill, New York (1993).

PART 11

IN-SPACE COMPUTING

Philippe Guyot

11.49 INTRODUCTION

In-space computing really started in the 1970s with the availability of the very first processors able to survive the space environment.

The need for on-board intelligence comes from the specificity of space systems, for which commandability, operability, and observability are major design drivers.

The increase of on-board capability demand has followed a quasi-Moore's law. In fact, predefined state machine hardwired systems would have become a major show stopper in the early days of space adventure had in-space computing capabilities not effectively started.

As we observe today, on-board processing is a key issue in space system design to such an extent that coming requirements specify flexibility to cope with new mission requirements not defined at the time of the overall software design and lately introduced in the system through the telecommanding path after the launch.

We can find software everywhere and at very different stages of the system. It can be divided into two categories:

1. Ground software:
   - Software which is required to design and manufacture the system (satellite and environment models, design tools, manufacturing tools, etc.)
   - Software used to operate the satellite (system database processes, telecommand and telemetry handling, data archiving and retrieving processes)
2. On-board software:
   - Software embedded in dedicated units to produce specific functions for which hard/soft trade-off has concluded that software introduction would ease design and production and decrease nonrecurring and recurring costs.
   - Software embedded in main computers in charge of operating a large part of the system (spacecraft software—communication or science software)

This Part will focus on the second category.

On-board software is a very critical item not only because the reaction time from ground stations is not sufficient to cover abnormal satellite behavior due to bugs in software design, but also and mainly because it cannot be completely tested on the ground.

These facts necessitate a great number of very dedicated processes to design, produce, and validate overall behavior. To ensure completeness and effectiveness of the validation build-up, stringent quality requirements are imposed on the whole cycle (development logic, software language, observation tools).

The space environment imposes stringent requirements on the hardware in terms of:

- Mechanical constraints during launch
- Thermal constraint caused by frequent hot/cold cycling
- Electromagnetic constraint due to high electrical and magnetic fields, especially in telecommunications satellites
- Radiation constraint due to space particles
- Reliability

A space system must be guaranteed to operate with very high reliability during the complete lifetime of the system, which is as much as 18 years for geostationary missions. Therefore, a specific architecture has to be implemented, including complex redundancy schematics without any single point failure. There is nothing comparable in automotive or aircraft avionics, in which repair is always possible after a specific mission barely longer than a few days.

In addition to the specific redundant single-point failure-free architecture, all the components and processes used in these designs must be qualified to withstand the severe environment (thermal, mechanical, electromagnetic) in orbit. At the parts level, only high-reliability parts can be used. These parts follow a very strict and severe screening program in order to meet the individual reliability requirements. For the report process (soldering, hybrids, etc.), all these elementary processes must prove completely qualified with respect to the mechanothermal cycling they will experience in flight.

In-space computing is now widespread in all functions required for operating the spacecraft (space avionics):

- Electrical power supply (EPS): conditioning, storing, and distributing energy to users. The primary power source is solar arrays, and the secondary source is batteries. The type of the regulation of the power bus is a function of the mission.
- Attitude and orbit control (AOCS): sensing and actuating (force and torque) to position, maintain spacecraft location in orbit, and provide to users (instruments, antenna) a defined attitude.
- Data handling (DHS): acquiring and distributing data between ground and spacecraft, platform management, failure detection, isolation, and recovery, in order to ensure spacecraft autonomy in both nominal scenario and failure cases.

Overview of Avionics Constraints

The two series of constraints to be supported by avionics are driven by mission requirements, ground–spacecraft communication interface, and space environments.

Mission Requirements. The requirements for avionics related to mission constraints are:

- *Power supply:* Eclipse duration and periodicity drive the design of the subsystem.
- *Attitude control:* Mission requirements fall under pointing requirements and orbit operations, including navigation and guidance for deep space missions.
- *Data handling:* Telemetry and telecommand protocols and coding, level of automation on board of different applications (thermal, battery, instruments, communication payload, and autonomy)

Ground–Spacecraft Communication Interface

Earth Orbits

- Low/medium earth orbit (LEO/MEO): For these orbits, observability of the system is not directly achievable throughout the orbit, which mandates the generation of on-board automatic processes to store history information of the system behavior and execute predefined sequences (typical rates > 100 kbps)
- Geostationary orbit (GEO): Telecommand access is available all of the time as the spacecraft remains in a predefined rectangle in the geosynchronous orbit above the ground station. Many of the missions using this orbit are for communications; almost as many are for meterorology. Communications missions require continuous operations in all conditions, which requires stringent requirements on-board failure diagnosis and recovery (typical rates 1 to 10 kbps).
- Deep space orbits: Scientific missions fall under these categories. Constraints combine the previous categories. Deep space missions, due to cost of communication in terms of on-board resources (added to important transmission delay and feedback), require a higher level of autonomy (typical rates <100 bps).

Space Environments

Mechanical Environment. Due to severe mechanical constraints during the launch and deployment phases, the hardware has to be qualified for high sinusoidal, random, and shock levels. This implies some restrictions in use of parts and hybrids and also in the use of specific mounting techniques.

A specific mechanical test campaign must also be conducted in order to prove that the complete unit is able to withstand this level in orbit.

Specific components are very sensitive to mechanical environment levels such as relays and oscillators. These components are frequently used in computer hardware and should be carefully selected, qualified, tested, and screened. Some specific mounting technology should be developed and qualified in order to soften the levels transmitted to the unit and then to the parts.

Thermal Environment. In low orbit, but also in geo-orbit, the satellite and all the units inside experience high thermal gradients and cycling due to solar variations. This causes some thermoelastic constraints and stresses, especially to highly integrated circuits like hybrids and ASICs. The mounting techniques should be carefully selected and qualified in order to withstand these cycles for the entire lifetime without inducing unacceptable stresses. This means that all the techniques cannot be used on flight, especially without a long and costly qualification campaign. The report process is limited with respect to chips or hybrid sizes, number of pins, etc.

Another thermal-related constraint has to do with the reliability requirements. All the parts follow the Arrhenius law and have a degraded reliability figure increasing with average in-orbit temperature. For each specific parameter, derating figures must be applied in order to minimize the stress during the complete lifetime, and these derating factors are directly related to the temperature experienced by the chip or the module during that lifetime. Therefore, a specific heat sink must be implemented in order to reduce the average temperature experienced by these parts. In space, only conductive and small radiative exchanges are possible. No convective processes are used, therefore the thermal control of these parts is much more severe

and difficult than for the equivalent ones on earth. Combined with the mechanical and other thermal stresses, these processes represent a great challenge in achieving a reliable, industrial product, especially concerning high dissipating parts like microprocessors, memory, etc.

Reliability Constraints. The use of qualified and testable parts and processes reduces the availability and complexity of technology that can be used. Not all the available parts on the commercial market can be used, nor the corresponding mounting processes. For instance, the need for powerful microprocessors, memory, and ASICs leads to a huge increase in size and pin number and thus to a change in packaging. Unfortunately, the most promising mounting processes, such as BGA (ball grid array) packages, cannot be used yet due to the lack of qualification and the great difficulty in testing them after the soldering process has been performed. Nevertheless, these technologies need to be developed for space application in order to answer very demanding integration technology, which is one of the major challenges for the forthcoming generation.

Space Radiation Environment. The space radiation environment is composed mainly of three different sources:

Van Allen Belts. Electrons and protons are trapped within the Earth's magnetic field up to a distance of nine earth radii.

Solar Flares. Solar activity is on an 11-year periodic cycle: 7 years of solar maximum and 5 years of solar minimum. During solar maximum activity, very energetic proton solar flares can occur. The most significant recent solar flares were in August 1972, October 1989, and July 2000.

Galactic Cosmic Rays. Galactic cosmic rays are very energetic ions coming from our galaxy, accelerated by supernova explosion. All ion species from hydrogen to uranium are available. Ion fluxes are steady, attenuated by the dynamic solar magnetic field.

Space Radiation Effects on Electronic Devices. Various particle types may induce one specific effect. Table 11.21 shows sources versus effects.

Total Dose Effects. The dose unit used is rads (radiation absorbed dose). 100 rad(Si) is equal to 1 joule deposited within 1 kg of silicon. Typical deposited dose levels in three different orbits are given in Table 11.22.

TABLE 11.21 Space Radiation Sources versus Effects

	Sources			
	Van Allen belts		Solar flares	Cosmic rays
Effects	Electrons	Protons	Protons	Ions
Total dose	X	X	X	
Single event				
Upset		X	X	X
Transient		X	X	X
Latchup		X	X	X
Burnout		X	X	X
Gate Rupture		X	X	X
Displacement Damage	X	X	X	

TABLE 11.22 Total Dose Levels and Total Dose Parts Behavior

Mission	Altitude km	Incl.°	Lifetime years	Dose, level Krad(Si)	Technologies (silicon)	Total dose behavior Krad(Si)
GEO	36,000	0°	18	15	MOS	5–20
LEO	1,300	53°	9	30	BIPOLAR	10–3000
POLAR	800	98°	6	5	Hardened MOS	100–300

GEO: geostationary: LEO: low earth orbit; and POLAR: polar orbit.

Future Trends

Total Dose Effects. Most recent commercial parts have total dose behavior in the range of 10–35 Krad(Si), which is compatible with most current space missions. Because shielding is an issue, no specific problems are expected in the next decade from this effect. Today, mass memory for space applications is widely using commercial 64 and 256 Mbit SDRAMs, with a total dose behavior of 30 Krad(Si).

Single-Event Upset Effects. SEU effects are observed in avionics systems, induced by secondary neutrons produced by cosmic rays interacting with the high-altitude atmosphere. Today, with the shrinking of technologies and increase of digital electronics at ground level (Davari 1996), SEU effects are now an issue. SEU may be induced by radioactive atoms located in the part case and by atmospheric neutrons. Soft error rate is given in FIT (one failure every 10^9 hours) and normalized to the current technology feature size of 0.25 μm. Major parts manufacturers are taking into account this SEU effect at ground level for the next generation of parts, because a factor of 10^3 in soft error rate is expected for the next technology of 0.13 μm.

Electronic systems in the space environment have to be tolerant of various radiation effects. The trend is to use up-to-date parts (that is, parts based on new technologies), that not necessarily hardened. Therefore, hardness assurance activities have to be implemented at the satellite design level:

- Design electrical radiation-tolerant applications.
- Simulate accurately the space radiation stress at the part level in order to optimize mechanical design (shielding).
- Perform very accurate testing in order to get a real estimation of the part degradation in space. Concerning single-event upset effects, the sensitivity at ground level will be a major chance for space applications because commercial parts will be hardened to soft error.

11.50 PHYSICAL ARCHITECTURES OF AVIONICS

Performance, availability, and the operational costs of a spacecraft platform are determined by the avionics system, its sensors, its computer system, and associated software. Fast-evolving capabilities of sensor and real-time computing systems enable the fulfillment of new requirements of the system functions, such as, for telecommunications missions:

- Ground interface via telecommand and telemetry requires autonomy, safety and high data rates (typical: 15 days of autonomy, TC: 10 kbps; TM: up to 750 kbps).
- Attitude and orbit control must ensure attitude pointing error of less than a few hundredths of a degree.
- Thermal regulation complexity follows the increase of mass and complexity of the spacecraft (typically: 600 temperatures and 300 heater lines)
- Power management
- Failure detection and recovery
- Payload control and monitoring

Classical Avionics Architecture

In the early satellite generations (up to Intelsat V and ARABSAT 1), a single unit was in charge of all TM/TC functions (RF section interface, TC decoding and emission to the end users, TM acquisition, multiplexing, and formatting). There was no system data bus. Then the increasing number of parameters to be handled led to the introduction of data buses, with a few terminals connected (two to four remote terminal units). This architecture was applied on most of the platforms.

The increasing payload size (hence the number of connected units) made these remote terminal units much more complex, with detrimental mass/cost impacts. Interconnection harness followed the same trend, leading to heavy bundles and layout problems.

Nowadays the classical avionics architecture is composed of:

- Computer system (one or two computers, one devoted to data handling and the other to attitude control)
- AOCS sensors: sun sensor, gyrometers, earth sensor, star tracker, navigation sensors provided by GPS, GLONASS, and Galileo constellations
- AOCS actuators: momentum and/or reaction wheels, thruster, and magneto torquers
- Power distribution and interface units which ensure temperature acquisition
- Power distribution to the heater, pyrotechnics devices, discrete ON/OFF commands
- The electrical power management unit, which ensures battery charge/discharge and primary power bus regulation

Integrated Modular Avionics

IMA is a recommended approach to optimize the avionics mass and cost. It was used extensively in American military programs (Pave Pillar) in the 1980s.* A similar program was initiated in Europe with EUCLID.†

*The Pave Pillar program, by the U.S. Air Force, created generic conceptual architecture, completed in 1989. Aircraft using this standard include the F-15, F-22, and RAH-66.

†EUCLID (European Cooperation for the Long Term in Defence) was started in 1989 to develop a modular integrated avionics standard. No real standard is yet available in Europe.

The IMA approach consists of developing a set of functional blocks which can be used for different functions or subsystems. The goal is to minimize the development cost through a top-down allocation of the functional needs to certain modules.

The main modules are:

- Data processor module or computer core, which must reflect the system architectural approach: centralized or distributed architecture. The data processor module includes on-board communication services and on-board time management. In addition to the data processor module, the computer core can include telemetry and telecommand interface management, mass memory, etc.
- Power supply module, which provides regulated voltage to the units from the primary power bus.
- Discrete input/output modules, which provide command/acquisition functions

These modules are standard hardware and software components which can be integrated into different units according to the avionics architecture and system needs.

Functional modules are connected together via standard interfaces available in the backbone of the units: power supply, communication bus, and synchronization signals.

Computer Subsystem

The computer is the brain of an avionics system. Modern integrated electronics and powerful processors have enabled the designer to choose from three fundamental types of architecture: centralized, federated, and distributed.

Centralized Computer. Architecture is characterized by data processing of all the avionics functions fitting in a single computer unit, which simplifies redundancy management and equipment validation. Centralized architecture allows for function interfacing minimization and communication data rate optimization. On-board software (OBSW) development is ensured or coordinated by one team, which ensures the complete SW validation. The main disadvantages of centralized architecture are that data bus management becomes more complex with each new added function, and OBSW modification and maintenance are more complex because a complete nonregression test must be performed for each new version.

A centralized computer has been used in space programs like APOLLO, Ariane 1 and Ariane 5 launch vehicle. New-generation telecommunications satellites rely on such architecture thanks to a new generation of powerful processors.

Federated Computer. Federated computer architecture involves each major subsystem, such as AOCS or DHS, sharing input or output (telecommands, temperatures) and their computed results over data buses. This approach enables several teams to develop different subsystems in parallel. It was used in the 1980s in space programs like Spacebus telecommunications satellites (one AOCS and one DHS computer) and Atlas 2AS (one flight control computer and one propulsion control computer).

Distributed Computer Subsystem. The architecture of the distributed computer subsystem is characterized by multiple processors enabled by the recurrent cost of processor and memory parts for space applications. Each processor is dedicated to more and more sophisticated data processing and allows for generic approach and customization approach of complex sensors (star trackers, Earth sensors, inertial and gyroscopic measurement unit, etc.) and of the mission management computer (payload mission control plan).

The advantages of the distributed architecture comes from the simplification of the data bus management and associated software in each processor and the capability of faster program execution.

The main disadvantage is related to the avionics validation process that must be carried out on a fully representative avionics set (commonly called ''iron bird'' in aeronautics). A space station is a typical application of distributed architecture.

Such an architecture requires a standardization of interfaces to ensure modularity and versatility and allow future extension. This is achieved mainly by means of communication buses. Two solutions are possible for that purpose:

- Proprietary standard is expensive but guarantees feasibility and perennially.
- Open standard benefits from developments in higher-volume markets (PC world, PCI, VME, VXI, etc.).

Emergence of Commercial Standards. Space budget reductions have pushed industry to adopt a ''components off-the-shelf'' (COTS) approach for recent programs. This approach allows HW development to be minimized through intensive use of commercial standards: power PC processor and VxWorks operating system, PCI or VME bus, etc.

For new systems, functional density continues to increase. This means higher performance within a smaller circuit card or chip area for lower cost (system on chip).

Use of complex COTS (e.g., Power PC and VxWorks) poses a significant challenge for system designers. Certification for critical application requires additional validation. System development and validation schedules are longer than the COTS product life.

Recent experiences with better, faster, cheaper programs based on COTS component have been mixed: Mars Pathfinder was a great success for its two-week mission, but problems experienced by IRIDIUM have shown the difficulty of applying commercial standards to space applications.

Processors. The processor is the key element in the development of space software. Performances have moved in the last decade from below 1 MIPS (0.4 MIPS SPOT/ARIANE 1-4) to above 10 MIPS and even more in architecture making use of commercial parts. This trend reflects a constant increase in performance. The next limitation in sight is in memory transfer time, which could become, at least for a while, a bottleneck in the evolution to higher performance. In terms of needs, if no major mission requirements require a significant increase in processor performance, introduction of new language and automatic code generation will require more power than is available today.

Processor technologies have moved from CISC architecture to RISC (reduced instruction set computer) and from 8 bit to 16 and then 32 bit registers. It is then possible to fit high-performance processors inside reduced-performance (in number of gates, speed, power) ASICs.

Example of Existing Processors. A few processors were developed on the base of the U.S. MIL-STD-1750 instruction set standard:

- MA31750 (GEC Plessey)
- PACE 1750 A (Performance SC)

RISC processors are now available based on SPARC. In Europe, processors based on the SPARC instruction set include:

- ERC 32 (Temic) (single chip will be qualified Q4 99)
- FLAME (Honeywell/Dassault Electronique)

and in the United States, the RH 3000 (Honeywell).

SPARC Architecture

Solutions using COTS and coming from the world of PCs include:

- Intel 80486 (Intel + SEI) Power PC processor
- Power PC603 (Motorola)
- Power PC603 (Thomson)
- Power PC microcontroller (Spar)

Dedicated architecture is devoted to signal processing, in particular the 21020 proposed by TEMIC as the ADSP 21020.

Other rad-hard processors are available but with a limited access because the development is being carried out through U.S. military programs.

The current trend leads us to consider the processor as a library of functions which is validated (intellectual property, IP) and to merge this library with other functions dedicated to a specific usage. It is then possible to get the processor and its peripherals onto a single chip, limiting new developments and costs: this concept is called system-on-a-chip (SOC).

The definition of a standard interface is the major problem to be solved in the IMA approach, which relies on data buses (avionics data communication). Data buses are the nervous system of an avionics system. Multiplexing the transmission and reception of multiple signals over a common path is one of the cornerstones of an integrated digital avionics system. It allows the sharing of data and computation results, ensuring that all connected subsystems are using a consistent set of data while reducing the weight of the wiring. A bus provides means of communication between several units. It is based on a number of communication layers (Figure 11.97), from the physical layer to the application layer.

Most bus architectures (Figure 11.98) are based on a single controller managing all exchanges on the bus, but buses exists that extend performance to decentralized architectures, allowing each processing unit to control the bus when requested (although the bus is not used by another controller). These are often based on ring architectures (FIP).

Two decades of experience with data buses have resulted in the commonly used data bus standards such as MIL-STD-1553, ESA OBDH bus, and Aeronautics Radio Inc. (ARINC) specification 429.

In parallel, development studies of miniature remote terminals have been conducted by industry and space agencies. They have opened the way to the latest

AVIONICS AND ASTRIONICS

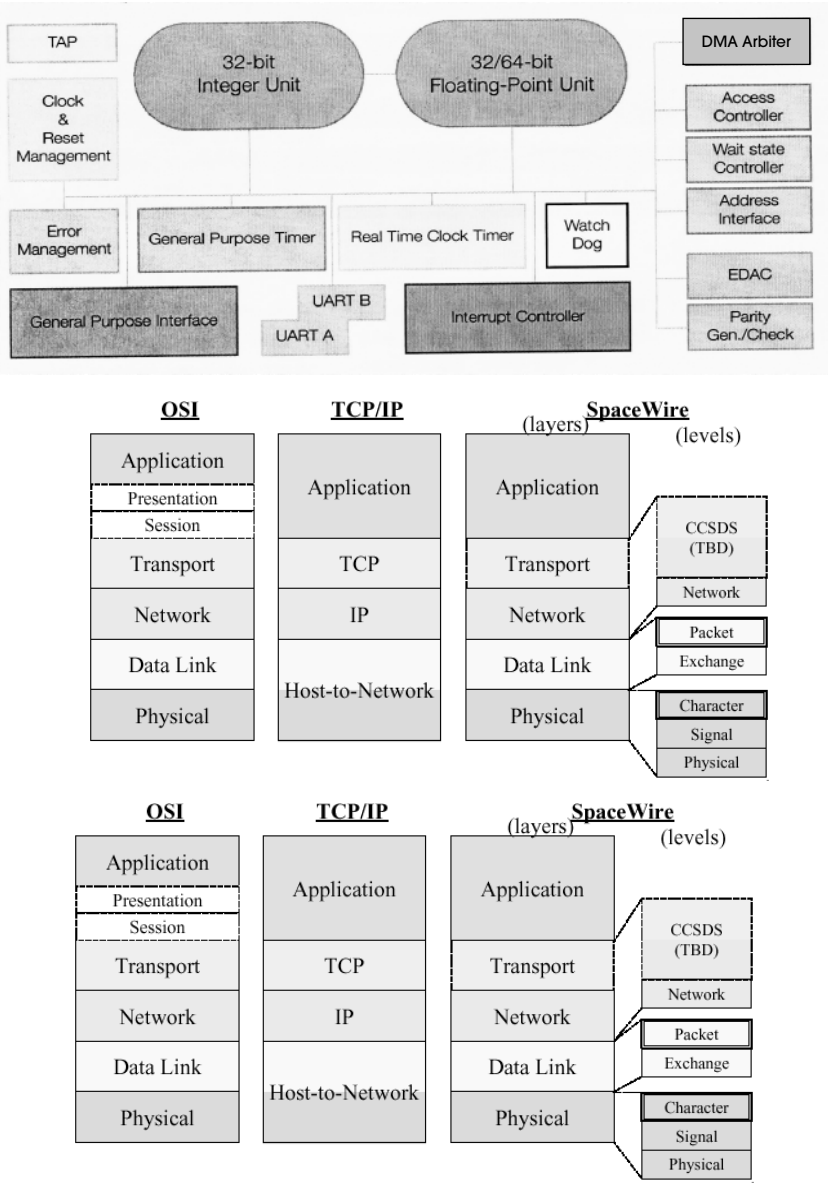


FIGURE 11.97 Communication layers (SpaceWire).

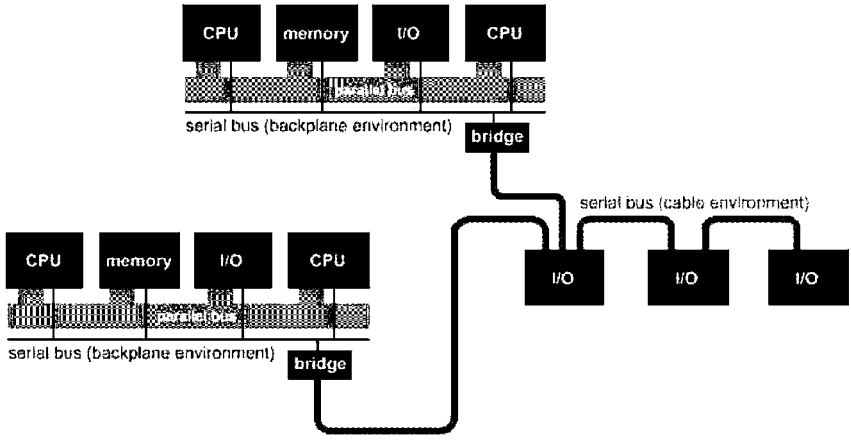


FIGURE 11.98 Bus architecture (IEEE 1394).

stage of remote terminal generation, taking advantage of improved technologies for higher integration. The terminals developed for the new platform are housed in a single small hybrid, which is incorporated in each active TM/TC user equipment. End-to-end connections are limited to passive or very simple units, while systematic use of data buses minimizes the need for dedicated terminal units and makes interconnection harness simple and standard.

Advanced concepts make use of:

- Communication data bus (RS485 and 1553B) for all units having several analog TM and/or serial communication interface and / or several TM statuses.
- No cross-strap at harness level to avoid extra I/O pin density at RTU connector level. The impact on reliability of this concept is in considering connector pins, wires, and eventually a part of the interface electronics as contributing to the user unit reliability.
- Use of matrix command and matrix acquisition for all on/off commands, relay commands, power on/off status, and relay status.

Standard communications protocol mapped on heterogeneous bus networks may come to be composed of classical standards like 1553 and emerging standards like CAN bus, used extensively in automobiles.

11.51 ON-BOARD SOFTWARE

Main Functions

The on-board software (OBSW) or flight software ensures data processing and data management related to mission management, spacecraft control, and payload management.

- Mission management handles autonomous phase transitions that must be engaged on different conditions: on-board time, spacecraft position, and telecommands sent by the ground segment.
- Spacecraft control includes attitude and orbit control, data handling functions like telecommand processing, telemetry acquisition and formatting, active thermal regulation, and battery management.
- Payload management provides high-level services to set up the payload in the configuration required for the mission with respect to the mission plan stored in the on-board computer memory. For example, it is possible to define a mission plan which will power-on an observation instrument just before a flyby of an asteroid.

The OBSW follows the evolution of on-board computing capabilities. The new generation of embedded computers can provide more than 10 MIPS and several megabytes of memory. This allows new functions to be introduced in the on-board SW, such as failure detection and recovery, on-orbit propagator, etc.

Furthermore, several functions can be refined to optimize subsystem performance, including automatic calibration and thermal regulation thresholds selected depending the mission phase.

In summary, OBSW is becoming more and more flexible to allow for recent adaptations to avionics units and spacecraft configurations.

Design Constraints

Flight software is a major critical component of avionics and must be error free. Any remaining bug can lead to interruption of a mission or loss of a spacecraft. Development and validation processes are based on very strict methods, which tend to prevent design errors and guarantee complete test coverage.

Although validation effort is intensive, error-free OBSW cannot be guaranteed before launch, and OBSW design provides several options to allow recovery of the mission in case of error:

- In case of major failure, the central computer is restarted with the backup OBSW, which provides the minimum subset of the OBSW functions required to engage the safe mode.
- Any piece of the OBSW can be dumped and patched to allow corrections or workaround solutions.
- OBSW maintenance must be possible during the full mission duration (up to 18 years in telecom satellites).

OBSW Architecture

Software Composition. To minimize the impact of software evolution, the design is based on an object approach, which allows complete modularity and is structured in a multilayered way to allow for parallel development. The different layers of the software can be defined in several ways. Essentials are:

- To create an application layer as independent as possible from target processor and real-time constraints coming from the processor board interfaces
- To organize a layer providing all the hardware/software-dependent routines (handlers)
- To use a real-time kernel

One example of a multilayer concept is depicted in Figure 11.99.

The major technical advantages of this decomposition are:

- The I/O drivers are defined as a set of elementary routines to interface the processor board.
- The I/O services are defined with regard to applications software constraints. This software layer is:
  - In charge of introducing for each I/O service the real-time aspects, such as preemption, protection, rendezvous, and so on and is in charge of offering high-level services
  - Based on the I/O drivers and real-time kernel services

Software Languages

Before the 1970s, COBOL, FORTRAN, and Assembler were the only available languages. Writing a program in Assembler is a task for an expert, the program is not easy to maintain. COBOL is well adapted for programs dealing with data management (banks, etc.) and FORTRAN for scientific computation. These languages do not include the concept of encapsulation or abstraction.

With the development of coding activities, new languages emerged (C in 1972 and Ada in 1983). These are generalist languages with a first level of abstraction. They have allowed the development of object-oriented design methods.

Subsequently, new languages such as C++, Ada95, and the last-generation Java have benefited from experience in previous language usage and allow high levels of abstraction and encapsulation. Moreover, Java introduced two new concepts: the virtual machine and the applet. A Java-compiled program will not be recompiled if the processor on which it should run is changed, but only the virtual machine that should become standard on any workstation.

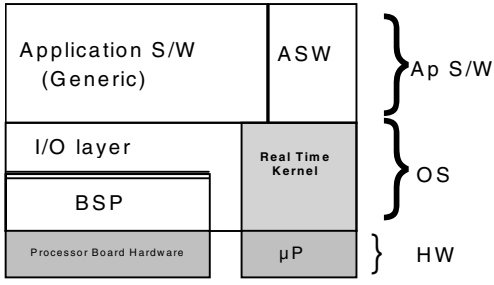


FIGURE 11.99 On-board software composition.

An applet is an application on a central server that can be downloaded across a network onto client stations in order to be executed on a local client. This download is done without intervention of a local client station user. This concept was the basis for the deployment of the Internet.

In the future, commercial interpreters will enable direct download during operations in orbit, validation being limited to the applet itself.

Real-Time Kernel

The basic reason for introducing schedulers is the constant increase in size with the increased risk and development time that implies.

The necessary following step is to ensure, from an existing design, a plug-and-play logic for any newly introduced application.

For embedded software, several schedulers have been developed, such as COTS. Ada runtimes include VxWorks, LynxOS, RTOS, T SMART, and RAVEN. Their characteristics include:

- Size of 180 kilobytes.
- Exhaustive validation cannot be guaranteed.
- Long-term maintenance (>15 years) cannot be guaranteed.

Moreover, these schedulers have different complementary features (TCP/IP access, serial line access, etc.) and can be targeted on different processors (PowerPC, ERC32, 68K, Intel, etc.). But because long-term maintenance of a COTS is not always guaranteed, its deterministic behavior is not always guaranteed, and size and CPU performance are not compatible with needs, major space companies have developed a privately owned scheduler (real-time kernel, RTK). Today, the RTK providers try to develop minimal real-time kernel compatibles with the constraints of safety-critical embedded software, such as Raven.

Alarm/Interrupt Processing

Embedded systems are subject to real-time constraints because they interact with their environment. If in some cases it is enough to use a simple polled loop of input signals, in more complex embedded systems, when asynchronous events are numerous or if many periodic actions are necessary, processing must use interrupts coming from devices (IO controllers, timers).

This evolution has led software developers to use real-time operating systems (RTOS). Several in-house operating systems were developed first, and in the early 1980s several commercial products were available on the market.

At first, these operating systems were restricted to the kernel functionalities, i.e., a task scheduler, generally preemptive, and a memory handler.

The spectrum of schedulers ranges from a simple cyclic executive to the many full-featured priority-based preemptive schedulers and to even more sophisticated schedulers. The choice of scheduling approach is related to the class of space applications.

In safety-critical applications, it is important that the execution system be predictable at all times.

Schedulability analysis is performed in order to verify analytically that the timing constraints applied to real-time applications are attainable. Analyses are based on techniques determined from fixed-priority scheduling theory.

The goal of this activity is to verify that all tasks are schedulable and so to guarantee that deadlines can be satisfied. Figure 11.100 gives an example of a space low earth orbit application.

BITE Function

Built-in test (BITE) functions have been introduced due to the complexity of onboard electronics, necessity of standardization, and performance increase of space electronics with regard to their operational use. This practice comes, in fact, from aeronautic applications, where maintenance is a real issue and where direct human intervention is possible during the operational lifetime of the system.

For space applications, which by definition allow limited capability for ground intervention and are limited to operations through the telecommand path, BITEs were limited to specific activities such as booting the computer, the other tasks being part of the failure detection, isolation, and recovery functions.

During boot operations, the system is generally copying the executable code from PROMs/EEPROMs to RAM for access time end power constraints. In order to prevent wrong software loading, the boot contains several dedicated tests to check correctness of the read-and-write process in the RAM areas.

The need to generalize BITE practices has come from the introduction of constellations, where on ground production required quick determination of the behavior of the different units on board without interfering with the hardware and software configuration.

Failure Detection, Isolation, and Recovery (FDIR) Functions

Most space systems require autonomy requirements either to protect the system against feared events or to minimize duration of outages (in the case of commercial applications). For instance, typical specifications for telecommunications satellites are:

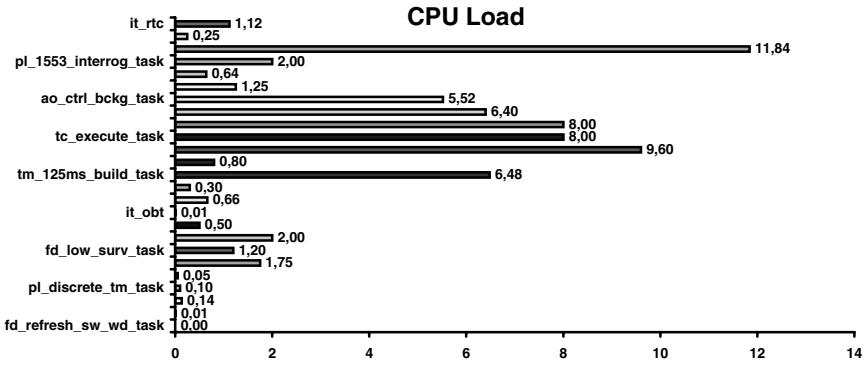


FIGURE 11.100 CPU load.

- Full operational performances shall be ensured from 14 and up to 28 days without ground contact and without failure.
- The satellite shall survive 72 hours without ground contact in case of failure.

Whatever the system design is, several systematic ways of approaching FDIR remain:

1. Object oriented view: A unit is monitored by checking the status of dedicated parameters (analog or digital) as a function of the assumed state. Through this process, a unit can be declared unhealthy and isolated from the rest of the spacecraft by switching it off, and full operation can be resumed if a unit of the same type can be used (redundant).
2. Functional view: The operational output of units that characterized the assumed status of the system can be monitored. For instance, observing the rotational rates around the satellite axis given by the gyroscope gives information about the effective capability of the spacecraft to control its attitude.
3. Computer-specific FDIR: As it runs the software related to FDIR, the specific FDIR computer needs specific control mechanisms to protect the system from its own errors. The basic idea is a watchdog mechanism; particularities as a function of the mission requirements are laid in the parameters used to trigger the watchdog.
4. System: Object and functional views do not cover all the failure cases, especially those related to combined hardware/software failures. It is therefore interesting to find hard-wired global data indicating that the system is healthy. For instance, hard-wired Earth presence for telecommunication missions is a very good candidate for ensuring that the system is in bad shape if this signal turns to Earth presence lost.

A hierarchy is defined associating level of FDIR treatment and probability of feared event (see Figure 11.101).

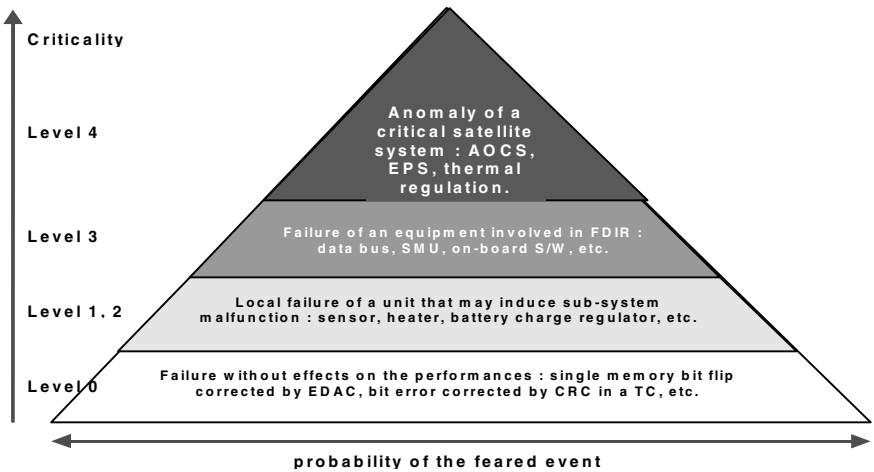


FIGURE 11.101 FDIR hierarchy.

Progressive fault contention is looked to minimize outage and ensure failure isolation.

FDIR Design Method. The preferred method is a top-down analysis of the feared events defined at the system level:

FE-1: satellite body depointing
FE-2: erroneous orbit
FE-3: curtailment of mission
FE-4: loss of power
FE-5: thermal degradation
FE-6: TM/TC link loss
FE-7: loss of RF mission
FE-8: launcher damage

For each feared event, a fault tree analysis is performed to identify elementary functional failures (see Figure 11.102).

In the 1990s artificial intelligence used to provide a fully autonomous system was considered a very promising technique. The complexity of the anomalies observed in flight, attributed to system software engineering, has moved the track to simpler state machine techniques describing unambiguously the possibilities of reacting in case of failure declaration.

Nowadays, tools exist to animate these FDIR representations by state machine providing very early in the system design feedback on assumed behavior of the system combining logical state transition and real timing constraints.

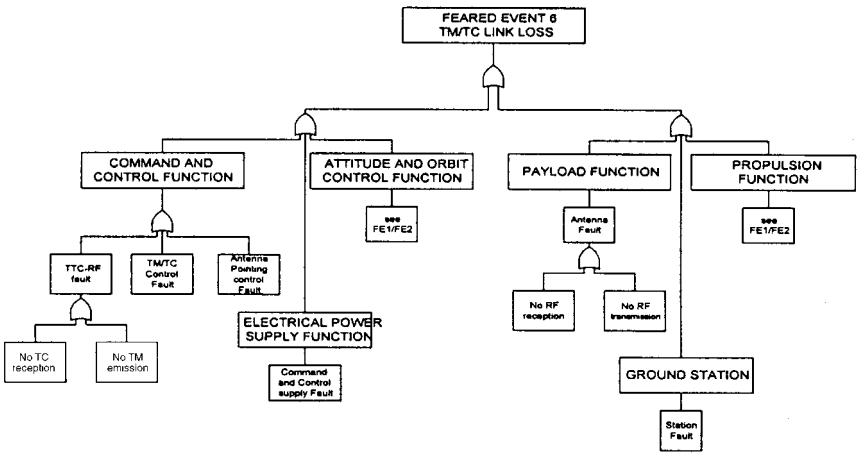


FIGURE 11.102 Fault tree.

OBSW Complexity and Sizing Assessment

OBSW complexity and sizing are required very soon in system development to justify the compatibility of the OBSW requirement with the computer system performance and assess the OBSW development costs.

OBSW sizing is based on a combination of methods: modeling, prototyping, and interpolation from previous similar applications.

Typical allocations are shown in Table 11.23. This allocation table must be refined and confirmed throughout the development process. In case of oversizing, simplifications or optimizations have to be made.

Real memory size can be assessed thanks to metrics verified by representative benchmarks. These metrics provide the number of bytes for one line of code. This ratio can vary significantly with the processor and the compiler: one line of Ada code requires 8 bytes on the 1750 A processor (16 bits), 12 bytes on MC 68020 (CISC 32 bits), 18 bytes on SPARC (RISC 32 bits).

11.52 DEVELOPMENT OF AVIONICS SYSTEM

General Approach

The classical approach is based on the well-known process in V, which can be presented in the scheme shown in Figure 11.103. This approach guarantees traceability of the evolution and guarantees formal appropriation of the OBSW function by the OBSW development team.

Driving Requirements

Many of the requirements which influence the avionics system design are not directly related to in-flight performance. Overall performance requirements such as recurrent cost, mass, planning, power consumption, safety, maintenance, mean time

TABLE 11.23 Allocations

Major functions	Lines of code	Period	CPU load
AOCS	16,000	100 ms	20%
AOCS sensor management	8,000		1%
Thermal regulation	1,500	30 s	1%
Battery management	2,500	4 s	1%
Payload management	5,000		1%
Platform unit management	1,500		1%
Telemetry	3,500	500 ms	10%
Telecommand processing	1,500	200 ms	1%
Failure detection and recovery	3,000	100 ms	5%
OS and communication and protocols	10,000	Asynchronous	15%
Total	52,500		<60%

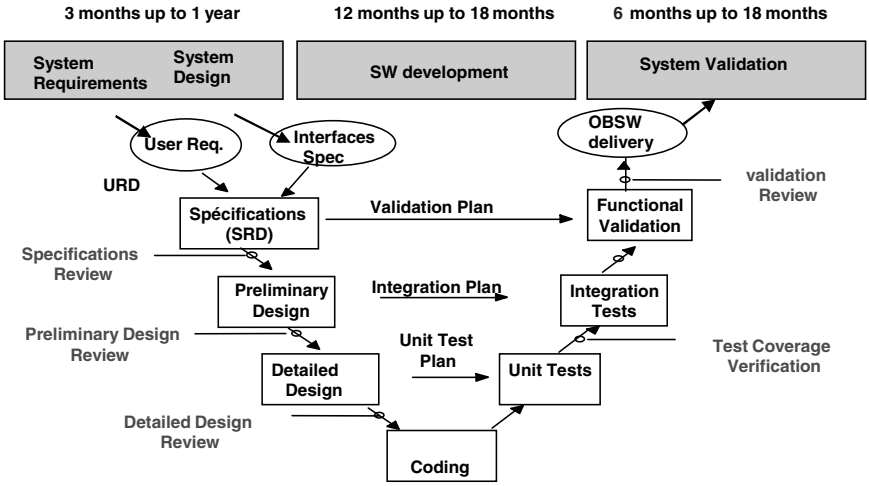


FIGURE 11.103 Avionics development.

to failure (MTTF), observability, and operational constraints must be taken into account in the development logic.

Classical Development Logic

Avionics design and development logic is based on the following steps:

- The mission analysis is used to define the avionics subsystem requirements, including mission profiles, functional data flow diagrams, and function allocation.
- The avionics architecture design phase defines the avionics architecture to support avionics functions requirements and interfaces between the different subsystems.
- On-board software development.
- Avionics integration and testing.

Avionics development is commonly based on a decomposition of the avionics in three subsystems:

- Electrical power subsystem (EPS)
- Attitude and orbit control subsystem (AOCS)
- Data handling subsystem (DHS)

In a classical approach, each subsystem is developed independently by different subcontractors and interfaces between them are minimized to reduce the development risks. In such an approach, the global design (recurrent cost and mass) cannot be optimized. Some functions have to be duplicated in the three subsystems. Each subsystem requires telecommand processing, telemetry acquisitions, and failure detection and recovery, which requires an on-board computer. In order to complete AOCS validation at the subsystem level, a dynamic bench test is used.

AOCS Development

On board software-critical applications are mainly those related to determining and controlling the attitude of the spacecraft throughout its lifetime and the command-and-control processes. The example given here is an application for geostationary missions. It provides information on this type of on-board synchronous applications and insight into one of the most difficult issues to solve regarding software development and validation of components.

The described function is designed to provide a three-axis stabilization for the transfer maneuvers as well as for the operational life in geostationary orbit.

Hardware Architecture. Attitude determination is mainly based on a star tracker (STR) able to deliver three-axis attitude measurements by using several stars' position information in its field of view, and an Earth sensor, providing roll and pitch spacecraft attitude angles directly with respect to the local orbital frame.

The STR is a ''star light in–quaternion out'' device. This means the equipment itself is able to perform star recognition by comparison of the star seen in its field of view with its star catalog. From this comparison, it autonomously computes and provides the quaternion representing the attitude of the equipment in a given inertial reference frame (J2000).

To get the attitude with respect to the local orbital frame, an on-board orbit propagator is implemented and provides the local orbital frame orientation with respect to the inertial frame.

Coarse Sun sensors and gyrometers are also used for rate damping and Sun acquisition. Sun acquisition is engaged after launcher separation and is used as a waiting mode between consecutive maneuvers in transfer orbit or as a safe mode in transfer orbit and geostationary orbit.

The actuators are basically a set of four reaction wheels, used without momentum bias, thus providing high angular momentum storage capability. Angular momentum bias is no longer necessary thanks to the permanent three-axis measurement provided by the STR.

AOCS commands the propulsion subsystems (chemical or electrical) in order to achieve the following goals:

- Perform attitude control torque when wheels are not used or when their control torque is not sufficient
- Maintain the wheel speed in the nominal range
- Perform the commanded velocity increments for orbit raising and station keeping

When electrical propulsion is used, electric thrusters are mounted on thrusters' orientation mechanisms. The orientation mechanisms are then also controlled by AOCS in order to perform wheel speed control.

Functional Architecture. AOCS is organized in functional modes. A dedicated mission and a corresponding set of equipment usages characterize a mode.

A simplified typical sequence of events in case of GTO transfer, involving the relevant modes, is shown in Figure 11.104.

The mode functionality and required equipment are summarized in Table 11.24:

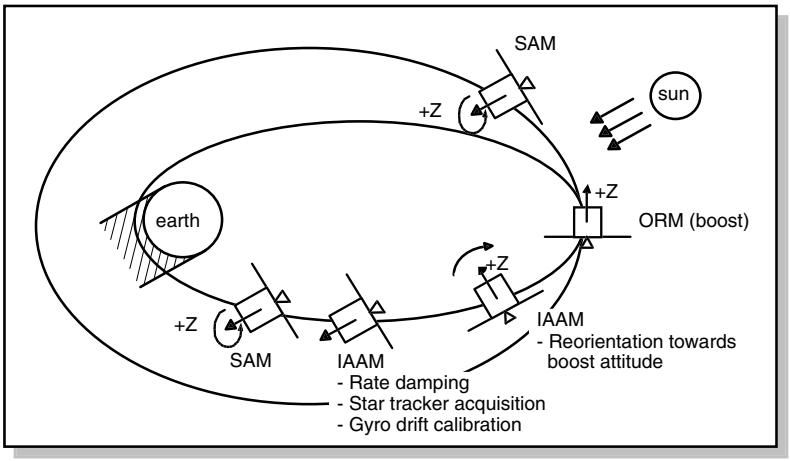


FIGURE 11.104 Sequence of events/modes used in transfer orbit.

TABLE 11.24 AOCS Mode

Mode	Function
SBM	AOCS initial configuration after separation
SAM	Sun acquisition mode Sun search and Sun pointing
IAAM	Inertial attitude acquisition mode Star acquistion and satellite pointing
ORM	Orbit raising mode. Injection into geostationary orbit
EPM	Earth pointing mode Earth pointing and transition to NM
NM	Normal mode and orbit corrections (E/W and N/S) Highly accurate satellite and antenna pointing
SKM	Station-keeping mode for final station acquistion and drift phase or satellite relocation (not used with payload ON)

Design and Validation Logic

Early Design Phases. In the early design phases, the attitude control studies are performed using tools such as Matlab/Simulink, Xmath/Systembuild, and Scilab/Scicos. During this phase, the spacecraft dynamics are modeled either with a simulink library developed internally or with simple rigid-body models. The sensors and actuators are modeled with very simple models, generally a perfect sensor with Gaussian noise and bias. This design simulator allows the performing of initial trade-offs to prevalidate the different operating mode concepts with preliminary control law design and linear stability analyses. The simulator is afterwards constantly updated during the conception and validation phases. It is used to prevalidate

the design modifications, to perform the controller gain adjustments for each satellite and also for all stability analyses.

Advanced Design Phases: Conception and Validation. During and after the advanced design phases, a numerical dynamic study simulator is developed in a high-level language. This study simulator is developed based on the internally produced library. This library implements several facilities to model the spacecraft dynamics (rigid bodies, sloshes, flexible modes, etc.), the orbital environment, and the sensors and actuators and provides powerful services scheduling and simulation management (I/O, event sequence, etc.).

The goal of this simulator is to provide the designers with the most accurate simulation of the spacecraft in order to validate functionally and numerically the performance and robustness of the attitude control system. For this purpose, some of the sensors and actuators models are provided directly by the equipment supplier.

An important feature of this simulator is that it allows the designer to act on actuators and sensors delay to check phase margins and global robustness. In this simulator, the attitude control designers implement all the control laws and functions used in the different attitude control modes of the satellites. This phase of conception is iterative and incremental. The design reports of the different modes and on-board functions are constantly improved in accordance with the simulator until a compliant design has been found. Once a compliant design has been found, it is described in user requirement definition documents. These documents are intended to the software development team for on-board software development.

The final validation of the design is obtained by performing several temporal simulations in nominal and degraded conditions. This allows the performance of the satellite and its robustness to be checked.

These modes and functions are implemented in order to be numerically and functionally representative of what will finally be implemented in the real satellite. This part of the simulator is called the *simulated on-board software*. To achieve optimal numerical representativity, it is developed using two dedicated on-board software libraries. One implements basic mathematical functions such as trigonometric function, matrices, vectors, and quaternion manipulation. The other implement basic functions dedicated to attitude control, such as filters, deadband, and attitude determination. These libraries can be updated during the design and conception phases according to the needs of the designers.

The user requirements documents describe the attitude control without any numerical algorithm description. They describe only state charts and calls to these libraries. Thus, the attitude control laws architecture and data flows in the simulated on-board software may be slightly different from the final on-board software, but all numerical treatments are identical. This way it is guaranteed that the robustness simulations performed with the study simulator truly validate the design of the attitude control.

To improve the level of confidence in this validation by numerical simulation, study simulator has to be instrumented with numerical analysis tools. Once instrumented, the simulated on-board software performs all its floating-point numerical treatments with stochastic arithmetic. This allows the designer to determine the relative accuracy of the different floating-point variables in the data flow. It is particularly useful to isolate sudden loss of accuracy caused by misconditioned computations. It allows the designers to numerically optimize their design and to use double precision only where it is really mandatory.

This phase functionally and numerically validates the design. The next step is to validate the implementation of this design on the on-board target and its interfaces with the rest of the spacecraft. For this purpose, the simulation tools team provides the software team with reference runs at different steps of the development. Reference runs are provided to check first the numerical nonregression of mathematical and elementary AOCS functions libraries between the simulation environment and the target environment. At the next step, reference runs are provided to check the good implementation of all mathematical libraries in the development of the different AOCS functions (e.g., controller attitude determination, wheel management). These tests are open-loop tests.

Note: the FDIR (failure detection and recovery, associated timing, etc.) is designed by state chart (Statemate tool) with respect to the URD in order to validate its behavior and simulate test scenarios. The more representative test scenarios are executed on the avionics test bench during the functional validation.

In parallel, a worst-case analysis is performed to determine what are the worst entry conditions in the different operating modes and FDIR can lead to. The study simulator is then used to check that these entry conditions will always lead to successful achievement of the emergency mode.

Software Development

Life Cycle. Before software development begins, a preliminary step is mandatory, where the users of the computer-based system define their needs in terms of requirements and describe the environment (hardware and software) in which the software will execute. Then the software follows a development life cycle that is divided into general phases, starting with the specification of the software requirements and with the software validation phase. The classical representation of this cycle is the V model, the successive phases of which are explained below.

The software requirements phase provides the detailed specifications of the software, as complete and error-free as possible. In this phase, the test strategy is studied (verification and validation activities to be done throughout the development)

The objective of the architectural design phase is to define the software architecture (static and dynamic) that will satisfy the requirements. In the meantime, the software integration is planed and prepared.

The detailed design phase, based on the software decomposition defined in the previous phase, consists of refining the software design (algorithms, data structures).

The aim of the coding and unit-testing phase is to produce the code and test it at the lowest breakdown level (the unit).

During the software integration phase, the software is progressively integrated from the unit level until the overall software and each level of integration is tested.

The objective of the software validation phase is to perform tests on the whole software to ensure that the software fulfills its requirements.

Finally, an acceptance phase demonstrates that the software fulfills the user's requirements. For embedded systems, this acceptance phase is preceded by hardware and software integration at the system level.

For space embedded software, which requires a strong quality level throughout the development activities, each phase ends with a formal review or an internal key point that states the results of the activity and allows the next phase to be started.

Precise software documentation is required for all activities (requirements, design, code, tests, etc.) and must be kept consistent during the development.

The configuration management of the software items is one of the key factors in successful development. This includes tracing precisely the progressive software production from requirements to design and validation tests, taking into account the changes that can affect the product.

Following the software acceptance phase, the maintenance phase takes place, its purpose being to ensure that the product continues to meet the real needs of the end user.

The development cycle described previously is somewhat theoretical, and the reality is too complex to be covered by the V model. In fact, the actual software development models often used in the space projects, and derived from the V model, are:

- The incremental model: Delivery of successive versions of the software, each of them offering a subset of validated main functions, allowing early step of verification at the system level. This model involves, for each release, rigorous non-regression verifications of the software.
- The evolutive model: Unlike the incremental model, the evolutive model produces successive versions of software, which address subsets of the detailed requirements (more and more, until the complete product is reached). This model allows the software development to be performed even if all the requirements are not available during the first steps of the project. For this model, emphasis must be put on the software configuration management.

In practice, a mix of these two models is applied in order to answer the problem of late requirements availability against early need for the executable software asked by the system team.

Method and Tools. Space software constitutes a critical component of the whole space system:

- Normal operation for long duration (up to 20 years)
- Limited maintenance capability (access by telecommand)
- Unacceptable consequences for software design errors
- Limited test coverage on ground (space conditions not fully simulated on the spacecraft)

Consequently, all the activities taking place during the software development life cycle must be supported by methods and tools contributing to the delivery of a high-quality software product.

The presentation of a typical environment used for embedded space software development is described below. During the software requirement phase, coherent, clear, and unambiguous items of all the fields of requirements must be produced. For this purpose, well-known methods based on functional decomposition, such as SADT and SART, are widely used, supported by software tools (e.g., Teamwork) which offer efficient graphical interface, central data repository, consistency checks on data flow and control flow of the requirements representation.

However, today the challenge of iterative developments in a highly concurrent design process with delivery time constraints and growing complexity of the soft-

ware, has led to the adoption of methods based on object-oriented analysis and design. Mainly based on the UML notation, these methods (supported by tools like Rhapsody, Tau, and Rose) allow the behavior of the software by means of various representations (use case diagram, class diagram, state transition diagram), then reducing the risks on development.

Whatever the method, tools (e.g., DOORS) producing traceability between user requirements documents and software requirements documents play an important part in this process.

The architectural design phase may be carried out following a functional approach (supported by tools like Teamwork) or with an object-oriented approach (e.g., HOOD method, initiated by the European Space Agency at the beginning of the 1980s, supported by tools, e.g., STOOD).

In the case of methods based on UML notation used for requirement analysis, the software architecture is designed by refinement of the software modelization in continuity with the requirement phase.

For the detailed design phase activity, which applies at the software unit level, it is usual to associate closely the detail design description (pseudo-code, internal data exhaustive description, etc.) to the source code itself, generated by using text editors (or syntactic editors, able to check the correctness of use of the programming language).

The code production step makes use of text editors, high-level language (Ada, C) native and cross-compilers (targeting the actual processor used in the flight computer, e.g., SPARC ERC32), linker, assembler, and libraries (run time routines, mathematical functions) in order to obtain an executable code, free from compilation errors and ready for unit tests. Otherwise, up-to-date object-oriented design environments provide the possibility of automatic code generation, which can lead to significant reduction in the coding and unit tests effort.

Before the unit tests are performed, the inspection of the source code takes place, in which proper application of the programming rules is checked, supported by tools like Logiscope RuleChecker.

The unit tests, which are based on the execution of the code, may be run in different environments according to the objectives of the tests:

- Workstation (native environment) for first functional debugging
- Processor simulator as close as possible to the embedded code

The tools needed for this test activity are:

- Symbolic debugger (e.g., AdaProbe, gdb) able to give information on the steps of execution, values of software variables, state and values of the processor hardware resources.
- Test tool (e.g., ATTOL UniTest) able to generate and execute test programs, then analyze execution data and produce test reports.
- Test coverage tool (e.g., Logiscope Testchecker, ATTOL Coverage) that can produce coverage reports (instructions, decisions paths, etc.) related to a set of unit tests execution (with code instrumentation for that purpose)

Typically, the unit test process includes first the functional testing of the software units (‘‘black box’’), and then the tests are completed by structural tests (‘‘white box’’) in order to reach the test coverage level required (this level depends on the

criticality of the software). Other verifications performed on the software units consist of static measures of quality features like complexity level (supported by tools like Logiscope Audit).

The integration test of the software has two parts:

1. Software/software integration, the objective of which is to build major functions of the software. This activity involves the tools already used during the unit test step, complemented by specific test programs, and relies on progressive substitution of the stubs (i.e., simulation of the missing software units) by the actual tested units.
2. The hardware/software integration, which lead progressively (i.e., basic software + service layer + each function of the application layer) to the overall software, running in real time on the target computer, in a simulated system environment. This activity needs to make use of:
   - A breadboard functionally representative of the flight computer, with connections to debug and analysis devices (e.g., logical analyzer, debugging probe)
   - A set of observation and trace and debug tools
   - Test equipment able to simulate and acquire in real time the data exchanged with the system interfaces (simplified simulation of the other equipment, allowing open-loop tests)

When the integration phase is achieved, the software can be delivered to the independent validation team for the validation phase.

All the technical activities described above are supplemented by transverse activities (project management, software product assurance, configuration management, change control, documentation management). These activities follow specific methods and rules, are supported by adequate tools (e.g., ClearCase for software configuration management), and contribute to the production of maintainable high-quality embedded software.

11.53 ADVANCED DEVELOPMENT METHODS

Because time to market and development cost are key drivers for commercial applications and system complexity is permanently increasing, new ways of development have been investigated.

Emerging Approach: Automatic Code Generation

Increasing complexity requires system engineers to rely on modeling techniques. The goals of these methods are to help system designers verify the consistency of the functions to be ensured by the avionics (H/W and S/W). The goal is to describe the system behavior required by the system designer.

Use of modern tools such as unified modeling tools allows the complete development process to be supported. First, system design is described in an architecture model which will be provided to the OBSW development team. Architecture design is used by the system design team to specify the validation test procedures. The ''architecture model'' is then refined by the OBSW team to define a OBSW design

model which will be used to support the detailed design, coding, and OBSW validation. SW may be partially generated automatically from the OBSW design model.

Automatic code generation is intensively used in the industrial world. Airbus Industries makes intensive use of SCADE tools (Telelogic) for SW embedded in airplanes, and mobile phone companies are making use of Rhapsody for the new generation of mobile phones.

Application of such tools is in progress for critical on-board software embedded on board a satellite and will be a technical challenge for the coming years.

System Analysis

The main objective is to find a way of producing structured information that can be designed and later on produced and tested but can also be animated to verify very early in the process that users will be satisfied by the way the requirements have been met. This has induced the use of simulation tools very closed to the specification language to give confidence that the proofs obtained by simulation tools are compatible with what is effectively stated in the requirement documents.

In the analysis phase, the studies are performed using modeling techniques such as UML. This language implies the definition of a new analysis engineering process. The major principles of the process are:

The process is iterative and incremental: it consists of successive iterations. The system is analyzed during the first iteration. It is then decomposed into subelements, and each element is analyzed during the following iterations, and so on until the entire functional need is covered by the product of the iteration (system, segment, or equipment, for instance).

The analysis is based on the architecture. The main architectural components are already known (system composed of segments or subsystems) at the beginning of the analysis.

The process facilitates concurrent work by the system engineering teams. In the first time, system engineers can work in parallel on specific elements to produce elements modeling. Then those element models are used to consolidate the system model.

The approach is top-down for specific elements but flexible in terms of starting point: because the decomposition is already known for the highest levels, the analysis of a lower level element can start in parallel with system analysis.

The new process has three main phases:

- Requirement analysis
- Preliminary design
- System model packaging

Requirement Analysis

This phase is used to understand customer needs and allows the functional decomposition of the system to be defined. The UML approach describes the functions expected by the users and proposes an operational presentation of the functionality based mainly on a use case diagram (Figure 11.105). Use cases have semantics very similar to the functional requirements definition.

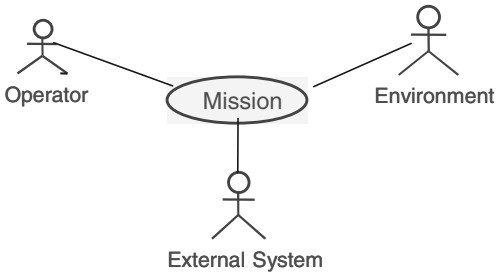


FIGURE 11.105 Use case diagram.

One of the main points of the analysis is to collect the requirements, structure them, and identify where they are mapped on the engineering models.

Preliminary Design

The purpose of preliminary design is to define the logical and dynamic aspects of the element under analysis. The logical aspects represent the object type (classes), contents, and static relationship. The dynamic aspects represent the object dynamics. The logical and dynamic analyses of an element are interlaced and are performed together.

Preliminary design is supported by the following activities:

- Logical modeling, which is illustrated by class diagrams
- Dynamic modeling, which is illustrated by sequence diagrams (Figure 11.106) (exchanges between classes) and state diagrams (Figure 11.107) (internal class behavior)

System Model Packaging

Packaging activities support coordination of system engineering activities. The architecture will structure the UML items into packages. For each iteration, the engineering teams build a reference model related to the system model. The official change of reference model version is made under the system engineering responsible agreement.

Development and Production Phase

Efforts have been made in the past 10 years to structure the different phases of software development based on object-oriented techniques and associated tools.

Validation Phase

Validation of complex systems working in a very specific and dynamic environment, with all on-board operations done in an automatic way, is a crucial issue.

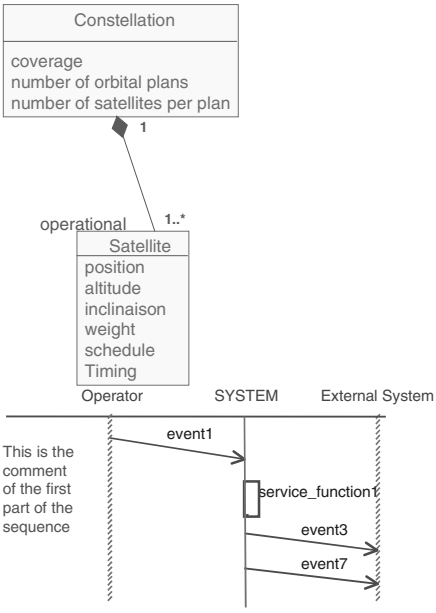


FIGURE 11.106 Sequence diagram.

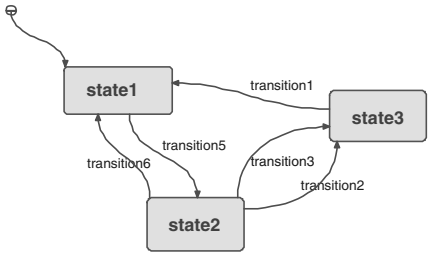


FIGURE 11.107 State diagram.

Software Validation Phase

The following validation measures are:

- Starting the development of test procedures earlier to have 50 to 70% of the test available at the beginning of validation phase.
- Having two sets of test equipment available. That has resulted in a schedule reduction of 40% for this phase.
- Specifying a functional test language which allows the syntax and the semantics of test procedures (precompiler) to be checked and to have test procedures independent to reference language of test equipment (layer approach as for onboard software).

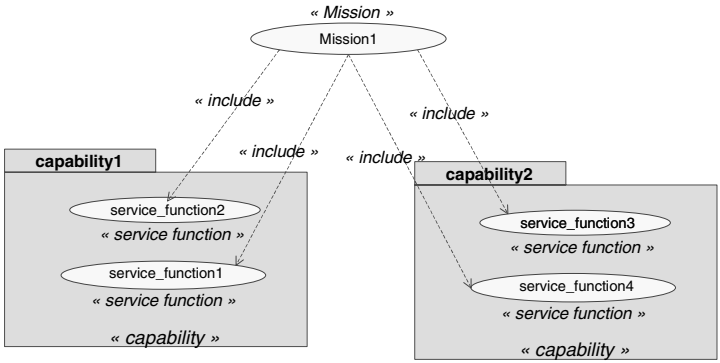


FIGURE 11.108

- Formalizing the test specification syntax so that a test procedures skeleton can be developed automatically.

Functional Validation on Hardware Test Benches

The final validation phase is the functional validation on hardware avionics test benches. The final on-board software is implemented on the target and tested in a hardware environment representative of the avionics implementation in the satellites. This means that all interfaces and wiring are identical or at least representative and several hardware functional or qualification models are present in the loop.

To achieve closed-loop simulation, the avionics test bench must provide a dynamic simulation of the environment and of absent hardware. In the particular case of attitude control, this dynamic simulation is obtained by migration of all the study simulator except the simulated on-board software in the real-time environment of the test bench.

To comply with this real-time environment and several functional validation requirements, the study simulator is slightly modified. In particular, the sensor and actuator models are enriched with functional aspects of the equipment, whereas the study simulator focuses on the performance of the model, and interfaces are added to allow direct interaction with electrical ground support equipment (EGSE) and buses.

This real-time simulator is used for closed-loop tests of the avionics with and without real equipment in the loop. The final validation of the implementation of the attitude control is obtained by comparison of reference runs performed with the study simulator and the simulated on-board software with tests cases performed on the test benches.

All these simulation tools are developed under the version management system.

Satellite Validation

The software and functional validation is just a part of the global satellite validation. To reduce the global satellite time schedule, is necessary to start the other steps of validation in parallel with the software validation.

The global satellite validation is decomposed into four independent validation steps:

1. On-board software validation on a computer representative model in a simulated environment
2. Functional chain validation on an avionics representative chain on a table (computer representative model, plus engineering model of AOCS equipment)
3. Assembly, integration, and test phases on satellite model
4. Ground segment validation

Due to concurrent engineering practice, these validation steps can be completely performed in parallel.

To allow the progress of each validation, computer representative model and on-board software are needed. Therefore, early in the development, an adaptation (test mode) and subsets of on-board software are defined for all these validation needs.

11.54 FUTURE AVIONICS FUNCTIONS

It seems that the complexity of software development and validation is now becoming the driving factor in terms of cost and planning, while the hardware platform, in terms of processing power and memory, will not be an obstacle in the years to come. This new fact will certainly drive a new type of hardware architecture, in which the electronics will be highly integrated, in small modules including processors, memory, and bus drivers and glue logic. Each of these modules will be equivalent to today's computer hardware and therefore a higher reliability scheme could be envisaged, perhaps allowing the severity of parts screening to be decreased, and therefore the associated cost. These modules could be interconnected by an internal high-speed standard serial bus (like Firewire or equivalent), which has the great advantage of minimizing cross-strap problems with respect to parallel buses. Theses internal buses would be connected to medium-rate external buses (medium-rate due to the number of users and length of the cables). The type of bus that could be used on board can range from optical link to RF links passing through data buses using power buses as support. Early studies have shown that the last possibility (using power buses as support) seems to be the most promising in terms of cost savings at the system level and also minimizing the AIT tasks.

This will also be driven by the progress of integrated technology in terms of integrated circuits, ASICs, and hybrids, but also in terms of report processes and PCBs. Memory multimodules (MCMs) are very promising despite the problem of thermal and reliability drawbacks yet to be solved.

In terms of software, the complexity and cost/schedule impacts come from the high flexibility of this subsystem, allowing very late freezing of specifications and usually very late changes in the same specifications. The high flexibility offered by the embedded software is not only an advantage but also a drawback in terms of schedule and cost. Whether the more demanding requirements on software subsystems are worthwhile in terms of cost and development strategy must be investigated.

Another option could be to use ASICs for more functions in order to oblige to freeze very early in the design process the specifications, the software being reduced to the high demanding part of flexibility of too complex part of state machines.

This could lead to substantial cost savings in the future, in which several hundred ASICs are much less expensive than several hundred software versions.

The future will tell if the hardware will return as a cost-effective solution with respect to highly flexible, but very costly, software solutions.

REFERENCES

Barth, Janet L. 1997. "Modeling Space Radiation Environment," IEEE 1997 Nuclear and Space Radiation Effects Conference, Short Course notebook, section II.

Cohen, N., Sriram, T. S., Leland, N., Moyer, D., Butler, S., and Fatley, R. 1999. "Soft Error Considerations for Deep-Submicronic CMOS Circuit Applications," *IEDM Technical Digest 1999,* pp. 315–318.

Dyer, C. S. 1998. "Space Radiation Environment Dosimetry," IEEE 1998 Nuclear and Space Radiation Effects Conference, Short Course notebook, section II.

FURTHER READING

Davari, B., "CMOS Technology Scaling, 0.1 μm and Beyond," *IEDM Technical Digest 1996,* pp. 555–558.

Department of Defense, *Ada Reference Manual,* ANSI/MIL-STD-1815 A, February 1983.

Dressendorfer, P., "Basic Mechanisms for the New Millennium," IEEE 1998 Nuclear and Space Radiation Effects Conference, Short Course notebook, section III (1998).

Fleetwood, D. M., "A First Principle Approach to Total Dose Hardness Assurance," IIIE 1995 Nuclear and Space Radiation Effects Conference, Short Course notebook, section III (1995).

HOOD Working Group, *HOOD Reference Manual,* WME/89-173/JB Issue 3.0, September 1989.

Larson, W. J., and Wertz, J. R., *Space Mission Analysis and Design,* 2d ed., Microcosm, Torrance, CA, and Kluwer Academic, Dordrecht (1999).

Petersen, E. L., "Single Event Analysis and Prediction," IEEE 1997 Nuclear and Space Radiation Effets Conference, Short Course notebook, section I (1997).

SECTION 12
AIRCRAFT SYSTEMS

Dieter Scholz

12.1 INTRODUCTION

Aircraft Systems—General

What Are Aircraft Systems? Broadly speaking, an aircraft can be subdivided into three categories:

1. The airframe (the aircraft structure)
2. The power plant (the engines)
3. The aircraft systems (the equipment)

This Section will deal with the last of these categories.

The airframe provides the aircraft with its (relative) rigidity. It also enables the generation of lift through its aerodynamic shape. A glider flies without a power plant, but in order to maintain weather-independent sustained level flight, a power plant is necessary to produce thrust to overcome the drag.

The airframe and power plant might seem to be all that is needed, but this is not so. Even the earliest aircraft needed more. Some means to steer the aircraft (flight controls) and to handle it on the ground (landing gear) were needed. These aircraft systems play a key role today and must be considered in the very early stages of aircraft design. A fuel system was also needed from the beginning of the history of powered flight. With aircraft flying longer distances, navigation and communication systems became important; with aircraft flying higher and taking passengers on board, cabin systems such as air conditioning and oxygen systems were introduced.

The above gives a general idea of what aircraft systems are. A more rigorous definition of the term will be given below.

Significance of Aircraft Systems. Aircraft systems account for one-third of the aircraft's empty mass. Aircraft systems have a high economic impact: more than one-third of the development and production costs of a medium-range civil transport craft can be allocated to aircraft systems, and this ratio can be even higher for military aircraft. The price of the aircraft is driven in the same proportion by aircraft

systems. Aircraft systems account for roughly one-third of the direct operating costs (DOC) and direct maintenance costs (DMC).

Historical Trends. Aircraft silhouettes and general design concepts have been stable since the 1960s. Nevertheless, remarkable progress has been made since that time. Just as aerodynamics, structures, and power plants have been optimized, aircraft systems have been gradually improved in economics, reliability, and safety. This has been made possible by constant evolution and optimization through in-service experience, research, and development and by employment of new technologies.

Probably the most important factor in the changes has been made by digital data processing. Today computers are part of almost every aircraft system in larger aircraft. Computers also play a key role in the design and manufacturing process of aircraft systems. The evolution of aircraft systems has not come to an end yet. Modern achievements in computer technology will continue to make their way into aircraft.

Striving for improved safety, economics, and passenger comfort will demand even more sophisticated technologies and complexity. The airlines have been reluctant to accept the ever-increasing complexity, since it does not make troubleshooting the aircraft any easier. The aviation industry has taken the approach that technology has to buy its way onto the aircraft—i.e., only if new technologies can prove their overall benefit will they be considered in new aircraft design.

The separate tasks of the structure, the engines, and the systems are being more and more integrated to handle the tasks together. Here are some examples:

- Electronic flight control systems stabilize a fighter aircraft with an unstable layout or stabilize aircraft structural or rigid body modes.
- A gust load alleviation system as part of the flight control systems helps reduce the design loads for the wing structure.
- A highly reliable yaw damper system enables the aircraft to be built with a fin smaller than would otherwise be required.
- Engine parameters are changed in accordance with air conditioning demands.

To achieve an overall optimum in aircraft design, it is no longer possible to look at the structure, the engines, and the aircraft systems separately. Today's challenge lies in optimizing the aircraft as a whole by means of multidisciplinary design optimization (MDO).

The Industry. Aircraft systems are defined by the aircraft manufacturer. This commonly takes place in joint teams with engineers from specialized subcontractors. The subcontractors work on the final design, manufacture the system or component, and deliver their parts to the aircraft manufacturer's final assembly line. The trend is for aircraft manufacturers to select major subcontractors who are made responsible for designing and manufacturing a complete aircraft system. These subcontractors may even become risk-sharing partners in the aircraft program. Aircrafts are maintained by dedicated maintenance organizations. Maintenance is done on and off aircraft. Off-aircraft maintenance is performed on aircraft components in specialized shops.

Scope of This Section. Section 12 provides background information and describes the general principles of transport category aircraft systems. The Airbus A321 (Figure 12.2) from the family of Airbus narrow-body aircraft is used to provide an

example of the systems under discussion. *At no time should the information given be used for actual aircraft operation or maintenance. The information given is intended for familiarization and training purposes only.* Space in this handbook is too limited for all aircraft systems to be covered in depth. For some aircraft systems only the definition is given and the reader is referred to other parts of the handbook that also deal with the subject. For other aircraft systems the definition is given together with selected views on the Airbus A321. Emphasis is put on selected major mechanical aircraft systems. The References and Further Reading show the way to actual design work and detailed studies.

Definitions

The term *system* is frequently used in engineering sciences. In thermodynamics, for example, a system is characterized by its defined boundary. The definition of the term with respect to aircraft is more specific.

The World Airlines Technical Operations Glossary (WATOG) defines:

- *System:* A combination of inter-related items arranged to perform a specific function
- *Subsystem:* A major functional portion of a system, which contributes to operational completeness of the system

The WATOG also gives an example together with further subdivisions of the system and subsystem:

- *System:* auxiliary power unit
- *Subsystem:* power generator
- *Component:* fuel control unit
- *Subassembly:* valve
- *Part:* seal

Note that these definitions refer to civil aircraft. With respect to military aircraft, instead of *aircraft systems* the term is *aircraft subsystems*. In the example above, the auxiliary power unit hence would be considered a subsystem.

In dealing with aircraft systems, all categories of aircrafts need to be considered. ICAO defines:

- *Aircraft:* Any machine that can derive support in the atmosphere from the reaction of the air (ICAO Annex 2)
- *Aircraft category:* Classification of aircraft according to specified basic characteristics, e.g., aeroplane, glider, rotorcraft, free balloon (ICAO Annex 1)

Combining the above definitions, a definition for aircraft systems might be:

- *Aircraft system:* A combination of interrelated items arranged to perform a specific function on an aircraft

This section deals with aircraft systems in powered heavier-than-air aircraft. Although aircraft systems in gliders, rotorcrafts, and free balloons have to take into account the specifics of their respective categories, they are not fundamentally different from aircraft systems in aeroplanes.

Breakdown

Aircraft systems are distinguished by function. It is common practice in civil aviation to group aircraft systems according to Specification 100 of the Air Transport Association of America (ATA) (ATA 100), which thoroughly structures aircraft documentation. According to ATA 100,[1] aircraft equipment is identified by an equipment identifier consisting of three elements of two digits each. The identifier 29-31-03 points to system 29, subsystem 31, and unit 03. The aircraft systems—or, in ATA terms, *airframe systems*—are listed in Table 12.1 together with their system identifiers. It is common practice to refer to just the system identifier *ATA 28,* instead of to the "fuel system." Furthermore, *Chapter 28* is often referred to, because that is the chapter allocated to the fuel system in any aircraft documentation showing ATA conformity.

Autopilot, communications, navigation, and indicating/recording systems (ATA 22, 23, 34, 31 [, 44, 45, 46]) are electronic systems, known in aviation as *avionic systems,* and are characterized by processing information (compare with SAE 1998).

Other systems provide fuel, power, and essential comfort to crew and passengers. These nonavionic systems are the *general* or *utility systems.* Today there is an increase in the number of electronic control units within the utility systems; nev-

TABLE 12.1 Aircraft Systems[a] (ATA 100)

Identifier	Name of system
21	air conditioning
22	auto flight
23	communications
24	electrical power
25	equipment/furnishings
26	fire protection
27	flight controls
28	fuel
29	hydraulic power
30	ice and rain protection
31	indicating/recording systems
32	landing gear
33	lights
34	navigation
35	oxygen
36	pneumatic
38	water/waste
49	airborne auxiliary power

[a]Not included in this table are Chapters 37, 41, 45, and 46 from ATA 100, which are not of relevance here. Also not included here are new Chapters 44 and 50 from ATA 2200.

[1]Recently ATA 100 became part of the new ATA 2200. ATA 2200 has introduced minor changes and updates to the definitions of aircraft systems. This text uses the well-established ATA 100 and presents differences to ATA 2200 in footnotes.

ertheless, the primary purpose of these systems remains some kind of energy transfer (Moir and Seabridge 2001)

Secondary power systems include the nonpropulsive power generation and transmission. They include electrical power, hydraulic power, pneumatic, and auxiliary power (SAE 1998) (ATA 24, 29, 36, 49). Secondary power systems provide power to other aircraft systems.

The *environmental control system* (ECS) is an engineering system that maintains the immediate environment of an organism within defined limits of temperature, pressure, and gaseous composition suitable for continuance of comfort and efficiency (AGARD 1980). The air conditioning system and oxygen system (ATA 21, 35) are assigned these tasks.

Other aircraft systems are grouped and assigned a specific name often without a formal definition.

Hydraulic systems comprise all systems that apply hydraulic power. In general, these are hydraulic power, flight controls, and landing gear (ATA 29, 27, 32).

Electric systems comprise all systems that apply electric power. In general, these are electric power (ATA 24) and all systems with major electrical consumers. Electrical systems are characterized by electrical power generation, distribution, and consumption and have to be distinguished from avionic systems.

Pneumatic systems comprise all systems that apply pneumatic power. In general, these are pneumatic and other systems with pneumatic components (ATA 36, 21, 30).

Cabin systems[2] comprise all systems with an impact on the cabin of the aircraft and hence with an influence on the passenger (ATA 21, 25, 35, 38, and partially 23, 26, 31, 33).

These groupings depend to a certain extent on the system technologies applied in the aircraft being considered.

Certification

After one or several prototype aircraft are designed and manufactured, they go through a series of *certification tests* in order to show compliance with the *certification requirements*. Compliance with the requirements may be shown by analysis, ground, or flight test, depending on the requirements or negotiations with the *aviation administration*. System tests are a substantial part of the certification program. In Europe, certification of large aeroplanes is based on the Joint Aviation Requirements (JAR-25), and in the United States it is based on the Airworthiness Standards: Transport Category Airplanes (FAR Part 25). Large aeroplanes are those aircraft with a maximum takeoff mass of more than 5,700 kg. JAR and FAR are very similar; the basic code for JAR-25 is FAR Part 25, and further harmonization of the requirements is in progress. The certification of one or several prototype aircraft leads to a *type certificate* being issued. Aircraft in series production have to show *airworthiness* and *conformity with the prototype aircraft.* In service the aircrafts have to be maintained according to an agreed maintenance schedule to prove continuous airworthiness.

[2] Following the new ATA 2200, ''Cabin Systems (ATA 44)'' are defined as ''Those units and components which furnish means of entertaining the passengers and providing communication within the aircraft and between the aircraft cabin and ground stations. Includes voice, data, music and video transmissions.''

JAR-25 and FAR Part 25 are grouped into several subparts (the following is based on JAR-25).

Subpart F, ''Equipment,'' contains many requirements for aircraft systems.

Subpart E, ''Power plant,'' contains requirements for power plant-related systems.

Also Subpart D, ''Design and Construction,'' contains requirements for aircraft systems.

Subpart J, ''Gas Turbine Auxiliary Power Unit Installation,'' contains requirements for airborne auxiliary power—i.e., the auxiliary power unit (APU).

General information on aircraft systems can be found in section 1301 ''Function and installation'' and section 1309 ''Equipment, systems and installations'' of JAR-25 and FAR Part 25. Section 1309 provides information on safety requirements, loads, and environmental conditions. Table 12.2 provides access to the certification requirements for large airplanes when specific information related to a particular aircraft system is needed.

Interpretative material to most paragraphs is provided:

- FAR: Advisory Circulars (AC) (especially in AC 25-17 and AC 25-22)
- JAR: Advisory Circular Joint (ACJ) (ACJ-25) and Advisory Material Joint (AMJ) (AMJ-25)

Safety and Reliability

Safety and reliability considerations of aircraft systems are an integral part of the safety and reliability considerations of the whole aircraft (see Section 17 of this handbook). Modern sophisticated aircraft depend very much on the proper functioning of their aircraft systems, so that safety and reliability considerations of aircraft systems have become highly important in their own right. For this reason an aircraft systems-specific approach to the topic is presented here.

Safety is a state in which the risk is lower than a permissible risk. The risk is defined by the probability of a failure and the expected effect.

The *effect* of failure describes the consequences of the failure (damage or injury).

The *probability of failure,* $F(t)$, is equal to the number of failures within a given period of time divided by the total number of parts in a test.

The *safety requirements* for aircraft systems are stated in section 1309 of the certification requirements JAR-25 and FAR Part 25 and are listed in Table 12.3. The probability of a failure in a system increases with the time period of operation and is specified for an operation time of one flight hour (FH). Obviously, the higher the effect of a failure is on aircraft operation, passengers, and the aircraft itself, the lower the permissible probability of such a failure has to be.

The *reliability* is the probability of survival, $R(t)$. It is an item's ability to fulfill defined requirements for a specific period of time under specified conditions. A statement referring to the reliability of a system can only be made if the failure criteria are precisely defined.

The reliability or *probability of survival,* $R(t)$, can also be defined as the number of parts surviving within a given period of time divided by the total number of parts in a test:

$$R(t) + F(t) = 1$$

Although referring to the reliability $R(t)$, mostly the value of the probability of

TABLE 12.2 Selected Certification Requirements for Aircraft Systems Based on JAR-25

Identifier name of system	Applicable sections
21 Air conditioning	831–833: Sections under the heading ''Ventilation and heating'' 841–843: Sections under the heading ''Pressurisation'' 1461: Equipment containing high energy rotors
22 Auto flight	1329: Automatic pilot system 1335: Flight director systems
23 Communications	1307: Miscellaneous equipment (radio communication) 1457: Cockpit voice recorders
24 Electrical power	1351: General 1353: Electrical equipment and installations 1355: Distribution system 1357: Circuit protective devices 1359: Electrical system fire and smoke protection 1363: Electrical system tests
25 Equipment/furnishings	771–793: Sections under the heading ''Personnel and cargo accommodations'' 819: Lower deck service compartments (including galleys) 1411: General (under heading safety equipment) 1413: Safety belts 1415: Ditching equipment 1421: Megaphones
26 Fire protection	851–867: Sections under the heading ''Fire protection'' 1181–1207: Sections under the heading ''Powerplant fire protection'' 1307: Miscellaneous equipment (portable fire extinguishers) A1181–A1207: Sections related to APU fire protection
27 Flight controls	671–703: Sections under the heading ''Control systems''
28 Fuel	951–981: Sections under the heading ''Fuel system'' 991–1001: Sections under the heading ''Fuel system components'' A952–A999: Sections related to the APU fuel system
29 Hydraulic power	1435: Hydraulic systems
30 Ice & rain protection	1307: Miscellaneous equipment (including: windshield wiper) 1416: Pneumatic de-icer boot system 1419: Ice protection
31 Indicating/recording systems	1303: Flight and navigation instruments 1305: Powerplant instruments 1321: Arrangement and visibility 1331: Instruments using a power supply 1333: Instrument systems 1337: Powerplant instruments
32 Landing gear	721–X745: Sections under the heading ''Landing Gear''

TABLE 12.2 Selected Certification Requirements for Aircraft Systems Based on JAR-25 (*Continued*)

Identifier name of system	Applicable sections
33 Lights	812: Emergency lighting 1322: Warning, caution, and advisory lights 1381: Instrument lights 1383: Landing lights 1385, 1387, 1389, 1391, 1393, 1395, 1397: Position lights 1401: Anti-collision light system 1403: Wing icing detection lights
34 Navigation	1307: Miscellaneous equipment (radio navigation) 1323: Airspeed indicating system 1325: Static pressure system 1326: Pilot heat indication system 1327: Magnetic direction indicator 1459: Flight recorders
35 Oxygen	1439: Protective breathing equipment 1441: Oxygen equipment and supply 1443: Minimum mass flow of supplemental oxygen 1445: Equipment standards for the oxygen distributing system 1447: Equipment standards for oxygen dispensing units 1449: Means for determining use of oxygen 1450: Chemical oxygen generators 1451: Fire protection for oxygen equipment 1453: Protection of oxygen equipment from rupture
36 Pneumatic	X1436: Pneumatic systems—high pressure 1438: Pressurisation and low pressure pneumatic systems
38 Water/waste	1455: Draining of fluids subject to freezing X799: Water systems
49 Airborne auxiliary power	Paragraphs in Subpart J—Gas turbine auxiliary power unit installations

failure $F(t)$ is given (10^{-7}) because the reliability yields values more difficult to handle (0.9999999).

The *hazard rate function,* $z(t)$, is a measure of the probability that a component will fail in the next time interval, given that it has survived up to the beginning of that time interval. If the hazard rate function is constant (which is often assumed), it is called the *failure rate,* λ. Failure rates of mechanical components are listed in Rome (1985), and failure rates for electric and electronic equipment can be estimated using MIL-HDBK-217. The failure rate has units of one per flight hour (1/FH). The inverse of the failure rate, called the mean time between failures (MTBF), is often used in reliability and maintenance circles.

$$\text{MTBF} = 1/\lambda$$

The *failure to removal ratio* (FTRR) is a maintenance quantity. It shows the ratio of faults found in a component during a shop visit, divided by the number of

TABLE 12.3 Safety Requirements for Large Airplane's Systems

Effect on aircraft and occupants	Normal	Nuisance	Operating limitations Emergency procedures	Significant reduction in safety margins Difficult for crew to cope with adverse conditions Passenger injuries	Large reduction in safety margins Crew extended because of workload or environmental conditions Serious injury or death of small number of occupants	Multiple deaths, usually with loss of aircraft
Category of effect	Minor	Minor	Minor	Major	Hazardous	Catastrophe
Probability of a failure according to JAR-25 (per flight hour)	Frequent 10^{0}–10^{-2}	Frequent 10^{-2}–10^{-3}	Reasonably probable 10^{-3}–10^{-5}	Remote 10^{-5}–10^{-7}	Extremely remote 10^{-7}–10^{-9}	Extremely improbable $<10^{-9}$

Source: ACJ-25.

component removals. Unfortunately, the FTRR is especially low in case of electrical components (0.6–0.7) and electronic components (0.3–0.4). Hydraulic components (0.8–0.9) and mechanical components (1.0) show better values. The product of MTBF and FTRR yields the maintenance cost driver, the *mean time between unscheduled removals* (MTBUR).

$$\text{MTBUR} = \text{MTBF} \cdot \text{FTTR}$$

The reliability and the probability of failure can be calculated from the failure rate

$$R(t) = e^{-\lambda t}, \quad F(t) = 1 - e^{-\lambda t}$$

For low failure rates, which are common in aviation, the probability of failure calculated for a period of one hour ($F(t)$/FH) equals almost exactly the failure rate, λ.

Systems are a combination of many components either in parallel, in series, or in a combination of both. The reliability of a *series system* is equal to the product of is component values.

$$R_S(t) = R_1(t)R_2(t)R_3(t) \ldots$$

The failure rate of a series system is approximately the sum of the failure rates of its (reliable) components.

$$\lambda_S \approx \lambda_1 + \lambda_2 + \lambda_3 \ldots$$

The probability of failure of a *parallel system* is equal to the product of is component values.

$$F_P(t) = F_1(t)\, F_2(t)\, F_3(t) \ldots$$

The failure rate of a parallel system is approximately the product of is (reliable) component values.

$$\lambda_P \approx \lambda_1 \lambda_2 \lambda_3 \ldots$$

Systems can be depicted by *reliability block diagrams* (RBDs). The analysis of large systems is carried out in successive stages. At each stage a small number of components connected either in parallel or in series is combined with equations as shown above. In this way the complexity of the system can be reduced step by step. The *fault tree analysis* (FTA) is an alternative method to deal with complex systems. Parallel systems are combined by an OR gate symbol. Series systems are combined by an AND gate symbol. Top events are shown in a rectangle and basic failure causes are shown in circles. Software tools exist that support a FTA or the analysis of a RBD. Systems might show cross-linkages so that some units are in more than one subsystem. One way of dealing with this problem is to use a theorem on conditional probability or to apply a truth table (Davidson 1988).

These approximate equations for series and parallel systems are quite useful in day-to-day business. The last equation also shows the ability of parallel systems to achieve low failure rates and thus high reliability. For example, three components combined in parallel with a failure rate of 10^{-3} 1/FH each, yield an overall failure rate of 10^{-9} 1/FH. This is a failure rate that could not have been achieved by a single component no matter how carefully this component was manufactured and tested. This thought leads us to the concept of redundancy, which is so typical in safety critical aircraft systems.

Redundancy is the existence of more means for accomplishing a given function than would simply be necessary. It is divided into

- Homogeneous redundancy (the multiple means are identical) and
- Inhomogeneous redundancy (the multiple means are of different type)

Inhomogeneous redundancy is divided into:

- Dissimilar redundancy or
- Diversitary redundancy

Safety-critical aircraft systems often show *triplex* subsystems. The system architecture of safety-critical computers may be even of *quadruplex* or *duo duplex* type.

The subsystems of a system with built-in redundancy may all work together. If one subsystem fails, the others will just have to cope with a somewhat higher load. These systems are called *active-active* systems. Other systems may be of the *active-standby* type and need to perform a changeover in case of a failure. If the standby subsystem is constantly waiting to be activated, it is on *hot standby;* otherwise it is on *cold standby.* The changeover should not be dependent on a changeover unit, because this unit with its own limited reliability might fail and prevent the changeover. If an active-standby concept is applied, the subsystems should take turns doing the job. This could be achieved with a planned changeover before every takeoff. If the same subsystem stays in standby all the time, it may show an (undetected) *dormant failure* and hence will not be able to take up the job in case of failure of the first subsystem. Systems with a potential of dormant failures need regular maintenance checks and should be avoided.

An assumption has been made in the calculation of parallel systems that the failures of individual subsystems are independent of each other, that is, that two or more subsystems do not fail simultaneously from precisely the same cause (except purely by chance). However, most systems have the potential of having more than one failure due to a common cause. These failures are called *common cause failures* (CCFs). They tend to arise from errors made during design, manufacture, maintenance, operation, or environmental effects. For example, loss of power supply could cause both a running and a standby pump to fail (design error), or an empty fuel tank could cause all engines to quit (error in operation). Because these failure modes may appear to be outside the system being assessed, they can easily be overlooked, leading to too-optimistic assessments. Methods to avoid common cause failures in the design stage are the application of

- Inhomogeneous redundancy (see above)
- Segregation in the rooting of redundant wires, pipes, and ducts
- Separation of redundant components
- Placement of safety-critical components in safe areas
- Design of redundant components or software programs by independent teams with different (software) tools

An aircraft should not only be safe to fly, it should also show very few errors that need the attention of maintenance personnel. In this respect we face a problem with high safety requirements. High safety requirements lead to the application of redundancy and hence more subsystems. The probability of a failure leading to the loss of the overall function can be reduced by redundancy, but the probability of

occurrence of any failure anywhere in the system is increased. Two subsystems with a failure rate of 10^{-3} 1/FH each yield an overall probability of failure of about 10^{-6} and a probability of any failure of $2 \cdot 10^{-3}$ (based on a one-hour operation). Three subsystems yield an overall probability of failure of 10^{-9} and a probability of any failure of already $3 \cdot 10^{-3}$. The level of safety during flight can only be achieved if all subsystems work properly before takeoff, but, as we have seen, the probability for any failure increases with an increased number of subsystems. These thoughts lead to what is called availability and dispatch reliability.

The *steady state availability* is defined as the probability that a system will be available when required, or as the proportion of total time that the system is available for use. Therefore, the availability of a system is a function of its failure rate λ and of its repair rate $\mu = 1/\text{MTTR}$, where MTTR is the mean time to repair:

$$A_{SS} = \frac{\text{MTBF}}{\text{MTBF} + \text{MTTR}} = \frac{\mu}{\lambda + \mu}$$

The *instantaneous availability,* or probability that the system will be available at time t, is

$$A_I = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-(\lambda+\mu)t}$$

Often it is more revealing to consider system unavailability, $U = 1 - A$. The instantaneous availability of an aircraft at the moment of dispatch from the gate is called *dispatch reliability.* Dispatch reliability, for technical reasons, primarily depends on the combined dispatch reliability of the aircraft systems. The airlines monitor their fleets' dispatch reliability very carefully because high *dispatch unreliability* leads to delays and cancellations of flights and incurs delay and cancellation costs (see below). Dispatch reliability depends on the maturity of an aircraft program and is on the order of 0.99. A method to increase dispatch reliability is the introduction of *built-in test equipment* (BITE) into electronic systems. Though this adds complexity and might result in spurious failure indications, it can greatly reduce maintenance times by providing an instantaneous indication of failure location. Another method is to provide *extra redundancy* above the level required for safety reasons. This would than allow to dispatch with one subsystem inoperative. Components that are not needed for takeoff may be known as *flying spares.* The pilot gets a clear indication about which subsystems or components need to be available at takeoff from the *minimum equipment list* (MEL), written by the airline on the basis of the master minimum equipment list (MMEL) provided by the manufacturer and approved by the authorities.

Reliability assurance during the aircraft system design applies a couple of different methods, including:

- Drawing a fault tree for a *fault tree analysis* (FTA) (see above) starts from consideration of system failure effects, referred to as top event. The analysis proceeds by determining how these can be caused by lower-level failures. In this way it is a top-down approach.
- The *reliability apportionment* breaks an overall system reliability requirement down into individual subsystem reliabilities. This is common in large systems when different design teams of subcontractors are involved. Clearly it follows a top-down approach.

- In contrast, the *failure mode, effects, and criticality analysis* (FMECA) (MIL-STD-1629) follows a bottom-up approach. It considers each mode of failure of every component of a system to ascertain the effects on system operation and defines a *failure mode criticality number.*
- The *zonal safety analysis* (ZSA), rather than looking at an aircraft from a functional point of view, looks at the components' location. The ZSA checks installation rules and checks the effects of events originating within the zone, in other zones, or on the outside.

Software defies the above calculations and methods. However, information can be drawn from RTCA/DO-178B, which deals with *software considerations* in airborne systems and equipment. *Environmental conditions* for airborne equipment are presented in RTCA/DO-160D.

Mass

Mass estimation of aircraft systems is part of the mass (or weight) estimation of the whole aircraft, which is covered under Weight Estimation in Section 13 of this handbook.

The mass of all the aircraft systems m_{SYS} amounts to 23–40% of the aircraft's empty mass m_{OE}, where m_{OE} is the mass related to the operational empty weight (OEW). The figure 23% is true in case of a modern long-range airliner, whereas 40% is about right for a smaller aircraft such as business jet. Hence, for civil jet transport we may write

$$\frac{m_{SYS}}{m_{OE}} \approx 0.23\text{–}0.4$$

On average this ratio comes to ⅓, as stated above. Taking into account the ratio of the aircraft's empty mass m_{OE} and the maximum takeoff mass m_{MTO}, the mass related to the maximum takeoff weight (MTOW).

$$\frac{m_{SYS}}{m_{MTO}} \approx 0.11\text{–}0.23$$

Figure 12.1 shows the mass of aircraft systems of selected civil jet aircraft as a function of their maximum takeoff mass. We follow a *top-down approach* and fit a curve to these data to obtain

$$m_{SYS} = 0.92\, m_{MTO}^{0.85} \qquad \text{for } m_{SYS} \text{ and } m_{MTO} \text{ in kg}$$

This function is shown in Figure 12.1. The average relative mass of the individual systems of civil jet aircraft is given in Table 12.4.

Some aircraft systems, like the landing gear system (ATA 32) and the equipment and furnishings (ATA 25), account for a large percentage of the total aircraft system mass. The avionic system relative mass is 6% on average, but this figure depends on aircraft size because the amount of avionics needed in jet aircraft tends to be nearly constant. For this reason, the relative mass of avionic systems of business aircraft may be as high as 14% and as low as 5% in case of a large civil transport. As can be seen in Table 12.4, a number of systems are of minor importance for aircraft system mass predictions.

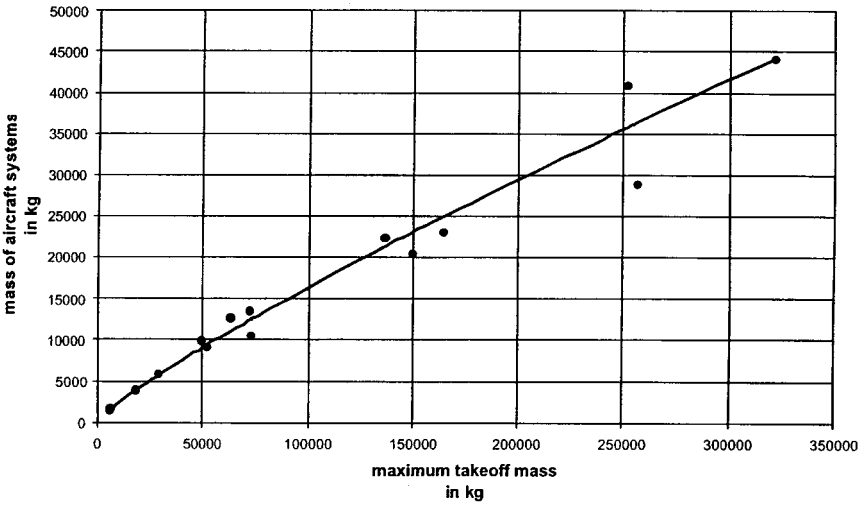


FIGURE 12.1 Mass of aircraft systems of selected civil jet aircraft plotted against their maximum takeoff mass.

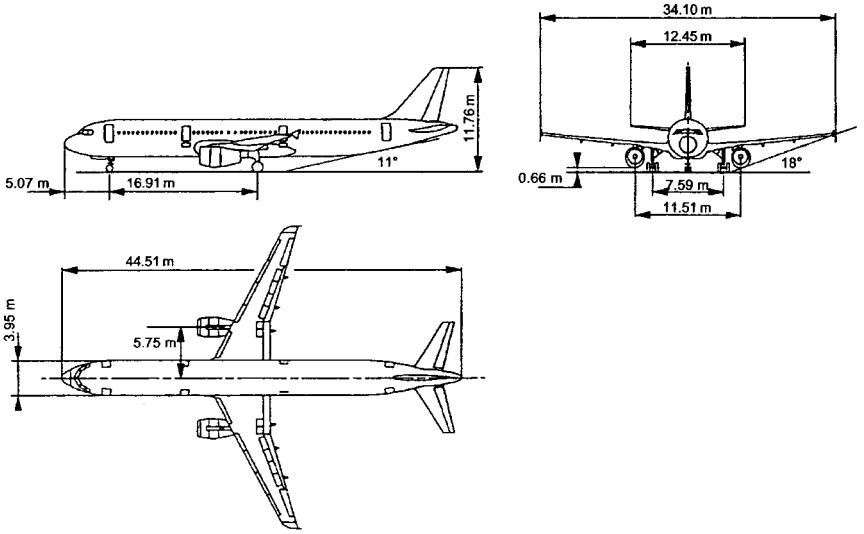


FIGURE 12.2 The Airbus A321 is used throughout this section to provide aircraft system examples. One hundred eighty-six passengers in two-class layout, MTOW: 83,000 kg, M_{MO} = 0.82, maximum FL 390.

TABLE 12.4 Average Relative Mass of Aircraft Systems of Civil Jets

Identifier	Name of system	Average relative mass of system
21	Air conditioning	6%
22	Auto flight	1%
23	Communications	2%
24	Electrical power	10%
25	Equipment/furnishings	24%
26	Fire protection	1%
27	Flight controls	8%
28	Fuel	3%
29	Hydraulic power	7%
30	Ice and rain protection	<1%
31	Indicating/recording systems	<1%
32	Landing gear	27%
33	Lights	2%
34	Navigation	3%
35	Oxygen	1%
36	Pneumatic	2%
38	Water/waste	1%
49	Airborne auxiliary power	2%

Alternatively, it is also possible to follow a *bottom-up approach.* This statistical technique uses system parameters to predict the mass of the system. Equations are given in Raymer (1992), Roskam (1989), and Torenbeek (1988). In addition, the knowledge gathered in papers from the Society of Allied Weight Engineers should be tapped (see SAWE 2002).

Statistics of aircraft system mass have to take as many aircraft into account as possible in order to broaden the statistical base. This, however, is really possible only if mass data are based on comparable and detailed mass breakdowns. Unfortunately, there are many quite different breakdowns in use, and it is found that system boundaries overlap from one method to another or are not well defined in the first place. So in the present situation it is very difficult to use and compare mass data and mass equations based on one of these breakdowns in another setting. This situation adds to the difficulties that exist with statistical methods anyhow and explains why statistical mass equations for systems or subsystems do not provide particularly reliable data.

Boeing has used a *breakdown format* called Weight Research Data 1 (WRD1). In the literature, breakdowns very similar to WRD1 can be found. Airbus uses so-called Weight Chapters. Another approach is given with MIL-STD-1374. Above we have used a mass breakdown according to the ATA 100 chapter numbering. ATA 100 also includes a widely accepted mass breakdown for weight and balance manuals. This breakdown, however, provides only as much detail as needed in aircraft operation but not enough detail for aircraft system design.

Note that aircraft system *mass predictions* deteriorate in *accuracy* when the level of detail is increased. For its old class I weight prediction method, Boeing estimates the prediction of single systems to be off by as much as $\pm 90\%$. In contrast, the resultant mass of all systems combined is claimed to be off by not more than $\pm 16\%$

(Boeing 1968). This is because many inaccuracies combined fortunately cancel out to a certain extent.

Detailed system mass predictions are also necessary for *center of gravity* (CG) *calculation* for the aircraft. The main landing gear accounts for about 87% and the nose landing gear for the remaining 13% of the complete landing gear mass. With known positions of nose and main landing gear, this information can be fed into the CG calculation of the aircraft. The CG of the other systems can roughly be assumed at a point 40–50% of the fuselage length aft of the aircraft nose.

Practical mass predictions will look like this: In the early design stage, statistical methods are used. The aircraft manufacturer can also use the information contained in the mass database of older aircraft for the new design. In a later design stage a subcontractor will offer a system or an item of equipment. The subcontractor probably has quite a good idea what the item's mass will be from a comparison with similar items already built. If the required size of equipment is different from an older one, a mass estimate may be obtained from scaling. In the final development stage, mass accounting can be based on the actual mass of components that are already delivered to the manufacturer.

There is another virtue in mass predictions: the system mass has been used for rough *cost calculations.* This is possible when, from statistics, costs per unit mass are known and costs are assumed to be proportional with mass. Evidently, the concept of calculating costs from mass fails if expensive mass reduction programs are being applied. The concept also fails if highly sophisticated technologies are applied to reduce mass that are not considered in the established cost per unit mass.

Power

Gliders use the energy of up-currents, while solar-powered vehicles use the energy from the sun. Human-powered flight has also been demonstrated. Propulsive power for any other ''down to earth'' flying depends on fuel. This fuel is used in the aircraft main engines. *Secondary power* systems (hydraulic power, electrical power, pneumatic power) in turn draw on engine power to supply their client systems with *nonpropulsive power* in all those cases where functions are not directly actuated by the pilot's muscles. This is the simple picture of the aircraft power management. However, there is more to it, due to safety requirements and the need for autonomous operation of the aircraft on the ground with engines shut down.

Various secondary power sources are available in the air and on the ground. Secondary power loads may be grouped into two major categories. Power conversion transforms secondary power from one form into another.

An *auxiliary power unit* (APU) (see above) is used to produce power from fuel independent of the main engines. An APU is a gas turbine engine. Most often it produces electrical power and pneumatic power. A *ram air turbine* (RAT) (see Subsection 12.8) is used to produce hydraulic or electrical power from the kinetic energy of the air passing by the aircraft. This is possible even without fuel and without the main engines running—at least as long as the aircraft soars down consuming its potential energy. Except for the pilot's own energy, the *aircraft batteries* are the last and very limited source of energy on board.

Ground power may be available on the apron or in the hangar. The aircraft may be supplied directly with electricity, high-pressure hydraulic fluid, pressurized air, and/or air conditioned air. Human power could work a hand pump in the hydraulic system. If only electrical ground power is available, the aircraft depends on its

secondary power conversion capabilities to activate the hydraulic and pneumatic system. Without ground equipment and with engines shut down, the aircraft may operate autonomously if it is equipped with an auxiliary power unit (APU).

First of all, *secondary power loads* may be grouped into:

- Technical loads consumed by equipment required to operate the aircraft safely
- Commercial loads consumed by equipment required to increase passenger comfort and satisfaction, given the airline's need to provide these services

Power conversion among different[3] secondary power systems is used to increase overall system reliability. If we consider electrical power, hydraulic power, and pneumatics:

- Six different unidirectional conversions are possible. Examples are:
  - Electrical to hydraulic power conversion: electric motor-driven pump
  - Pneumatic to hydraulic power conversion: air turbine motor-driven pump
  - Hydraulic to electrical power conversion: hydraulic motor-driven generator
- Three different bidirectional conversions are possibilities that allow a two-way power conversion among two different secondary power systems within one conversion unit.

For many years hydraulic, pneumatic, and electrical power supply in commercial aircraft had been sufficient to meet the demands from technical and commercial loads. System design emphasized reliable, lightweight solutions. From fuel input to system output, very low overall efficiencies were accepted in exchange.

In recent years it has been observed that aircraft face increasing technical loads. Also, market trends together with increasing flight durations have resulted in higher commercial loads, caused, for example, by today's standards in in-flight entertainment. Possibilities for power off-takes do not increase proportionally with aircraft size. Large modern civil aircraft are therefore likely to face limitations of cost effectiveness, geometry, or weight with present-day technologies in an attempt to meet these new power load levels. The aerospace industry has identified a potential deadlock, where power needs will exceed the maximum available power supply.

In the future a move towards electrical power as a single source to meet secondary power demands is expected to be a solution to the problem. The last aircraft generation brought steering by wire. The next generation of aircraft might bring power by wire.

Costs and Trade-off Studies

Trade-off studies play an important roll in aircraft system design. Trade-off studies try to find the best among several system design proposals. Safety aspects allow no compromise because certification regulations have to be closely followed. Also,

[3] Power conversion is even applied within one type of secondary power system: the hydraulic system. Transport category aircraft apply several independent hydraulic systems. Among pairs of these hydraulic systems unidirectional or bidirectional hydraulic power transfer without the interchange of hydraulic fluid can be desirable. For this purpose, power transfer units (PTU) (ARP 1280) are used. They are built by coupling a hydraulic motor and a hydraulic pump via a connecting shaft.

performance aspects leave little room because usually only as much performance as necessary to do the job will be allowed for. More powerful aircraft systems will unnecessarily produce costs that add to the overall costs of the aircraft. Clearly, costs need to be reduced as much as possible to come up with a viable product. Therefore, it is the costs aspect that is usually decisive in trade-off studies of which system design will get on board the aircraft.

At the aircraft system level, evaluations are done in the early design stage by looking separately at various aspects:

- Mass
- Maintainability
- Reliability
- System price
- Other specific criteria depending on the aircraft system in question

Based on these separate evaluations, the simplest way to come up with one single figure of merit for a proposal is to define subjectively a *weighted sum* of the results based on the individual criteria.

In contrast to the above approach, at the aircraft level an evaluation is traditionally based primarily on one single figure: the direct operating costs (DOC) (see Section 13 of this handbook). DOCs take account of criteria such as mass, maintainability, and aircraft price, but combine these separate parameters unambiguously by calculating their economical implications. Subjective manipulations of the results are largely avoided in this way.

Unfortunately, aircraft DOC methods cannot be taken as is for applying this advantage to an aircraft system evaluation. In contrast to aircraft DOC methods, a DOC method on the systems level must incorporate many system-specific parameters. Therefore, a *DOC method for aircraft systems* called $\mathrm{DOC_{SYS}}$ has been developed (Scholz 1998) which follows the principles of aircraft DOC methods as closely as possible while taking aircraft system peculiarities into account as much as necessary.

$$C_{\mathrm{DOC,SYS}} = C_{\mathrm{DEP}} + C_F + C_M + C_{\mathrm{DEL}} + C_{\mathrm{SH}}$$

where C_{DEP} = depreciation of the system (a function of system price)
C_F = fuel costs caused by the system
C_M = direct maintenance costs caused by the system (see Airframe Maintenance in Section 16 of this handbook)
C_{DEL} = delay and cancellation costs caused by the system
C_{SH} = capital costs caused by necessary system spare parts on stock (spare holding)

The fuel costs, C_F, are due to:

- Transportation of the system's mass (fixed or variable during flight) (taking into account the lift-to-drag ratio of the aircraft and the specific fuel consumption of the engines)
- Power off-takes from the engines (by electrical generators or hydraulic pumps)
- Bleed air off-takes (for the pneumatic system)

- Ram air off-takes (e.g., for the air conditioning system)
- Additional drag caused by the presents of aircraft systems, subsystems, or single parts (e.g., due to drain masts)

In contrast to Scholz (1998), who combines various system aspects to U.S. dollars, Shustrov (1998) combines system mass effects and effects related to the system's energy consumption to a quantity called *starting mass*.

Proprietary methods for the evaluation of aircraft systems are in use at aircraft manufacturers and subcontractors.

12.2 AIR CONDITIONING (ATA 21)

Air conditioning as defined by ATA 100:

> Those units and components which furnish a means of pressurizing, heating, cooling, moisture controlling, filtering and treating the air used to ventilate the areas of the fuselage within the pressure seals. Includes cabin supercharger, equipment cooling, heater, heater fuel system, expansion turbine, valves, scoops, ducts, etc.

Fundamentals

Impact of Atmospheric Parameters. In the troposphere, the air temperature decreases with increasing altitude. In the stratosphere above 11,000 m (36,089 ft), the air temperature is at constant −56.5 °C. The air pressure also decreases with altitude (see Section 16 of this handbook). Although oxygen amounts to approximately 21% independent of altitude, the partial pressure[4] of oxygen drops with increasing altitude. Our body is used to a partial oxygen pressure of about 0.21 times sea level pressure. If we want to survive at high altitudes, either (a) the oxygen fraction has to be increased (using an oxygen system), or (b) the total pressure has to be maintained close to sea level pressure (using a pressurization system). For civil aircraft generally option (b) is applied; flights in nonpressurized cabins[5] without supplemental oxygen are limited to an altitude of 10,000 ft. Military aircraft use a combination of (a) and (b); cabin altitude[6] does not exceed about 20,000 ft.

Purpose of Air Conditioning Systems. The purpose of the air conditioning system is to make the interior environment of the aircraft comfortable for human beings. Depending on the type of aircraft and altitude of operation, this may involve only *ventilation* of the cabin by supplying a flow of fresh air using air vents. If the temperature must be adjusted, some method of *heating* or *cooling* is required. At

[4] Partial pressure: "The pressure exerted by one gas in a mixture of gases; equal to the fraction . . . of one gas times the total pressure" (AIR 171).

[5] Nonpressurized cabin: "An airplane cabin that is not designed . . . for pressurizing and which will, therefore, have a cabin pressure equal to that of the surrounding atmosphere" (SAE 1998).

[6] Cabin altitude: "The standard altitude at which atmospheric pressure is equal to the cabin pressure" (SAE 1998).

high altitudes the aircraft can fly above most of the weather conditions that contain turbulence and make flight uncomfortable. Additionally, the fuel efficiency of the aircraft is increased. *Pressurization* is necessary if the aircraft is operated at these high altitudes. In some parts of the world the relative humidity[7] is quite high. Water extractors are therefore used for *dehumidification* of the cabin air. This is necessary to prevent damage to electrical and electronic equipment, aircraft insulation and structure. Reduced humidity also limits window and windscreen misting. At an altitude of 40,000 ft the relative humidity is quite low (1–2%) compared to the comfort level for crew and passengers (30%). Nevertheless, *humidification* of the cabin air would be impractical for the other reasons named and for the costs involved in carrying that water (AIR 1609).

The air conditioning system is a safety-critical system because passengers and crew depend on its proper function. Transport category aircraft will have two independent subsystems to meet these safety requirements. The certification requirements include minimum standards. The aircraft manufacturer may choose higher standards in order to increase passenger comfort.

Ventilation

- Under normal conditions 4.7 l/s (10 ft^3/min $\approx$ 0.6 lb/min) are required for each crew member (JAR-25 section 831(a)).
- Manufacturers will typically provide a minimum of about 7.8 l/s (1.0 lb/min) for each person in the aircraft.
- In case of a failure (with a probability of not more than 10^{-5} 1/FH) the supply of fresh air should not be less than 3.1 l/s (0.4 lb/min) per person excluding supply from the recirculation system (JAR-25 section 831(c)).
- In order to avoid drafts, the air velocity in the cabin should be limited to 0.2 m/s (40 ft/min) in the vicinity of the passengers (AIR 1168/3). Individual air outlets, however, show air velocities of about 1.0 m/s. Conditioned air may enter the cabin through cabin outlets at not more than 2.0 m/s.

Temperature Control

- The temperature control from the cockpit may typically be possible in the range between 18 °C and 30 °C.
- Heating and cooling requirements have to be met as specified for various steady state and transient scenarios. Here are some lessons learned:
  - During cruise cooling almost always is required (an exception is flights without passengers).
  - Cooling loads on the ground on a hot day with passengers on board are higher then in flight.
  - Transient scenarios will probably determine the heating and cooling performance of the air conditioning system in civil subsonic aircraft:
    - Heating the cabin of a cold-soaked airplane from −32 °C to 21 °C in 30 minutes (no internal heat loads, doors closed).
    - Cooling the aircraft from 46 °C to 27 °C in 30 minutes (full passenger load, doors closed) (ARP 85).

[7] Relative humidity: "The ratio, expressed as percentage, of the amount of water vapor . . . actually present in the air, to the amount of water vapor that would be present if the air were saturated with respect to water at the same temperature and pressure" (SAE 1998).

- Cooling requirements for high-speed aircraft are driven by kinetic heating. Kinetic heating occurs when the aircraft skin heats up due to friction with air molecules. In the flight range below Mach 2, the skin temperature is equal to the recovery temperature:[8] $T_{skin} \approx T_{ambient}\ (1 + 0.18\ M^2)$

Pressure Control (for aircraft with a pressurized cabin)[9]

- Under normal conditions the cabin altitude in pressurized cabins must not be more than 2,440 m (8,000 ft) (JAR-25 section 841(a)).
- In case of a failure (with a probability of not more than 10^{-5} 1/FH), cabin altitude must not be more than 4,570 m (15,000 ft) (JAR-25 section 841(a)).
- For passenger comfort, the cabin rate of climb should not be more than 2.5 m/s (500 ft/min) and the cabin rate of descent should not be more than 1.5 m/s (300 ft/min) (ARP 1270).
- The flow rate of air for cabin pressurization shall be enough to account for cabin leakage (allowing for an in-service increase of 10–15%) and cabin repressurization with 1.5 m/s (300 ft/min) (ARP 85).

Heating Systems

The simplest type of heating system, often employed in light aircraft, consist of a *heater muff* around the engine exhaust, an air scoop to draw ram air into the heater muff, ducting to carry the heated air into the cabin, and a valve to control the flow of heated air. Alternatively to the heater muff, a portion of the exhaust gases could also be fed to a *heat exchanger* to heat the ram air or the recirculated air from the cabin.

In larger aircraft *combustion heaters* are often employed. The heater burns fuel in a combustion chamber, and airflow around the chamber is heated and carried through ducts into the cabin.

Turbine engine-powered aircraft with a nonpressurized cabin normally make use of hot pressurized air tapped from the turbine engine compressor. This air is call *bleed air*. Temperature control is achieved by mixing the bleed air with ambient or recirculated air before it enters the cabin.

A pressurized aircraft cabin is usually heated by regulating the temperature of the air used to pressurize the cabin. This again is combined with an effort to cool the cabin. The combined process will be addressed in the following subsections.

Cooling Systems

There are several *heat sources* that cause a need for cooling. *External* heat sources include heat transfer through cabin walls and heat received through solar radiation. *Internal* heat sources include passengers and crew, heat generated by electronic, electric, and mechanical equipment.

Cooling systems require *energy* for their operation. This energy may come from ram air, engine bleed air, an engine-driven compressor, or the auxiliary power unit.

[8] Recovery temperature: "The equilibrium temperature of an object placed in a flow . . . always less than the total temperature" (AGARD 1980).

[9] Pressurized cabin: "An airplane cabin that is constructed, sealed, and equipped with an auxiliary system to maintain a pressure within the cabin greater than that of the surrounding atmosphere" (SAE 1998).

Cooling may apply different *heat sinks* to get rid of the heat: ram air, engine fan air, cabin exhaust air, fuel, or expendable cooling media (water or liquid hydrogen). Note that any ambient air taken aboard is at total temperature.[10]

The *cooling air* (ram air) may be moved by a fan driven by an electric or hydraulic motor, the air cycle machine, or an ejector pump.

The above means may be combined in systems applying two basic cooling principles. These systems are known as:

- The vapor cycle system, in which the heat of vaporization is lost by evaporating a liquid refrigerant.
- The air cycle system, which is based on the reduction of heat by the transformation of heat energy into work.

Combination of both principles is possible.

The *vapor cycle system* (Figure 12.3) is what is used in refrigerators. The cooling process is best explained starting at the *compressor,* where the refrigerant (a special fluid) is in gaseous form. The compressor increases pressure and temperature of the refrigerant and pushes it through the entire system. A heat exchanger called a *condenser* extracts heat from the compressed refrigerant and carries the heat overboard. The refrigerant cools down a little and changes into liquid form. Still under pressure, the refrigerant goes past the *expansion valve,* where it is sprayed into little droplets. Behind the expansion valve, pressure is low. With reduced pressure the temperature is also considerably reduced. The *evaporator* is the second heat exchanger in the system. The refrigerant, in the form of cold droplets, cools the air destined for the cabin that goes past the evaporator. By taking up the energy from the passing air in the evaporator, the refrigerant changes to gaseous form again. It

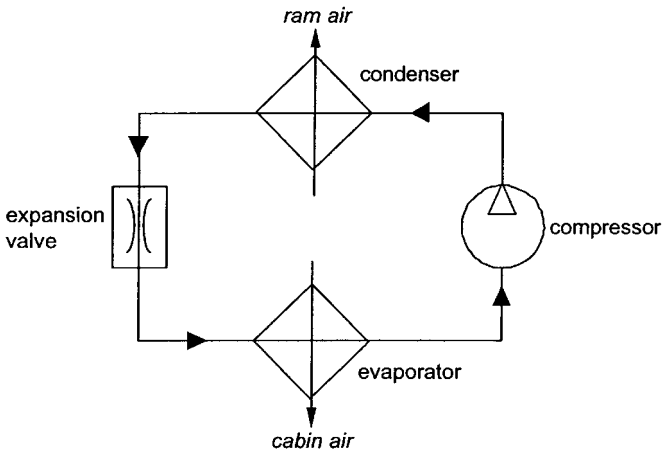


FIGURE 12.3 Vapor cycle system.

[10] Total temperature = stagnation temperature: "The temperature which would arise if the fluid were brought to rest adiabatically" (AGARD 1980).

now enters the compressor, where the cycle starts anew. Example: Dassault Falcon 10.

The vapor cycle is a closed cycle that works with a phase change from gas to liquid and vice versa. The latent heat[11] involved in the phase change makes the vapor cycle very efficient.

If we substitute air for the refrigerant and a turbine for the expansion valve, we basically get a *closed air cycle system*. In aircraft air conditioning, however, the cold air leaving the turbine is used directly as cabin air, forming an *open air cycle system*. Various air cycle systems have been conceived. The discussion here is limited to three open air cycle systems: the *basic air cycle systems,* the *bootstrap system,* and the *three-wheel system*.

In the *open basic air cycle system* (Figure 12.4), bleed air is cooled in a *heat exchanger* with ram air. The bleed air drives a *turbine,* using the pressure differential between bleed and cabin pressure. The bleed air is cooled during the expansion in the turbine. The work extracted from the turbine drives a *fan* that augments the airflow through the heat exchanger. In the cold air behind the turbine, water is condensed in form of minute drops (fog). A *low-pressure water separator* extracts this water. A *bypass valve* is used for temperature regulation and to prevent ice buildup in the water separator. Example: Lockheed C-130.

The turbine can also be used to drive a compressor that further increases the pressure of the air supplied to the cooling turbine. A higher pressure ratio leads to a higher temperature drop across the turbine and hence an improved performance. An air cycle system with a turbine coupled to a compressor is called a bootstrap system.

The *open bootstrap air cycle system* (Figure 12.5) directs bleed air through a *primary heat exchanger.* The air is compressed and then passed through a secondary

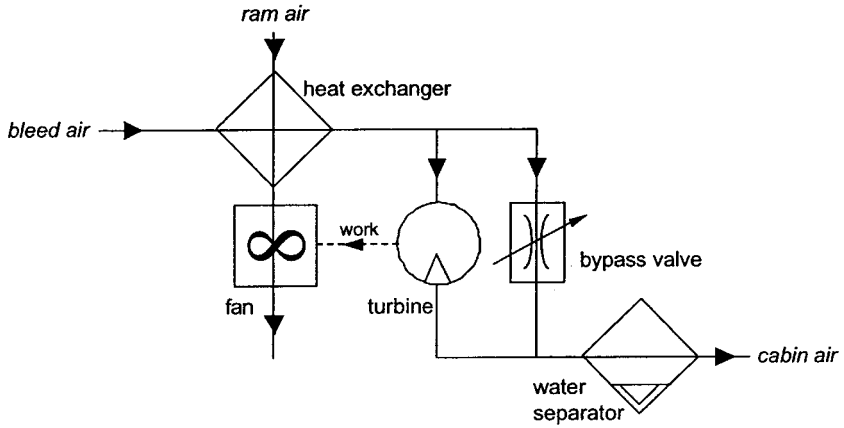


FIGURE 12.4 Open basic air cycle system.

[11] Latent heat: "The unit quantity of heat required for isothermal change in a state of a unit mass of matter" (SAE 1998).

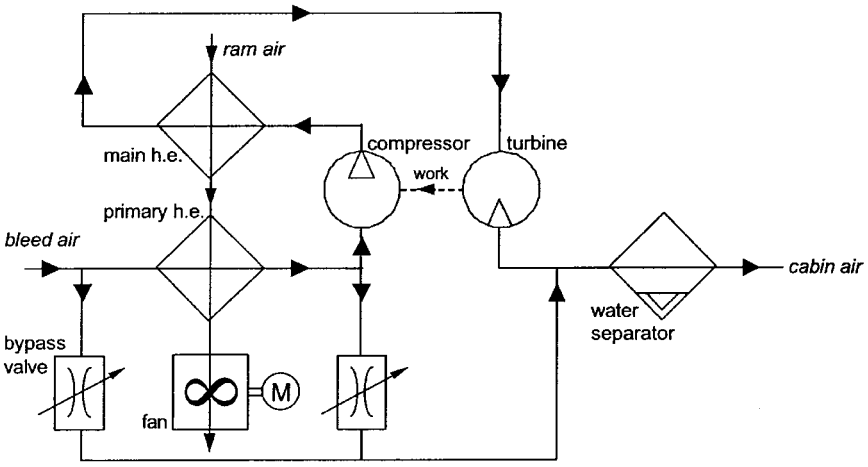


FIGURE 12.5 Open bootstrap air cycle system.

heat exchanger (or *main heat exchanger*). The air then enters the *turbine,* where it is expanded to cabin pressure. A *low-pressure water separator* reduces the water content. Heat energy is converted into shaft work and used to drive the bootstrap *compressor.* The primary and main heat exchangers are cooled by ram air. The *fan,* used to augment the airflow through the heat exchangers, may be driven by an electric motor. Bypass lines are integrated for temperature control. Example: Boeing 727.

Two types of *water separators* exist. So far we have seen the application of a low-pressure water separator that is installed behind the turbine and limits cabin air to temperature above 0 °C. In contrast, a high-pressure water separator is installed before the turbine. Separating the water before the turbine requires at least one more heat exchanger: a condenser or a condenser and a reheater. The advantage of the high-pressure water separator is that the air may be cooled down to temperatures of −50 °C. This results in higher temperature differences at the heat exchangers and higher efficiency of the system.

More recent transport category aircraft use the *open three-wheel air cycle system* with a high-pressure water separator (examples: B757, B767, A320). The three-wheel system is a bootstrap system where the turbine drives not only the compressor but also the fan. Figure 12.6 shows this configuration.

Pressurization Systems

As we saw above, pressurization is necessary to fly at high altitudes (compare with Figure 12.8). The use of pressurization is found in aircraft ranging from light single-engine aircraft up to big turbine-powered transport aircraft. Although the basic controlling mechanisms for each of these types are the same, the *sources of pressure* and details of the system vary. Pressure generation and distribution are the respon-

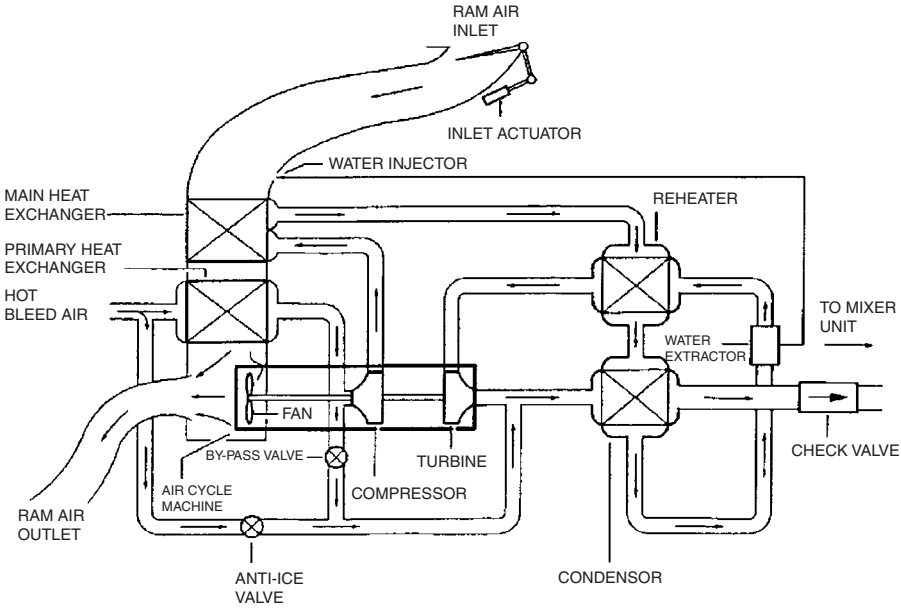


FIGURE 12.6 A321 air cooling in the pack.

sibility of the pneumatic system and are discussed in Subsection 12.13 in more detail. Reciprocating engines can supply pressure from a supercharger, a turbocharger, or an engine-driven compressor. Turbine-powered aircraft usually use bleed air as a source for compressed air. Bleed air is air that is tapped from the compressor section of the turbine engine.

Heating and cooling with an open air cycle system provide conditioned air to the cabin that is used at the same time for pressurization. Heating, cooling, and pressurization all have to be integrated in such a way that an optimum overall system solution results.

The flow of air into the cabin is approximately constant. *Pressure control* is hence achieved by varying the amount of flow out of the cabin. This is done with a regulated outflow valve. The outflow valve may be operated directly, by pneumatic pressure, or by electric motors.

An aircraft must have enough *structural strength* to withstand the stresses caused by a pressurized cabin. The limiting factor in how high an aircraft can operate is the maximum allowed cabin differential pressure, i.e., the difference between the cabin pressure and the pressure at maximum altitude for which certification is sought: $\Delta p = p_{\text{cabin}} - p_{\text{max,alt}}$. Aircraft are not intended to fly with a cabin pressure below ambient pressure.

Safety valves are used to safeguard against unauthorized positive or negative differential pressure. A pressure relief valve opens automatically if the cabin differential pressure gets above permitted limits. An automatic negative pressure relief valve opens automatically if the negative cabin differential pressure gets above

permitted limits. A dump valve is used to release remaining cabin differential pressure when the aircraft lands. Note that one pressurization control valve may serve more than one function in a specific aircraft design.

Example: Airbus A321

The Airbus A321 has two air conditioning packs which are open three-wheel air cycle systems. Figure 12.6 shows an air conditioning pack with the air cycle machine, the heat exchangers, and a high-pressure water separator.

The cabin temperature can be adjusted by computer individually in three different cabin zones (Figure 12.7). The air conditioning packs (Figure 12.6) deliver air at a temperature to satisfy the zone with the lowest temperature demand. Air from the packs is delivered to the mixing unit. Also, recirculated air from the cabin enters the mixing unit through filters and cabin fans. The recirculated air amounts to 40% of the total air supplied to the cabin. Recirculated air restores some humidity into the cabin. Trim air valves mix hot bleed air with the air from the mixing unit to attain the individually requested zone temperatures.

The pressurization control system includes two cabin pressure controllers. Operation may be fully automatic, semiautomatic, or manual. The outflow valve is equipped with three electrical motors. Two safety valves avoid excessive positive (593 hPa = 8.6 psi) or negative (−17 hPa = −0.25 psi) differential pressure (compare with Figure 12.8).

Figure 12.9 shows the air distribution in the cabin.

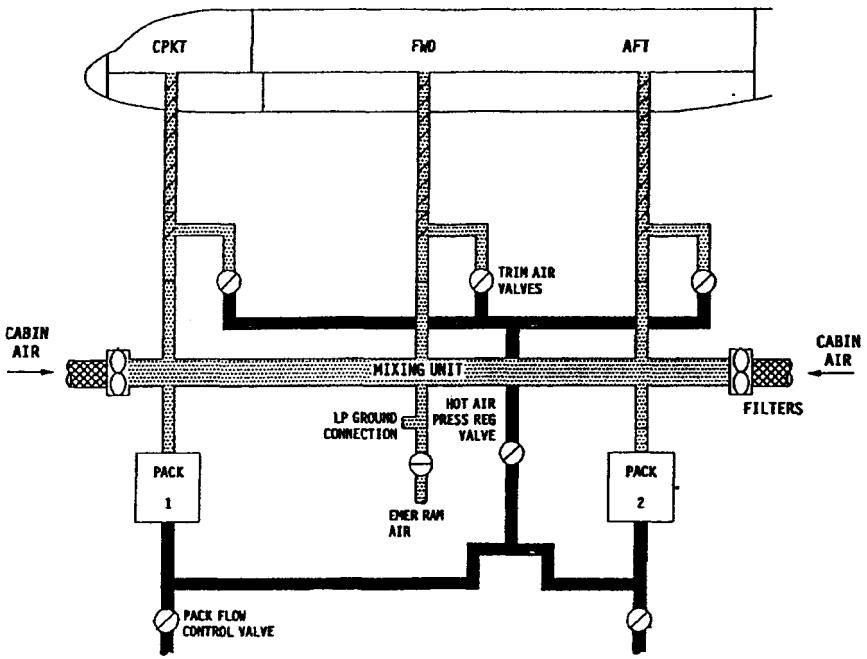


FIGURE 12.7 A321 air conditioning.

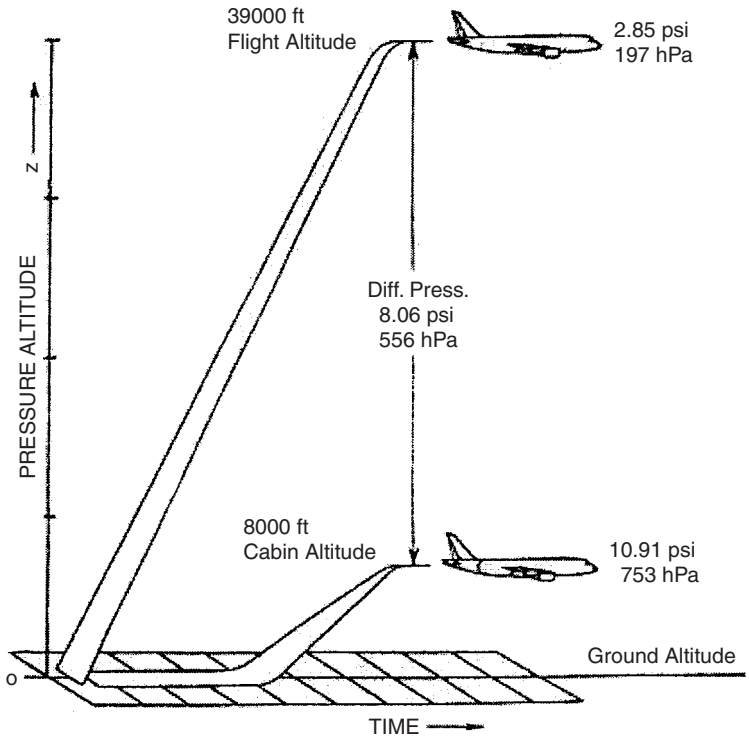


FIGURE 12.8 A321 pressure control.

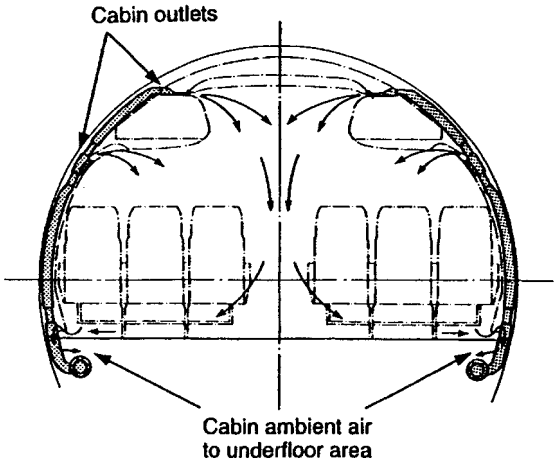


FIGURE 12.9 A321 cabin air distribution.

12.3 ELECTRICAL POWER (ATA 24)

Electrical power as defined by ATA 100:

> Those electrical units and components which generate, control and supply AC and/or DC electrical power for other systems, including generators and relays, inverters, batteries, etc., through the secondary busses. Also includes common electrical items such as wiring, switches, connectors, etc.

System Classification

Electrical power includes (ATA 100):

- Power generation:
  - Generator drive systems: constant speed drives (CSD)
  - Alternating current (ac) generation
  - Direct current (dc) generation
  - External power
- Power distribution:
  - Alternating current (ac) electrical load distribution
  - Direct current (dc) electrical load distribution

Power Generation

Power is generated with different electrical components. Light aircraft use 14 V or 28 V direct current (dc) generators or alternators. Large aircraft employ generators that produce an alternating current (ac) of 115 V at 400 Hz. Compared to a 28 V dc system, a higher-voltage ac system will develop several times as much power for the same weight and hence provide a great advantage where heavy electrical loads are imposed.

Aircraft *dc generators* have for the most part been replaced by dc alternators on modern aircraft. Although generators and alternators are technically different, the terms alternator and generator are used interchangeably.

A *starter-generator* is a combination of a dc generator and a dc motor in one housing. Starter-generators are typically employed on small turboprop and turbine-powered aircraft.

There are two major types of *alternators* currently used on aircraft: the dc alternator and the ac alternator. Dc alternators are most often found on light aircraft where the electric load is relatively small. Ac alternators are found on large commercial airliners and many military aircraft.

Both ac and dc alternators for aircraft show a construction with a rotating field (supplied with current from the outside via slip rings) and a stationary armature. The aircraft alternator is a three-phase unit having three separate windings 120° apart.

Light airplanes use an alternator with a three-phase full-wave rectifier to produce dc power. The rectifier is built into the alternator, so that dc current leaves the alternator with a nominal voltage of either 14 V for a 12 V battery system or with 28 V for a 24 V battery system.

Transport category aircraft use three-phase ac alternators with Y-connected stator windings. (Note: high output ac alternators are mostly called ac generators. If they are of a design without slip rings, they are called *brushless generators*.) The output frequency depends on the drive speed of the generator. The required constant frequency of 400 Hz requires the use of a constant speed drive (CSD). The integrated drive generator (IDG) contains both, the CSD and the generator in one unit. Details of this state-of-the-art system are explained using the Airbus example below.

Advantages of ac high-voltage systems include:

- Weight savings
- Voltage transformation possibilities
- Low current, low power losses in the wiring

Electrical power generation systems on large aircraft show a range of typical components:

- A *generator control unit* (GCU) is a solid-state device that carries out voltage regulation, current limiting, and frequency control.
- An *inverter* is a device for converting direct current into alternating current at a demanded frequency (400 Hz). A *static inverter* achieves this with standard electric and electronic components.
- A *transformer rectifier* (TR) unit is a device for converting alternating current into direct current.
- A *variable-speed constant-frequency* (VSCF) system employs a generator driven directly from the engine without a constant-speed drive (CSD). The generator is driven at variable engine speeds, thus producing a variable-frequency output. A generator converter control unit converts the variable frequency into a constant frequency of 400 Hz. A VSCF system is found on the Boeing 737.

Power Distribution

The design of the power distribution system depends on

1. The size of the aircraft and hence upon its system complexity
2. On the type of primary power generation applied (ac or dc)

A simple power distribution system consists of a *bus bar* or *bus*. The bus is a conductor designed to carry the entire electrical load and distribute that load to the individual power users. Each electric power user is connected to the bus through a circuit breaker. Simple distribution systems like this are found on small single-engine aircraft.

More complex power distribution systems consist of bus bars, bus tie breakers, and various solid-state controllers such as *generator control units* (GCUs).

Electrical power distribution systems on large aircraft show a range of typical components:

- *Bus tie contactors* (BTCs) (also known as bus tie breaker) are electric solenoids used to connect two bus bars.
- *Generator line contactors* (GLCs) (also known as generator breakers) are similar to BTCs but connect the generators to the buses.

- *Bus power control units* (BPCUs) are supplied with information from all parts of the distribution system. Taking this information into account, BPCUs will ensure the appropriate distribution system configuration. In some architectures, the GCUs include the BPCU functions. The BPCUs enable *reconfiguration* of the power distribution between individual busses. For example, if a generator fails or a bus shorts to ground, the appropriate BTCs and GLCs must be set to the correct position. In the event of a system overload, the controller must reduce the electrical load to an acceptable level. This is called *load shedding*. The aircraft's galley power is usually the first nonessential load to be disconnected.

Figure 12.10 shows the two principal distribution systems with

1. Primary ac generation and dc generation through transformer-rectifiers
2. Primary dc generation and ac generation through inverters

Three different power distribution systems exist for large aircraft, all of which apply primary ac generation:

1. The split-bus system
2. The parallel system
3. The split parallel system

The *split-bus system* (Figure 12.11) contains two completely isolated power-generating systems. Each system contains its own ac generator. The generator 1 (GEN 1) and generator 2 (GEN 2) power their respective loads independently of other system operations. In the event of a generator failure, the remaining operating generator is connected to both buses AC 1 and AC 2, or the APU generator (APU GEN) may be employed to carry the electrical load of the inoperative generator. The major advantage of a split-bus system is that the generators operate independently, so that generator output frequencies and phase relationships need not be so closely regulated. A split-bus system is used on the Airbus A321 and

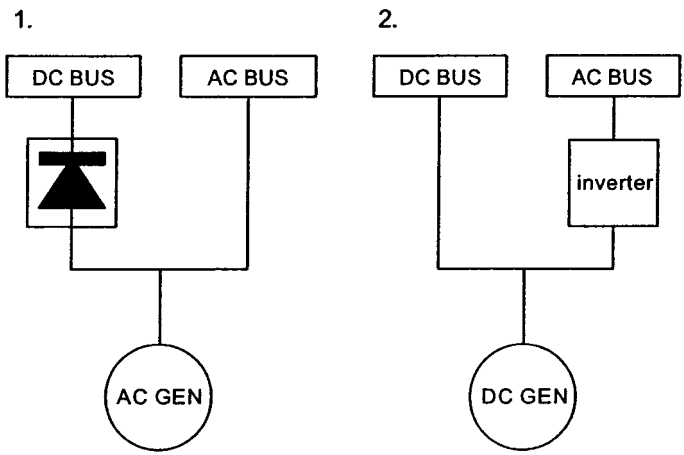


FIGURE 12.10 Distribution systems with 1. primary ac power generation; 2. primary dc power generation.

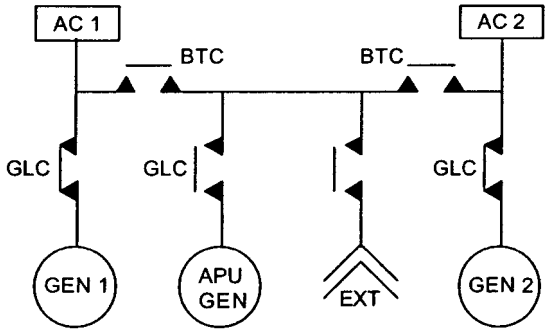


FIGURE 12.11 General layout of a split-bus system.

most other modern twin-engine transport category aircraft. The A321's electrical system diagram is shown below in more detail.

In a *parallel system* (Figure 12.12), all ac generators are connected to one tie bus. This type of system maintains equal load sharing for three or more ac generators. Since the generators are connected in parallel to a common bus, all generator voltages, frequencies, and their phase sequence must be within very strict limits to ensure proper system operation. If one generator fails, the generator is isolated from its load bus. Nevertheless, that load bus still continues to receive power while connected to the tie bus. A parallel system is used on, for example, the Boeing 727.

A *split parallel system* (Figure 12.13) allows for flexibility in load distribution and yet maintains isolation between systems when needed. The ac buses are paralleled through the bus tie breakers (BTB) and the split system breaker (SSB). When

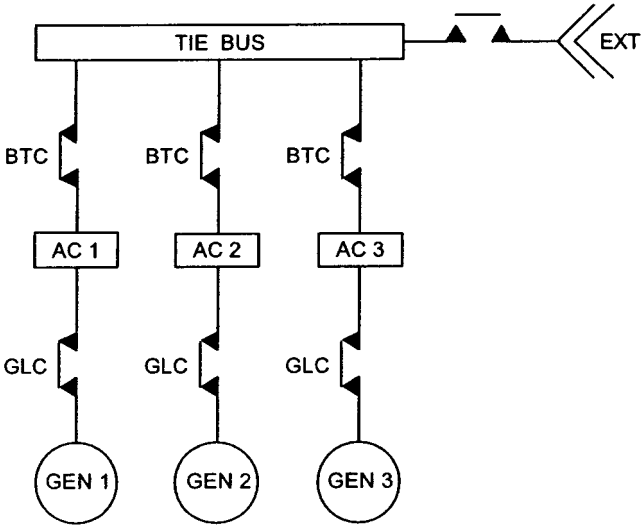


FIGURE 12.12 General layout of a parallel system.

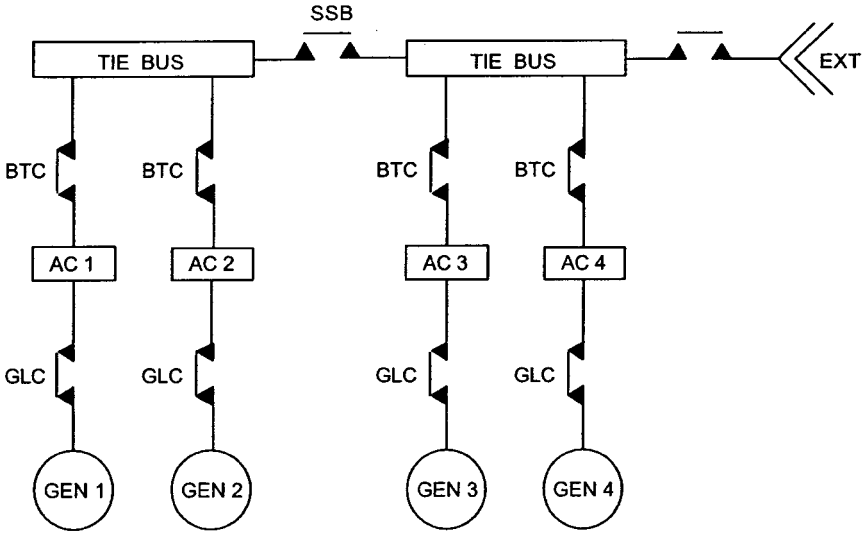


FIGURE 12.13 General layout of a split parallel system.

the SSB is open, the right system operates independently of the left. With this system any generator can supply power to any load bus (AC 1, AC 2 . . .), and any combination of the generators (GEN 1, GEN 2 . . .) can operate in parallel. A split parallel system is used on the Boeing 747-400.

Let's look at the *dc distribution systems* on aircraft with primary ac power generation. Transformer rectifiers (TRs) powered by an ac bus, feed their main dc bus bars. In the event of a complete generator system failure, the aircraft's batteries would supply the essential dc power. An inverter would also be powered from the batteries in an emergency situation to operate all essential ac loads.

The aircraft electrical system is designed with a *power distribution hierarchy.* The system is designed so that the most critical components are the least likely to fail. The generators feed their respective bus AC 1, AC 2 . . . The least critical ac loads are powered by these busses. The critical ac loads are powered by the essential ac bus (AC ESS). The same is true for the dc busses: the least critical dc loads are powered by the DC 1, DC 2 . . . buses, which are fed by their respective transformer rectifier (TR 1, TR 2 . . .). The next-most critical systems are powered by the essential dc bus (DC ESS), which can be powered by any transformer rectifier. The most critical loads are powered by the battery bus (BAT BUS).

Example: Airbus A321

In the A321, primary ac power generation is applied, where ac is converted to dc by means of transformer rectifiers (TRs). The distribution system is a split-bus system and consists of two separated distribution networks. Normally, one main generator supplies each network. The two distribution networks may be connected when the aircraft is on external power, APU power, or if one main generator fails. Under no circumstances may two generators be connected.

A321 *power generation* encompasses primary ac power generation in flight and on the ground, dc power generation, and ac power generation from dc. The location of related components in the aircraft is shown in Figure 12.14.

In flight, two engine-driven generators (GEN 1 and GEN 2), also known as integrated drive generators (IDGs), supply the aircraft electrical power system. A third APU-driven generator (APU GEN) can replace one engine-driven generator. In the event of a major failure, a unit consisting of a constant-speed hydraulic motor coupled to a generator (constant-speed motor/generator, CSM/G) is able to supply the most essential parts of the electrical systems. The CSM/G is powered by the ram air turbine (RAT) via the Blue hydraulic system.

On the ground, an external electrical ground power unit (GPU) can supply the aircraft. Alternatively, the APU generator can serve as an independent source for electrical power supply on the ground.

All the power sources named above supply the distribution network with ac power. Dc power is supplied by transformer rectifiers (TR). Two batteries are used as a dc emergency power source and for APU start in flight and on the ground.

Essential ac power can be obtained in an emergency situation from the batteries through a static inverter.

A321 *power distribution* encompasses (Figure 12.15):

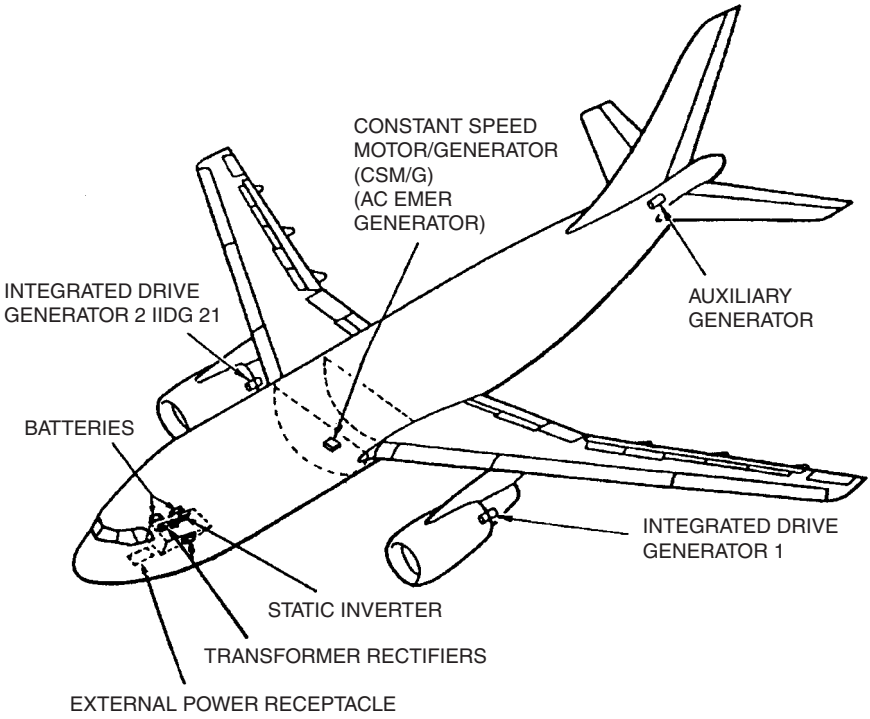


FIGURE 12.14 A321 electrical power sources and their location in the aircraft.

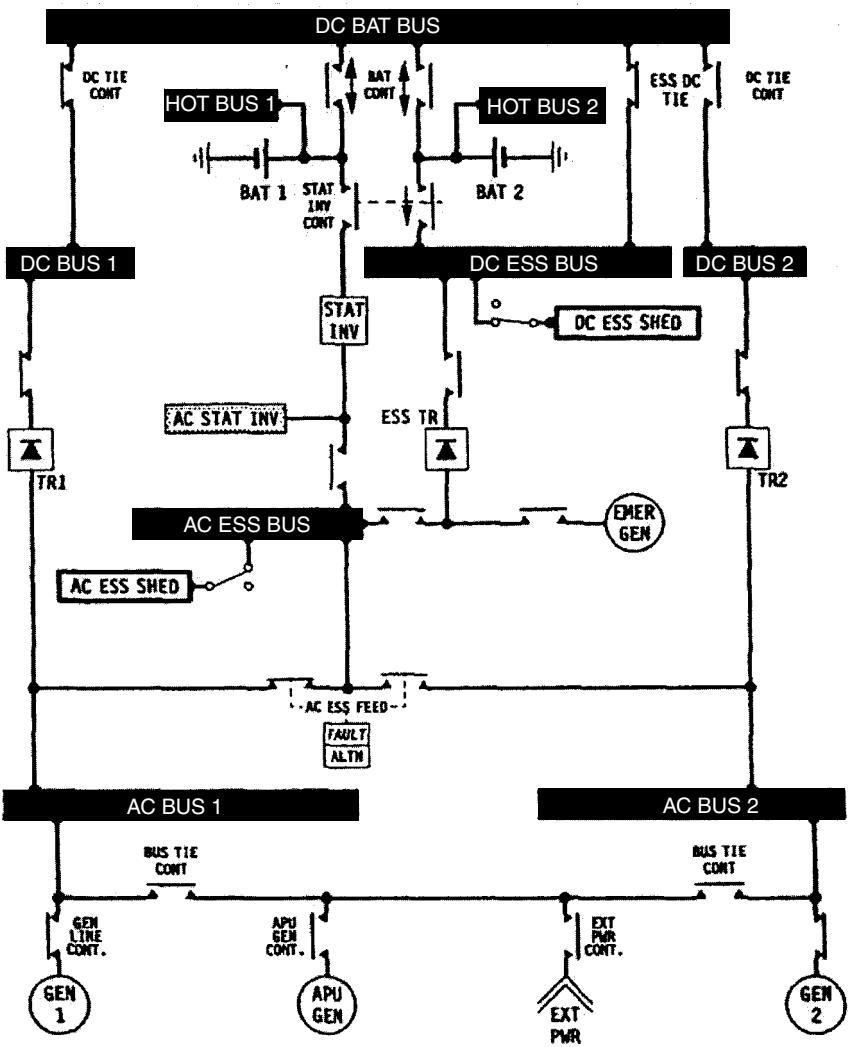


FIGURE 12.15 A321 electrical system diagram.

- The distribution network 1, which consists of AC BUS 1, AC ESS BUS, AC ESS SHED. The AC ESS SHED may be shed due to a lack of power in an emergency.
- The distribution network 2, which consists of AC BUS 2.
- The transformer rectifier 1 (TR 1), which is powered from the AC BUS 1 supplies through its contactor: DC BUS 1, DC BAT BUS, DC ESS BUS, DC ESS SHED. The DC ESS SHED may be shed due to a lack of power in an emergency.
- Two batteries, which are associated with the DC BAT BUS.

- The transformer rectifier 2 (TR 2), which is powered from the AC BUS 2 supplies through its contactor the DC BUS 2.
- A third essential transformer rectifier (ESS TR), which can be powered from the AC BUS 1, or the emergency generator (EMER GEN) may supply the DC ESS BUS and the DC ESS SHED through its contactor only in certain failure cases.

In failure cases, various possibilities for reconfiguration exist.

Each engine high-pressure stage drives its associated *integrated drive generator* (IDG) through the accessory gearbox (Figure 12.16). The drive speed varies according to the engine rating. The IDG provides a 115/200 V, three-phase, 400 Hz AC supply. The IDG consists of two parts: the constant-speed drive (CSD) and the generator. The hydromechanical CSD drives the ac four-pole generator at a nominal speed of constant 12,000 rpm.

The *constant-speed drive* (CSD) consists of a mechanical differential gear that transmits power to the generator of the IDG. The output speed of the differential gear is modified by two mechanically coupled twin hydraulic subassemblies: a pump and a motor. Each subassembly includes a hydraulic swashplate: the pump is equipped with a variable-angle swashplate, and the motor is equipped with a fixed swashplate. A governor controls the CSD output speed by the swashplate angle of the pump (Figure 12.17).

The *generator* is a three-stage assembly that includes three machines connected in cascade. The first machine is a 12-pole permanent magnet generator (PMG). The second machine is a 10-pole stator and receives its field excitation from the first machine via the voltage regulator in the generator control unit (GCU). Its dc output feeds the rotating field of the third machine (the main alternator). The main alternator has a three-phase star-connected stator winding. The three phases and star point are taken to the generator output terminal block.

12.4 EQUIPMENT/FURNISHINGS (ATA 25)

Equipment and furnishings as defined by ATA 2200:

> Those removable items of equipment and furnishings contained in the flight and passenger compartments. Includes emergency, galley and lavatory equipment. Does not include structures or equipment assigned specifically to other [systems].

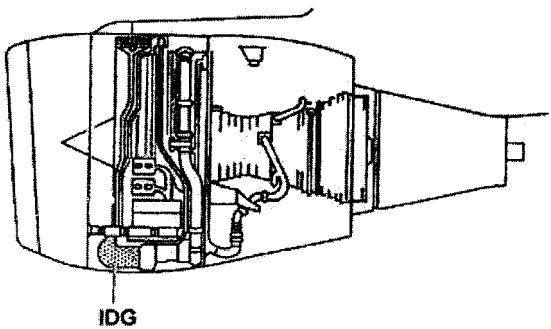


FIGURE 12.16 A321: location of the integrated drive generator (IDG).

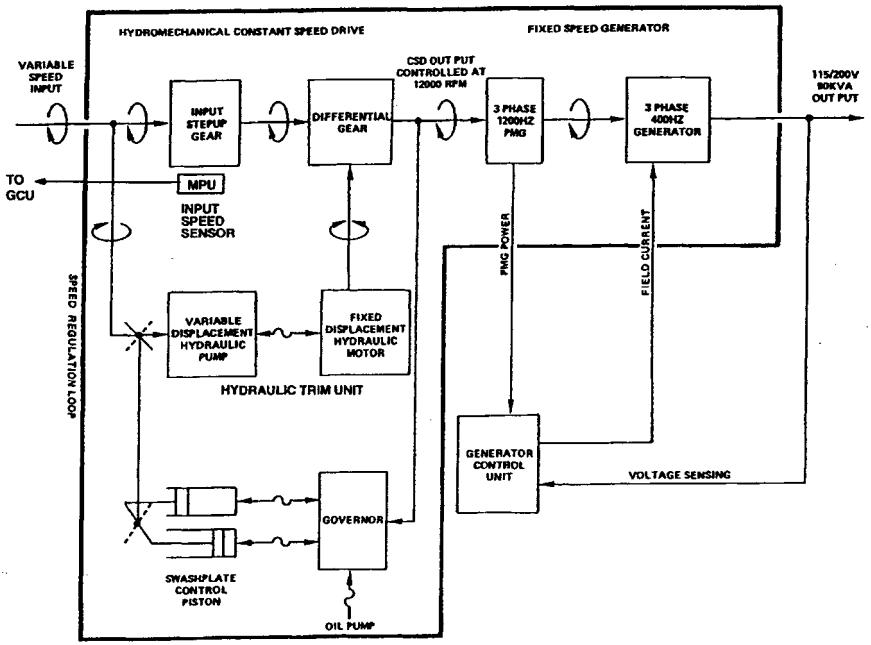


FIGURE 12.17 A321 integrated drive generator (IDG): speed conversion and power generation.

Elements of Equipment

Equipment and furnishings include items in several parts of the aircraft. Examples of such equipment include:

- In the flight compartment: flight crew seats, tables, wardrobes, electronic equipment racks, and stowage facilities for manuals and other equipment.
- In the passenger compartment: seats, overhead storage compartments, wall coverings, carpets, wardrobes, movable partitions.
- In buffets and galleys: cabinets, ovens, refrigerators, coffee maker, electrical outlets and wiring, trolleys, garbage containers.
- In the lavatories: mirrors, seats, cabinets, dispensing equipment, electrical outlets and wiring (the wash basin and the closets are part of the water/waste system).
- In the cargo compartment:[12] equipment used to load and unload the aircraft; includes restrains and latches, rollers, and drive systems.
- In all parts of the aircraft, thermal insulation[13] minimizes the losses of heat from the fuselage, stops the formation of condensation, and reduces the noise level in the fuselage. Thermal insulation is dimensioned in conjunction with the design of the air conditioning system.

[12] Under the new ATA 2200, allocated to "Cargo and Accessory Compartment (ATA 50)."

[13] Under the new ATA 2200, allocated to "Cargo and Accessory Compartment—Insulation (ATA 50-60)."

Some aircraft, especially very large commercial transports, also offer space for additional equipment in the *under floor area.* The space can be used for crew rest facilities, galleys, a bar, or an exercise room. The need might arise to incorporate an elevator (lift) in multideck aircraft in order to move goods or passenger.

Emergency equipment includes items for use in emergency procedures, such as evacuation equipment, life rafts, jackets, crash ax, flashlights, megaphone, protective gloves, emergency locator transmitters, underwater locator devices, first aid kits, and supplementary medical equipment. Fire extinguishers and oxygen equipment are part of their respective systems. Evacuation equipment facilitates passenger and crew evacuation. These procedures are explained below.

Cabin Design

The cabin is the place where the paying customer has to be satisfied. Much attention is given to its design, starting during aircraft design, where an optimum cabin cross-section has to be found. Designers have to find ways to create an aesthetically pleasing impression and a suggestion of spaciousness within the always limited dimensions of an aircraft (Figure 12.18). These design activities have an influence on the shape of ceiling panels, sidewall panels, stowage compartment doors, and passenger service units (PSUs) located underneath the stowage compartment. Cabin lighting design is also part of this effort. The airlines would like to see their corporate design reflected not only outside but inside the aircraft. They may choose their own material, pattern, and texture for panel coverings, dividers, curtains, and seats and will select a suitable carpet. All cabin materials have to fulfill requirements related to fire, wear, and cleaning.

FIGURE 12.18 Boeing 717: the result of a thorough cabin design (Granzeier 2001).

Passenger Seats

Passenger seats are probably the most important single item of equipment in the cabin. They should provide comfortable seating for many hours during normal flights and the best protection during a crash. Elements of a seat are shown in Figure 12.19. Not visible in the figure are the literature pocket and the folding table on the back of the seat. Seats are installed on seat tracks in the cabin floor structure. This allows flexibility in spacing the seats.

Seat pitch is a comfort measure for seat spacing. It is the distance between corresponding points on two seats installed one in front of the other. The seat pitch is internationally given in inches. Seats in first, business, and economy class feature different levels of comfort, and the seat pitch also varies among these classes. Typical values today are:

- First class: 62 in. (1.57 m)
- Business class: 40 in. (1.02 m)
- Economy class: 32 in. (0.81 m)
- High density: 30 in. (0.76 m)

These numbers are not fixed, but change with product policy of the airlines. During the last decades seat pitch has increased in first class, but decreased in economy class in a fight for low fares.

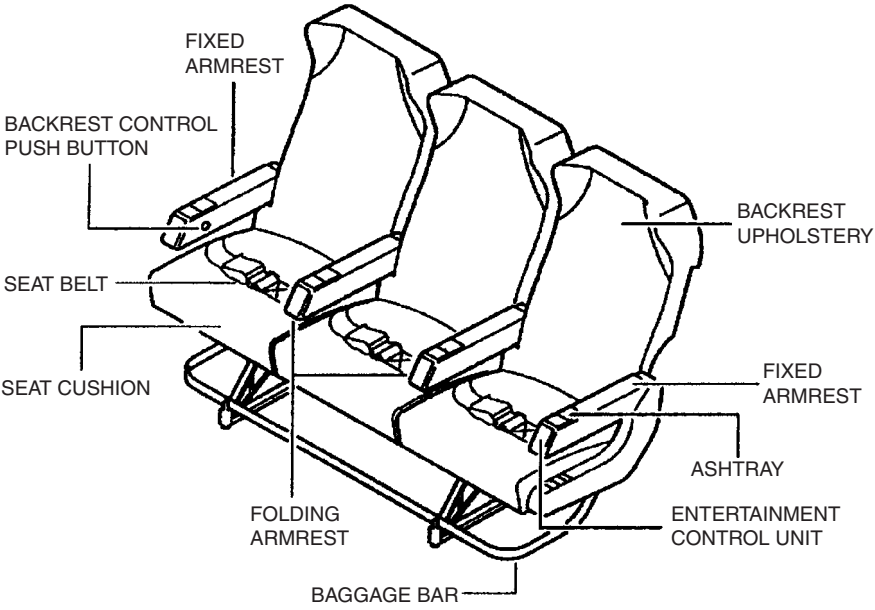


FIGURE 12.19 Economy class passenger seats (A321).

Seats are bought by the airline from specialized seat manufacturers as buyer-furnished equipment (BFE) and are then installed by the aircraft manufacturer in the new aircraft.

Emergency Evacuation

Rapid evacuation of passengers and cabin crew has to be possible in case of a crash landing. For airplanes with 44 passengers or more it must be shown that passengers and cabin crew can be evacuated to the ground within 90 seconds, with up to 50% of the emergency exits blocked (JAR-25, section 803; AC 25.803). In an emergency, passengers usually leave the aircraft through *emergency exits* (these can also be the normal passenger doors) via inflatable escape slides (Figure 12.20).

Evacuation of flight crew from commercial aircraft designed to be achieved through passenger emergency exits, through a *hatch,* or by using an escape rope to slide down from the flight deck through the opening side *windows.*

Evacuation of crew from military combat aircraft is usually achieved with *ejector seats* that allow the crew to abandon their aircraft at all flight conditions, ranging from high speed, high altitude to zero speed and zero height. The ejector seat is mounted in the aircraft on a slide rail and is propelled out of the aircraft by a rocket motor. After a predetermined time, the seat detaches from the person, who is

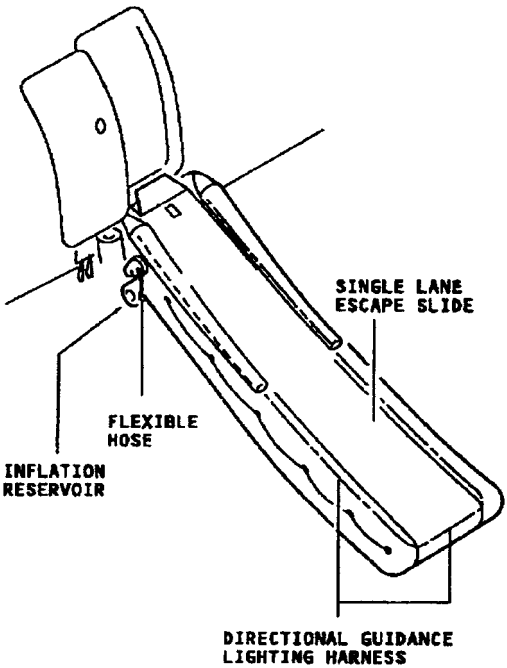


FIGURE 12.20 Escape slide (Airbus A321).

brought to the ground by parachute. In some multicrew combat aircraft the crew are evacuated in an *escape module* that is jettisoned and parachuted to the ground.

Example: Airbus A321

Equipment and furnishings give comfort and safety to passengers in the cabin and to the crew in the cockpit. Equipment is also used for handling of cargo in the cargo compartments.

The cockpit is equipped with adjustable seats for two crew members (Figure 12.21). The A321 has a fly-by-wire flight control system steered with a side stick. The side stick armrest located on the outboard side of the seat can be adjusted in height and tilt angle so that the pilots can rest their respective arm in an optimum position with respect to the side stick controller. A third occupant seat and a folding seat for a fourth occupant are also available.

The cabin also includes the galleys (Figure 12.22) and lavatories (Figure 12.23), in addition to the passenger seats (Figure 12.19).

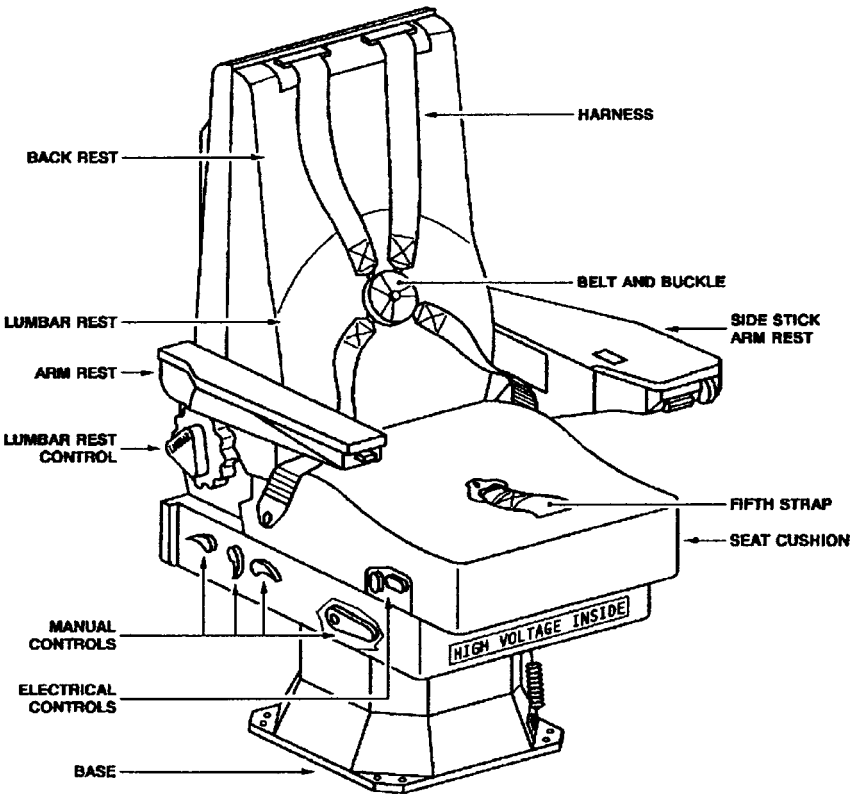


FIGURE 12.21 A321 captain/first officer seat.

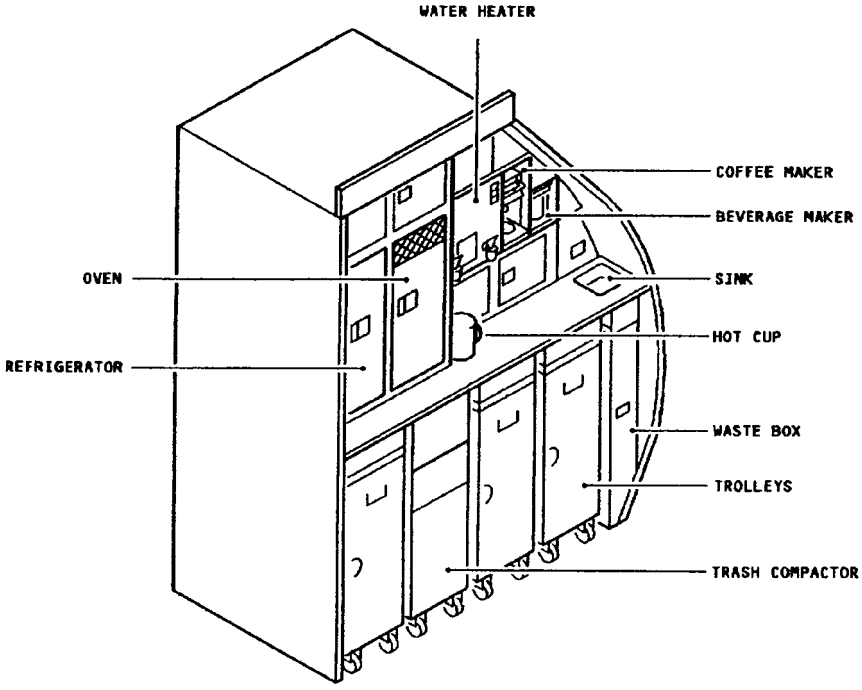


FIGURE 12.22 Galley equipment (A321).

12.5 FIRE PROTECTION (ATA 26)

Fire protection equipment as defined by ATA 100:

> Those fixed and portable units and components which detect and indicate fire or smoke and store and distribute fire extinguishing agent to all protected areas of the aircraft; including bottles, valves, tubing, etc.

Detection Fundamentals

Fire detection includes that part of the fire protection system which is used to sense and indicate the presence of overheat, smoke, or fire (ATA 100).

There are various ways in detecting a fire, including:

- Direct observation by cockpit and cabin crew (optical indication, sensing of heat or smell)
- Overheat detector
- Smoke detector
- Rate-of-temperature-rise detector

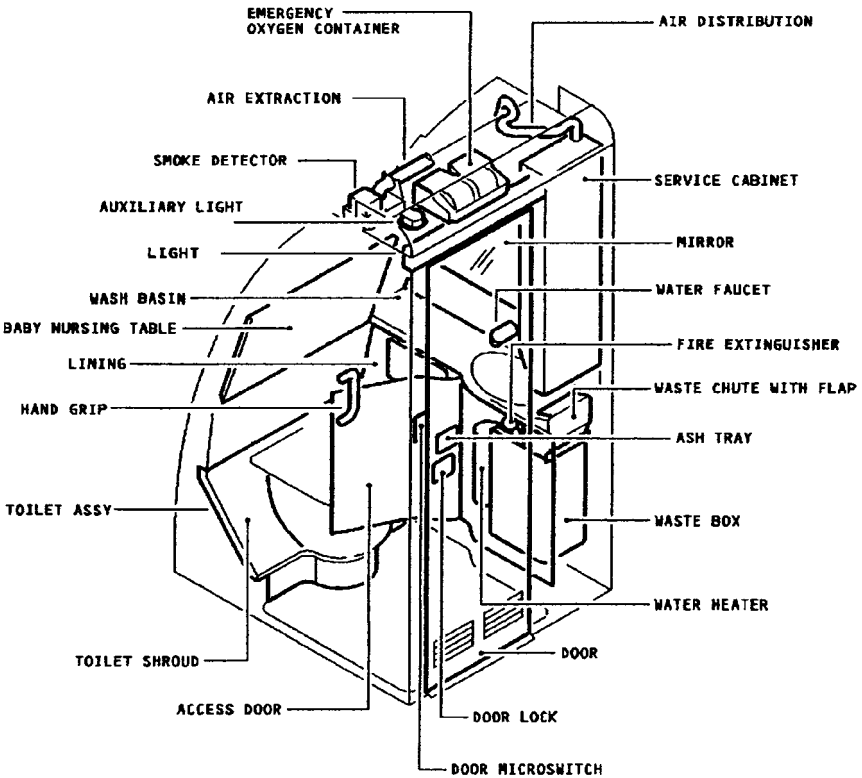


FIGURE 12.23 Lavatory equipment (A321).

- Inspection by video camera
- Fiberoptic detectors
- Thermal imaging devices
- Radiation sensing devices
- Ultraviolet aircraft fire detection system
- Detection of combustion gases like CO or CO_2

Designated fire zones must be equipped with fire detection and extinguishing equipment. Designated fire zones are (JAR-25, FAR Part 25):

- Power plant compartment (section 1181)
- Auxiliary power unit (APU) compartment (section A1181)
- Combustion heater chamber (section 859)

Fire detection and extinguishing equipment is required for cargo compartments according to the *cargo compartment classification* (section 857, JAR-25, FAR Part 25):

- Class A compartments are accessible in flight. A fire in the compartment would be easily discovered by a crew member while at his station.
- Class B compartments provide access in flight to enable a crew member to use a hand fire extinguisher. The compartments are equipped with a smoke or fire detector.
- Class C compartments are equipped with a smoke or fire detector and a built-in fire extinguishing system.
- Class D compartments are able to confine a fire completely without the safety of the aircraft being endangered.

Lavatories must be equipped with a smoke detector system, and lavatories must be equipped with a built-in fire extinguisher for each disposal receptacle for towels, paper, or waste located within the lavatory (section 854, JAR-25, FAR Part 25).

Other areas equipped with fire detectors may include the avionic compartment or the landing gear bay.

Fire detectors are generally either overheat detectors or smoke detectors. From the beginning until today, these and other fire-detection devices for aircraft have been developed by only a few U.S. companies: Walter Kidde, Fenwal, and Systron-Donner. Their component designs will be presented here (Hillman et al. 2001). The roadmap to the following discussion of the most widely used detection devices is presented in Figure 12.24.

Overheat Detection

In the 1940s, overheat detection coverage in the engine nacelle was done with thermal switches or thermocouples. Several of these switches were positioned in

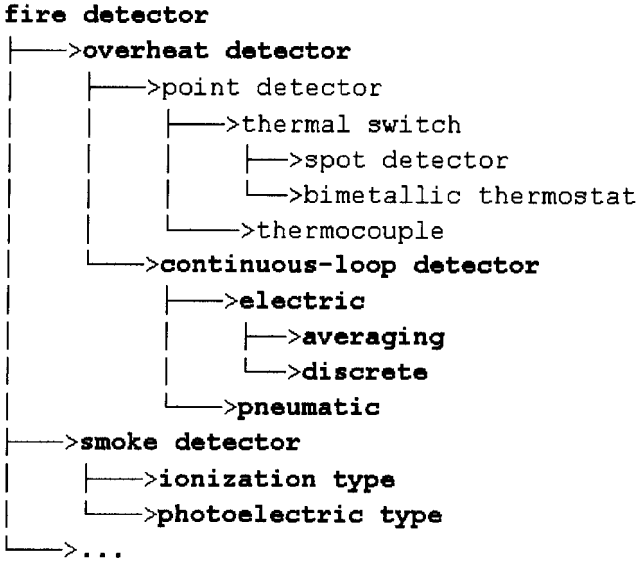


FIGURE 12.24 Roadmap to the most widely used detection devices.

parallel at different places around the engine. A fire alarm was activated if one of the switches was triggered. However, it was recognized that these *point detectors* were very limited with regard to area of coverage. The placement of the point detector therefore became the most critical factor in how successful the detection system would be.

In the early and mid 1950s, *continuous-loop detectors* were introduced in the aircraft industry. This technology became the most popular detection approach for aircraft engines and has remained so to this day. Continuous-loop detectors are either electric or pneumatic continuous-loop detectors. Electric continuous-loop detectors are of either averaging type or the discrete type (Figure 12.25).

Some versions of electric continuous-loop detector depend on the amount of element heated to reach their alarm threshold level. These have been termed *averaging electrical continuous-loop detectors.* Their alarm threshold averages the temperature over its entire length. These detectors monitor either changing electrical resistance alone or resistance and capacitance in conjunction. Electrical continuous-sensing elements have one or two internal wire conductors embedded in a ceramic-like thermistor material contained in a metallic outer tube. As the surrounding temperature increases, the resistance between the inner conductor and the outer tube conductor decreases while the capacitance increases. When two internal wire conductors are embedded in the sensing element, the resistance change between these two wires is typically measured. When the resistance between the internal conductor and the external sensing element tube drops to some predetermined level (and/or the capacitance increases) corresponding to the desired alarm temperature, a monitoring control unit issues a hazard signal. When the hazard condition is eliminated and the temperature returns to normal, the resistance increases and the

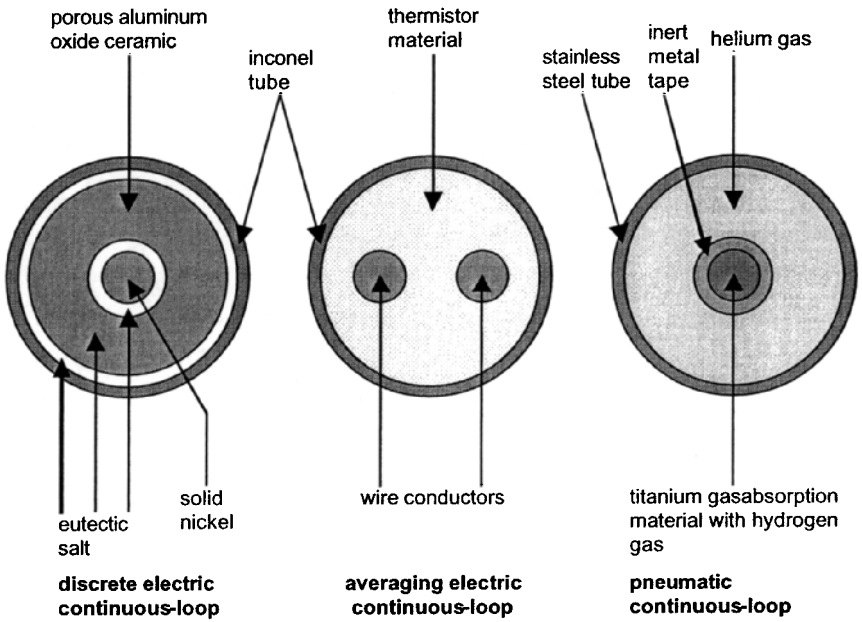


FIGURE 12.25 Cross-section of continuous-loop detectors.

capacitance decreases, thereby canceling the alarm. Multiple trip resistance/capacitance settings can be used when multiple thresholds are pursued to indicate fire versus overheat.

Shortly after the first averaging-type detection systems, *discrete electrical continuous-loop detectors* were introduced (Figure 12.26). To achieve its alarm threshold, the discrete system utilizes sensing elements that are essentially independent of the length of element heated. These systems employ a sensing element which, as in the electrical averaging systems, has either one or two internal wire conductors embedded in a ceramic-like core material surrounded by a metallic outer tube. The ceramic core is impregnated with eutectic salt. The salt melts at its eutectic melt temperature, even when only a very short length of element is heated. When this occurs, the electrical resistance between the inner conductor and the outer tube very rapidly breaks down (also, the capacitance increases), and a monitoring control unit signals a fire or overheat, depending on which is appropriate for the intended application. The characteristics of the discrete type are paramount for reliable early warning of small, discrete overheat events, such as bleed air duct failures. By its nature, the discrete type cannot provide multiple alarm thresholds or any kind of analog temperature trend information.

Pneumatic-based continuous-loop detectors rely on increasing gas pressure to achieve the alarm threshold. These sensing elements have a hydrogen-charged core surrounded by helium gas, contained in a metallic outer tube. As the surrounding temperature increases, the helium gas pressure increases, closing a pressure switch and thereby issuing an alarm. As the temperature returns to normal, the pressure decreases and the alarm is canceled. If a localized high-temperature event is present, the hydrogen core also outgasses its hydrogen gas, increasing the internal pressure and closing the pressure switch. As the sensing element cools, the hydrogen absorbs back into the core so that the internal pressure decreases, removing the alarm output. A leak in the detector can be discovered with an integrity switch opening due to a loss of pressure (Figure 12.27).

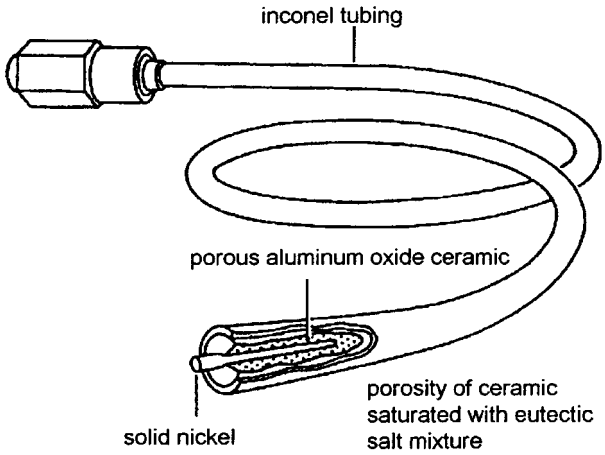


FIGURE 12.26 Discrete electric continuous-loop detector (A321, pneumatic system, leak detection).

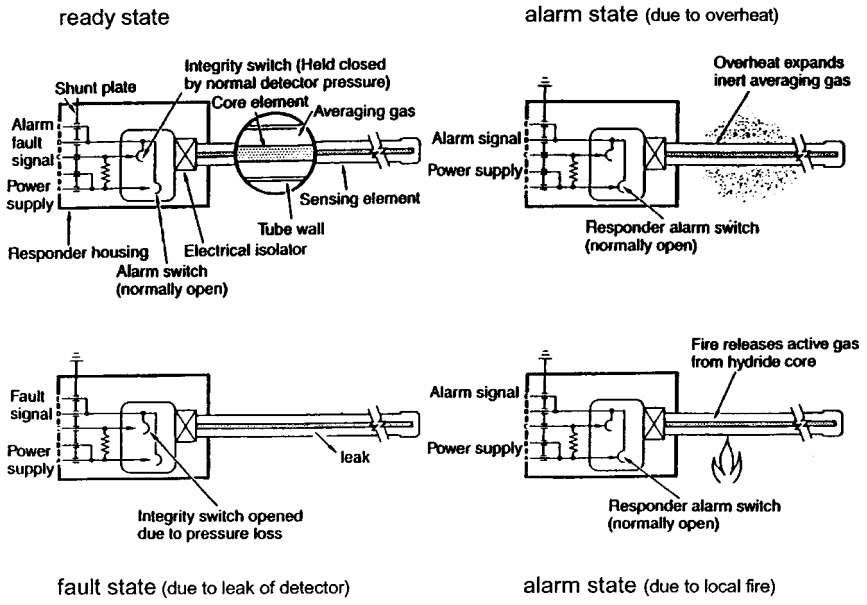


FIGURE 12.27 Principle of pneumatic continuous-loop detector (A321).

Overheat detection may be applied in the areas of the engine, auxiliary power unit (APU), bleed air ducts, and the landing gear bay.

Smoke Detection

Smoke detection systems are the primary means of fire detection used in cargo compartments. This has not changed much over the last 50 years. While solid state electronics and new optics and new processing algorithms have been introduced, the basic mechanism that these detectors operate under has remained the same. There are two basic designs of smoke detectors: ionization and photoelectric.

Ionization-type smoke detectors monitor ionized combustion byproducts as they pass through a charged electrical field. Photoelectric detectors measure light attenuation, reflection, refraction, and/or absorption of certain wavebands. Ionization smoke detectors have been used from the early years. The typical approach was to use a radioactive isotope as the source to charge the combustion products (Figure 12.28). However, this source may also charge everything else, including dust and fine water droplets, and can make ionization-type detectors unreliable. Ionization-type smoke detectors have been used by the commercial aviation community primarily in lavatories and cargo compartments.

Photoelectric-type smoke detectors have become the industry standard. This is not to imply that photoelectric-based detectors have been free from false alarms. These detectors, too, have been quite troublesome over the years. Most cargo compartment applications use aerospace-quality photoelectric-type smoke detectors that rely on scattered or reflected light radiation caused by a particulate matter between

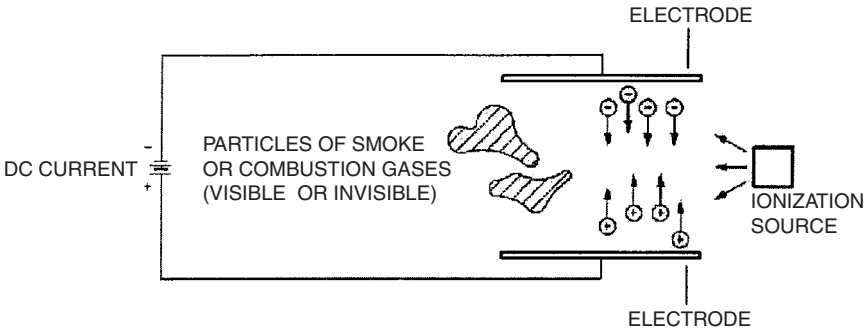


FIGURE 12.28 Principle of ionization-type smoke detector (A321).

a radiation-emitting source and a detector device. Solid state photoelectric smoke detectors use a long-life light-emitting diode (LED) as the source of light.

Smoke detectors still have many limitations. Their operational success depends highly on their placement with respect to where a fire event is. But there are also problems with other detectors. Since one cannot count on visual line of sight of a cargo bay fire, future cargo-detection technologies cannot rely on the use of video camera or thermal imaging devices. Deep-seated fires and/or fires inside LD3 containers will still be hidden. This makes standalone thermal-based systems impractical. While combustion gases, such as CO or CO_2, could be monitored, these gases may be introduced from sources other than fires.

Smoke detection can be applied in the cargo compartment, lavatories, galleys, and avionic compartments.

Extinguishing Fundamentals

Fire extinguishing includes that part of the fire protection system using fixed or portable systems used to extinguish a fire (ATA 100).

A *fire classification* includes three types of fire relevant to aircraft application:

- Class A: Fires involving ordinary combustible solid materials, such as wood, paper, rubber, and many plastics
- Class B: Fires involving flammable liquids, oils, greases, paints, lacquers and flammable gases
- Class C: Fires involving energized electrical equipment

Each of these types of fire requires its own suitable *type of extinguisher:*

- Water extinguishers are used on Class A fires only. Water must never be used on Class C fires and can be counterproductive on Class B fires.
- CO_2 extinguishers are specifically used to combat Class C fires. A hand-held CO_2 extinguisher includes a megaphone-shaped nozzle that permits discharge of the CO_2 close to the fire. Be aware that excessive use of CO_2 extinguishers robs a closed area of oxygen. In an aircraft, this could affect passengers.

- Dry chemical fire extinguishers can be used on Class A, B, or C fires. Use of such an extinguisher on the flight deck could lead to temporary severe visibility restrictions. In addition, because the agent is nonconductive, it may interfere with electrical contacts of surrounding equipment.
- Halon has almost exclusively been in use in portable aircraft fire extinguishers.

In the late 1940s, the very effective halogenated hydrocarbon (later termed *halon*) fire extinguishing agents were introduced. The primary agents used for fixed fire extinguishing systems were methylbromide (Halon 1001) and bromochloromethane (Halon 1011). Halon 1011 eventually displaced Halon 1001 for engine extinguishing systems primarily because of lower toxicity and corrosion.

The halons introduced in the early 1950s were less toxic than Halon 1011. Over the next 30 years, the higher-vapor pressure bromotrifluoromethane (Halon 1301) essentially displaced most of the Halon 1011. Because of the high vapor pressure of Halon 1301, the use of elaborate spray nozzles and spray bars was no longer required. The new Halon 1301 extinguisher systems were designed to discharge at a very high rate. This concept was called the *high rate discharge* (HRD) concept. The high rate discharge systems utilized halon pressurized to 600 psig (40 bar).

Hand-held dibromofluoromethane (Halon 1211) and/or water extinguishers have been the approved approach for accessible firefighting.

In recent years, due to international agreement on banning the production and use of ozone-depleting substances, including all the halons, the need for alternative extinguishing agents to the halons has arisen. However, the use of halons is still permitted for essential applications, such as aircraft, until a suitable replacement agent can be developed, approved, and certified for aircraft use. Until that time comes, existing stocks of halon, recovered from decommissioned fire protection systems, are sufficient to support many years of aircraft production and use. Upon review of alternative agents, it is evident that there is no clear winner with respect to a replacement for Halon 1301 in fire suppression systems that will use similar hardware and architecture. Each candidate has at least one characteristic that makes it inferior to Halon 1301.

Engine and APU Extinguishing

First step: The engine is shut down and combustible fluid entry (jet fuel, hydraulic fluid, and engine oil) into the engine compartment is stopped. This is necessary for the engine extinguisher to be effective. If the engine were not shut off, the fire would probably just relight after the extinguishing agent dissipated. Because of this practice, only multiengine aircraft utilize extinguishing systems.

Second step: The extinguishing agent flows from a pressure vessel through rigid pipes and is sprayed in the engine-protected zones.

Third step: If after some time (30 s) the fire warning still remains on, extinguishing agent from a second pressure vessel (if still available for that engine) may be used for further fire extinguishing.

The extinguishing agent is stored in high-pressure vessels commonly called *bottles*. A spherical-shaped pressure vessel design represents the most weight- and volume-efficient geometrical configuration for containing the greatest amount of agent. It is also the optimum shape with respect to stress levels in the vessel's

material. The spherical pressure vessel is the most popular design (Figure 12.29). Other details of the design are stated in section 1199 of JAR-25 and FAR Part 25.

APU fire extinguishing is technically similar to engine fire extinguishing, but the APU may only be equipped with one bottle.

Cargo Extinguishing and Inerting

Cargo compartments have traditionally been protected with hand-held fire extinguishers if the compartment was accessible and with a fixed Halon 1301 fire extinguishing/inerting system if the compartment was not accessible.

Like engine extinguishing systems, a cargo compartment suppression system is required to provide an initial peak volumetric agent concentration to knock down the fire. Since complete fire extinction cannot be assured, a cargo suppression system is required to maintain a lower concentration for some extended period of time. The compartment is thus inerted to prevent the fire from reigniting or growing. The

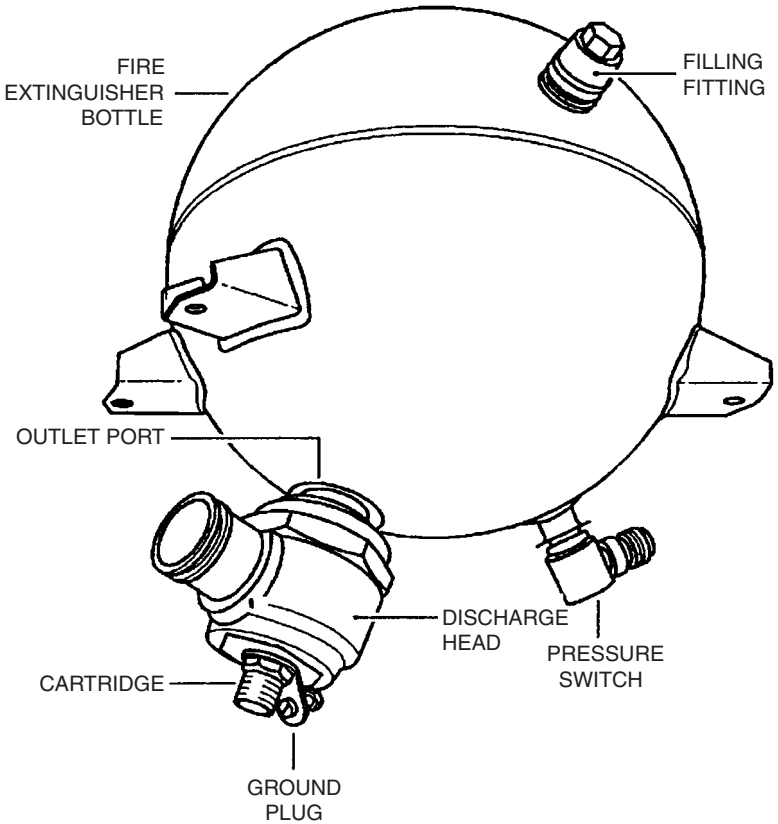


FIGURE 12.29 Fire extinguishing bottle (A321).

typical time period for keeping the compartment inert against flaming combustion is 60 minutes. In case of extended range twin-engine operations (ETOPS), inerting periods are much higher.

A typical cargo fire-suppression system will consist of two fire extinguishers connected to single or multiple cargo compartments by distribution plumbing. The knock-down or high rate discharge (HRD) extinguisher provides the initial high volumetric concentration, and the second low rate discharge (LRD) extinguisher provides the metered lower inerting concentration.

Passenger Compartment Extinguishing

Fires that could occur in an aircraft cockpit or cabin are Class A, B, and C. The number of hand-held fire extinguishers to be carried in an aircraft is determined by section 851 of the certification regulations (JAR-25, FAR Part 25).

For airplanes with a passenger capacity of 20 or more, each lavatory must be equipped with a built-in fire extinguisher for each disposal receptacle for towels, paper, or waste, located within the lavatory. The extinguisher must be designed to discharge automatically into each disposal receptacle upon occurrence of a fire in that receptacle (section 854, JAR-25, FAR Part 25).

Example: Airbus A321

For each engine, two fire extinguisher bottles contain fire extinguishing agent. The fire extinguisher bottles are connected to the extinguishing lines. The lines are routed in the pylon, leading to the outlet nozzles around the engine. The agent from the second bottle can be used if, after application of the first bottle, the fire warning remains on. The fire extinguisher bottles are controlled from the cockpit by pressing the DISCH (discharge) button. This supplies 28 V dc to two filaments in the cartridge on the bottle (see Figure 12.30). The filaments ignite 400 mg of explosive powder, which in turn causes rupture of the frangible disk in the cartridge and frees the agent with a high discharge rate.

12.6 FLIGHT CONTROLS (ATA 27)

Flight controls (the flight control system) are addressed in two other Sections of this handbook. Please consult Sections 10 and 11.

Flight controls as defined by ATA 100:

> Those units and components which furnish a means of manually controlling the flight attitude characteristics of the aircraft, including items such as hydraulic boost system, rudder pedals, controls, mounting brackets, etc. Also includes the functioning and maintenance aspects of the flaps, spoilers and other control surfaces, but does not include the structure.

Flight controls extend from the controls in the cockpit to the control surface actuators. The definition reads ‘‘means of *manually* controlling’’; this sets the flight control system apart from the auto flight system. Thus, the flight control system is

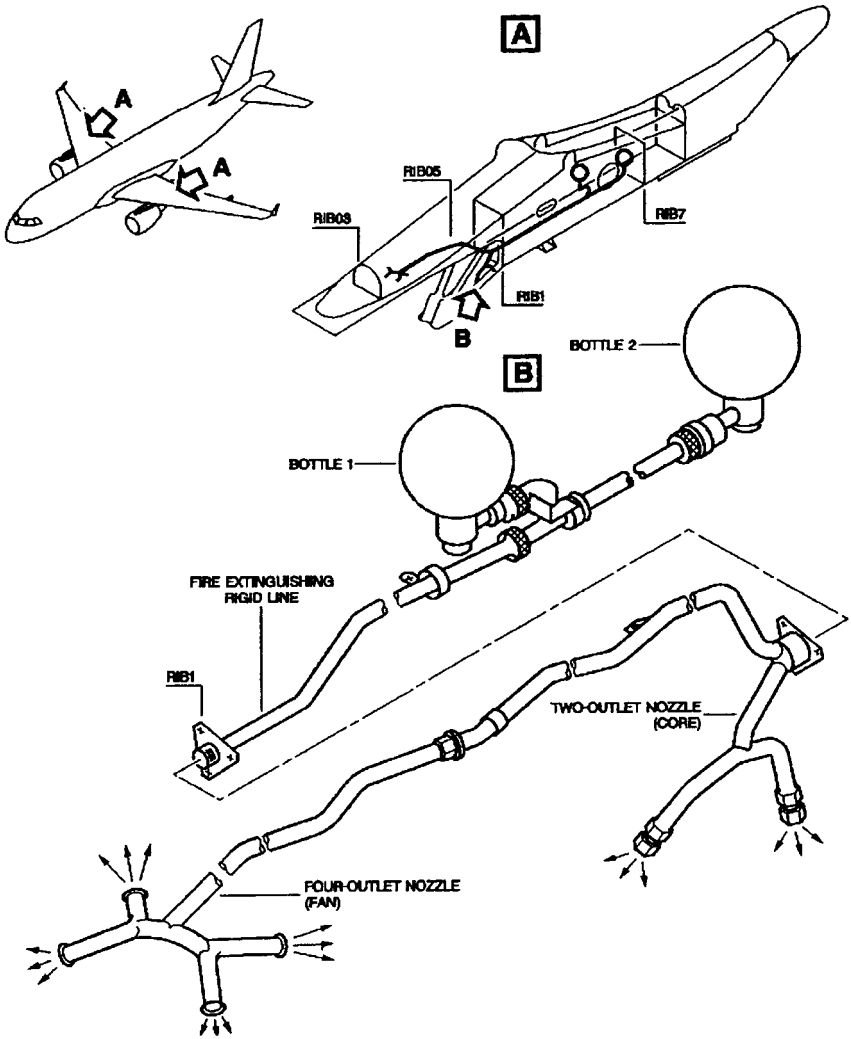


FIGURE 12.30 A321 engine fire extinguishing distribution system.

concerned only with direct inputs from the pilot via control column, rudder pedals, or other such control devices and the transformation of these inputs to adequate control surface movements.

Flight controls are subdivided into the mechanical aspects of the system and—in case of fly-by-wire (FBW) aircraft—the electronic (avionic) part.

Following ATA 100, the *mechanical subsystems* include:

- The ailerons
- The rudder

- The elevator
- The spoilers
- The horizontal stabilizer
- The high-lift system
- Gust locks and dampers

The *electronic* (avionic) *subsystem* is the Electronic Flight Control System (EFCS).

Even in modern FBW aircraft there exist many mechanical parts because in the end control surfaces have to be moved against heavy air loads in limited time. The high-lift systems (flaps and slats) also show a considerable amount of mechanical parts.

See the References and Further Reading for more on the mechanical aspects of modern flight control systems design.

12.7 FUEL (ATA 28)

The fuel system as defined by ATA 100:

> Those units and components which store and deliver fuel to the engine. Includes engine driven fuel pumps for reciprocating engines, includes tanks (bladder), valves, boost pumps, etc., and those components which furnish a means of dumping fuel overboard. Includes integral and tip fuel tank leak detection and sealing. Does not include the structure of integral or tip fuel tanks and the fuel cell backing boards which are [part of the structure], and does not include fuel flow rate sensing, transmitting and/or indicating, which are covered [by the powerplant systems].

Fuel—General

The purpose of the fuel system is to provide reliably the proper amount of clean fuel at the right pressure to the engines during all phases of flight and during all maneuvers.

The fuel system includes (ATA 100) all components necessary to achieve

- Fuel storage (tanks, components for tank ventilation, over-wing filler necks and caps)
- Fuel distribution (all components from the filler to the tank and from the tank to the engine quick disconnect: plumbing, pumps, valves, and controls)
- Fuel dump (all components used to dump fuel overboard during flight)
- Indicating (all components used to indicate the quantity, temperature, and pressure of the fuel)

Without fuel supply, powered sustained flight would not be possible. For this reason, the fuel system, together with the flight control system and the landing gear, can be considered the most essential systems of an aircraft. This fact is also reflected in the many sections of the certification requirements dedicated to the fuel system: For transport category aircraft these are sections 951 through 1001 of JAR-25 and FAR Part 25.

All aircraft use hydrocarbon fuels. Piston engine aircraft use a high-octane number *gasoline.* Common for these aircraft is AVGAS 100LL. Jet engine aircraft use *kerosene.* Depending upon the application (civil or military), various grades are utilized. Common jet fuel for civil applications is JET A-1. Table 12.5 contains some fuel data relevant to aircraft fuel systems.

Kerosene has a sufficiently high *flashpoint.* At sea-level pressure and normal temperatures, kerosene can be considered a safe fuel. Gasoline, in contrast, could easily ignite and needs to be handled especially careful.

When fuel in the fuel lines is heated enough to cause it to vaporize, a bubble of fuel vapor appears, blocking the fuel from flowing to the engine. Such a situation is called vapor lock and must obviously be avoided. The *vapor pressure* is a measure showing if a fuel is prone to vapor lock.

Fuel contains a certain amount of energy per unit mass known, as specific heat or *heating value H*. The fuel tank offers a limited fuel volume *V*. Hence, fuel mass *m* and fuel energy *E* in the fuel tank vary with fuel density ρ.

$$m = \rho \cdot V \qquad E = m \cdot H \qquad E = \rho \cdot V \cdot H$$

Since fuel density decreases with increasing temperature, so do storable fuel mass and energy. For aircraft operation, the amount of energy on board is of importance. Accordingly, *indicating fuel mass* to the pilots does make sense (in contrast to indicating fuel volume). The drawback: not only measurements of fuel level and hence fuel volume are required, but additionally, measurements of fuel density.

Water may be contained in the fuel dissolved, entrained, or free. As fuel is taken from the tank, air (at given humidity) enters the space above the fuel in the tank. With decreasing temperature, water condenses from this air and enters into the fuel. During flight at high altitudes and low temperatures, ice crystals can form that clog fuel filters. To prevent clogging, the fuel may be passed through a fuel heater prior to entering the filter. Fuel systems must be capable of sustained operation with a specified amount of free water under critical conditions for icing (section 951). With the aircraft at rest, water in the fuel collects in the fuel tank sump that is the lowest part of the fuel tank. This happens because density of water (1,000 kg/m^3) is greater than fuel density. ''Each fuel tank sump must have an accessible drain'' (section 971). Water drain valves are used to extract the water.

Microorganisms, bacteria or fungi, may grow in jet fuel tanks. These organisms live and multiply in the water contained in the fuel and feed on the hydrocarbons. The buildup of microorganisms not only interferes with fuel flow and quantity

TABLE 12.5 Fuel Characteristics Related to Aircraft Fuel Systems

	AVGAS 100LL gasoline for piston engines	**JET A-1** kerosene for jet engines
Flashpoint (at standard sea level pressure)	−40 °C (−40 °F)	+38 °C (+100 °F)
Vapor pressure (Reid standard conditions)	500 hPa (7.25 psi)	10 hPa (0.145 psi)
Density (at 15 °C)	720 kg/m^3 (6.0 lb/U.S. gal)	810 kg/m^3 (6.7 lb/U.S. gal)
Heating value	43.5 MJ/kg (18700 BTU/lb)	42.5 MJ/kg (18300 BTU/lb)

indication but can start electrolytic corrosion. The organisms form a dark slime on the bottom of the lowest parts of the fuel tank, especially near water drain valves. Regularly draining water from the fuel together with fuel additives may solve the problem of microbial growth.

Unintended ignition of fuel must be prevented. Therefore section 954 of JAR-25 and FAR Part 25 reads: ''Each fuel system must be designed and arranged to prevent the ignition of fuel vapour'' by lightning strikes or other effects at outlets of the vent and jettison systems or directly through the structure.

Fuel Storage

Fuel *tank location* can be in the wing, fuselage, horizontal stabilizer, or fin. Tanks can be permanently attached or mounted onto the wing tip (tip tank). In the case of combat aircraft, additional tanks can be under-wing mounted, over-wing mounted, or belly mounted. Transport aircraft often use the center section of the wing for a center tank (Figure 12.33). These aircraft may trade payload versus fuel capacity (i.e., maximum range) by using part or all of the cargo compartment for additional center tanks (ACTs).

''Fuel tanks must have an *expansion space* of not less than 2% of the tank capacity. It must be impossible to fill the expansion space inadvertently with the aeroplane in the normal ground attitude'' (section 969). A 2% expansion is equivalent to an increase in fuel temperature of 20 °C.

Fuel initially filled into the empty tanks cannot practically be expected to be taken out again ''to the last drop'' under all operating conditions. The amount of fuel that remains in the tank is called *unusable fuel*. ''The unusable fuel quantity for each tank and its fuel system components [is] the quantity at which the first evidence of engine malfunction occurs under the most adverse fuel feed condition for all intended operations and flight manoeuvres'' (section 959). Aircraft manufacturers try to reduce the unusable fuel volume as much as possible. So-called scavenge pumps are used to collect fuel from different areas of the tank.

The fuel in the fuel tanks can be used for center of gravity (CG) control. Supersonic aircraft may use *CG control* to minimize trim drag that is caused by the rearward shift of lift at supersonic speeds. The Concorde uses trim tanks in the forward part of the wing for CG control. Subsonic aircraft may use a trim tank in the empennage to maintain an optimum rearward CG in cruise. An aft CG reduces trim drag and thus enhances aircraft performance. The Airbus A340 applies a trim tank in the horizontal tail to move the CG in cruise back to approximately 2% mean aerodynamic chord (MAC) forward of the certified aft limit.

The weight of fuel in the wings directly balances lift. This reduces *wing-bending* moments and allows for the design of a lighter structure. In order to make as much use as possible of this phenomenon, fuel is preferably taken from the center tank or an inner wing tank first, whereas the fuel in outboard wing tanks is used only during the last part of the flight. During the last part of the flight, lift is already reduced anyway due to a reduction of aircraft weight as a result of fuel consumption.

Fuel tank construction can be divided into three basic types: rigid removable, bladder, and integral.

A *rigid removable fuel tank* is one that is installed in a compartment designed to hold the tank. The tank must be fuel-tight, but the compartment in which it fits is not fuel-tight. The tank is commonly made of aluminum components welded

together or composites. Rigid fuel tanks are used on small aircraft or as additional center tanks (ACT). ACTs inside the fuselage must be double-walled.

A *bladder tank* is a reinforced rubberized bag placed in a non-fuel-tight compartment designed to structurally carry the weight of the fuel. Bladder tanks are found on medium- to high-performance light aircraft or inside a rigid ACT structure to produce a double-walled tank.

An *integral fuel tank* is a tank that is part of the basic structure of the aircraft. Integral fuel tanks, e.g., in the wing, use structural members of the wing and sealing materials where members join to form a fuel-tight tank. Tank access panels seal the oval cutouts in the lower wing surface used for tank inspection. Baffles are frequently installed inside fuel tanks to reduce fuel sloshing. Baffles may have check valves that open in the inboard direction only. These check valves keep fuel in the inboard part of the tank, where the pumps are located.

The *tank vent system* ''maintains acceptable differences of pressure between the interior and exterior of the tank'' (section 975) under all operating conditions, including:

- Cruise (fuel burn)
- Maximum rate of climb and descent (change of outside pressure)
- Refueling and defueling

Overpressure and underpressure in the tanks can cause structural damage. Underpressure can cause engine fuel starvation. The vent system for a light aircraft may be as simple as a hole drilled into the fuel cap. Large aircraft connect each main tank via vent pipes with a vent surge tank for tank venting (Figure 12.34). The vent surge tanks take up any overflow fuel from the main tanks and direct it back to these tanks through vent float valves. The vent surge tanks are each connected to the outside via a NACA air intake, which achieves a pressure in the fuel tank slightly above ambient pressure.

Aircraft fuel is also used as a *heat sink.* The hydraulic system and the air conditioning system, especially of jet fighters, use fuel for cooling purposes. It is obviously important to monitor the fuel temperature carefully in order to avoid overtemperatures.

Fuel Distribution

The fuel distribution system may consist of:

- The engine feed system
- The fuel transfer system
- The crossfeed system
- The refuel/defuel system

Engine feed, i.e., fuel flow to the engines, may be either gravity feed or pressure feed.

In the case of *gravity feed,* the fuel flows by gravity to the engine. This is possible if the tank is located sufficiently above the engine. Gravity feed is used on small high-wing aircraft and on large aircraft in emergency cases with system fuel pumps inoperative (suction provided from engine fuel pumps).

In the case of *pressure feed,* fuel pumps are used to move fuel through the fuel system. For turbine-engine fuel systems there must be one *main pump* for each engine (section 953) and one *emergency pump* (section 991) immediately available to supply fuel to the engine if the main pump fails. Various fuel pump principles exist including: vane pump, centrifugal pump, and ejector pump.

The *centrifugal pump* (Figure 12.31) draws fuel into the center inlet of a centrifugal impeller and expels it at the outer edge. Fuel can flow through the pump when the pump is not in operation. This eliminates the need for a bypass valve.

An *ejector pump* (Figure 12.32) is used to scavenge fuel from other areas of the fuel tank or from adjacent fuel tanks. This type of pump has no moving parts. Instead it relies on the fuel flow from a main pump.

Fuel *selector valves* provide means to select a tank from which to draw fuel in a multiple-tank installation, transferring fuel from one tank to another and directing fuel to one or more engines. A *shutoff valve* (section 1189) disconnects fuel flow to an engine. The shutoff valve is also closed by the fire handle in case of engine fire. "There must be a fuel strainer or filter" (section 997).

The *fuel transfer system* allows fuel to be pumped from one tank into another. The main feature of the crossfeed system is its fuel manifold. Fuel is supplied form the tanks to the crossfeed manifold. Crossfeed valves on the crossfeed manifold can be set such that each engine can be fed from all tanks.

There are two basic refuel procedures for aircraft: over-wing refueling and pressure refueling. In addition, some aircraft are able to use in-flight refueling.

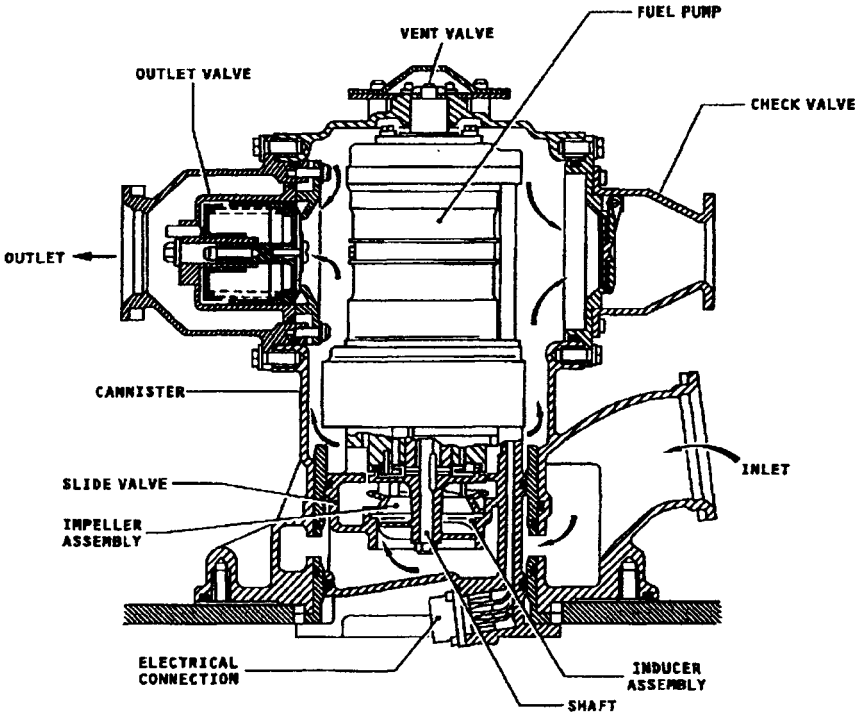


FIGURE 12.31 Centrifugal fuel pump (A321).

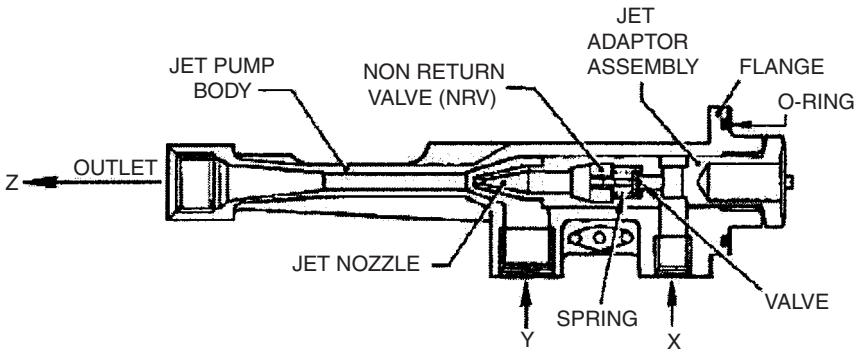


FIGURE 12.32 Jet pump (A321).

X = input from main pump
Y = suction input
Z = output

The historical form of refueling an aircraft from above simply by gravity is called *over-wing refueling.* Small aircrafts apply this simple method. It is slow, and depending on aircraft size and wing location it may be difficult to reach on top of the wing.

Pressure refueling uses pressure from the fueling station or truck to force fuel into the aircraft tanks. This is usually done through a fueling coupling located under the wing at the right wing leading edge. Pressure refueling is fast and the refuel coupling is in easy reach.

During *in-flight refueling,* a military aircraft is supplied with fuel in the air from a tanker aircraft. Tanker aircraft are converted large civil transports. The connection between the receiving and providing aircraft can be established with a flexible hose or a rigid boom. In-flight refueling was first used for fighter aircraft to extend their limited range capabilities. Later, in-flight refueling was applied to cover large distances in global conflicts or to maintain constant combat air patrol.

Defueling is the opposite of refueling: fuel is pumped out of the aircraft fuel tanks and back into the station or truck. During *fuel ground transfer,* fuel is pumped from one aircraft tank into another tank. Defueling and fuel ground transfer may become necessary prior to tank maintenance.

Fuel Jettison

Fuel weight amounts to a large fraction of aircraft gross weight, especially at the beginning of a long-range flight (with a long-range aircraft). If an *emergency* occurs shortly after takeoff, the aircraft may be forced to return and land as soon as possible. In such a situation, the present aircraft weight will still be considerably above maximum landing weight. An overweight landing might unduly stress and endanger the aircraft, and in the case of a discontinued approach, the heavily laden aircraft will not be able to fly a successful go-around maneuver with sufficient climb rate (section 1001).

A fuel *jettison system* (fuel dump system) helps to solve the situation. The fuel jettison system allows dumping of all but some reserve fuel overboard in not more

than 15 minutes. This now brings the aircraft weight down quickly as a prerequisite for a successful emergency landing.

Two fuel-jettison principles have been used: systems can work with gravity or with pump pressure. A *gravity jettison system* is equipped with long dump chutes that are deployed at both wing tips. The long chutes produce the necessary pressure differential for the flow. A *pump jettison system* is equipped with dump nozzles at both wing tips.

Indicating

Quantity, temperature, and pressure of the fuel can be measured for the fuel system. Other fuel parameters are measured by the engine.

A *fuel quantity indicator* can be a mechanical quantity indicator, a resistance quantity indicator, or a capacitance quantity indicator.

A *capacitance quantity indicator* is a condenser installed in the tank so that the condenser is immersed in the fuel. Fuel respectively air in the tank serve as dielectric material for the condenser. When the probe is dry, its capacitance value is low, but as fuel moves up the probe its capacitance value increases. A controller monitors the capacitance value and converts it into a fuel volume.

In addition to the fuel quantity indicator, which is primarily used in flight, it is desirable to have an alternative provision to determine the fuel quantity visually. On light aircraft this may be accomplished by viewing the fuel surface through the fuel filler cap opening, but on large aircraft this would be extremely difficult. For this reason, calibrated hollow fiberglass *dripsticks* have been used that are unlocked and slowly lowered from under the wing. The position of the stick when it drips marks the fuel level inside the tank.

More sophisticated are *magnetic level indicators* (MLIs). MLIs are also unlocked and lowered from under the wing. A magnetic float on the fuel surface gets hold of the magnetic top of a stick. The position of the stick attached to the float determines the fuel level.

Example: Airbus A321

The Airbus A321 has three *fuel tanks* (Figure 12.33): the left wing tank, the right wing tank, and the center tank. The total usable fuel capacity of these tanks is 23,700 l. The total unusable fuel capacity is 89.7 l. This is less than 0.4%.

The *vent surge tanks* (Figure 12.34) do not normally contain fuel. They are connected to the wing tank and center tank through the stringer vent duct and the center tank vent pipe. The vent surge tanks can vent these tanks because they are open to the external air through a vent duct. The vent duct contains a vent protector with a flame arrestor and an ice protector. The vent duct is connected to a NACA intake on the bottom of the tank. The vent surge tanks are also a temporary reservoir for the fuel that could enter through the vent pipes. This fuel is drained back to the wing tanks through vent float valves (clack valves). In case of an obstruction in the vent duct, the overpressure protector ensures that the pressure in the vent surge tank does not exceed specified limits.

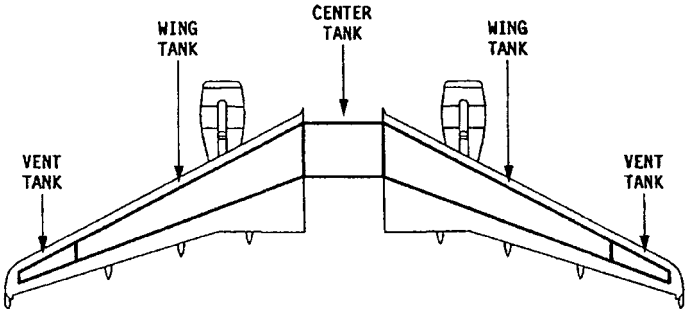


FIGURE 12.33 A321 fuel tanks.

The *fuel distribution system* of the A321 is shown in Figure 12.35:

- The engine feed system takes fuel from the wing tanks and supplies it to the engines. Two main pumps (Figure 12.31) are located in each wing tank.
- The main transfer system enables transfer of fuel from the center tank to the left and right wing tank. This fuel transfer is a normal procedure necessary to make use of the fuel in the center tank. Fuel transfer is achieved with ejector pumps (jet pumps). The jet pumps in the center tank are driven by fuel from the main pumps.
- The crossfeed system connects the left and right fuel feed system. The engine feed line has a crossfeed valve that permits the isolation or interconnection of the left (engine 1) and right (engine 2) fuel supply system. Under normal conditions, the crossfeed valve is closed.
- The refuel/defuel system:
  - Refueling: Fuel is supplied to the fuel tanks via the refuel coupling in the right wing. A second refuel coupling in the left wing is optionally available.
  - Defueling: Fuel is pumped out of the tanks by way of the refuel coupling. The defuel transfer valve is open.
  - Fuel transfer: The system may be used to transfer fuel from one tank into any other tank. The defuel transfer valve is open.
- The APU feed system takes fuel from the engine feed line and supply fuel to the auxiliary power unit (APU).

12.8 HYDRAULIC POWER (ATA 29)

The hydraulic system as defined by ATA 100:

> Those units and components which furnish hydraulic fluid under pressure (includes pumps, regulators, lines, valves, etc.) to a common point (manifold) for redistribution to other defined systems.

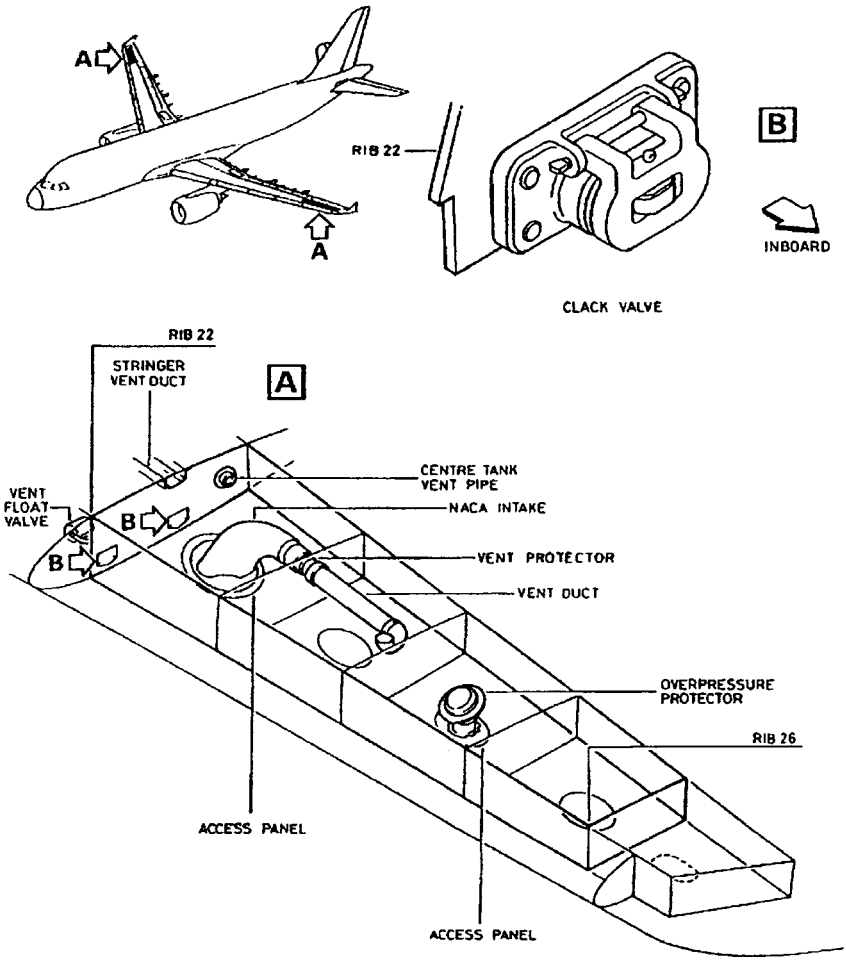


FIGURE 12.34 A321 vent surge tank.

Purpose

The purpose of the hydraulic system is to assist the pilot in accomplishing mechanical tasks that would otherwise be impractical or impossible because of the level of force, work, or power required. On smaller aircraft the flight control surfaces are moved by pilot force. On larger and faster aircraft this becomes impossible and so hydraulic power is applied. A total failure of the flight control system evidently has a catastrophic effect. Consequently, a failure of the hydraulic power supply of large aircraft has to be extremely improbable. This required level of safety is achieved with redundancy through three or even four independent hydraulic subsystems.

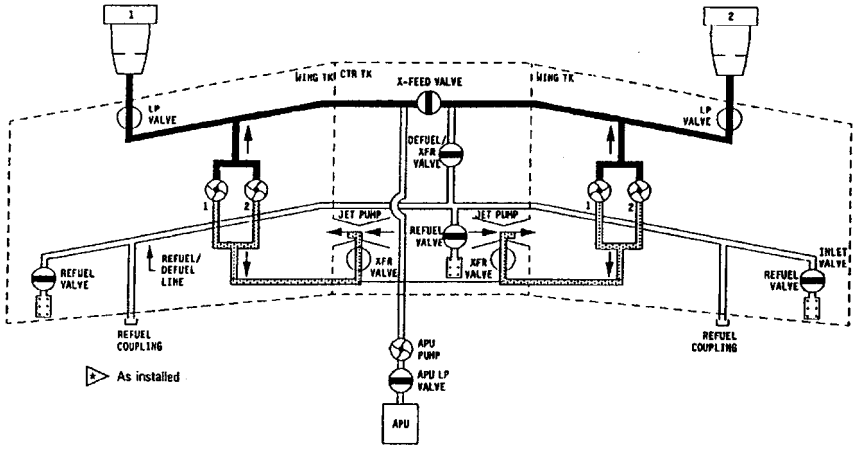


FIGURE 12.35 A321: Overview of the fuel distribution system.

black lines: engine feed system
gray lines: main transfer system
white lines: refuel/defuel system and APU feed
X-FEED: crossfeed
XFR: transfer

Principle

Figure 12.36 shows the principle of a hydraulic system. Hydraulic fluid is contained in a *reservoir.* Through a suction line the *pump* draws fluid from the reservoir and puts it at a higher pressure. Today aircraft hydraulic systems are typically designed to a nominal pressure of 206 bar (3,000 psi). The trend is toward higher system pressure: 345 bar (5,000 psi). An *accumulator* serves as temporary energy storage and is able to store or redistribute surplus high-pressure fluid. A *pressure-relief valve* is able to shortcut the high-pressure line to the reservoir in case of a system malfunction leading to higher pressure than specified. The pressure differential supplied by the pump is used by hydraulic consumers. The example shows a typical consumer in the flight control system. An *actuator* piston rod has to move in and out in order to deflect a control surface (not shown). The actuator piston is moved through hydraulic fluid that enters the left actuator chamber and fluid that leaves the right actuator chamber (or vice versa). A valve schedules the required fluid flow. Shown is a *servo valve.* The valve has four connections to hydraulic tubes: one connection to each of the two actuator chambers, one connection to the high-pressure line, and one connection to the return line. The valve may be moved into one of three positions that lead to piston rod extension, piston rod retraction, or no piston rod movement. In the case of a flight control consumer, it is necessary that the valve move gradually from one position into the other to allow a proportional control of the surface. In the case of landing gear extension and retraction, a *selector valve* would be used. The selector valve allows three distinct valve positions without any intermediate positions.

During system design the complete circuit, including hydraulic power generation, distribution, and consumption, has to be analyzed. According to the ATA

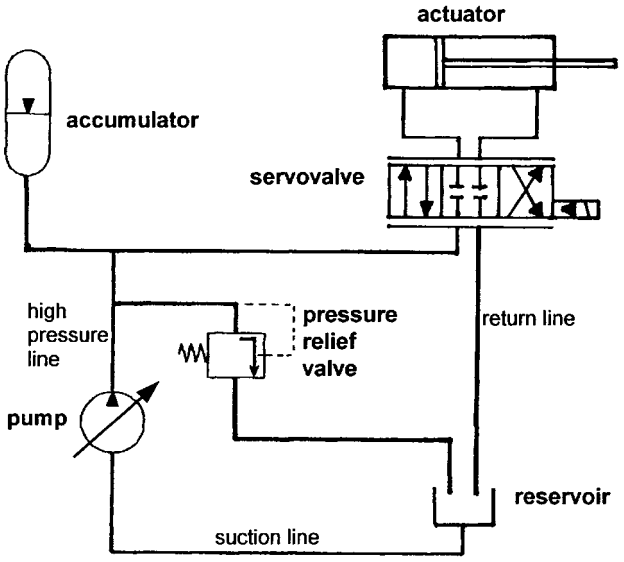


FIGURE 12.36 A basic hydraulic system.

breakdown, the consumers with their valves are allocated to their respective system. ATA 29 deals only with power generation and distribution.

Three types of *hydraulic fluids* exist: vegetable based, mineral based, and synthetic or phosphate ester-based. Transport category aircraft use the purple-colored phosphate ester-based fluid—most commonly Skydrol® LD. Skydrol® shows good performance even at low temperatures, excellent flammability characteristics, and minimal effects on most common aircraft metals, but does react with certain types of paint and can be an eye and respiratory irritant.

Components

The *reservoir* acts as a storage tank for the system's fluid. Reservoirs can be broken down into two basic types, *in-line* and *integral,* and these can be further classified as *pressurized* and *unpressurized*. Integral reservoirs, found on small aircraft, are combined with the pump. Aircraft that operate at low altitudes could use unpressurized reservoirs that vent the reservoir to the atmosphere. Other aircraft positively pressurize the reservoir with air from the pneumatic system, hydraulic pressure (bootstrap reservoir) (Figure 12.37), or a spring. In a *bootstrap reservoir,* high-pressure (HP) fluid acts on a small plunger that is coupled with a large plunger that in turn acts on the low pressure (LP) fluid in the reservoir. Commonly, air pressure is used for reservoir pressurization. The air pressure usually needs to be reduced by a pressure regulator. It then enters the airspace above the fluid in the reservoir.

Commonly used are axial multiple-piston *pumps.* The two principles applied are *constant displacement* and *variable displacement*. The shaft can be driven by the aircraft engine, by an electric motor, or through a device powered by the pneumatic

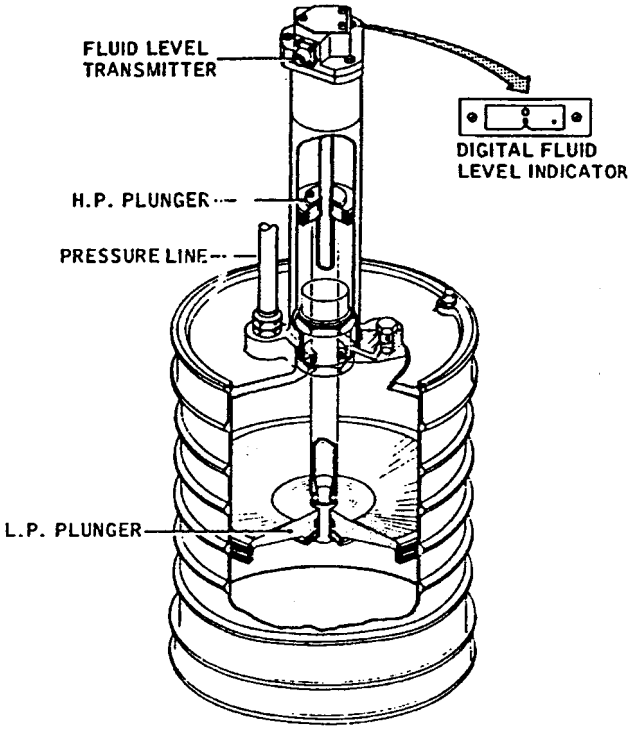


FIGURE 12.37 Hydraulically pressured reservoir known as bootstrap reservoir (VFW 614).

system. The shaft turns the cylinder block with the pistons. Whenever an elevated piston is pushed into the cylinder block, fluid is ejected into an out port. Accordingly, during the other half of the revolution on its way back to the elevated position, the piston draws fluid from an in port into the cylinder block. Constant displacement pumps deliver exactly the same amount of fluid every revolution and must incorporate a pressure regulator. Most widely used, however, are *variable displacement axial multiple-piston pumps* (Figure 12.38). The variable displacement is achieved by a swashplate. The angle of the swashplate is adjusted by a pressure controller. At highest swashplate angle, the pump achieves its maximum flow rate, and at zero angle there is no fluid flow.

For minor tasks, *hand pumps* may be applied. A *ram air turbine* (RAT) (Figure 12.43) may be turned into the free stream of air to power a hydraulic pump. This is done in the event of an engine failure or a major electrical system failure.

Three types of *accumulators* are known: the *diaphragm-type accumulator,* the *bladder type accumulator,* and the *piston type accumulator* (Figure 12.39). The diaphragm, bladder, or piston divides the fluid chamber from the nitrogen chamber of the accumulator. Hydraulic fluid is allowed to flow freely into and out of the fluid chamber of the accumulator. The compressible nitrogen acts like a spring against the hydraulic fluid. The accumulator acts as a high-pressure and fluid storage and eliminates shock waves from the system.

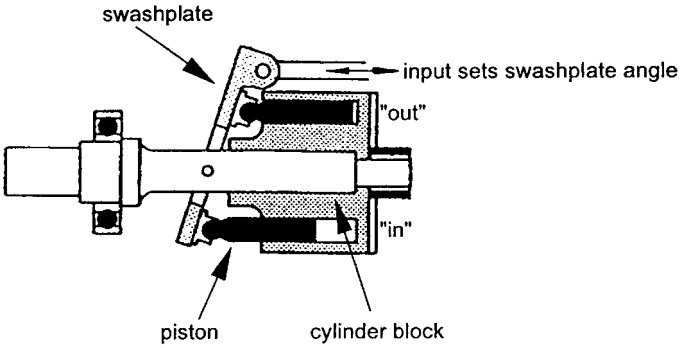


FIGURE 12.38 Variable displacement axial multiple-piston pump (TUHH).

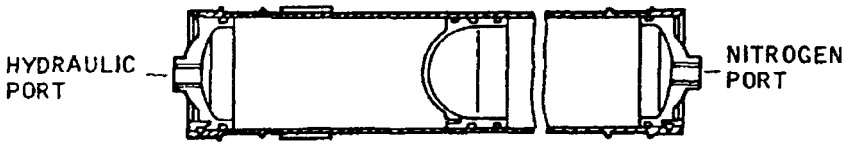


FIGURE 12.39 Piston-type accumulator (VFW 614).

Filters are installed in the high-pressure and the return line. Three filter types are in use: micron, porous metal, and magnetic. *Micron filters* contain a treated paper element to trap particles as the fluid flows through the element. *Porous metal filter* are composed of metal particles joined together by a sintering process. *Magnetic filters* attract metal particles. Filters consist of a head assembly that contains the fluid line connections and a bypass valve to prevent the system from becoming inoperative should the filter become clogged, a bowl assembly, and the filter element. Fluid enters through the head into the bowl and leaves through the filter element and out of the head (Figure 12.40).

Two principal types of *valves* are used in the hydraulic system: flow control valves and pressure control valves. *Flow-control valves* route the fluid through the system. Examples are s*elector valves,* which permit the user to channel the fluid selectively, and *servo valves,* as explained above. *Check valves* permit flow only in one direction. A *hydraulic fuse* is a safety valve that prevents fluid flow in the event of a serious system leak. Examples of *pressure-control valves* are the *pressure-relief valve* and the *pressure regulator*. A *priority valve* is mechanically identical to a pressure relief valve, set to an opening pressure below nominal pressure. The priority valve is closed at low pressure and allows flow to secondary consumers only if a minimum system pressure has been reached. In this way it gives priority to primary consumers located upstream of the priority valve.

Hydraulic *fluid lines* are classified as rigid or flexible. *Rigid lines* are made of either aluminum for return and suction lines or of stainless steel for high-pressure lines. *Flexible lines* are hoses typically wrapped with stainless steel braid. *Fittings* are used to connect fluid lines with other hydraulic components.

"The *power transfer unit* (PTU) is a device which uses some of the hydraulic power in one hydraulic system to supplement the hydraulic power in a second system without interchange of fluid between the systems" (ARP 1280).

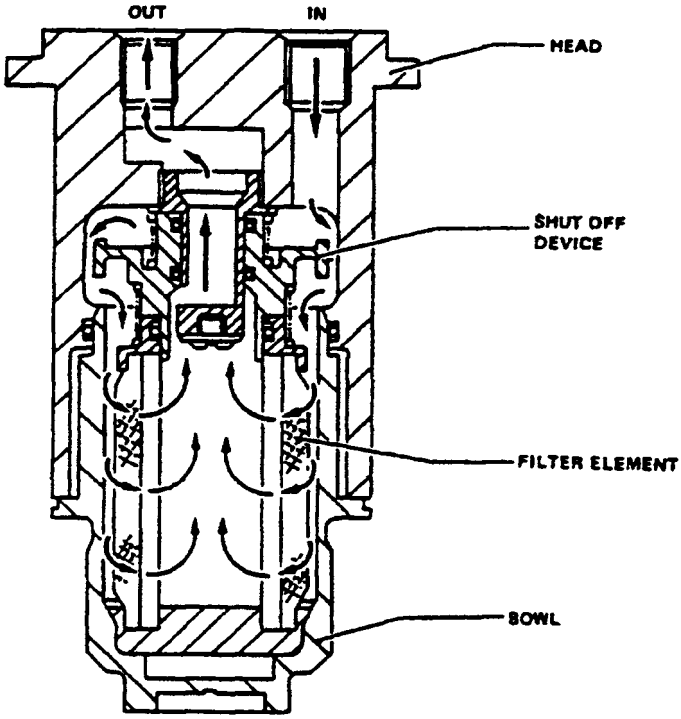


FIGURE 12.40 Low-pressure filter (A321).

PTUs can be designed either to transfer power from one system to a second system in one direction only (unidirectional PTU) or to transfer power in either direction between two systems (bidirectional PTU) (Figure 12.41). The basic concept consists of a hydraulic motor driving a pump, mounted back-to-back. The displacement of each of these may be the same or different. Accordingly, PTUs can be used as pressure reducers, as pressure intensifiers, or to maintain the same pressure in both systems. If bidirectional operation is required, both the pump and the motor reverse their functions. That unit which was previously the pump will operate as motor and vice versa. If the pressure relationship between the two systems must remain the same in both directions of operation, at least one of the units must be of a variable displacement design.

Example: Airbus A321

The Airbus A321 has three main hydraulic (sub)systems (Figure 12.42):

- The Green system
- The Blue system
- The Yellow system

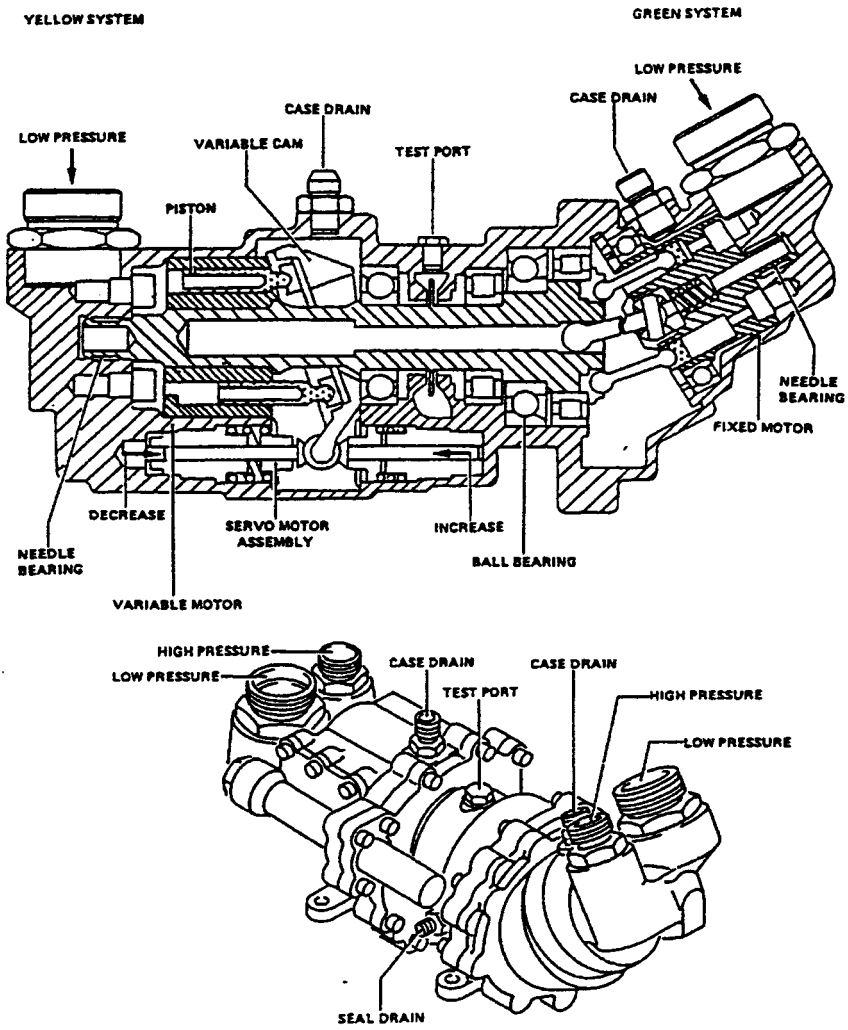


FIGURE 12.41 Bidirectional power transfer unit (PTU) (A321).

Together they supply hydraulic power at 20.7 MPa (3,000 psi) to the main power users. These include:

- Flight controls
- Landing gear
- Cargo doors
- Brakes
- Thrust reversers

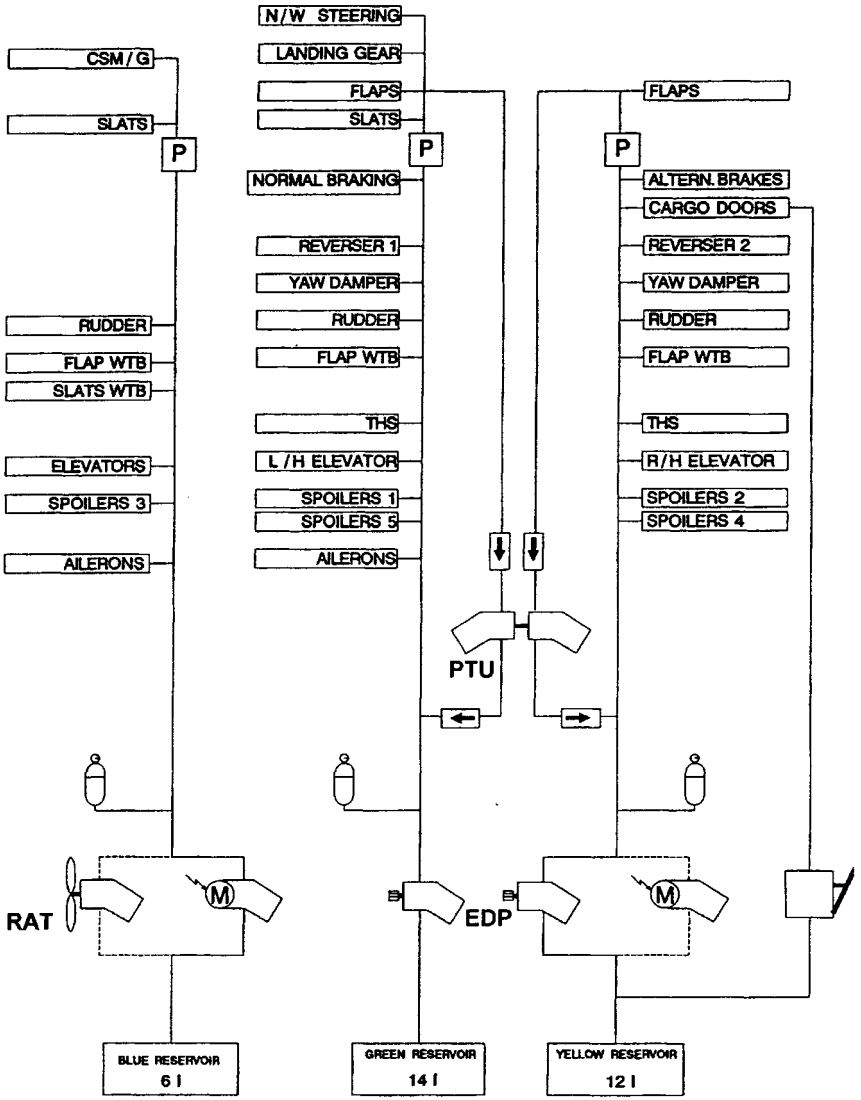


FIGURE 12.42 A321 hydraulic system schematic.

EDP: Engine-driven pump
M: Electric pump
RAT: Ram air turbine
PTU: Power transfer unit
P: Priority valve
→: Check valve (indicating flow direction)
CSM/G: Constant-speed motor/generator (emergency generator)
THS: Trimmable horizontal stabilizer (horizontal tail)
WTB: Wing tip brake (in high-lift system)

Main system pumps are the engine-driven pumps (EDPs) in the Green and Yellow system as well as the electric pump in the Blue system. The EDP of the Green system is connected to the left (No. 1) engine. The EDP of the Yellow system is connected to the right (No. 2) engine. The three main systems automatically supply hydraulic power when the engines operate. The two EDP are connected directly to their related engine (through the accessory gearbox), and the Blue electric pump operates when any one of the two engines starts. The three system main pumps are usually set to operate permanently. If necessary (because of a system fault, or for servicing), the pumps can be set to off from the flight compartment.

If the main pumps cannot be used, it is possible to pressurize each hydraulic system with one or more of the *auxiliary system pumps.*

- The Green system can also be pressurized by the power transfer unit (PTU).
- The Blue system can also be pressurized by the ram air turbine (RAT) (Figure 12.43).
- The Yellow system can also be pressurized by the Yellow electric pump or the power transfer unit (PTU).

Pressurization of the hydraulic systems *on the ground* is possible as follows:

- Yellow system—with the Yellow electric pump
- Green system—with the Yellow electric pump (through the PTU)
- Blue main system—with the Blue electric pump

For maintenance, all of the systems can be pressurized from a ground hydraulic supply. Connectors are installed on the ground service panels of the three systems. The cargo doors can also be operated with a hand pump in the Yellow system.

12.9 ICE AND RAIN PROTECTION (ATA 30)

Ice and rain protection as defined by ATA 100:

> Those units and components which provide a means of preventing or disposing of formation of ice and rain on various parts of the aircraft. Includes alcohol pump, valves, tanks, propeller/rotor anti-icing system, wing heaters, water line heaters, pitot heaters, scoop heaters, windshield wipers and the electrical and heated air portion of windshield ice control. Does not include the basic windshield panel. For turbine type power plants using air as the anti-icing medium, engine anti-icing is [part of the powerplant].

System Classification

Ice and rain protection may be classified as follows:

- Nontransparent surfaces: ice protection (leading edges, radome, inlets, etc.):
  - Pneumatic boot systems
  - Thermal ice protection systems:
    - Hot air systems
    - Electrical resistance systems

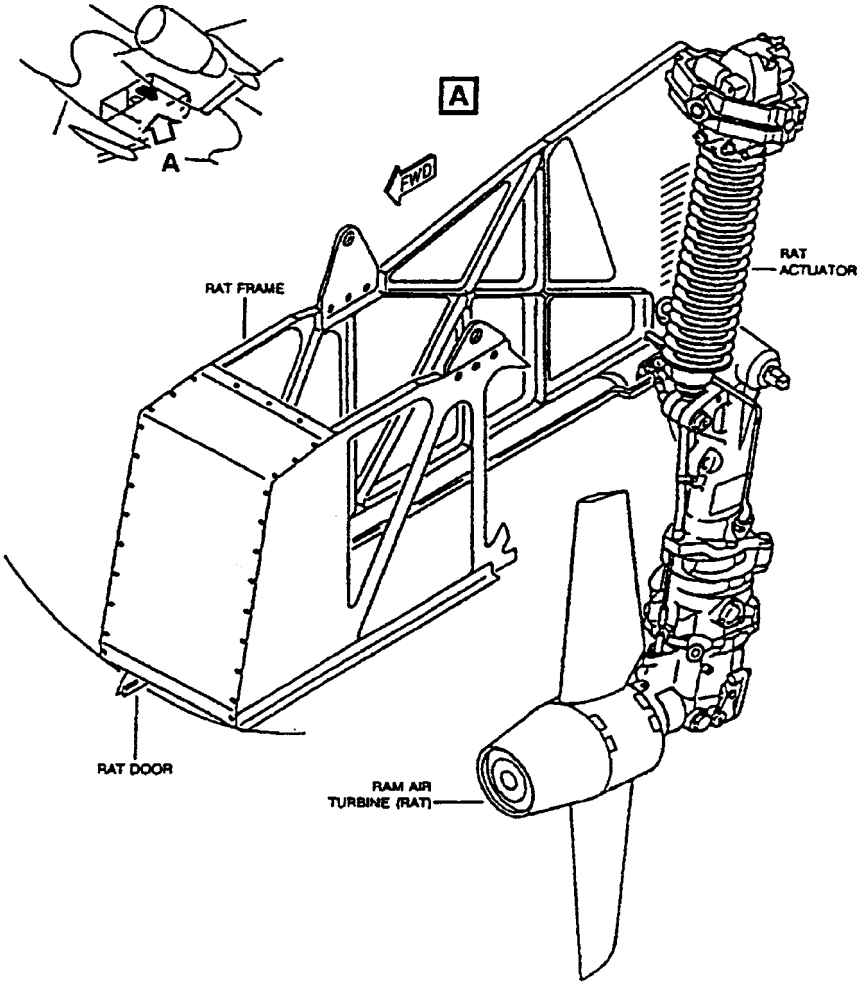


FIGURE 12.43 A321 ram air turbine (RAT).

  - Fluid systems
  - Electroimpulse deicing (EIDI) systems
  - Microwave systems
- External components: ice protection (antennas, sensors, drain masts, etc.)
- Internal components: ice protection (water lines, etc.)
- Windshield: ice and fog protection
- Windshield: rain removal
- Ice detection

External and internal components are generally protected against icing by electrical resistance systems. Some technical solutions for windshield ice protection serve at the same time for windshield rain removal.

The two main ice *protection principles* are deicing and antiicing. Various technical solutions exist. Some ice protection *technical solutions* can perform both deicing and antiicing. Other technical solutions only manage deicing (Table 12.6).

The terms deicing and antiicing are defined in AIR 1168/4:

- *Deicing* is the periodic shedding, either by mechanical or thermal means, of small ice buildups by destroying the bond between the ice and the protected surface.
- *Antiicing* is the prevention of ice buildup on the protected surface, either by evaporating the impinging water or by allowing it to run back and freeze on noncritical areas.

Icing Fundamentals

From our daily experience we know that water freezes to ice below 0 °C (32 °F) and melts again above 0 °C. When it comes to aircraft icing, we learn that this need not be so. Small droplets can still be in the liquid phase below 0 °C! Most droplets will have turned to ice below −20 °C (−4 °F), though very small and pure droplets may reach temperatures as low as −40 °C (−40 °F) and still remain liquid. Below −40 °C finally all water in the air will be frozen. ''Liquid'' water below 0 °C is called *supercooled water*. Supercooled water can exist because the water has been totally undisturbed during cooling—nothing has caused it to turn to ice. When an aircraft hits the droplet, however, the droplet receives the necessary input for the phase change and turns to ice. (The phase change from water to ice *usually* requires some latent heat extraction, but when the droplets are supercooled water, the heat extraction has already taken place.) The ice will be slightly warmer than the supercooled water was just a second earlier. Summing up: Supercooled water turns instantly to ice due to the interaction with the aircraft. The result will be *ice accretion* on the aircraft surface if the surface is below 0 °C.

Aircraft icing is thus possible if

1. The air contains water (clouds are an indication of water in the air).
2. The air temperature is below 0 °C.
3. The air temperature is above −40 °C
4. The aircraft surface is below 0 °C.

TABLE 12.6 Ice Protection Technical Solutions and Protection Principles

Ice protection technical solutions	Deicing	Antiicing
Pneumatic boot systems	x	○
Hot air systems	x	x
Electrical resistance systems	x	(x)
Fluid systems	x	(x)

x solution usually applied to protection principle
(x) solution usually not applied to protection principle
○ solution can not be applied to protection principle

There are *other icing mechanisms* besides the standard one just discussed:

- Icing will occur during descent from high altitudes if the aircraft encounters humid air even above 0 °C. The aircraft surface will be below 0 °C due to a long flight at high altitudes. The fuel in the wings will also be below freezing. The fuel is in close contact with the skin as a consequence of integral fuel tank design (see Subsection 12.7). The fuel does not warm up quickly and is likely to remain below 0 °C until landing.
- Carburetor icing can occur at temperatures between −7 °C (20 °F) and +21 °C (70 °F) when there is visible moisture or high humidity. Carburetor icing is caused by cooling from vaporization of fuel, combined with the expansion of air as it flows through the carburetor.
- Water and slush that the aircraft picks up during taxi out can freeze at higher altitudes with detrimental effects to the aircraft.
- Frost, ice, and snow that have settled on an aircraft on the ground have to be removed before takeoff. Ground deicing equipment and procedures have been developed (see AC 135-16).

The two *basic forms of ice build-up* on the aircraft surface are clear ice and rime ice (Figure 12.44).

- Clear ice forms between 0 °C and −10 °C, usually from larger water droplets or freezing rain, and can drastically change the form of the leading edge. It can spread over the surface.
- Mixed ice forms between −10 °C and −15 °C. A mixture of clear ice and rime ice have the bad characteristics of both types and can form rapidly.
- Rime ice forms between −15 °C to −20 °C from small droplets that freeze immediately when contacting the aircraft surface. This type of ice is brittle, rough looking, and colored milky white.

In order to *calculate the total water catch* of the wing, let us cut off a piece of a wing with a spanwise extension Δy and maximum thickness t. This piece of wing will fly at a speed v through a unit volume of air with a certain mass of supercooled water. The mass of supercooled water per volume is called liquid water content

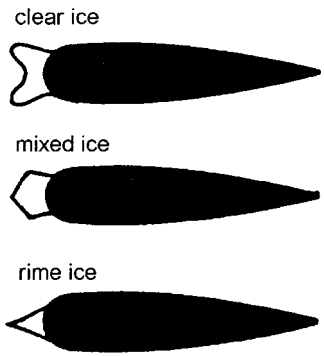


FIGURE 12.44 Ice shapes on the leading edge of airfoils (TÜV 1980).

(LWC) and is something like a density we name ρ_{LWC}. We consider $t \cdot \Delta y$ as the area of an imaginary sieve at an angle perpendicular to its flight path. The mass flow rate of supercooled water through the sieve would be $\dot{m} = vt\ \Delta y\ \rho_{LWC}$. The impingement of water on the leading edge of the wing will, however, be different from the flow through the sieve as shown in Figure 12.45. The air and with it very small droplets pass around the wing; only larger droplets hit the surface. This phenomenon is expressed by the water catch efficiency E_m. The imaginary sieve shows an efficiency $E_m = 1$. The total water catch of a piece of wing is calculated by including E_m:

$$\dot{m} = vt\ \Delta y\ \rho_{LWC}\ E_m$$

E_m is a function of aircraft speed and droplet size, airfoil shape and thickness, viscosity, and density of the air.

- High aircraft speeds and large droplet size cause an increase in water catch efficiency.
- High aircraft velocities, however, lead to aerodynamic heating of the leading edges. This reduces icing.
- Thin wings divert the flow less and increase the water catch efficiency.

AIR 1168/4 presents detailed methods to calculate E_m.

A *simplified method to calculate the water catch efficiency* E_m is presented here based on Figure 3F-3 of AIR 1168/4 as a function of aircraft speed v and wing thickness t (Figure 12.46):

$$E_m = 0.00324 \left(\frac{v}{t}\right)^{0.613} \qquad \text{for } v \text{ in m/s and } t \text{ in m}$$

This equation is based on typical airfoils with a relative thickness of 6–16% at an angle of attack $\alpha = 4°$. The mean effective drop diameter $d_{med} = 20\ \mu$m, altitude $h = 10{,}000$ ft. Other altitudes from sea level to $h = 20{,}000$ ft will result in an error less than 10%.

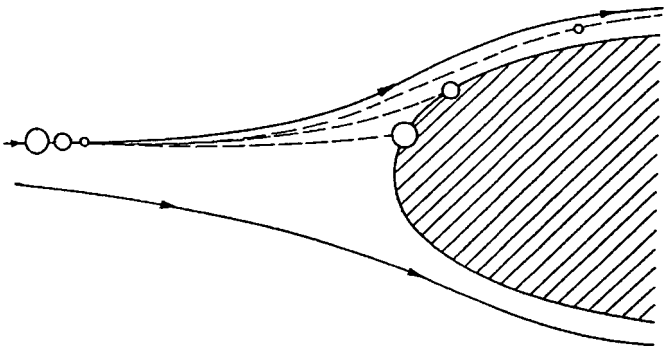


FIGURE 12.45 Flow around a wing leading edge: streamlines of dry airflow; trajectories of differently sized droplets (TUHH).

AIRCRAFT SYSTEMS

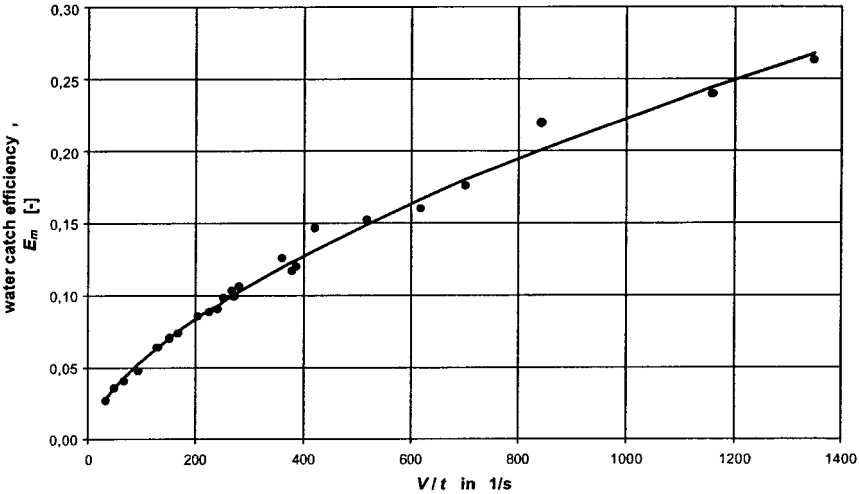


FIGURE 12.46 Water catch efficiency E_m as a function of aircraft speed v and wing thickness t for typical applications. The diagram is calculated from AIR 1168/4, Figure 3F-3.

ACJ 25, section 1419 assumes for certification a typical mean effective drop diameter $d_{\text{med}} = 20\ \mu\text{m}$. The *liquid water content* (LWC) that an aircraft is supposed to meet continuously in flight ranges from $\rho_{\text{LWC}} = 0.2\ \text{g/m}^3$ at $-30\ ^\circ\text{C}$ to $\rho_{\text{LWC}} = 0.8\ \text{g/m}^3$ at $0\ ^\circ\text{C}$.

The *detrimental effects of icing* on the aircraft are manifold. Ice can:

- Alter the shape of an airfoil. This can change the angle of attack at which the aircraft stalls, and cause the aircraft to stall at a significantly higher airspeed. Ice can reduce the amount of lift that an airfoil will produce and increase drag several-fold.
- Partially block control surfaces or limit control surfaces deflection.
- Add weight to the aircraft. The aircraft may not be able to maintain altitude. The stall speed is higher.
- Block the pitot tube and static ports.
- Cause the breakage of antennas on the aircraft.
- Cause a tailplane stall. The airplane will react by pitching down, sometimes in an uncontrollable manner.
- Reduce propeller efficiency. Ice that is hurled away from the propeller is a hazard to everything in its plane of rotation.
- Endanger the internal parts of a jet engine.

In order to protect the aircraft properly against these effects, ice protection may become necessary in areas shown in Figure 12.47.

The design of ice protection systems will always have to be based on the *certification requirements.* For transport category aircraft, the fundamental statement reads: ‘‘If certification for flight in icing conditions is desired, the airplane must be

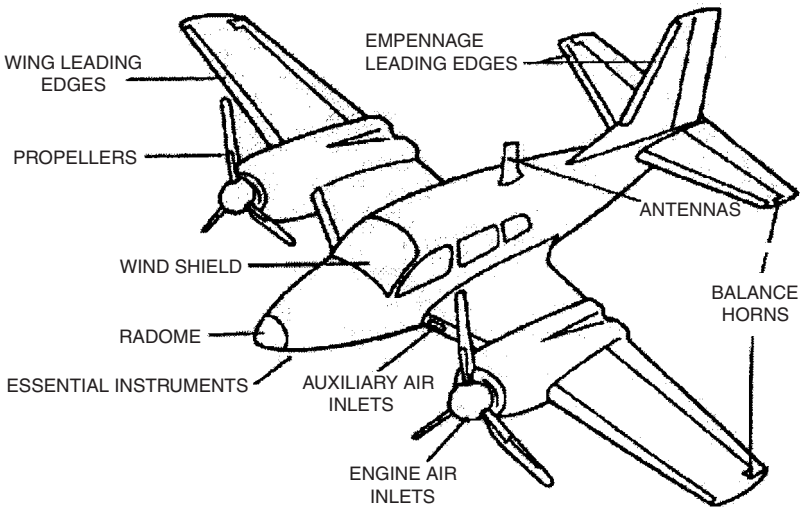


FIGURE 12.47 Areas of the airframe that may require ice protection (FAA 1993).

able to safely operate in the continuous maximum and intermittent maximum icing conditions'' (section 1419, FAR Part 25, JAR-25). Icing conditions are given in Appendix C of these documents.

Critical parts of the aircraft (like the wing) will probably need some kind of ice protection device. Other parts or the aircraft (like the empennage) may fulfill the requirements without being protected.

Pneumatic Boot Systems

Pneumatic boot systems have been the standard ice protection method for piston engine aircraft since the 1930s. The *boot surfaces* remove ice accumulations mechanically by alternately inflating and deflating tubes within a boot that covers the surface to be protected (Figure 12.48). Inflation of the tubes under the accreted ice breaks the ice into particles and destroys the ice bond to the surface. Aerodynamic forces and centrifugal forces on rotating airfoils then remove the ice. In principle, this method of deicing is designed to remove ice after it has accumulated rather than to prevent its accretion in the first place. Thus, by definition a pneumatic boot system cannot be used as an antiicing device. Conventional pneumatic boots are constructed of fabric-reinforced synthetic rubber or other flexible material. The material is wrapped around and bonded to the leading-edge surfaces to be deiced on wings or empennage. Total thickness of typical pneumatic boots is usually less than 1.9 mm (0.075 in.). Pneumatic boots require very little power and are a lightweight system of reasonable cost. The tubes in the pneumatic boot are usually oriented spanwise but may be oriented chordwise. The inflatable tubes are manifolded together in a manner to permit alternate or simultaneous inflation as shown in Figure 12.48. Alternate inflation is less commonly used.

In addition to the boots, the primary *components of a pneumatic system* are a regulated pressure source, a vacuum source, and an air distribution system. Miscellaneous components may include check and relief valves, air filters, control

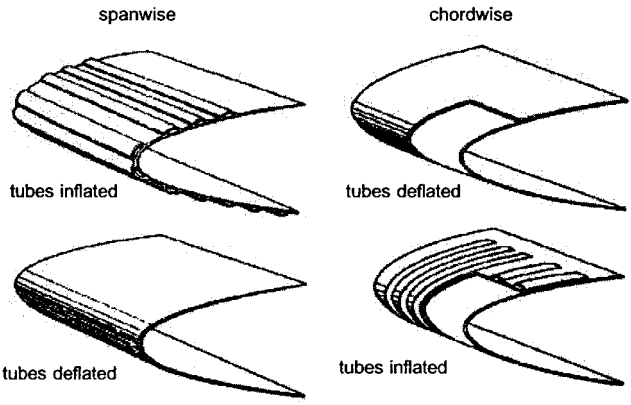


FIGURE 12.48 Inflatable deicing boots (FAA 1993).

switches and timer, and electrical interfaces, including fuses and circuit breakers. A regulated pressure source is required to ensure expansion of all tubes in the system to design limits and within design rise times. Pneumatic boots should inflate and deflate rapidly to function effectively. The time to reach full pressure should be about 5 to 6 seconds. If tube expansion is too slow, deicing effectiveness is lessened. The vacuum source is essential to ensure positive deflation and keep the tubes collapsed during nonicing flight conditions to minimize the aerodynamic penalty. Air pumps generally multiply the atmospheric pressure by a fixed factor, so the pressure delivered becomes a function of altitude. Therefore, for air pump systems, the pressure produced at service ceiling altitude is a design condition.

Some aerodynamic *drag penalty* is to be expected with pneumatic boot deicing systems on an airfoil, but it can be lessened by recessing the surface leading edge to offset the boot thickness. Pneumatic boot deicing systems have been in use for many years, and their repair, inspection, *maintenance,* and replacement are well understood. Pneumatic boot material deteriorates with time, and periodic inspection is recommended to determine the need for replacement. System *weight and power* requirements are minimal.

Ice bridging is the formation of an arch of ice over the boot, which is not removed by boot inflation. In the lore of flight of early piston-powered air transports, it used to be recommended that some ice should be allowed to accrete before the deicing system was turned on in order to avoid ice bridging. The aircraft flight manual (AFM) for modern aircraft now requires that the system be activated at the first sign of ice formation. Ice bridging for modern, properly functioning deicing boots has not been reported (FAA 1993).

Hot Air Systems

Hot air systems and electrical resistance systems are *thermal ice protection systems.* Thermal ice protection systems are classified into three groups:

1. *Evaporative antiicing systems* supply sufficient heat to evaporate all water droplets impinging upon the heated surface.

2. *Running-wet antiicing systems* provide only enough heat to prevent freezing on the heated surface. Beyond the heated surface of a running-wet system, the water can freeze, resulting in runback ice. For this reason, running-wet systems must be used carefully so as not to permit buildup of runback ice in critical locations. For example, a running-wet system may be used for a turbine engine inlet duct where the runback is permitted to enter the engine.
3. *Cyclic deicing systems* periodically shed small ice buildups by melting the surface–ice interface with a high rate of heat input. When the adhesion at the interface becomes zero, aerodynamic or centrifugal forces remove the ice.

An evaporative antiicing system uses the most energy of the three ice protection principles presented, cyclic deicing uses the least energy.

Hot air systems are used on most of the large jet transports because of the availability of hot air from the engines and the relative efficiency and reliability of these systems. Hot air is used to antiice or deice leading-edge wing panels and high-lift devices, empennage surfaces, engine inlet and air scoops, radomes, and selected components. Details of the hot air system are given below, using the Airbus A321 as an example.

Electrical Resistance Systems

Electrical resistance systems are thermal ice protection systems. They may also be classified as *evaporative, running wet,* or *cyclic*.

Electrical resistance systems have a wide range of application:

- *External components* protected with an electrical resistance system will use the evaporative technique.
- *Internal components* like water lines are simply heated above 0 °C to prevent freezing.
- *Nontransparent surfaces* will most probably use cyclic deicing because the electrical loads would otherwise become unbearable.

Electrical resistance systems use electrical resistance heaters in the form of foil, film, resistance wire, or mesh embedded in fiberglass, plastic, rubber, or metal to heat the surface or component.

Electrical deicing systems for nontransparent surfaces may use *parting strips* to divide the total protected area into smaller sequentially heated areas. The spanwise and chordwise parting strips must prevent any ice bridging from one shedding zone to another (Figure 12.49). Parting zones reduce the total instantaneous power requirement and maintain a stable load on the electrical system. Aircraft wings with about 30° or more sweepback will normally use only chordwise parting strips.

For efficient deicing protection, the correct *amount of heat* must be supplied. If there is too little heat, the ice may not shed as required, perhaps causing large chunks of ice to shed. If too much heat is supplied, there can be too much melting, resulting in undesirable amounts of runback ice. It has been found desirable to have a high specific heat input applied over a short period.

The *off-time* of a shedding zone depends upon the rate at which the surface cools to 0 °C (32 °F). It also depends upon the icing rate. The off-time may be tailored

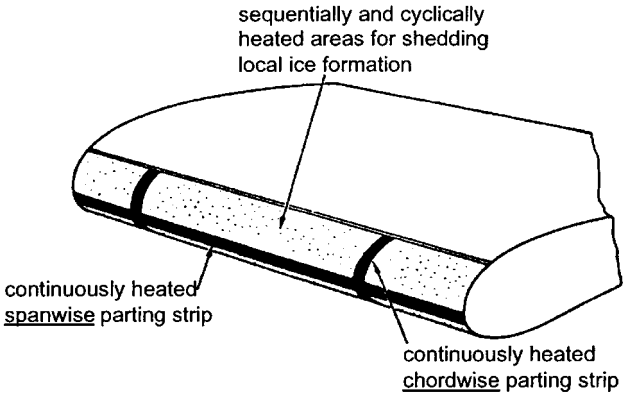


FIGURE 12.49 Arrangement of an area with an electric cyclic deicing systems (TÜV 1980).

to the maximum ice thickness allowed for the application and can be as long as 3 to 4 minutes for fixed-wing aircraft.

The biggest *disadvantage* of an electrical resistance system for large surfaces is the high power demand. If additional generators are installed just for the purpose of ice detection, the system will get very heavy.

Fluid Systems

Fluid ice protection systems operate on the *principle* that the surface to be protected is coated with a fluid that acts as a freezing point depressant (FPD). Current systems use a glycol-based fluid. When supercooled water droplets impinge on a surface, they combine with the FPD fluid to form a mixture with a freezing temperature below the temperature of the ambient air. The mixture then flows aft and is either evaporated or shed from the trailing edge of the surface. FPD fluid is distributed onto the surface leading edge by pumping it through porous material or spraying the fluid onto the surface.

The use of a freezing point depressant can provide antiicing or deicing protection. The antiicing mode is the normal mode of operation in light to moderate icing conditions. The deicing mode is a condition allowing ice to accumulate and bond to the wing surface. When the fluid ice protection system is turned on, a flow is introduced between the ice and the surface to weaken the bond so that the ice is shed by aerodynamic forces.

FPD fluid is stored in a tank. A pump meters the system's fluid flow requirements. Porous panels are constructed typically of sintered stainless steel mesh or laser-drilled titanium for the outer skin, a stainless steel or titanium backplate to form a reservoir, and a porous plastic liner to provide uniform control of panel porosity (Figure 12.50).

The principle *disadvantage* of the fluid protection system is the fluid storage requirement. The stored fluid weight may be significant when compared to other

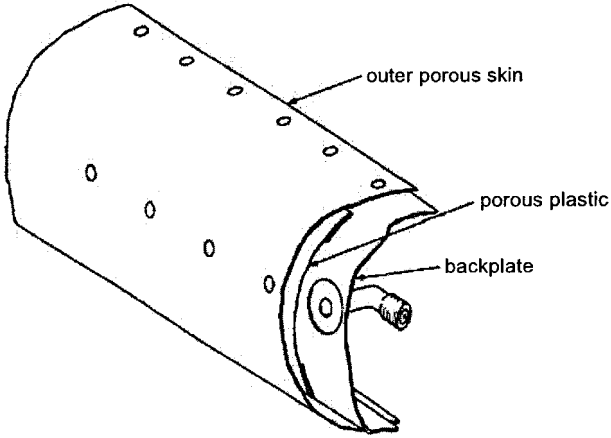


FIGURE 12.50 Construction of a typical porous panel (FAA 1993).

candidate ice protection systems. The system has a finite period of protection, dependent on fluid supply (FAA 1993).

Windshield Ice and Fog Protection

Windshield panels are usually provided with antiicing protection on those aircraft that are required to operate in all weather conditions. The most widely used system is an *electrical resistance system* for antiicing, whereby electric current is passed through a transparent conductive film or resistance wire that is part of the laminated windshield. The heat from the antiicing film or resistance wire also achieves internal defogging. Electrical heat may also be used to maintain the windshield layers of glass and plastic near the optimum temperature for resistance against bird strikes.

Where electric power seems not to be the adequate solution, an *external hot air blast system* can be an alternative. This system may also be used for rain removal.

Windshield Rain Protection

Rain-removal systems are designed to allow the pilots to have a clear view out of the cockpit at the airport and during departure and approach. The systems are not commonly used during flight at altitude. Rain may be removed by the use of windshield wipers. Alternatively, an external hot air blast can clear the windshield. In addition to either one of the two systems, a chemical rain repellent may be used.

Windshield wipers perform adequately, although their ability is limited. High oscillation rates are desirable to keep up with high rates of rain impingement during heavy rainfall. Sufficient blade pressure on the windshield must be maintained to produce satisfactory wiping when the aerodynamic forces are high at high aircraft speeds. Unfortunately, wipers also cause considerable aerodynamic drag.

An *external hot air blast* operates on the principle of blanketing the outside surface of the windshield with a protective wall of high-velocity, high-temperature

air. The air blast prevents water impingement by deflecting many of the incoming raindrops. Water on the surface that has penetrated the air blast will be evaporated.

Rain repellent may be sprayed on the windshield to form a transparent film that reduces the adhesive force between the water and the glass. The water draws up into beads that cover only a portion of the glass. The high-velocity slipstream continually removes the beads. Depending on the rain intensity, the rain impingement breaks down the repellent film, causing the window to return gradually to a wettable condition. Unless the windshield is wiped off frequently, the effectiveness or repeated repellent application decreases. Windshield wipers spread the repellent and improve its efficiency. Rain repellent used together with an external hot air blast is used in the critical landing phase when engine bleed air pressure is low and the jet blast is reduced.

Ice-Detection Systems

Some method of ice detection is necessary so that the ice protection system is operated only when necessary. Two methods exist: visual detection and electronic detection.

Visual detection is achieved by the flight crew monitoring such things as windshield wipers, wing leading edges, pylons, or landing lights that could serve as an ice datum. Those surfaces of the airplane directly exposed to stagnation flow conditions usually accumulate the largest quantity of ice. Wing and engine scan lights are used to monitor the engine intakes and the wing leading edges at night.

Electronic ice detectors consist of a probe extending into the free stream. The probe vibrates at a known frequency. When ice starts to build on the probe, the frequency will decrease. This will be detected by an attached controller. The controller will energize a heating element in the probe to remove the ice so that the probe can check again for icing conditions.

Example: Airbus A321

The ice and rain protection system lets the aircraft operate normally in ice conditions or heavy rain. Ice protection is given by the use of hot air or electrical power to make the necessary areas of the aircraft hot. The areas supplied by hot air are (Figure 12.51):

- The leading edge of slats 3, 4, and 5 on each wing
- The engine air intakes

The engine bleed air system supplies the hot air to the antiice system.

The items with electrical heaters are:

- The cockpit windshield and side windows
- The total air temperature (TAT) probes
- The angle of attack (alpha) probes
- The pitot and static probes of the air data system (ADS)
- The wastewater drain masts

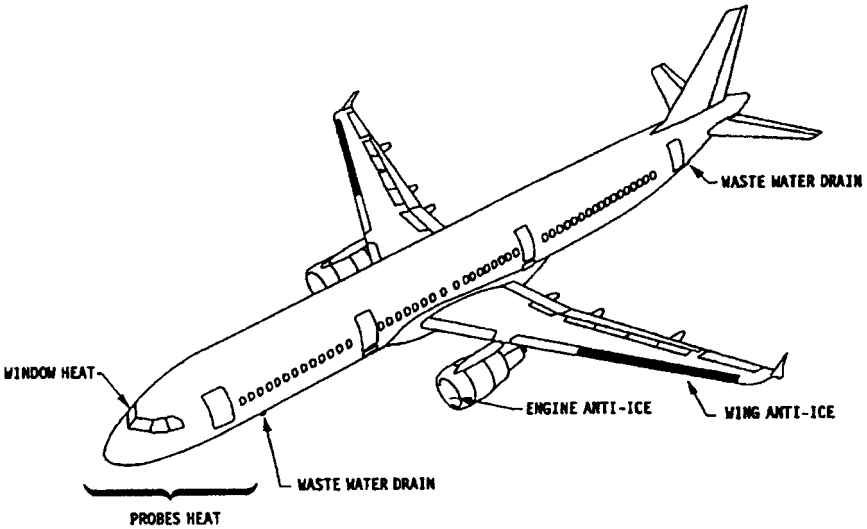


FIGURE 12.51 A321 ice and rain protection component locations.

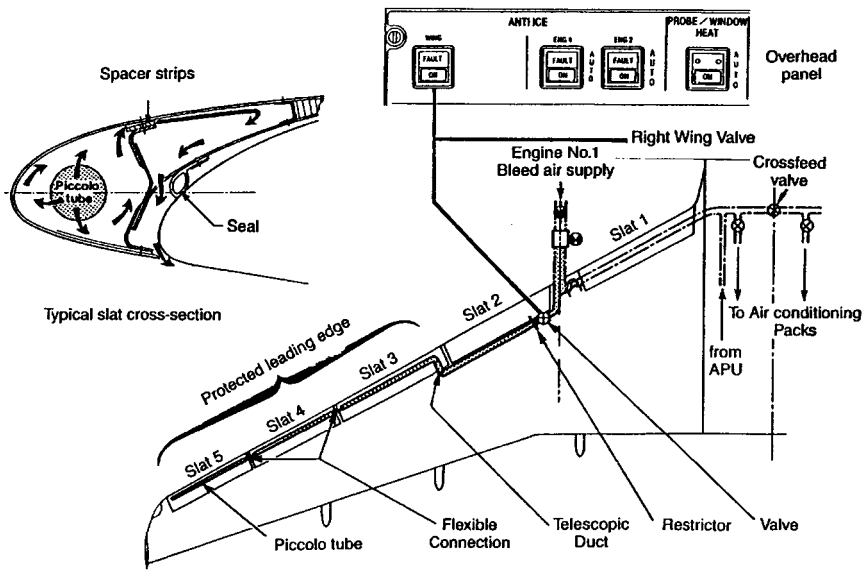


FIGURE 12.52 A321 wing anti-ice.

Rain is removed from the windshield with windshield wipers.

The A321 wing ice protection system is a hot air evaporative antiice system. Only slats 3, 4, and 5 on the outboard wing need to be ice protected. The hot air is bled from the engine. Each engine supplies its related wing. On both wings, an antiice valve isolates the anti ice system from the bleed air supply. When the cross-feed valve is open, it is possible to supply the two wings from only one engine bleed-air system. Lagged ducts connect the antiice valve to a telescopic duct at slat 3. A piccolo tube runs along slat 3, 4, and 5 and supplies the hot air to the leading edge. A piccolo tube is a tube with calibrated holes that ensures that hot air is evenly distributed along the leading edge, although bleed pressure decreases towards the wing tip. The bleed air in the slats is released overboard through the holes in the bottom surface of the slat. The operation of the antiice valve is controlled by the WING push-button switch on the ANTI ICE overhead panel in the cockpit.

12.10 LANDING GEAR (ATA 32)

Landing gear is addressed in Section 13 of this handbook. Landing gear as defined by ATA 100:

> Those units and components which furnish a means of supporting and steering the aircraft on the ground or water, and make it possible to retract and store the landing gear in flight. Includes tail skid assembly, brakes, wheels, floats, skids, skis, doors, shock struts, tires, linkages, position indicating and warning systems. Also includes the functioning and maintenance aspects of the landing gear doors but does not include the structure [of the doors].

Following ATA 100, the landing gear system may be subdivided into:

- Main gear and doors
- Nose gear and doors
- Extension and retraction system
- Wheels and brakes
- Steering system
- Position indicating and warning
- Supplementary gear (devices used to stabilize the aircraft while on the ground and prevent damage by ground contact)

Landing gear design has always been an integral part of aircraft design. The aircraft configuration cannot be laid out without due considerations given to the landing gear. Details of the steering system, the extension and retraction system, as well as the wheels and brakes may be the subject of separate studies.

Further Reading includes literature on landing gear design.

12.11 LIGHTS (ATA 33)

Lights as defined by ATA 100:

> Those units and components (electrically powered) which provide for external and internal illumination such as landing lights, taxi lights, position lights, rotating lights, ice lights, master warning lights, passenger reading and cabin dome lights, etc. Includes light fixtures, switches and wiring. Does not include warning lights for individual systems or self-illuminating signs.

Example: Airbus A321

Detailed requirements for instrument lights, landing lights, position lights, anticollision lights, ice-detection lights, and emergency lighting are laid down in the certification requirements sections 1381 to 1403 and 812. Much room for varying system designs is thus not permitted. Innovation has been brought in, however, through new lighting technologies and new circuit designs to control light intensities. The Airbus A321 lighting system provides illumination inside and outside of the aircraft. The system includes different parts.

The *cockpit lighting* consists of the following subsystems:

- General illumination of cockpit panels, instruments, and work surfaces
- Integral lighting of panels and instruments
- Test system for annunciator lights
- Dimming system for annunciator lights

The *cabin lighting* consists of the following subsystems:

- General illumination of cabin, galley areas, and entrances
- Illumination of the lavatories
- Passenger reading lights (customer option)
- Cabin lighted signs
- Work lights for the cabin attendants

The *cargo and service compartment lighting* provides illumination and power outlets for maintenance purposes. The system includes:

- Service area lighting for equipment and APU compartments
- Air conditioning duct and accessory compartment lights
- Cargo compartment lights
- Equipment compartment lights
- Wheel well lighting

The *external lighting* system illuminates the runways and/or taxiway, some aircraft surfaces, and gives an indication of the aircraft's position. The system (see Figure 12.53) consists of different lights:

AIRCRAFT SYSTEMS

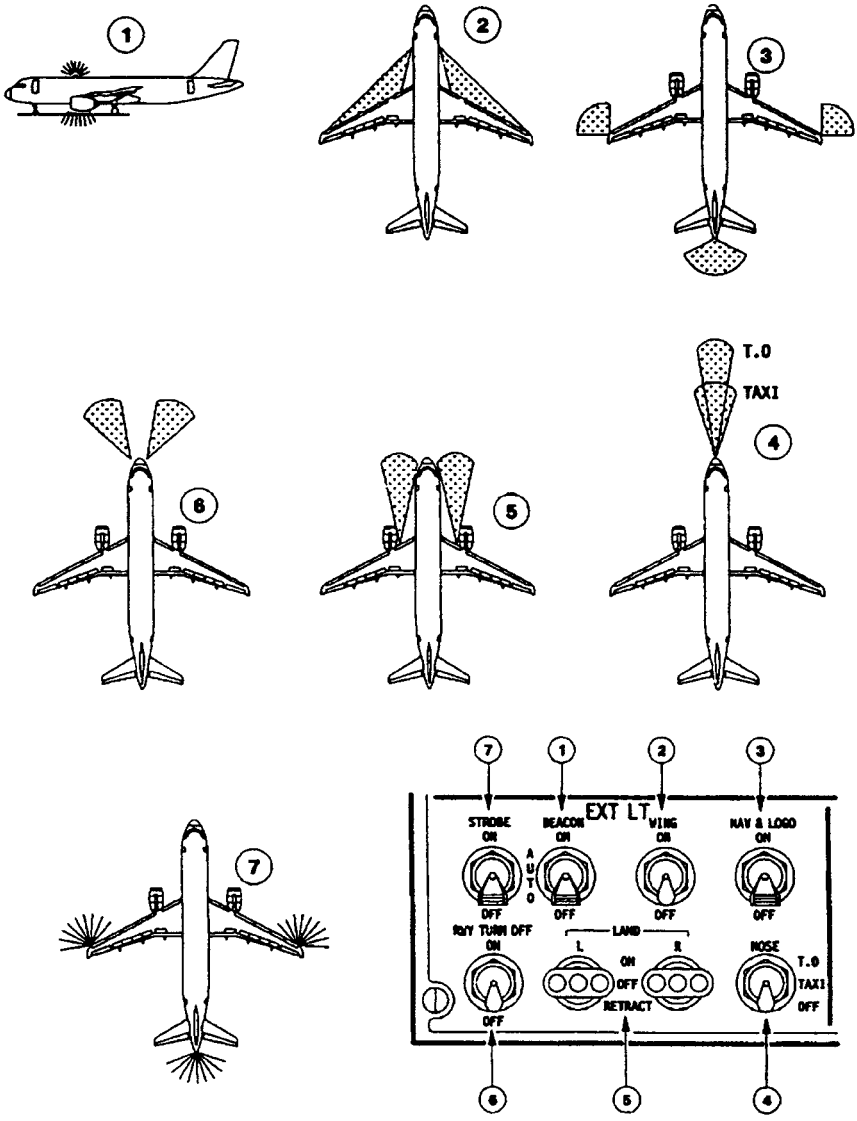


FIGURE 12.53 A321 external lights.

- Two anticollision beacon lights (1) which flash red, installed one at the top and one at the bottom of the fuselage
- Two wing and engine scan lights (2) installed one at each side of the fuselage to illuminate the wing leading edge and engine air intakes to detect ice accretion
- Three navigation lights (3), colored red (port), green (starboard), and white (tail), installed one at the tip of each wing and one aft of the fuselage
- Two logo lights (not shown) installed in the upper surface of each horizontal stabilizer to illuminate the company logo on the vertical stabilizer, provided the main gear struts are compressed or the flaps are extended
- One fixed-position takeoff light (4) (600 W) and one fixed-position taxi light (4) (400 W) installed on the nose landing gear
- Two retractable landing lights (5) (600 W) installed one under each wing
- Two fixed runway turnoff lights (6) installed on the nose landing gear
- Three synchronized strobe lights (7), one on each wing tip and one below the tail cone

The *emergency lighting* system provides illumination with batteries independently of the aircraft power supplies in the event of a failure of the main lighting system. Illumination is provided for:

- The cabin and the exit areas
- The exit location signs and the exit marking signs at all doors
- The door escape slides
- The marking system of the emergency escape path
- The lavatories

12.12 OXYGEN (ATA 35)

The oxygen system as defined by ATA 100:

> Those units and components which store, regulate, and deliver oxygen to the passengers and crew, including bottles, relief valves, shut-off valves, outlets, regulators, masks, walk-around bottles, etc.

Human Oxygen Requirements

The *human reaction to a lack of oxygen* depends on altitude. Normally, individuals living at sea level may become aware of the effects of altitude at about 3,048 m (10,000 ft). Above 10,000 ft, piloting skills are degraded. Up to 4,267 m (14,000 ft), the body is more or less able to compensate for the diminishing partial oxygen pressure by a higher breathing frequency. Above 14,000 ft, compensation is not possible anymore and hypoxia symptoms (headache, etc.) become apparent. Above 6,096 m (20,000 ft) unconsciousness and death are only a function of time. If a person is exposed to an altitude of 9,144 m (30,000 ft), unconsciousness may well set in after 1 minute. At an altitude of 15,240 m (50,000 ft), unconsciousness may set in after 10 s.

In order to compensate these effects, the partial oxygen pressure can be increased by breathing higher oxygen concentrations. The partial oxygen pressure[14] p at sea level (SL) is

$$0.21p_{SL} = 0.21 \cdot 1013 \text{ hPa} = 212.7 \text{ hPa}$$

If this partial pressure is to be maintained with altitude h, the required oxygen concentration x is

$$x = \frac{212.7 \text{ hPa}}{p(h)}$$

As can be seen from Figure 12.54, 100% (pure) oxygen is required at an altitude of about 37,000 ft. Beyond 37,000 ft, it becomes necessary to increase the pressure of the oxygen delivered to the mask in order to provide a sea-level equivalent environment. The lungs are in effect supercharged by the differential pressure between the mask and the surrounding pressure in the (nonpressurized) cabin.

It is evident that *cabin decompression* at high altitudes requires immediate action by the crew. Passengers and crew have to be provided with oxygen, and an emergency descent has to be initiated. The lower the aircraft gets, the longer the survival time. Circumstances are eased by the fact that even a big hole in the structure does not instantly lead to ambient pressure in the cabin.

Certification requirements for transport category aircraft (with pressurized cabins) state, e.g., "If certification for operation above 30 000 ft is requested, the dispensing units providing the required oxygen flow must be automatically pre-

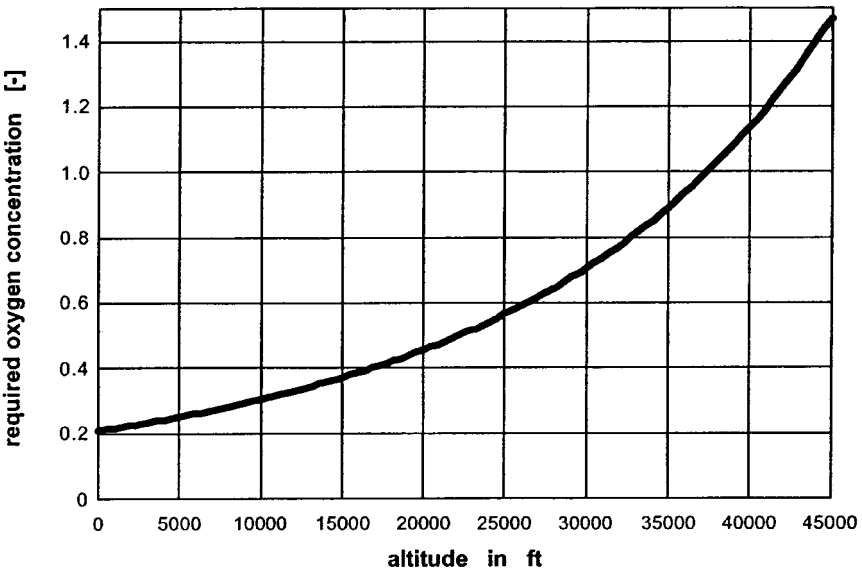


FIGURE 12.54 Required oxygen concentration with altitude.

[14] Many special terms relevant to the oxygen system are defined in Subsection 12.2.

sented to the occupants before the cabin pressure altitude exceeds 15 000 ft'' (JAR-25, section 1447).

System Classification

A classification of oxygen systems may take various aspects into account. We will look at classifications based on:

- Various reasons for oxygen supply
- Fixed versus portable oxygen equipment
- Oxygen regulator types
- Oxygen mask types
- Different oxygen sources
- The type of person supplied with oxygen:
  - Passenger oxygen system
  - Crew oxygen system

An *oxygen supply* may be necessary *for various reasons.* During high-altitude flights in nonpressurized cabins, normal oxygen supply is part of the normal flight procedures. In case of a failure of the normal supply, emergency oxygen is needed. In pressurized cabins emergency oxygen is supplied to all passengers and crew in case of cabin decompression. Provisions may have to be made for the supply of sustenance oxygen to a limited number of passengers after an emergency descent. Provisions also have to be made to supply first-aid oxygen to individual passengers for medical reasons. '' 'Supplemental oxygen' means the additional oxygen required to protect each occupant against the adverse effects of excessive cabin altitude and to maintain acceptable physiological conditions'' (JAR 1).

Oxygen equipment may be grouped into fixed and portable equipment. Fixed equipment is provided in those aircraft in which oxygen is frequently required or many passengers are involved. Additional portable equipment is used to allow the crew to move in the aircraft cabin under varying conditions. This could include the use of portable equipment when fighting small cabin fires. Portable equipment is also used for first-aid oxygen supplies to individual passengers. Small aircraft with nonpressurized cabins may not have a fixed oxygen system installed, so portable equipment is taken aboard whenever the situation arises due to planned high-altitude flights.

Oxygen taken from a bottle that provides a continuous flow via a supply hose directly into the mouth would technically be the easiest way to inhale. Although this was historically the first method applied, it has several disadvantages. The most apparent are:

1. Oxygen will be wasted during exhalation.
2. There is no need to inhale 100% oxygen at low altitudes.
3. There will be a need to hold the hose.
4. Communication will be hampered.
5. In a toxic environment (smoke) a face protection will be missing.

In order to overcome disadvantages 1 and 2, different *types of oxygen systems* based on the regulator design have evolved: the continuous flow system, the demand system, the pressure-demand system, the diluter-demand system, and the pressure-demand system with dilution at low altitudes. The most common systems in transport aircraft are the continuous flow system for passengers and the diluter-demand system for members of the flight crew.

Problems 3, 4, and 5 are addressed with the specific design of the *oxygen masks.*

Regulators

A *continuous-flow system* provides—as the name indicates—a continuous flow of oxygen to the mask. In order not to waste the volume of oxygen flowing towards the mask during exhalation, a flexible plastic or rubber reservoir is incorporated between the mask and the supply hose. The reservoir that is used to collect the oxygen has typically a volume of 0.5 to 1.0 l. During inspiration the stored oxygen can be used together with the oxygen currently flowing. Three valves are built into a continuous-flow mask: an exhalation valve and a nonreturn valve to the reservoir and a dilution valve. The exhalation valve opens the mask to ambient air during exhalation. At the same time, the nonreturn valve to the reservoir closes to prevent used air to enter the oxygen reservoir. When the reservoir has been emptied during the first part of the inhalation phase, the dilution valve opens and allows ambient air to dilute the already inhaled oxygen from the reservoir during the second part of the inhalation phase. The primary disadvantage of the constant-flow system is its inability to adjust itself automatically to various levels of physical activity. A regulator could, however, be provided for manual adjustment of flow to the reservoir. A constant-flow regulator provides automatic control of the flow depending on altitude. This capability evidently depends on the ability of the oxygen source to allow for varying flows. Varying the flow of oxygen is not always possible; the chemical oxygen generators commonly used in aircraft cabins do not allow flow control.

A *demand system* provides—as the name indicates—a flow of oxygen only on demand, i.e., during the inhalation phase, conserving oxygen during exhalation. A demand system requires a demand oxygen regulator for each user. The regulator may be panel-mounted, man-mounted, or seat-mounted. The regulator includes an outlet control valve that responds to minute changes in pressure. The slight negative pressure (compared to ambient cabin pressure) created within the mask at the onset of inhalation opens the valve and permits a flow of oxygen into the mask. At the end of the inhalation phase, the pressure has become slightly positive and the valve shuts off the flow. Masks for demand systems have to fit tightly. If the breather drew too much ambient air around the mask, the mask could not hold negative pressure and hence the regulator could not function properly.

A *pressure-demand system* is a demand system that has the ability also to supply oxygen under positive pressure (compared to ambient cabin pressure) to the mask. The principal components of the system are a mask that has the ability to hold positive pressure and an oxygen pressure regulator. A pressure-demand system is necessary for operation at altitudes above 10,668 m (35,000 ft) to maintain safe partial pressure for the user (compare with Figure 12.54).

The *diluter-demand system* is a demand system that has the ability to control the air–oxygen ratio automatically depending on altitude. The purpose of air dilution is to conserve the aircraft oxygen supply further and still maintain a safe partial

pressure. For safe operating conditions, dilution occurs up to 9,754 m (32,000 ft). At this altitude the dilution port in the diluter-demand oxygen regulator, which is automatically controlled, is shut off and the regulator delivers 100% oxygen. Besides an on-off-type supply lever, these regulators have an oxygen-selection lever to obtain 100% oxygen delivery throughout the whole altitude range. Some models are also provided with an emergency lever which, when actuated, will deliver a limited amount of positive pressure (safety pressure) for emergency toxic atmosphere protection (Figure 12.55).

Masks

Different oxygen masks exist. Apart from the differences resulting from the type of oxygen system for which they are used (see above), we may differentiate various types.

The *nasal mask* fits snugly around the nose and is intended for flights below 4,877 m (16,000 ft), where air intake through the mouth is acceptable. The *oronasal mask* fits completely over the mouth and nose. Provisions are made for the inclusion of a microphone for communication purposes. *Full-face masks* cover the mouth, nose, and eyes. These masks can meet protective breathing equipment requirements but cannot be used in a pressure-demand system because the eyes should not be exposed to a positive pressure. *Goggles* combined with an oronasal mask can meet both protective breathing and pressure-demand requirements.

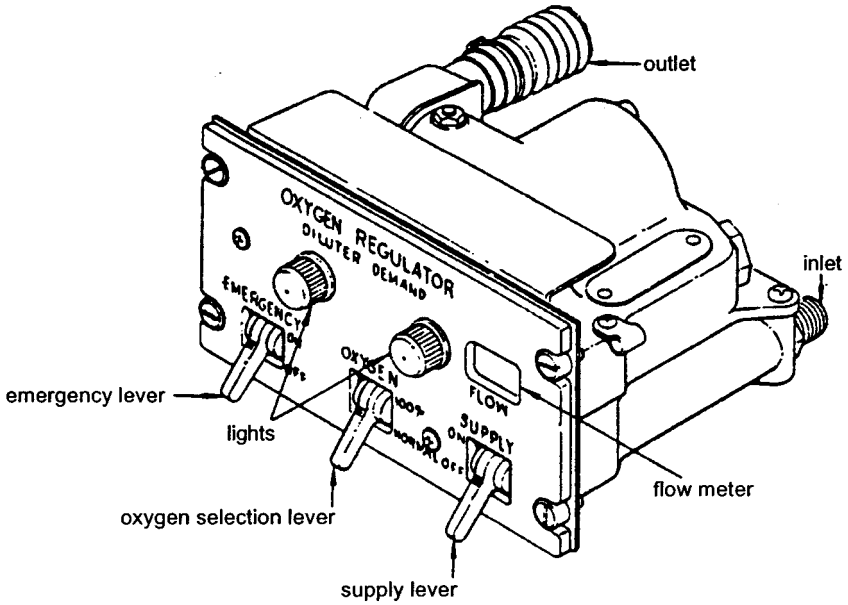


FIGURE 12.55 Basic panel-mounted diluter-demand oxygen regulator (VFW 614).

If certification for transport category aircraft is sought for operation above 25,000 ft, each flight crew member must be provided with a *quick-donning mask* (see Figure 12.57) that can be put on within 5 seconds (JAR-25, section 1447). Quick-donning masks are equipped with an inflatable harness. The crew member presses a side lever on the mask when passing the harness over the head. The side lever guides pressurized oxygen into the harness, causing the harness to stretch. When the side lever is released, the oxygen escapes from the harness and integrated straps pull the harness tightly to the head.

A *smoke hood,* a mask used to fight small cabin fires, protects the head and parts of the body and includes some type of oxygen supply.

Sources

Oxygen supply may be in the form of gaseous oxygen supply, liquid oxygen (LOX) supply, chemical oxygen supply, and on-board oxygen generation (OBOG).

Gaseous oxygen is stored in the aircraft in special oxygen cylinders. U.S. oxygen cylinders are colored green. They are properly marked and must only be filled with aviators breathing oxygen. Charge pressure is 12.8 MPa (1,850 psi). Oxygen cylinders are fitted with a combined flow-control and pressure-reducing valve as well as a pressure gauge. Two types of high-pressure cylinders exist: standard weight cylinders and lightweight cylinders. These cylinders are certified to Department of Defense (DOT) standards. They must regularly be checked and are life limited. Safety precautions have to be adhered to because of the general danger associated with such pressure vessels and the risk involved with handling oxygen. Crew oxygen systems on transport aircraft use gaseous oxygen.

Oxygen boils at sea-level pressure at −183 °C. The highest boiling point is −118 °C at 5.07 MPa. Hence, *liquid oxygen* has to be below that temperature. Liquid oxygen is stored in insulated tanks. Special equipment is required to convert liquid oxygen to gaseous oxygen on-board the aircraft. Liquid oxygen systems show weight and space savings compared to equivalent gaseous oxygen systems. Evaporation losses, however, can amount to 5% per 24 hours and need constant refilling in service. For these reasons, liquid oxygen systems are used on most combat aircraft but seem impractical for civil operation.

Chemical oxygen generation on aircraft is done with sodium chlorate. Sodium chlorate decomposes when heated to 478 °C into salt and oxygen:

$$2\ NaClO_3 + \text{heat} \rightarrow 2\ NaCl + 3\ O_2$$

The heat is generated with some kind of fuel, commonly iron. The chemical reaction is:

$$NaClO_3 + Fe \rightarrow NaCl + FeO + O_2 + \text{heat}$$

The overall mass balance of both equations combined: 100% sodium chlorate yields 45% oxygen by weight, 38% of which is delivered and 7% of which is used in oxidizing of the iron. The chlorate core is located in the center of the generator and is insulated against the outside steel housing. Nevertheless, the outside of the generator reaches temperatures of up to 260 °C, so that adjacent aircraft components need to be protected against the generator. The oxygen cools quickly and has reached normal temperatures when it arrives at the mask. The chemical reaction is

self-sustained and can be started mechanically (in most aircraft by pulling a lanyard) or electrically (Lockheed L-1011) with an adequate device on the generator. An outlet filter holds back particles and gaseous impurities. The reaction cannot be stopped once it is in progress. In case the outlet gets blocked, a pressure-relief valve averts an explosion of the generator. Figure 12.56 shows a cross-section of a chemical oxygen generator. Its diameter determines the flow rate and its length the duration of the supply. Generators are designed for a flow duration of about 15 minutes. The overall flow rate depends on the number of masks attached to the generator (1, 2, 3, or 4) and on certification requirements. The flow rate decreases over the duration of the supply. Most transport aircraft use chemical oxygen generation for the passenger oxygen system because of weight and maintenance savings compared with gaseous oxygen supply.

On-board oxygen generation systems (OBOGS) apply electrical power and bleed air to produce breathable oxygen from ambient air. Various techniques exist. Air can be processed through molecular sieve beds to provide oxygen-enriched breathing gas.

Example: Airbus A321

The aircraft has three separate oxygen systems: a flight crew oxygen system, a passenger oxygen system, and a portable oxygen system.

The *flight crew oxygen* system (Figure 12.57) supplies oxygen to the flight crew if there is a sudden decrease in cabin pressurization. It also supplies oxygen if there is smoke or dangerous gases in the cockpit. Each crew station has a quick-donning mask with a demand regulator installed. The oxygen is supplied from a high-pressure oxygen cylinder to the masks through a pressure regulator/transmitter assembly and a distribution circuit.

The *passenger oxygen* system provides emergency oxygen for passengers and cabin attendants (Figure 12.58). Emergency oxygen containers are installed:

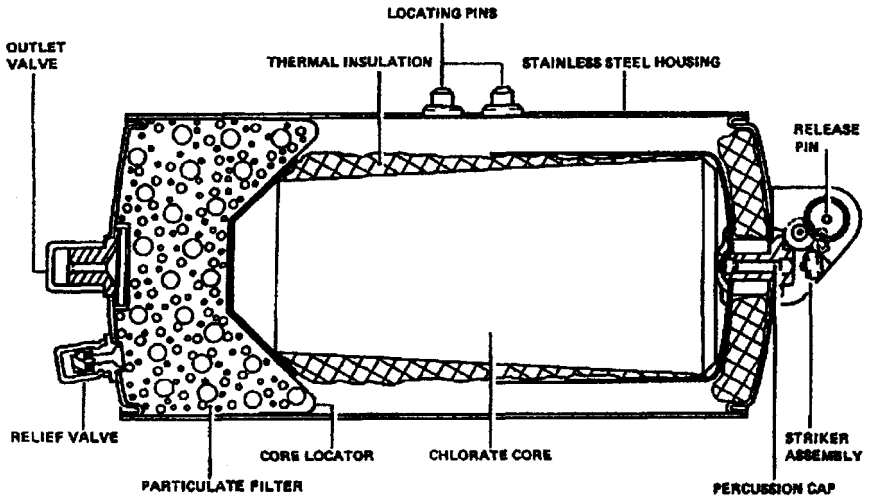


FIGURE 12.56 Chemical oxygen generator (Airbus A321).

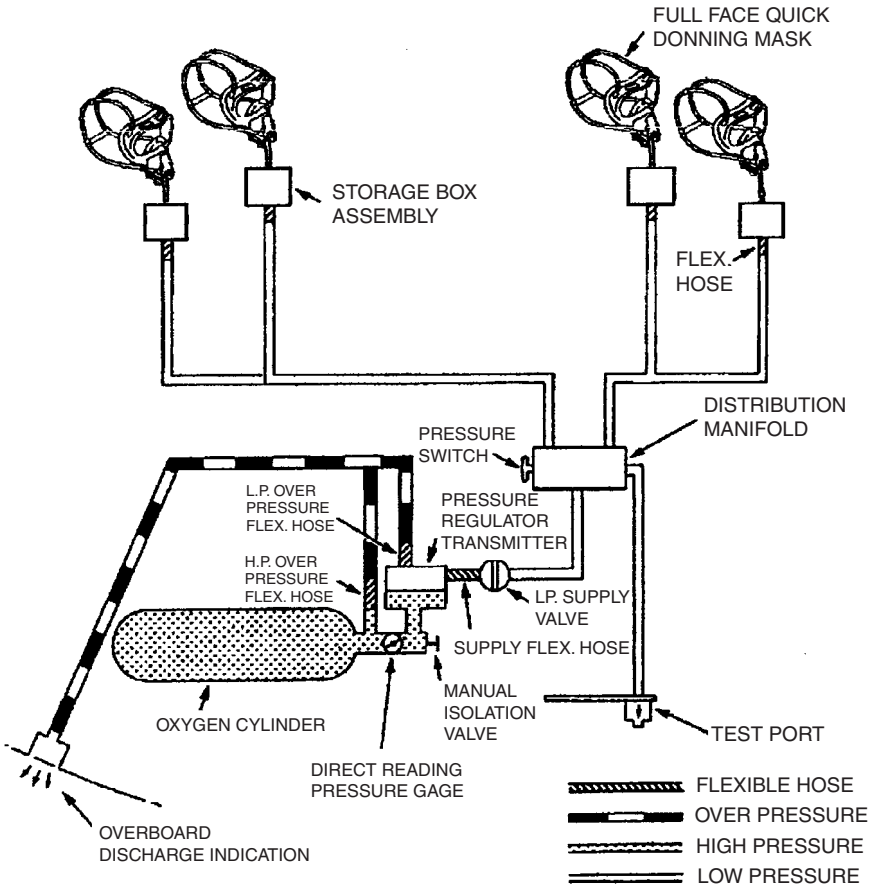


FIGURE 12.57 A321 crew oxygen system.

- Above the passenger seats
- In the lavatories
- At the cabin attendant stations
- In the galley working areas

Each container has a chemical oxygen generator and two or more continuous-flow oxygen masks, each with a flexible supply hose.

12.13 PNEUMATIC (ATA 36)

The pneumatic system as defined by ATA 100:

> Those units and components (ducts and valves) which deliver large volumes of compressed air from a power source to connecting points for such other systems as air conditioning, pressurization, deicing, etc.

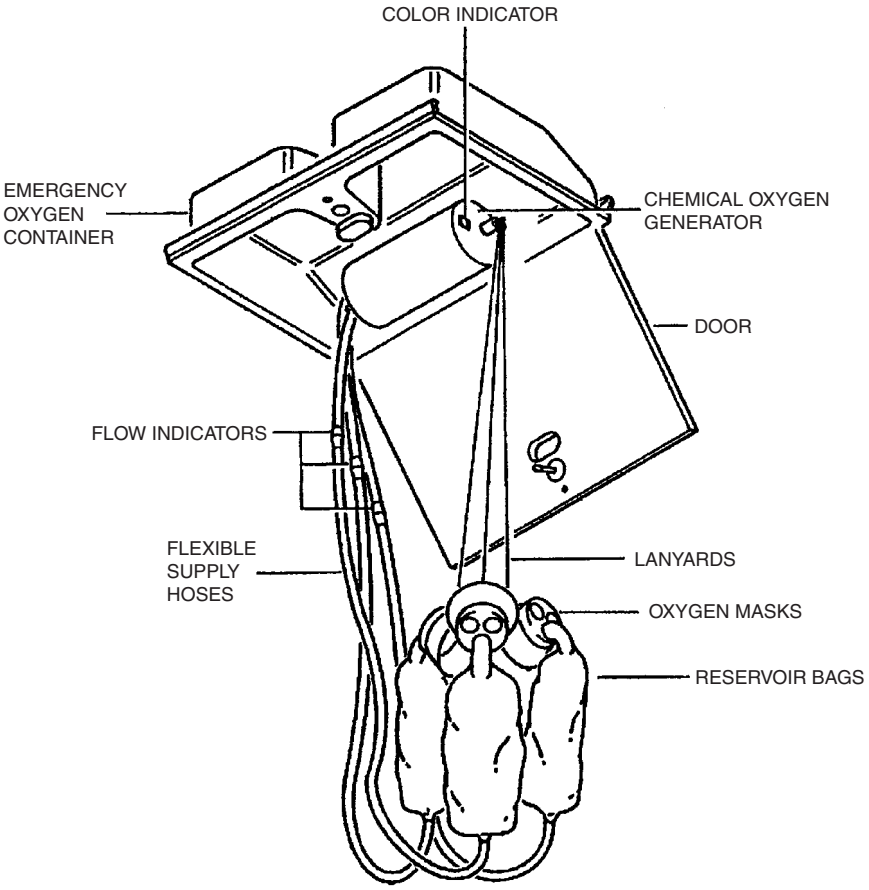


FIGURE 12.58 A321 emergency passenger oxygen container.

High-Pressure Pneumatic Systems

High-pressure pneumatic systems must be differentiated from low-pressure pneumatic systems. High-pressure pneumatic systems, much like hydraulic systems, may apply a nominal system pressure of 20.7 MPa (3,000 psi). In contrast, low-pressure pneumatic systems may operate at only 0.3 MPa (44 psi).

High-pressure pneumatic systems work very similarly to hydraulic systems. The difference is that in pneumatic systems compressible air is used instead of incompressible hydraulic fluid. Pneumatic systems do not need a reservoir because air is directly available from the operating environment. The air is put to high pressure in a compressor. The pneumatic pressure is stored in an air storage bottle. The bottle can provide a short-burst reserve flow for heavy operations, or limited emergency flow in case of compressor failure. The compressed air is routed through

tubes, filters, moisture separators, and valves to the consumer. After having done its duty at the consumer, the air is simply released. In a high-pressure system it is of the utmost importance that the air in the system be completely dry. Moisture in the system can cause freezing of units and thus interfere with normal operation. High-pressure pneumatics have been applied, e.g., for landing gear extension and retraction, nose wheel steering, as well as to wheel and propeller braking. The Fairchild Hiller FH-227 is equipped with such a high-pressure pneumatic system.

High-pressure pneumatics shows advantages and disadvantages compared to hydraulics in aircraft operation:

- *Advantages:*
  - Air is a readily available, nonaggressive, clean, and lightweight fluid.
  - There is no need for return lines.
- *Disadvantages:*
  - Due to compressibility of the air, pneumatic systems lack the instant response that hydraulic systems provide.
  - The rate of movement of pneumatic actuators is highly load-dependent.
  - An actuator position cannot easily be controlled since even when the flow has stopped, the actuator will move in response to load variations.
  - Pneumatic systems are inefficient in transmitting power because energy is lost in compressing the air.

The many more disadvantages than advantages explain why high-pressure pneumatic systems are rarely used. This is much different than the low-pressure pneumatics used extensively on most aircraft.

Low-Pressure Pneumatic Systems

Low-pressure consumers include:

- Air conditioning (including cabin pressurization)
- Wing and engine antiicing
- Engine starting
- Hydraulic reservoir pressurization
- Potable water pressurization
- Air-driven hydraulic pumps

One aircraft type will not necessarily use all these pneumatic functions.

Pressurized air is generated and used in aircraft ranging from light single-engine aircraft up to big turbine-powered transport aircraft. The simplest *source of pressurized air* is ram air. Reciprocating engines can supply pressure from a supercharger (driven by the engine primarily used to produce compressed air for the combustion process), a turbocharger (similar to a supercharger but driven by exhaust gases), or an engine-driven compressor. Turbine-powered aircraft usually use bleed air as a source for compressed air. The bleed air system will now be explained in more detail.

The engine *bleed air system* extracts pressurized air from one or more bleed ports at different stages of the engine compressor of each engine on the aircraft. The system controls the pressure and temperature of the air and delivers it to a

distribution manifold. The pressure is controlled by a pressure-regulating valve and the temperature is lowered in a precooler with fan air or ram air. Bleed air from alternate sources such as the auxiliary power unit (APU) or a ground cart is also connected to the distribution manifold. The consumers are supplied from the distribution manifold. Additional bleed air from each engine may be taken directly off the engine (independent from the pneumatic system) for engine demands such as engine intake antiice. Isolation valves and a crossbleed valve are required in the distribution manifold to maintain essential functions in the event of a failure in the supply or in a consumer. Check valves are required to prevent reverse flow. The Airbus A321 (Figure 12.59) shows all those elements that are typical for a conventional bleed air system.

Pressure control is set to the lowest level acceptable to all consumers. Engine bleed port switching is designed to use intermediate pressure (IP) bleed air during cruise. When intermediate stage bleed pressure is not adequate, the system switches automatically to off-takes from the high-pressure (HP) stage. A check valve prevents air from flowing back to the IP port. Pressure control may be pneumatic or computer controlled electropneumatic.

With modern high-bypass-ratio engines the *fuel burn penalty* of a given amount of engine bleed air has been decreased. However, high-bypass-ratio engines also show decreased total compressor airflow relative to engine thrust. Hence, less bleed air is available from these engines. The economic impact of the bleed air system is by no means negligible. An overall economically optimum solution has to take into account all aspects that were named in Subsection 12.1 under Costs and Trade-off Studies. These design details could be considered:

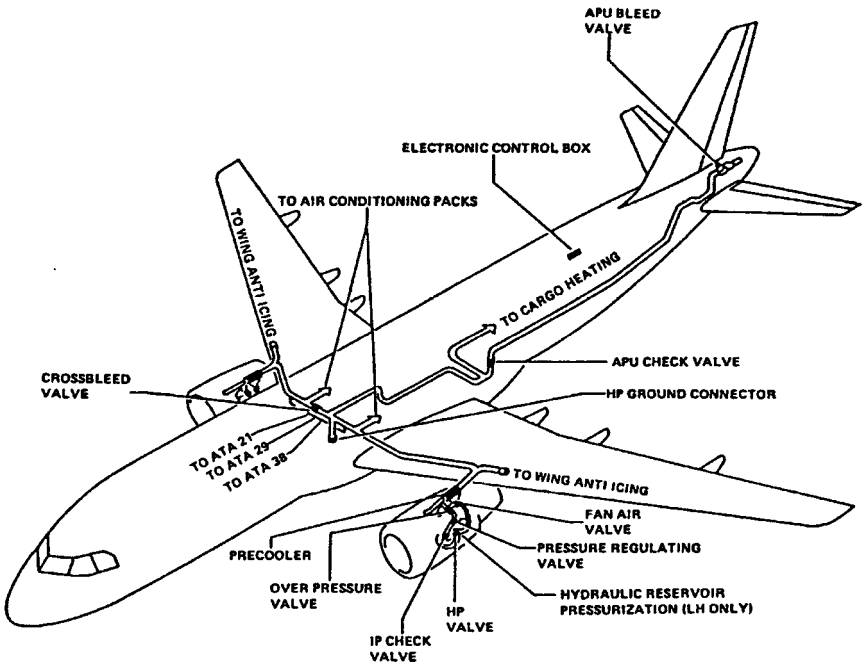


FIGURE 12.59 A321 pneumatic system overview.

- Use of lowest acceptable compressor stage bleed port
- Strict control of leakage from pneumatic systems
- Optimized precooler design with a trade-off among weight, price, and coolant air usage
- Optimum proportioning of bleed flows from multiengine installations
- Use of multiple bleed ports, i.e., tapping at more than the typical two compressor stages
- Consideration of alternate sources of compressed air (APU, mixing ejector, auxiliary compressor: engine driven, pneumatic, hydraulic, or electric driven)

At the airport, an external supply with pressurized air (in contrast to an APU supply) is environmentally more friendly and can also be more economical.

Example: Airbus A321

The A321 pneumatic system supplies high-pressure hot air to these *consumers:*

- Air conditioning
- Engine starting
- Wing antiicing
- Hydraulic reservoir pressurization
- Potable water pressurization

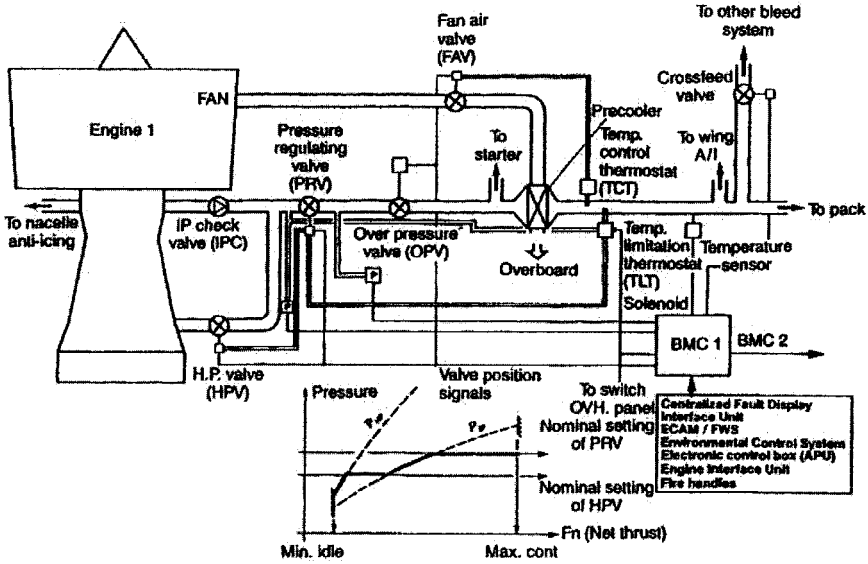


FIGURE 12.60 A321 schematic diagram of the pneumatic system.

There are two *engine bleed systems* (Figure 12.59): the left side (engine 1) (Figures 12.60 and 12.61) and the right side (engine 2). A crossbleed duct connects both engine bleed systems. A crossbleed valve mounted on the crossbleed duct allows the left and the right side to be either interconnected or separated. During normal operation, the crossbleed valve is closed and the systems are separated. There are two interconnected bleed monitoring computers. BMC 1 is used primarily for engine 1 bleed system, and BMC 2 is used primarily for engine 2 bleed system.

Air is normally bled from the IP valve. When IP pressure is not sufficient, the HP valve opens. This happens at low engine speeds, especially during descent, with engines at idle. *Pressure regulation* is done downstream of the junction of HP and IP ducting with the pressure-regulating valve (PRV), which acts as pressure regu-

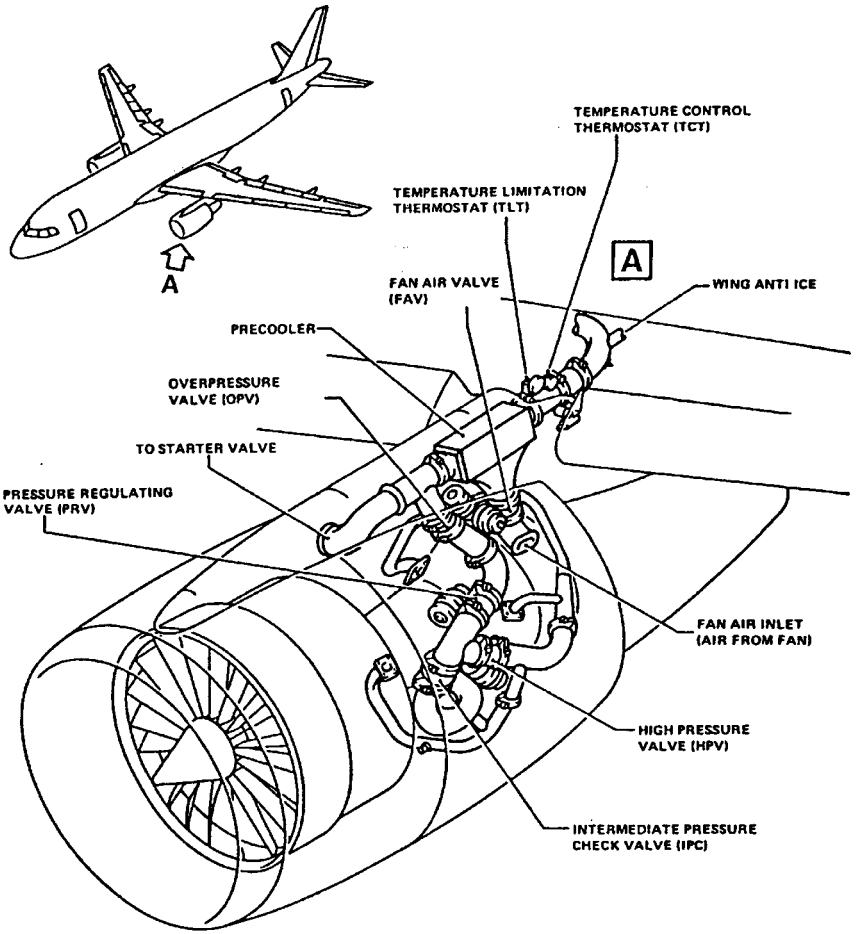


FIGURE 12.61 A321 engine bleed air supply components.

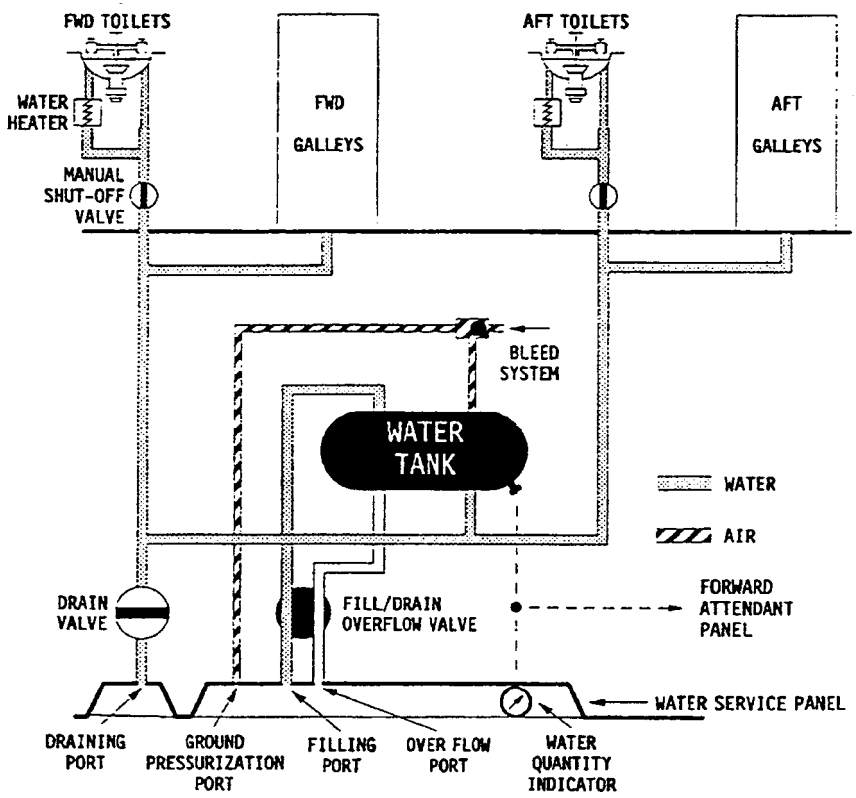


FIGURE 12.62 A321 potable water system.

lator and shut-off valve. Delivery pressure is regulated to 0.3 MPa (44 psi). When pressure is excessive in a failure case, an over-pressure valve (OPV) closes.

Temperature regulation of the bleed air is achieved with a fan air valve (FAV) and an air-to-air crossflow tubular heat exchanger called a precooler. The precooler uses cooling air bled from the engine fan to regulate the original bleed air with a temperature of up to 400 °C down to a delivery temperature of 200 °C.

12.14 WATER/WASTE (ATA 38)

The water/waste system as defined by ATA 100:

> Those fixed units and components which store and deliver for use, fresh water, and those fixed components which store and furnish a means of removal of water and waste. Includes wash basins, toilet assemblies, tanks, valves, etc.

System Classification

The water/waste system may be divided into three subsystems:

1. The potable water system is used to store and deliver fresh drinking water.
2. The wastewater drain system disposes the wastewater from lavatory washbasins and galley sinks.
3. The toilet system gives sanitary facilities to passengers and crew.

Potable Water Systems

The potable water system delivers drinking water to faucets and coffee makers in the galleys and to faucets and (in some cases) toilet bowls in the lavatories.

The water is stored in tanks made from composite material. Sensors on the tank measure the water quantity. The distribution system delivers the water through lines to the consumers. In critical areas, lines and valves are protected against freezing by insulation material and electrical heating elements. Nevertheless, water must be drained from the potable water system if the aircraft is parked overnight at temperatures below freezing.

If water left the tank just by gravity, the exit pressure would be very low. For this reason, gravity dispensing is applied only on small aircraft. On most aircraft, potable water tanks located below the cabin floor are pressurized with air. The pressurized air exerts a pressure on the water surface in the tank and thus enables water distribution at a higher pressure. The tanks may be pressurized with bleed air from the engines or the APU. Alternatively, air could be pressurized with a dedicated compressor. On the ground it is also possible to pressurize the tanks from an external pressure source.

In-service measurements have shown an average water consumption of about 0.2 l per passenger (pax) per hour in aircraft with a vacuum toilet system. This amount is made up of:

- 0.11 l/pax/hr consumed in the washbasin
- 0.07 l/pax/hr used for toilet rinsing
- 0.02 l/pax/hr consumed in the galley

Wastewater Systems

The wastewater system disposes the wastewater from lavatory washbasins and galley sinks. Commonly, wastewater is drained overboard through drain valves via drain lines to drain masts on the lower side of the fuselage. The drain masts are electrically heated to prevent water from freezing on exit. The drain valve in the drain line prevents leakage of cabin air through the drain line. *Note:* toilet waste is never drained overboard.

Principally, the wastewater could also be disposed into the waste tanks together with toilet waste. This technique, however, would increase aircraft weight compared with draining the wastewater. The wastewater could also be reused on board for flushing of vacuum toilets. This would save potable water taken on board and would therefore reduce aircraft weight.

Toilet Systems

Two types of toilet systems are in use: the chemical toilet system and the vacuum toilet system.

Waste tanks of recirculating liquid *chemical toilet systems* are precharged with a dye–deodorant–disinfectant chemical flushing liquid. Sensors on the tank measure the waste quantity. A tank-mounted motor/pump/filter assembly develops pressure to flush the toilets. A flush signal is generated when the flush control lever on a toilet is pressed. This signal is electronically processed and opens the flush valve. Subsequently, pressurized and filtered flushing liquid rinses the toilet bowl. The waste and the flushing liquid enter the waste tank. The waste tanks are vented overboard.

Simpler chemical toilet systems are operated with a toilet-mounted foot pedal that is connected to a mechanical pump.

The *vacuum toilet system* (Figure 12.63) is described in the Airbus example.

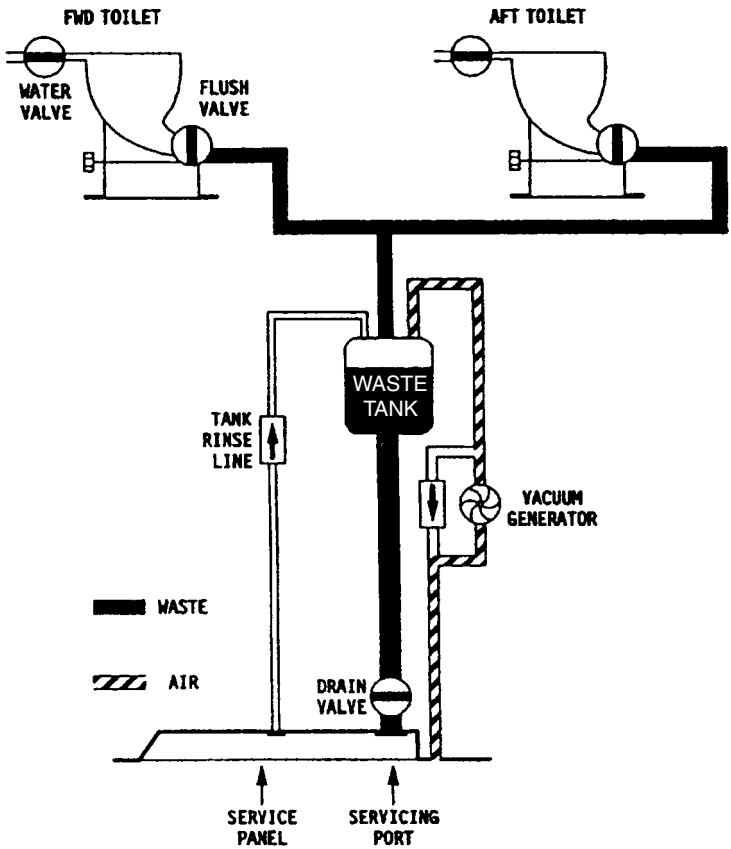


FIGURE 12.63 A321 vacuum toilet system.

Example: Airbus A321

The *potable water system* supplies water from a water tank (200 l) through a distribution system. Potable water is supplied to water faucets in the galleys and lavatories. The system also supplies potable water to the water heaters, which are located below the lavatory washbasins, and to the toilet bowls for rinsing. Water lines in cold areas of the aircraft are insulated and heated to avoid freezing. Air pressure is used to pressurize the potable water system. The air is supplied from the bleed air system or the ground pressure connection.

The A321 is equipped with a *vacuum toilet system.* It removes waste from the toilet bowls through a vacuum drain to an under floor waste tank (170 l). Toilet wastes are flushed to the waste storage tank under the effect of differential pressure between the cabin and the waste tank. On ground and at low altitudes (below 16,000 ft) a vacuum generator produces the necessary differential pressure. At high altitudes (above 16,000 ft), ambient pressure alone ensures the differential pressure. A vacuum system controller (VSC) controls the operation of the vacuum generator. The system uses water from the aircraft potable water system to flush the toilet. A flush control unit (FCU) in each toilet controls the flush process. During ground service, the waste holding tank is emptied, cleaned, and filled with a prescribed quantity of sanitary fluid.

12.15 AIRBORNE AUXILIARY POWER (ATA 49)

Airborne auxiliary power as defined by ATA 100:

> Those airborne power plants (engines) which are installed on the aircraft for the purpose of generating and supplying a single type or combination of auxiliary electric, hydraulic, pneumatic or other power. Includes power and drive section, fuel, ignition and control systems; also wiring, indicators, plumbing, valves, and ducts up to the power unit. Does not include generators, alternators, hydraulic pumps, etc. or their connecting systems which supply and deliver power to their respective aircraft systems.

Fundamentals

An auxiliary power unit (APU) is a compact, self-contained gas turbine-powered unit delivering rotating shaft power, compressed air, or both. Rotating shaft power can be used to drive a generator, a hydraulic pump, and/or a load compressor. An APU includes the air intake and exhaust, the fuel and oil system, engine controls and indications, as well as ignition and starting equipment.

An APU may be used on the ground and in the air, or only on the ground. For the overall aircraft system safety concept it makes a difference if the APU is dependable or not. If overall safety depends on the APU, then the APU is essential; otherwise it is nonessential.

An *essential APU* is ''an APU which produces bleed air and/or power to drive accessories necessary for the dispatch of the aircraft to maintain safe aircraft operation'' (JAR-1).

A *nonessential APU* is ''an APU which may be used on the aircraft as a matter of convenience, either on the ground or in flight, and may be shut down without jeopardising safe aircraft operations'' (JAR-1).

An essential APU is necessary for dispatch. For the pilot this will be indicated on the minimum equipment list (MEL).

The APU is installed in the tail cone of most airplanes, isolated from flight-critical structure and control surfaces by a firewall. The APU is started by battery. When running, the APU is able to start the main engines with its pneumatic power supply.

The significance of APU power within the concept of the secondary power systems is explained in Subsection 12.1 under Power.

Example: Airbus A321

The A321 is equipped with an APU (Figure 12.64) to permit aircraft *ground operation* independent from external power supply, allowing the operator to service airports without adequate ground power facilities. The APU is also available in flight. This is of importance for flights under extended-range twin-engine operations (ETOPS) rules, where the aircraft flies on remote routes with no alternative airfield available within a flight time of up to 180 minutes.

The APU essentially generates shaft power. A load compressor is flanged to the shaft to generate *pneumatic power.* With APU pneumatic power it is possible to start the aircraft main engines and operate the air conditioning system.

The APU shaft also drives a 90 kVA generator via a gearbox to generate *electrical power.* The APU is regulated to a constant speed, so that the generator is able to produce 110 V ac at a constant frequency of 400 Hz. If an increase in demand to the aircraft systems is necessary, the supply of the electrical power has priority over the supply of bleed air.

The APU is fitted with a dc starter motor, which draws its power from the electrical system battery bus. The APU starts in flight up to an altitude of 7,620 m (25,000 ft) with the use of the aircraft batteries alone. The starter motor turns the engine to such speed that self-sustained engine operation becomes possible. The electronic control box (ECB) automatically controls and monitors the APU. Manual control of the APU is possible through the crew interfaces in the cockpit. The APU

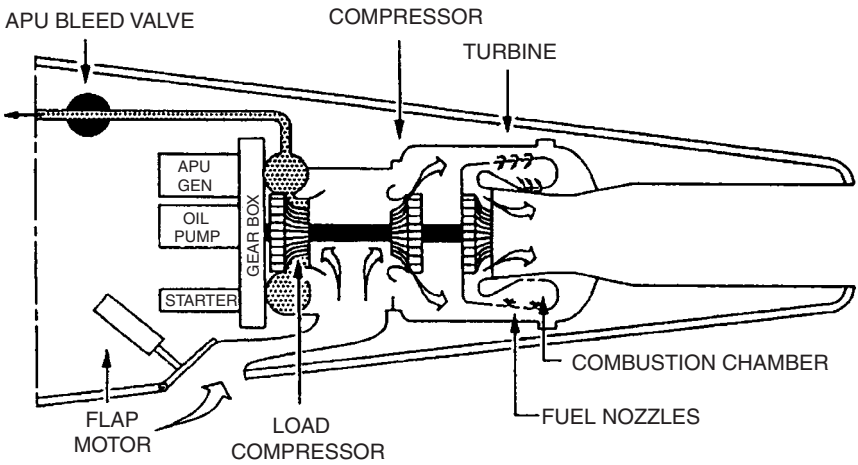


FIGURE 12.64 A321 auxiliary power unit (APU).

is supplied with fuel from the aircraft tanks. The APU compartment is equipped with a fire detection and extinguishing system.

12.16 AVIONIC SYSTEMS

Avionic systems are dealt with in Section 11 of this handbook. For the sake of completeness, definitions of the avionic system are given here in the same way as above for the nonavionic systems. Introductory information can also be obtained from the related literature given in Further Reading.

Auto Flight (ATA 22)

Details of the auto flight system are covered in Section 11. The auto flight as defined by ATA 100:

> Those units and components which furnish a means of automatically controlling the flight of the aircraft. Includes those units and components which control direction, heading, attitude, altitude and speed.

The most important parts of the auto flight system are the autopilot and the auto throttle (auto thrust) system.

The autopilot is (ATA 100):

> that portion of the system that uses radio/radar signals, directional and vertical references, air data (pitot-static), computed flight path data, or manually induced inputs to the system to automatically control the flight path of the aircraft through adjustment to the pitch/roll/yaw axis or wing lift characteristics and provide visual cues for flight path guidance, i.e.: Integrated Flight Director. This includes power source devices, interlocking devices and amplifying, computing, integrating, controlling, actuating, indicating and warning devices such as computers, servos, control panels, indicators, warning lights, etc.

and the auto throttle is

> that portion of the system that automatically controls the position of the throttles to properly manage engine power during all phases of flight/attitude. This includes engaging, sensing, computing, amplifying, controlling, actuating and warning devices such as amplifiers, computers, servos, limit switches, clutches, gear boxes, warning lights, etc.

Communication (ATA 23)

Details of the communication system are covered in Section 11. Communication systems as defined by ATA 100:

> Those units and components which furnish a means of communicating from one part of the aircraft to another and between the aircraft or ground stations, includes voice, data, C-W communicating components, PA [Passenger Address] system, intercom and tape reproducer-record player.

The communication system includes (ATA 100):

- Speech communication: Radio communication air-to-air, air to ground. HF, VHF, UHF radio communication, in-flight telephone, and satellite receiver
- Data transmission and automatic calling: Selcal (Selected Call) and ACARS (Aircraft Communicating Addressing and Reporting System)
- Passenger address and entertainment system:[15]
  - Entertainment: Audio, overhead video, in-seat video, interactive video, in-seat telephone, video on demand, Internet systems, and seat power supply system for passenger laptops
  - Passenger address system: The system to address the passengers from the cockpit or the cabin crew station, playback of automatic recordings, boarding music, or acoustic signs.
- Audio integrating: Controls the output of the communications and navigation receivers into the flight crew headphones and speakers and the output of the flight crew microphones into the communications transmitters; also includes the interphone, used by flight and ground personnel to communicate between areas on the aircraft
- Integrated automatic tuning of navigation transmitters and receivers
- Cockpit voice recorder

Indicating/Recording Systems (ATA 31)

The indicating/recording system deals primarily with the instrument panels and controls. This aspect is covered in Section 11 of this handbook. Indicating/recording systems as defined by ATA 100:

> coverage of all instruments, instrument panels and controls... Includes systems/units which integrate indicating instruments into a central display system and instruments not related to any specific system.

The indicating/recording system includes (ATA 100):

- The instrument and control panels (Figure 12.65)
- Independent instruments (not related to any other aircraft system)
- Flight data recorder, recorders for performance or maintenance data
- Central computers, central warning and display systems

[15] In ATA 2200 the passenger address and entertainment system has become the "cabin systems (ATA 44)" in its own right. Definition: "Those units and components which furnish a means of entertaining the passengers and providing communication within the aircraft and between the aircraft cabin and ground stations. Includes voice, data, music and video transmissions."

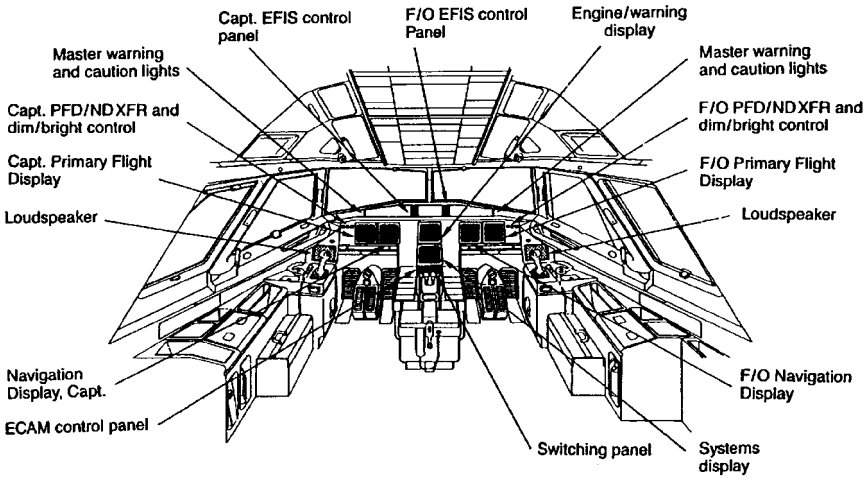


FIGURE 12.65 A321 general cockpit arrangement and instrument layout.

Navigation (ATA 34)

Details of the navigation system are covered in Section 11. The navigation system as defined by ATA 100:

> Those units and components which provide aircraft navigational information. Includes VOR, pitot, static, ILS, ... compasses, indicator, etc.

Data handling of the navigation system includes (ATA 100):

- Flight environment data (pitot/static system, rate of climb, airspeed, etc.)
- Magnetic data (magnetic compass)
- Independent data (inertia guidance systems, weather radar, Doppler, proximity warning, collision avoidance)
- Dependent data (DME, transponder, radio compass, LORAN, VOR, ADF, OMEGA, GPS)
- Data from landing and taxiing aids (ILS, marker)

ACKNOWLEDGMENT

All figures named "A321" are by courtesy of Airbus. They are taken from the *Aircraft Maintenance Manual* (AMM), the *Flight Crew Operations Manual* (FCOM), or other material prepared or used for flight maintenance training. *At no time should the information given be used for actual aircraft operation or maintenance. The information given is intended for familiarization and training purposes only.*

REFERENCES

AC 25-17. Federal Aviation Administration, Department of Transportation. 1991. *Transport Airplane Cabin Interiors Crashworthiness Handbook* (AC 25-17), available online from http://www.faa.gov.

AC 25-22. Federal Aviation Administration, Department of Transportation. 2000. *Certification of Transport Airplane Mechanical Systems* (AC 25-22), available online from http://www.faa.gov.

AC 135-16. Federal Aviation Administration, Department of Transportation. 1994. *Ground Deicing and Anti-Icing Training and Checking* (AC 135-16), available online from http://www.faa.gov.

AC 25.803. Federal Aviation Administration, Department of Transportation. 1989. *Emergency Evacuation Demonstration* (AC 25.803), available online from http://www.faa.gov.

ACJ-25. Joint Aviation Authorities. *Joint Aviation Requirements for Large Aeroplanes* (JAR 25), Section 2, Acceptable Means of Compliance and Interpretations (ACJ), available from the JAA (see http://www.jaa.nl).

AGARD. 1980. *Multilingual Aeronautical Dictionary,* Advisory Group for Aerospace Research and Development, Neuilly sur Seine, available online from NATO's Research and Technology Organisation, http://www.rta.nato.int.

AIR 1168/3. Society of Automotive Engineers (SAE). 1990. *Aerothermodynamic Systems Engineering and Design,* SAE, Warrendale, PA (AIR 1168/3).

AIR 1168/4. Society of Automotive Engineers. 1990. *Ice, Rain, Fog, and Frost Protection,* SAE, (AIR 1168/4).

AIR 1609. Society of Automotive Engineers (SAE). 1982. *Aircraft Humidification,* SAE, Warrendale, PA (AIR 1609).

AIR 171. Society of Automotive Engineers (SAE), *Glossary of Technical and Physiological Terms Related to Aerospace Oxygen Systems,* SAE, Warrendale, PA (2000) (AIR171D).

AMJ-25. Joint Aviation Authorities. *Joint Aviation Requirements for Large Aeroplanes* (JAR-25), Section 3, Advisory Material Joint (AMJ), available from the JAA (see http://www.jaa.nl).

ARP 85. Society of Automotive Engineers (SAE). 1981. *Air Conditioning Systems for Subsonic Airplanes,* SAE, Warrendale, PA (ARP 85E).

ARP 1270. Society of Automotive Engineers (SAE). 2000. *Aircraft Pressurization Control Criteria,* SAE, Warrendale, PA (ARP 1270),

ARP 1280. Society of Automotive Engineers (SAE). 1994. *Application Guide for Hydraulic Power Transfer Units*. SAE, Warrendale, PA (AIR 1280A).

ATA 100. Air Transport Association of America (ATA). 1999. *Manufacturers' Technical Data* (ATA Spec 100), ATA, Washington, DC, available from ATA, http://www.airlines.org.

ATA 2200. Air Transport Association of America (ATA). 1999. *Information Standards for Aviation Maintenance* (ATA iSpec 2200), ATA, Washington, DC, available from ATA, http://www.airlines.org.

Boeing Company, Weight Research Group. 1968. *Weight Prediction Manual—Class I,* The Boeing Company, Commercial Airplane Division, Renton, WA (D6-23201 TN).

MIL-STD-1374. Department of Defense. 1997. *Weight and Balance Data Reporting Forms for Aircraft* (MIL-STD-1374A), available online from http://www.sawe.org.

Davidson, J. 1988. *The Reliability of Mechanical Systems,* Mechanical Engineering Publications, London.

Federal Aviation Administration, Department of Transportation. 1993. *Aircraft Icing Handbook,* FAA Tech Report DOT/FAA/CT-88/8-2, updated sections available online from http://www.fire.tc.faa.gov.

FAR Part 25. Federal Aviation Administration, Department of Transportation. Part 25, Airworthiness Standards: Transport Category Airplanes, available online from http://www.faa.gov.

Granzeier W. 2001. "Flugzeugkabine Boeing B717-200," in *Flugzeugkabine/Kabinensysteme—Die naechsten Schritte Workshop DGLR S2.1/T8, Hamburg, 2001,* ed. D. Scholz, Deutsche Gesellschaft fuer Luft- und Raumfahrt, Bonn, pp. 79–87, available online from SAE, http://s2.dglr.de.

Hillman, T. C., Hill, S. W., and Sturla, M. J. 2001. *Aircraft Fire Detection and Suppression,* Kidde plc, URL: http://www.walterkidde.com (2002-02-28)

ICAO Annex 1. International Civil Aviation Organization (ICAO). 2001. *Convention on International Civil Aviation,* Annex 1: Personnel Licensing, 9th ed., ICAO, Montreal, available from ICAO, http://www.icao.int.

ICAO Annex 2. International Civil Aviation Organization (ICAO). 1990. *Convention on International Civil Aviation,* Annex 1: Rules of the Air. 9th ed., ICAO, Montreal, available from ICAO, http://www.icao.int.

Joint Aviation Authorities. Definitions and Abbreviations (JAR-1), available from the JAA (see http://www.jaa.nl).

Joint Aviation Authorities. *Joint Aviation Requirements for Large Aeroplanes* (JAR-25), Section 1, Requirements, available online from http://www.jaa.nl.

MIL-HDBK 217. Rome Air Development Center. 1985. *Reliability Prediction for Electronic Systems* (ADA 163900), available from the National Technical Information Service, http://www.ntis.gov.

MIL-STD-1629. Department of Defense. 1980. *Procedures for Performing a Failure Mode, Effects and Criticality Analysis* (MIL-STD-1629A), available from the National Technical Information Service, http://www.ntis.gov.

Moir, I. and Seabridge, A. 2001. *Aircraft Systems: Mechanical, Electrical, and Avionics Subsystems Integration,* AIAA Education Series, AIAA, Washington, DC.

O'Connor, P. D. T. 1991. *Practical Reliability Engineering,* John Wiley & Sons, Chichester.

Raymer, D. P. 1992. *Aircraft Design: A Conceptual Approach,* AIAA Education Series, AIAA, Washington DC.

Rome Air Development Center; Hughes Aircraft Company. 1985. *Nonelectronic Reliability Notebook,* Revision B. (ADA 163900), available from the National Technical Information Service, http://www.ntis.gov.

Roskam, J. 1989. *Airplane Design,* vol. 5, *Component Weight Estimation,* Roskam Aviation and Engineering Corporation, Ottawa, KS, available from DARcorporation, http://www.darcorp.com.

RTCA/DO-160D. Radio Technical Commission for Aeronautics. 2001. *Environmental Conditions and Test Procedures for Airborne Equipment*. RTCA, Washington, DC (RTCA/DO-160D Change 2) (see http://www.rtca.org). Document also available from the National Technical Information Service, http://www.ntis.gov.

RTCA/DO-178B. Radio Technical Commission for Aeronautics. 1992. *Software Considerations in Airborne Systems and Equipment Certification,* RTCA, Washington, DC (RTCA/DO-178B).

SAE Dictionary of Aerospace Engineering, ed. J. L. Tomsic. 1998. Society of Automotive Engineers, Warrendale, PA.

SAWE 2002. http://www.sawe.org (2002-02-28).

Scholz, D. 1998. DOCsys—A Method to Evaluate Aircraft Systems, in *Bewertung von Flugzeugen* (*Workshop: DGLR Fachausschuβ S2—Luftfahrtsysteme, München, 26./27. October 1998*), ed. D. Schmitt, Deutsche Gesellschaft für Luft- und Raumfahrt, Bonn, available online from http://www.ProfScholz.de.

Shustrov, Y. M. 1999. " 'Starting Mass'—a Complex Criterion of Quality for Aircraft On-Board Systems," *Aircraft Design,* vol. 1, pp. 193–203.

Torenbeek, E. 1988. *Synthesis of Subsonic Airplane Design,* Delft University Press, Delft.

VFW 614. *Schulungsunterlagen VFW614,* Vereinigte Flugtechnische Werke—Fokker GmbH, Germany.

TUHH. *Flugzeugsysteme* (Lecture notes), Technische Universität Hamburg—Harburg, Germany.

TÜV 1980. Luftfahrt Bundesamt, Bundesminister für Verkehr. *Grundlagen der Luftfahrzeugtechnik in Theorie und Prasix,* vol. 2, *Flugwerk,* TÜV Rheinland, Köln, Germany.

WATOG. Air Transport Association of America. 1992. *Airline Industry Standard, World Airlines Technical Operations Glossary* (WATOG), ATA, Washington, DC, available from ATA, http://www.airlines.org.

FURTHER READING

Aircraft Systems—General

Cundy, D. R. and Brown, R. S. *Introduction to Avionics,* Prentice Hall, Upper Saddle River, NJ (1997).

Federal Aviation Administration, Department of Transportation, *Airframe and Powerplant Mechanics Airframe Handbook,* AC 65-15A, FAA (1976), available online from http://www.faa.gov.

Kroes, M. J., Watkins, W. A., and Delp, F., *Aircraft Maintenance and Repair,* McGraw-Hill, Singapore (1993).

Lombardo, D., *Advanced Aircraft Systems,* TAB Books, McGraw-Hill, New York (1993).

Middleton, D. H., ed., *Avionic Systems,* Longman, Harlow (1989).

Roskam, J., *Airplane Design,* vol. 4, *Layout Design of Landing Gear and Systems,* Roskam Aviation and Engineering Corporation, Ottawa, KS (1989), available from DARcorporation, http://www.darcorp.com.

Wild, T. W., *Transport Category Aircraft Systems,* IAP, Casper, WY (1990).

Wilkinson, R. *Aircraft Structures and Systems,* Addison Wesley Longman, Harlow (1996).

Definitions and Breakdown

Society of Automotive Engineers (SAE), *Aerospace Landing Gear Systems Terminology,* SAE, Warrendale, PA (1994) (AIR 1489).

Society of Automotive Engineers (SAE), *Nomenclature, Aircraft Air Conditioning Equipment,* SAE, Warrendale, PA (1978) (ARP 147C).

Society of Automotive Engineers (SAE), *Terminology and Definitions for Aerospace Fluid Power, Actuation, and Control Technologies,* SAE, Warrendale, PA (1994) (ARP 4386).

Certification

Federal Aviation Administration, Department of Transportation. *Advisory Circular Checklist,* FAA (2000) (AC 00-2.13).

Safety and Reliability

Federal Aviation Administration, Department of Transportation. 1998. *System Design and Analysis,* FAA (AC 25.1309-1A), available online from http://www.faa.gov.

Power

Society of Automotive Engineers (SAE), *Aerospace Auxiliary Power Sources,* SAE, Warrendale, PA (1995) (AIR 744B).

Society of Automotive Engineers (SAE), *Power Sources for Fluidic Controls,* SAE, Warrendale, PA (1995) (AIR 1244A).

Air Conditioning

Department of Defense, *Environmental Control System, Aircraft, General Requirements for* (1986) (MIL-E-18927E), available from the National Technical Information Service, http://www.ntis.gov.

Society of Automotive Engineers (SAE), *Aerospace Pressurization System Design,* SAE, Warrendale, PA (1991) (AIR 1168/7).

Society of Automotive Engineers (SAE), *Aircraft Fuel Weight Penalty Due to Air Conditioning,* SAE, Warrendale, PA (1989) (AIR 1168/8).

Electrical Power

Eismin, T. K., *Aircraft Electricity and Electronics,* Glencoe/Macmillan/McGraw-Hill, New York (1994).

Pallett, E. H. J., *Aircraft Electrical Systems,* GB: Longman, Harlow (1998).

Equipment/Furnishings

Society of Automotive Engineers (SAE), *Crew Rest Facilities,* SAE, Warrendale, PA (1992) (ARP 4101/3).

Society of Automotive Engineers (SAE), *Galley Installations,* SAE, Warrendale, PA (1986) (ARP 695C).

Society of Automotive Engineers (SAE), *Lavatory Installation,* SAE, Warrendale, PA (1998) (ARP 1315C).

Society of Automotive Engineers (SAE), *Passenger Evacuation Devices—Civil Air Transport,* SAE, Warrendale, PA (1989) (ARP 495C).

Society of Automotive Engineers (SAE), *Performance Standard for Seats in Civil Rotorcraft, Transport Aircraft, and General Aviation Aircraft,* SAE, Warrendale, PA (1997) (AS 8049A).

Flight Controls

Raymond, E. T. and Chenoweth, C. C., *Aircraft Flight Control Actuation System Design,* Society of Automotive Engineers, Warrendale, PA (1993).

Schmitt, V. R., Morris, J. W., and Jenny G. D., *Fly-by-Wire: A Historical and Design Perspective,* Society of Automotive Engineers, Warrendale, PA (1998).

Scholz, D., "Development of a CAE-Tool for the Design of Flight Control and Hydraulic Systems," in Institution of Mechanical Engineers, *Avionic Systems, Design and Software,* Mechanical Engineering Publications, London (1996), pp. 1–22. [Introduction to the mechanical design aspects of fly-by-wire aircraft.]

Hydraulic Power

Federal Aviation Administration, Department of Transportation, *Hydraulic System Certification Tests and Analysis,* FAA (2001) (AC 25.1435-1), available online from http://www.faa.gov.

Green, W. L., *Aircraft Hydraulic Systems: An Introduction to the Analysis of Systems and Components,* John Wiley & Sons, Chichester (1985).

Guillon, M., *Hydraulic Servo Systems: Analysis and Design*. Butterworths, London (1968). [Translation of the French edition: *Étude et Détermination des Systèmes Hydrauliques,* Dunod, Paris (1961).]

Scholz, D., "Computer Aided Engineering for the Design of Flight Control and Hydraulic Systems," *SAE 1996 Transactions, Journal of Aerospace,* sec. 1, vol. 105, pp. 203–212 [SAE Paper: 961327].

Society of Automotive Engineers (SAE), *Aerospace—Design and Installation of Commercial Transport Aircraft Hydraulic Systems,* SAE, Warrendale, PA (1994) (ARP 4752).

Society of Automotive Engineers (SAE), *Hydraulic Systems, Aircraft, Design and Installation, Requirements for,* SAE, Warrendale, PA (1998) (AS 5440) [formerly MIL-H-5440].

Ice and Rain Protection

Federal Aviation Administration, Department of Transportation, *Aircraft Ice Protection,* FAA (1971) (AC 20-73), available online from http://www.faa.gov.

Federal Aviation Administration, Department of Transportation, *Certification of Transport Category Airplanes for Flight in Icing Conditions,* FAA (1999) (AC 25.1419-1), available online from http://www.faa.gov.

Federal Aviation Administration, Department of Transportation, *Effect of Icing on Aircraft Control and Airplane Deice and Anti-Ice Systems,* FAA (1996) (AC 91-51A), available online from http://www.faa.gov.

Landing Gear

Conway, H. G., *Landing Gear Design,* Chapman & Hall (1958).

Currey, N. S., *Aircraft Landing Gear Design: Principles and Practices,* AIAA Education Series, AIAA, Washington, DC (1988).

Department of Defense, *Landing Gear Systems* (1984) (MIL-L-87139), available from the National Technical Information Service, http://www.ntis.gov.

Pazmany, L., *Landing Gear Design for Light Aircraft,* Pazmany Aircraft Corporation, San Diego, CA (1986),

Society of Automotive Engineers (SAE), *Landing Gear System Development Plan,* SAE, Warrendale, PA (1997) (ARP 1598A).

Lights

Society of Automotive Engineers (SAE), *1994 SAE Aircraft Lighting Handbook,* SAE, Warrendale, PA (1994) [collection of *all* aerospace standards prepared by the SAE A-20 Committee].

Oxygen

Society of Automotive Engineers (SAE), *Chemical Oxygen Supplies,* SAE, Warrendale, PA (1991) (AIR 1133A).

Society of Automotive Engineers (SAE), *Introduction to Oxygen Equipment for Aircraft,* SAE, Warrendale, PA (2001) (AIR 825/1).

Society of Automotive Engineers (SAE), *Oxygen Equipment for Aircraft,* SAE, Warrendale, PA (1986) (AIR 825B).

Pneumatics

Department of Defense, *Bleed Air Systems, General Specification for* (1966) (MIL-B-81365), available from the National Technical Information Service, http://www.ntis.gov.

Society of Automotive Engineers (SAE), *Engine Bleed Air Systems for Aircraft,* SAE, Warrendale, PA (1987) (ARP 1796), available from SAE, http://www.sae.org.

Society of Automotive Engineers (SAE), *High Pressure Pneumatic Compressors Users Guide for Aerospace Applications,* SAE, Warrendale, PA (1996) (AIR 4994).

Airborne Auxiliary Power

Society of Automotive Engineers (SAE), *Commercial Aircraft Auxiliary Power Unit Installations,* SAE, Warrendale, PA (1991) (AIR 4204).

Availability of SAE documents:

Aerospace Information Reports (AIR) and Aerospace Recommended Practice (ARP) are listed together with a summary of the document on http://www.sae.org. In most cases, the documents may be ordered online.

SECTION 13
AERONAUTICAL DESIGN

Michael W. Jenkins

13.1 DEFINITIONS

Aeronautics is a science centered on flight in the atmosphere. It involves the design, construction, testing, and operation of lighter-than-air and heavier-than-air aircraft. These aircraft can be either mechanically driven (jets or propellers), rocket propelled, or not driven at all (gliders or balloons). All rely upon lift for successful operation. Fixed-wing aircraft lift is generated by the dynamic action of the air on the fixed wing. Rotary-wing aircraft (helicopters and autogyros) obtain lift from rotors placed horizontally above the fuselages. A helicopter differs from an autogyro in that the rotors are turned by motor power and there is usually no auxiliary propeller for forward motion. Balloons generate lift through density differences between the balloon gas (or hot air) and the local atmosphere. Flight regimes for fixed-wing aircraft include VTOL (vertical takeoff and landing), STOL (short takeoff and landing), V/STOL (vertical and/or short takeoff and landing), CTOL (conventional takeoff and landing, and STOVL (pronounced ''stoval'') (short takeoff and vertical landing).

13.2 INTRODUCTION

There are three phases in an aircraft design process: conceptual design, preliminary design, and final or detail design. These phases will overlap considerably, depending on the design organization. The final design and life cycle costs are almost entirely fixed by the conceptual design results, and therefore the conceptual design process is what will be discussed in this Section. In the conceptual design phase the aircraft takes shape and the influences on the design configuration of the mission and program requirements can be most readily observed. This phase traditionally relies on the use of already proven technology in the form of databases from previously successful aircraft concepts. Thus, during the conceptual studies a continual validation process is in place that lends credibility to the emerging configuration. If a radical configuration is to be developed, it is usual to precede the conceptual design with a technology base development, either through extensive wind tunnel testing or through the use of experimental aircraft specifically designed for the purpose.

13.3 OVERALL APPROACH

There are many ways to develop aircraft conceptual designs. Figure 13.1 shows one such approach. This approach relies upon the use of historical databases.

The mission requirements usually identify the aircraft type (fighter, bomber, transport), which then leads to the selection of an appropriate aspect ratio and a corresponding maximum lift-to-drag ratio, L/D, from a suitable database. By assuming a parabolic drag polar (a good assumption for conceptual design), the magnitude of allowable C_{D_0} can be estimated. Application of these data to the mission requirements will, with a suitable set of engine data and through the use of standard geometric relationships, permit the configuration to emerge. Life cycle costs can be developed and the resulting aircraft can be compared with the mission and program requirements in an iterative fashion to refine a final conceptual aircraft.

This approach requires a significant amount of historical data, collected in a format for easy and quick use. Many excellent books exist on aircraft design (see References and Further Reading), containing a vast array of very useful design data. No attempt will be made in this Section to duplicate these. It is understood

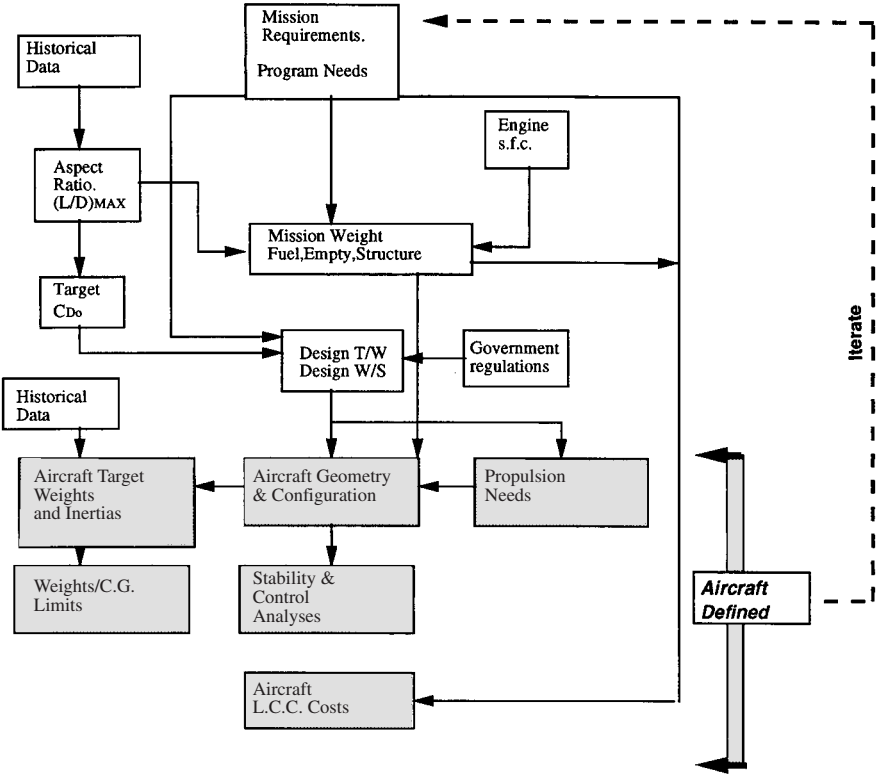


FIGURE 13.1 A conceptual design approach.

that many other data sources exist, both private and company confidential, that can supplement or replace those identified here. Table 13.1 summarizes some characteristics of a range of fixed-wing aircraft.

Design Aims

A successful aircraft design is the result of a compromise among sometimes conflicting requirements. The external shape is usually influenced by the particular philosophy of the design organization, resulting in different configurations emerging from different design organizations designing to the same requirements (Figure 13.2). The basic design principles, however, remain fairly consistent. These are listed in Figure 13.3 in terms of desirable minima and maxima. These are postulated to promote minimum acquisition cost, minimum operational cost, and maximum effectiveness as a functioning aircraft performing its mission.

Other factors specifically influence military aircraft configurations, such the desire for low infrared, low radar cross-section, and low noise signatures and cargo loading/unloading needs. Further special operational requirements also influence military aircraft configurations, such as the need to possess VTOL, V/STOL, STOVL, STOL, aircraft carrier compatibility, in-flight refueling, and supercruise capability. For civil or commercial aircraft the corresponding factors include airport gate compatibility (influencing wing span), passenger emergency evacuation requirements, the normal loading/unloading of passengers, and low operating costs. For very large military and civil/commercial aircraft the landing gear design and layout is a function of the thickness of the paved runway surface to be used as well as airframe stowage considerations.

The need for a balanced approach to the design function is best summarized, humorously, in Figure 13.4, where bias toward one technology or another has influenced the final configuration adversely. A good design avoids such bias.

Mission Establishment

The starting point for aircraft design studies is usually a new mission requirement or the need for the application of a new technology. Mission requirements usually stem from mission area analyses; new technologies can be in the form of a new engine, new materials, or the need to explore the influence of stealth for example. A request for proposal (RFP) or similar document issued by the customer contains specifics defining the needs for the new concept.

Mission Requirements

The specifics deduced from the RFP permit design-to criteria to be established in the form of such considerations as maximum speeds, rate of climb, range, endurance, payload, takeoff and landing distance requirements, and costs. Also identified are the various government specifications to which the proposed concept must be designed. These contain such considerations as the allowable symmetric load factor, the gust environment through which the aircraft must fly safely, and the definition of takeoff and landing distances. Government specifications are necessary to ensure the safe design and operation of aircraft in national and international airspace.

TABLE 13.1 Some Typical Aircraft Data

Manufacturer	Boeing	Dassault	Lockheed	Lockheed	Sukhoi
Name	F/A-18E/F	Rafael C	F-117A	F-22	Su 37
Gross Weight (lbs)	66,000	54,000	52,500	—	56,590
Empty Weight (lbs)	30,500	22,400	29.500	—	—
Span (ft)	44.9	33.4	43.3	44.5	48.2
Length (ft)	60.3	50	65.9	62.1	72.8
Wing Area (sq-ft)	500	490	913	740	667
Aspect Ratio	4.03	2.28	2.05	2.68	3.48
Power Plant	(2) GE F414 tf	(2) Snecma M88-2 tj	(2) GE F404 tf	(2) PW F119-PW-100 tf	(2) AL-37FU
Maximum Speed (Mach or MPH)	—	—	—	M 2.5	—
Cruise Speed (Mach or MPH)	—	M 1.8	—	—	—
Range (nm)	—	—	—	—	1,782
No. of Passengers	—	—	—	—	—
Payload (lbs)	—	—	—	—	3,090

Manufacturer	Northrop Grumman	EuroFighter	Boeing	Saab	Mapo
Name	F-14D	Typhoon	TAV-8B	JAS-39	MIG-31
Gross Weight (lbs)	74,349	50,264	31,000	27,557	90,000
Empty Weight (lbs)	41,780	24,239	14,223	12,900	47,000
Span (ft)	64.2/38.2	35.9	30.4	27.5	44.1
Length (ft)	61.9	52.4	50.3	46.3	74.5
Wing Area (sq-ft)	565	538	230	—	663
Aspect Ratio	7.29/2.58	2.40	4.02	—	2.93
Power Plant	(2) GE F110-GE-400 tf	(2) Eurojet 200 tf	(1) RR F402 or RR 406A tf	(1) Volvo Aero/GE RM12	(2) Soloviev D-30G
Maximum Speed (Mach or MPH)	M 2+	M 2 Class	M 0.9	M 2	M 2.8
Cruise Speed (Mach or MPH)	—	—	—	—	—
Range (nm)	—	—	—	—	1,630
No. of Passengers	—	—	—	—	—
Payload (lbs)	—	—	—	—	—

Manufacturer	Lockheed	Lockheed	Tupelov	Boeing	Antonov
Name	C-130J	C-5B	Tu 160	C-17A	An225
Gross Weight (lbs)	155,000	837,000	606,260	580,000	1,320,000
Empty Weight (lbs)	73,007	374,000	279,980	277,000	—
Span (ft)	132.6	222.8	182.6	169.8	290
Length (ft)	97.8	247.8	177.3	174	275.6
Wing Area (sq-ft)	1,745	6,200	—	3,800	9,743.70
Aspect Ratio	10.08	8.01	—	7.59	.63
Power Plant	(4) All. AE2100D3 tp	(4) GE TF39-1C tf	(4) NK-32 tf	(4) PW F117-PW-100	(6) Lotarev D-18T
Maximum Speed (Mach or MPH)	—	—	—	—	—
Cruise Speed (Mach or MPH)	340	563	M 2.3	M 0.77	430 kt
Range (nm)	—	3,000	7,640	3,000	3,800
No. of Passengers	—	270	—	3	6
Payload (lbs)	41,800	261,000	88,200	172,200	550,000

Manufacturer	Boeing	Illyushin	Boeing	Sukhoi	Northrop
Name	KC 135R	IL-78M	B-1B	SU-34	B-2A
Gross Weight (lbs)	322,500	462,966	477,000	99,400	336,500
Empty Weight (lbs)	119,212	204,088	186,234	—	—
Span (ft)	130.8	165.6	137	46.6	172
Length (ft)	128.8	152.8	145.8	76	69
Wing Area (sq-ft)	2,433	3,229.20	1,946	667.3	5,140
Aspect Ratio	7.03	8.49	9.64	3.25	5.76
Power Plant	(4) F108-CF-100	(4) D-30KP	(4) GE F101-GE-102	(2) AL-31F	(4) GE F118-GE-100
Maximum Speed (Mach or MPH)	—	—	—	—	—
Cruise Speed (Mach or MPH)	530 mph	M 0.80	—	M 1.8	—
Range (nm)	4,000	1,620	—	—	—
No. of Passengers	4	6	4	2	2
Payload (lbs)	—	105,820	—	17,640	40,000

TABLE 13.1 Some Typical Aircraft Data (*Continued*)

Manufacturer	Gulfstream	Airbus	Boeing Co.	Bombardier	Cessna
Name	Gulstream V	Airbus Corporate Jet	Boeing Business Jet	Global Express	Citation Excel
Gross Weight (lbs)	90,900	166,500	171,000	93,750	19,200
Empty Weight (lbs)	48,000	75,800	94,000	41,000	11,310
Span (ft)	93.5	111.8	112.6	94	55.7
Length (ft)	96.4	111	110.3	99.4	51.8
Wing Area (sq-ft)	1,136.5	1,318	1,342	1,022	370
Aspect Ratio	7.69	9.48	9.45	8.65	8.39
Power Plant	(2) BMW RR BR710-48	(2) CFM56-5B or IAE V2500	(2) CFM56-7B	(2) BMW RR BR700-710A2-20	(2) PWC PW545A
Maximum Speed (Mach or MPH)	M 0.89	M 0.82	M 0.82	M 0.89	M 0.77
Cruise Speed (Mach or MPH)	M 0.80	M 0.80	M 0.80	M 0.85	490 mph
Range (nm)	7,820	—	7,135	7,811	2,331
No. of Passengers	3–19	2–124	2–100	3/4–19	8–11
Payload (lbs)	—	—	—	—	—

Manufacturer	Dassalt	Raytheon	Raytheon	Cessna	Piper
Name	Falcon 2000	Hawker 800 XP	Beech 1900D Air-liner	Skyhawk/172R	PA46-350P Malibu Mirage
Gross Weight (lbs)	36,000	28,120	16,950	2,457	4,300
Empty Weight (lbs)	20,735	16,270	10,615	1,640	3,080
Span (ft)	63.4	51.4	57.8	36.1	43
Length (ft)	66.3	51.1	57.7	27.2	28.9
Wing Area (sq-ft)	527.6	374	310	174	175
Aspect Ratio	7.62	7.06	10.78	7.49	10.57
Power Plant	(2) CFE 738-1-18	(2) ASE TFE731-5BR M 0.80	(2) PWC PT6A-67D	(1) Lyc. 10-360-L2A	1 Lyc. TI0540-AE2A
Maximum Speed (Mach or MPH)	M 0.87	M 0.75	331 mph	142	267
Cruise Speed (Mach or MPH)	M 0.80	2,906	331 mph	141	246
Range (nm)	3,975	8–15	1,827	668	1,547
No. of Passengers	8–19	—	14–21	4	6
Payload (lbs)	—		—	—	—

Manufacturer	Airbus	Boeing	Boeing	Boeing	Bombardier
Name	A340-600	B767-400ER	B777-200ER	MD-11	CRJ 700 ER
Gross weight (lb)	804,700	450,000	656,000	630,500	72,500
Empty weight (lb)	560,000	227,300	314,000	295,600	43,500
Span (ft)	208.7	170.3	199.9	170.6	76.3
Length (ft)	245.9	201.3	209.1	202.2	106.7
Wing area (ft^2)	4,704	2,837	4,605	2,688	760
Aspect ratio	9.26	10.22	8.68	7.89	7.66
Power plant	(4) RR Trent 500	(2) GE CF6-80C2B or PW4062	(2) PW4000s, GE90's or RR Trents	(3) PW4460, PW4462, or GE CF6-80C2-DF1	(2) GE CF34-3C1
Maximum speed (Mach or Mph)	M 0.86	M 0.86	M 0.87	596	M 0.81
Cruise speed (Mach or mph)	M 0.83	M 0.80	M 0.84	M 0.83	M 0.77
Range (nm)	8,600	6,469	8,723	7,660	2,283
Number of passengers	380	261–304	301–440	298–410	70
Payload (lb)	—	53,700	125,000	98,500	5,500

Manufacturer	EMBRAER	Fairchild	British Aerospace	British Aerospace	British Aerospace
Name	ERJ-145 LR	728 Jet	BAe-146–300	AVRO RJ 100	Concorde
Gross weight (lb)	48,501	76,850	97,500	97,500	408,000
Empty weight (lb)	26,124	44,175	54,700	56,850	173,500
Span (ft)	65.8	87.4	86.4	86.4	83.8
Length (ft)	98	84.6	101.8	101.7	203.8
Wing area (ft^2)	551	807	832	832	3,640
Aspect ratio	7.86	9.46	8.97	8.97	1.93
Power plant	(2) All. AE3007-A1	(2) GE CF34-8D	(4) ASE ALF502R-5	(4) ASE LF507	(4)RR Olympus 593 Mk. 610
Maximum speed (Mach or	M 0.78	M 0.82	M 0.70	500 mph	M 2.0
Mph)	M 0.78	M 0.79	—	M 0.70	M 2.0
Cruise speed (Mach or mph)	1,771	1,843	1,300	2,310	4,083
Range (nm)	50	70–78	300	110	100
Number of passengers	2,646	—	8,370	8,370	—
Payload (lb)					

Source: *Aviation Week Source Book* 1999.

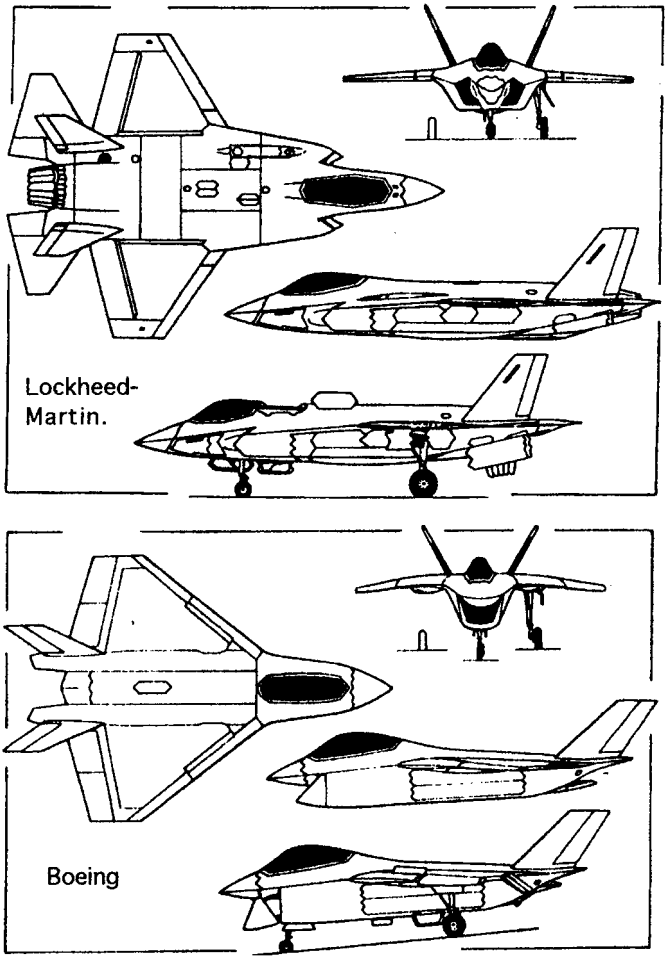


FIGURE 13.2 Joint strike fighter contenders (*Air International*).

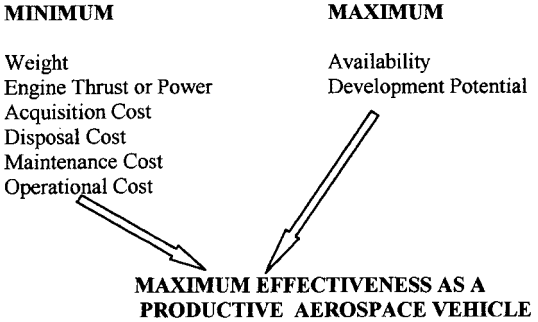


FIGURE 13.3 Generic overall design aims.

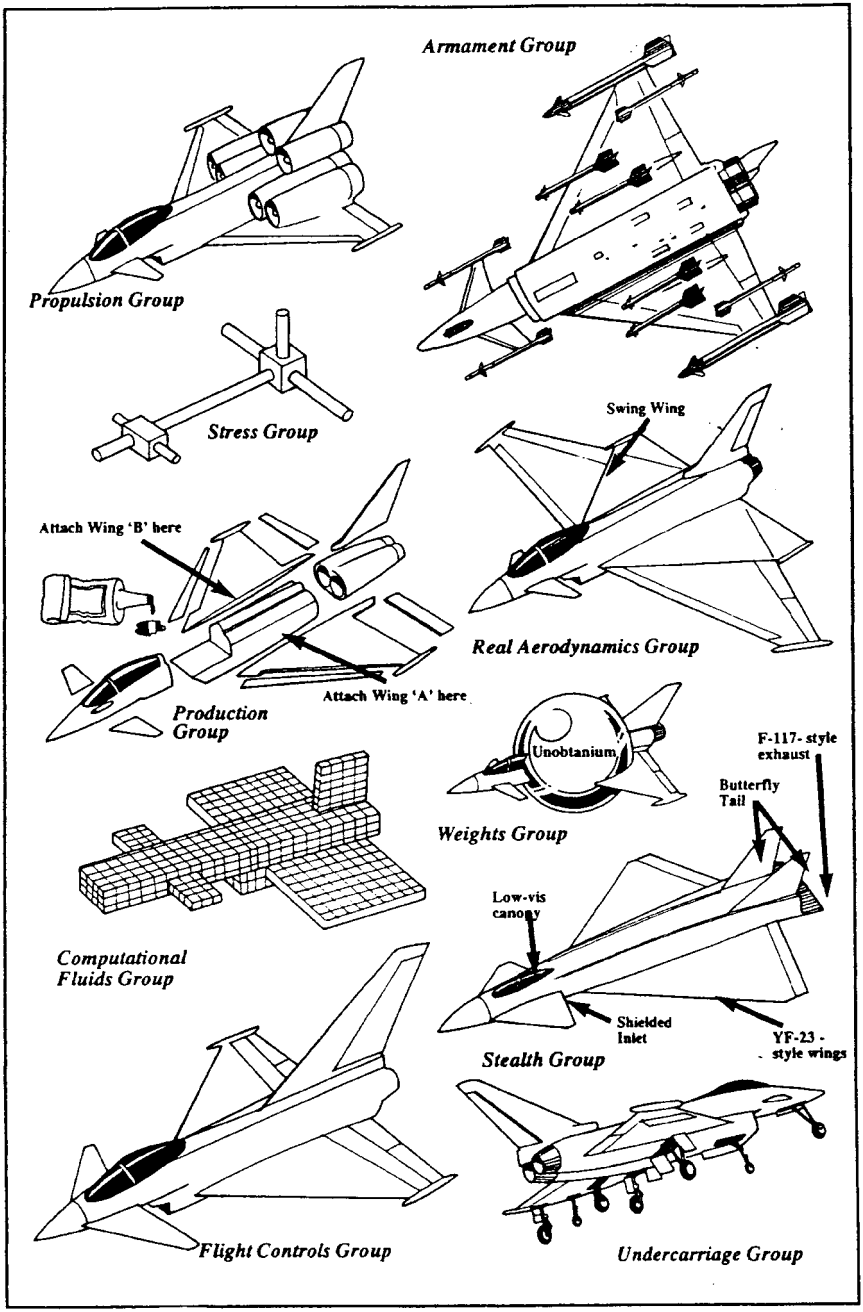


FIGURE 13.4 Light-hearted look at how specialists would have designed Eurofighter EF 2000 if given a free hand (Pete West/*Air International*).

Company Design Approach

A mission profile (Figure 13.5) is developed to reflect the major design needs. Initial sizing studies are performed, usually following the outline of Figure 13.1. The overall gross weight of the aircraft is established primarily through the use of the Breguet range and endurance equations as applied to each segment of the mission profile. Experience has shown that the combination of the minimum value of T/W and maximum value of W/S to meet all the mission requirements usually results in the minimum weight and minimum cost for the concept. This is the design aim for modern aircraft.

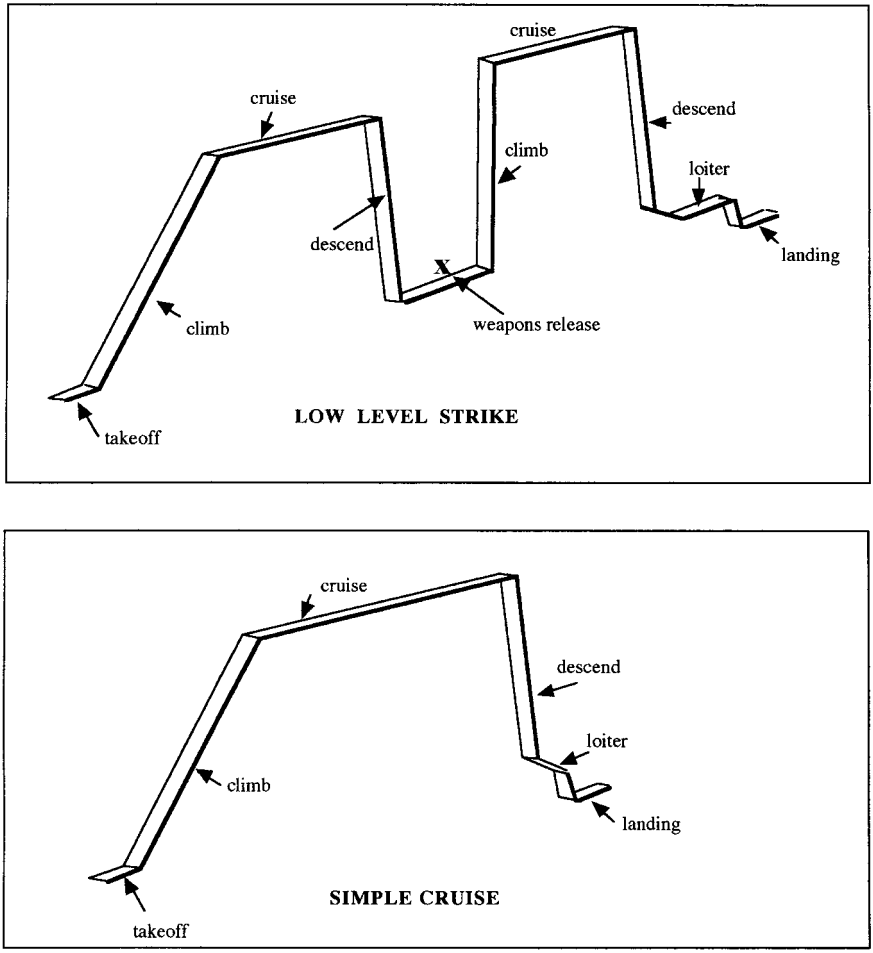


FIGURE 13.5 Some typical mission profiles.

Customer Needs

Sometimes a design configuration is significantly influenced by special customer needs. Military cargo transports sometimes require drive-on and drive-off capabilities (Figure 13.6), resulting in fuselage floors close to the ground and wings on top of the fuselage; increasing weight of larger transports dictate unusual landing gear configurations (Figure 13.7); and VTOL needs will influence engine (Figure 13.8) or wing design (Figure 13.9).

13.4 GOVERNMENT REGULATIONS

To be considered in the design process, these regulations are primarily related to the safety of the aircraft operation and of the aircraft structure. Aeronautical research and operational experience accumulated over many years have resulted in the generation of a set of requirements to which modern aircraft are now designed. The appropriate regulations define, for example, airplane maneuver loads, and stability and control requirements, which are a function of the intended use of the aircraft, and also define the gust environment through which all aircraft may eventually fly, irrespective of their intended use. Not only must the aircraft structure withstand the specified loads, but all internal equipment must be strong enough and mounted such as to withstand the resulting angular and linear accelerations. Specifically, engine gyroscopic induced loads must also be accounted for.

Commercial and civil aircraft are usually covered by FAR/JAR or similar regulations, and military aircraft by military specific regulations (MIL-SPECS). Figure 13.10 lists some of these. It can generally be assumed that the international civil

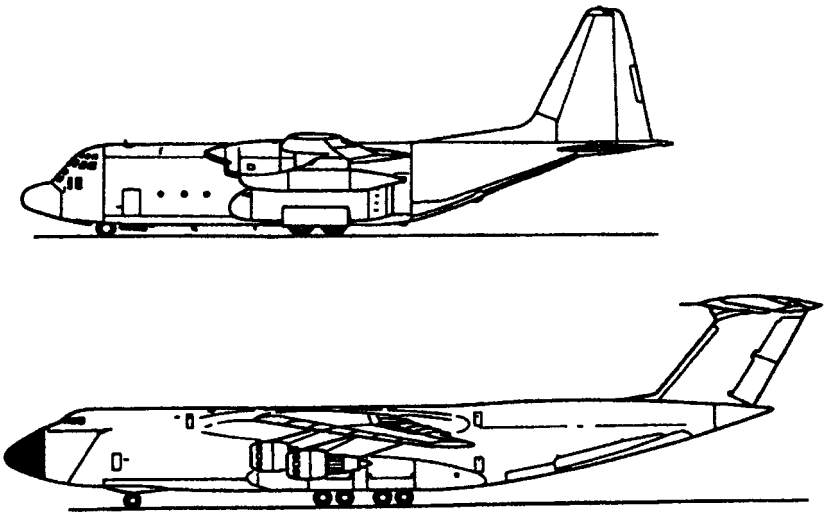


FIGURE 13.6 Military design requirements influence wing and fuselage location (LockheedMartin).

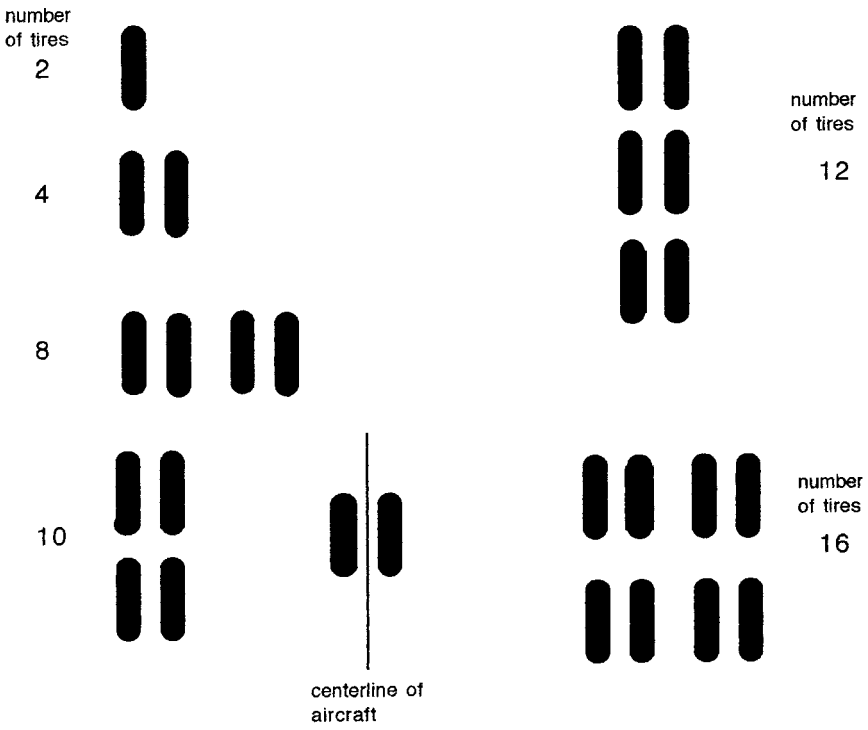


FIGURE 13.7 Some typical main gear wheel layouts per aircraft side.

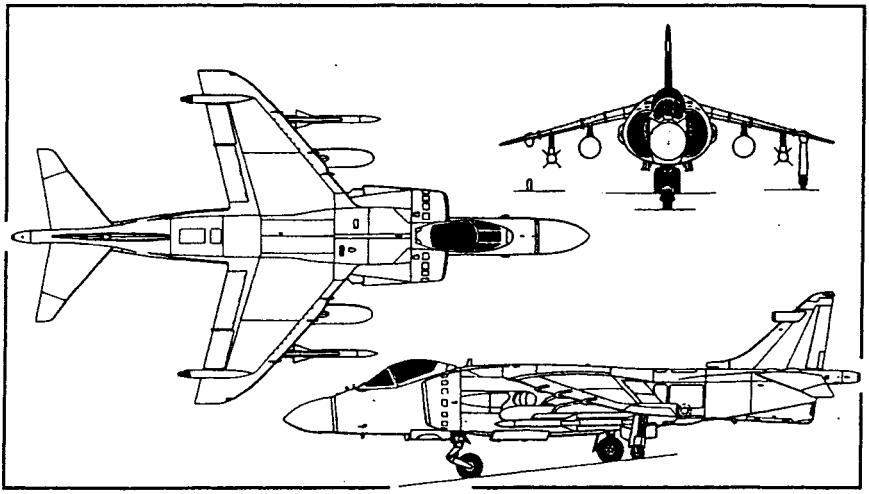


FIGURE 13.8 British Aerospace Sea Harrier (*Air International,* December 1999).

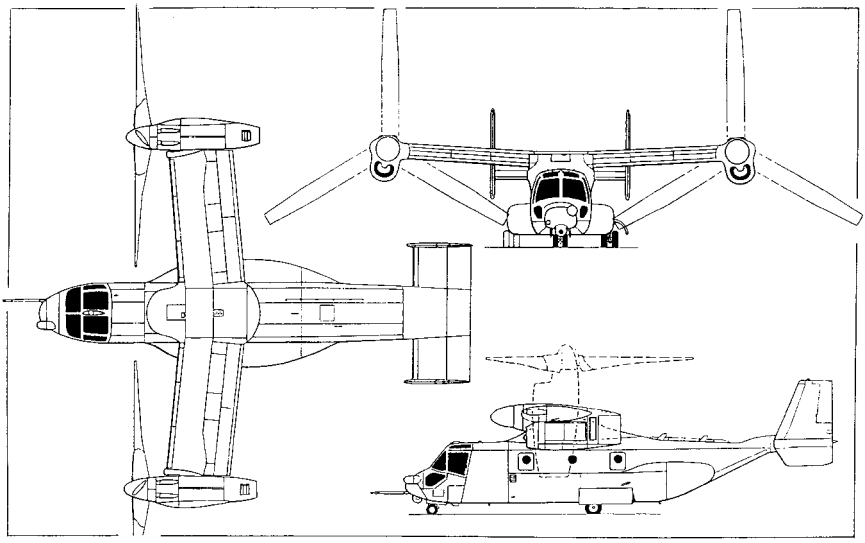


FIGURE 13.9 V-22 tilt rotor military transport (*Air International,* May 1989).

maneuver regulations are equivalent for FAR/JAR and that the gust environment, in terms of vertical gust magnitudes, is the same.

Commercial (FAR 25) and civil (FAR 23) aircraft are separated primarily by gross weight and propulsion type (Figure 13.11).

Maneuver loads and gust-induced loads are applied separately to the aircraft structure during the analysis stage of design. Upon initiation of production, at least one sample of each aircraft type is subjected to structural testing, using design loads, to validate the safety of the aircraft structure. This extensive testing is a demonstration of the structural safety of the aircraft and, because of the number of test cases involved this testing, can take up to a year to complete.

V-n Diagram—Maneuver Envelope (Figure 13.12). This defines the symmetric flight operational boundaries of an aircraft. The airplane must be structurally safe to fly on or within any point on this envelope which defines the structural, aerodynamic and propulsion limits of the concept. The maximum value of n, for military aircraft depends upon the function or mission of the aircraft for military operations (Table 13.2). For FAR 23 aircraft n_{max} is a function of aircraft type (Figure 13.13).

For FAR 25 commercial aircraft n_{max} is also a function of maximum takeoff gross weight (MTOGW). For aircraft up to 4100 lbs, $n_{max} = 3.8$ and at higher gross weights n_{max} varies as

$$n_{max} = 2.1 + \left[\frac{24{,}000}{\text{MTOGW} + 10{,}000}\right]$$

up to a gross weight of 50,000 lb, where n_{max}. becomes equal to 2.5 and thereafter remains constant for higher gross weights.

CHAPTER 1 – **FEDERAL AVIATION ADMINISTRATION**, DEPARTMENT OF TRANSPORTATION.– **FAR** – (Extract).

(FAR – Federal Aviation Regulations – U.S.)

Subchapter C – Aircraft.

Part	
21	Certification procedures for products and parts.
23	Airworthiness standards: normal, utility, acrobatic, and commuter category airplanes.
25	Airworthiness standards: traansport category airplanes.
27	Airworthiness standards : normal category rotorcraft.
29	Airworthiness standards: transport category rotorcraft.
31	Airworthiness standards: manned free balloons.
33	Airworthiness standards: aircraft engines.
35	Airworthiness standards: propellers.
36	Noise standards : aircraft type and airworthiness certification.
39	Airworthiness directives.
43	Maintenance, preventive maintenance, rebuilding, and alteration.

TYPICAL MAKE-UP OF **JOINT AIRWORTHINESS REQUIREMENTS.**

(JAR – EUROPEAN).

JAR – 25 , Large Aeroplanes.

Section 1 – Requirements.

Subpart A – General

Subpart B – Flight

Subpart C – Structure

Subpart D – Design and Construction

Subpart E – Powerplant

Subpart F – Equipment

Subpart G – Operating limitations and information

Subpart J – Gas turbine auxiliary power unit installation.

FIGURE 13.10 Some FAA, JAR, and U.S. Military specifications (after Nicolai 1984).

V-n Diagram—Gust Envelope. (Figure 13.14). This envelope describes the load factors imposed on a trimmed aircraft when encountering a vertical up-gust or down-gust in horizontal constant-speed flight.

A sharp-edged vertical gust is assumed to be present, and a gust alleviation factor to account for gust penetration and gust response is assumed to exist. The resultant total load factor induced as given by MIL-SPEC MIL-A-8861(AS) is

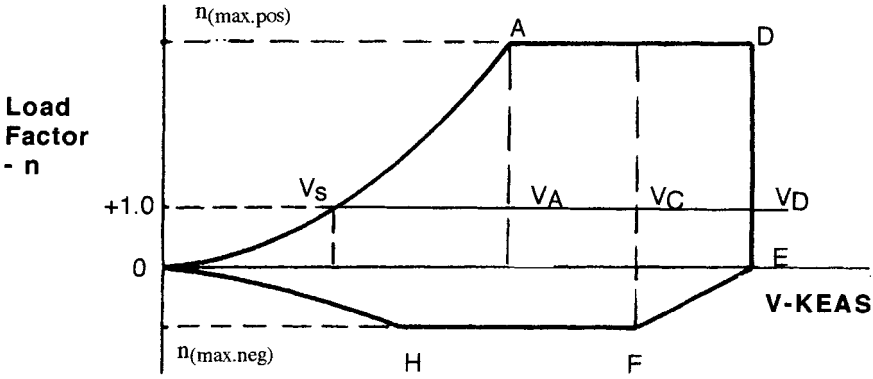
TYPICAL **MIL.SPEC. STANDARDS.** – U.S.

MIL-STD-1797A	Flying qualities of Piloted Aircraft
MIL-F-83300	Flying qualities of Piloted V/STOL Aircraft
MIL-F-9490	Flight Control Systems-Design, Installation and Test of Piloted Aircraft
MIL-S-8369	Stall/Post-Stall/Spin Flight Test Demonstration Requirements for Airplanes
MIL-C-18244	Control and Stabilization Systems: Automatic, Piloted Aircraft.
MIL-D-8708	Demonstration Requirements for Airplanes
MIL-A-8860 through 8864 and 8870.	Airplane Strength and Rigidity
MIL-P-26366	Propellers, Type Test of
MIL-I-8700	Installation and Test of Electronics Equipment in Aircraft
MIL-S-18471	Seat System, Ejectable, Aircraft
MIL-W-25140	Weight and Balance Control Data
MIL-STD-850	Aircrew Station Vision Req. for Military Aircraft
MIL-STD-757	Reliability Evaluation from Demonstration Data
MIL-C-5011	Charts; Standard Aircraft Characteristics and Performance
MIL-STD-881	Work Breakdown Structure (WBS)

FIGURE 13.10 (*Continued*) Some FAA, JAR, and U.S. Military specifications (after Nicolai 1984).

FAR 23	**FAR 25**
Propeller driven-generally	Jet or propeller driven
Less than 19 passengers	No minimum passenger limit
$W < 12{,}000$ lb generally	$W > 12{,}000$ lb
Categories	**Categories**
Normal Utility Acrobatic	Normal Cargo

FIGURE 13.11 Commercial and civil FAR divisions.



V_S = Level flight stall speed

V_A = Corner speed, or maneuver speed for military

V_C = Cruise speed,or max.level flight speed for military

V_D = Limit speed, or max. dive speed for military and civil

KEAS = Equivalent airspeed (knots)

FIGURE 13.12 Typical maneuver diagram.

TABLE 13.2 Typical Maximum Maneuver Load Factors

USAF airplane type	USN airplane type	Positive	Negative
Fighter		8.67	−3.00
Attack	Fighter, attacker, trainer	7.33	−3.00
	Observation	6.00	−3.00
Trainer		5.67	−2.33
Utility	Utility	4.00	−2.00
Small bomber		3.67	−1.67
Medium bomber, assault transport	Patrol, weather, Antisubmarine, reconnaissance	3.00	−1.00
Medium transport		2.50	−1.00
Heavy bomber, heavy transport		2.00	−1.00

Source: MIL-SPEC Mil-F-8860.

Normal Category

$$\textit{Positive} + n_{max} = 2.1 + \left[\frac{24{,}000}{MTOGW + 10{,}000}\right]$$

$$\textit{Negative} - n_{max\ neg} = 0.4(+n_{max})$$

Exceptions:–

	+nmax	nmax.neg
Utility	4.4	1.76
Acrobatic	6.0	3.0

FIGURE 13.13 Design maximum load factor for FAR 23 type aircraft (FAR 23).

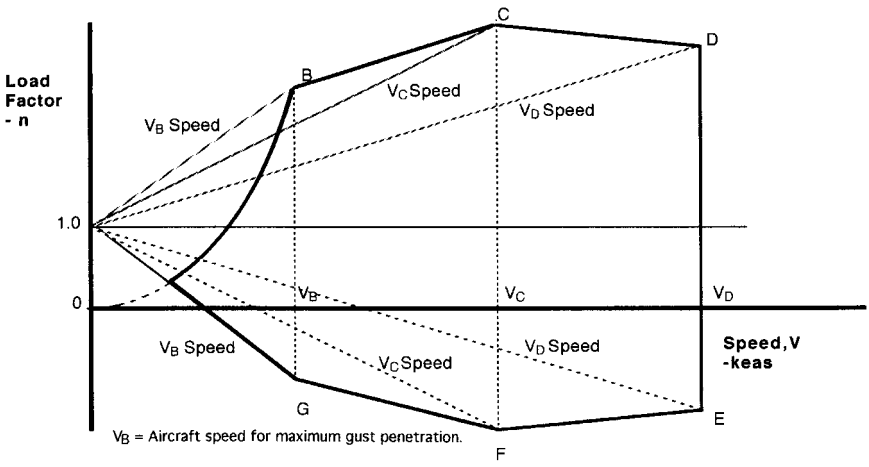


FIGURE 13.14 Typical gust design envelope.

$$n = 1 \pm \frac{K_g\, U_{ge}\, V_E\, a}{498\ (W/S)}$$

where $K_g = \dfrac{0.88\ \mu_g}{5.3 + \mu_g}$ = gust alleviation factor

$\mu_g = \dfrac{2\ (W/S)}{g\ \rho \bar{a} \bar{c}}$ = mass parameter

U_g = gust velocity (fps) (e.a.s.)

ρ = air density (slugs/ft^3)

(W/S) = wing loading (psf)
$\bar{c}$ = wing mean geometric chord (ft)
g = acceleration due to gravity (f/s^2)
V_E = aircraft equivalent airspeed (knots) (f/s^2)
V_E = aircraft equivalent airspeed (knots) (e.a.s.)
a = slope of airplane normal force curve (per radian)

Gust velocity magnitudes and their variation with altitude are given in Table 13.3. When the symmetric maneuver and symmetric gust envelopes are superimposed, one of these will establish a maximum value of symmetric load factor. For highly maneuvering aircraft a symmetric *maneuver* load factor will be a maximum, which is the case for fighter aircraft. Most commercial aircraft will exhibit a maximum symmetric load factor as generated by the *gust* envelope. The former is called a *maneuver critical* aircraft, the latter *gust critical*.

13.5 CONCEPTUAL DESIGN

Weight Estimation

The maximum design weight of an aircraft, W_0, can be written as follows (see Raymer, 1989; Roskam 1980):

$$W_o = W_e + W_f + W_c + W_{pl} \text{ (lbs)}$$

where W_e = empty weight
W_f = fuel weight
W_c = crew weight
W_{pl} = payload weight

This can be written as

TABLE 13.3 Standard Gust Velocities

Gust intensities (fps e.a.s.)		Maximum aircraft velocity (knots e.a.s.)
S.L. to 20,000 ft	20,000 ft to 50,000 ft	
±66	linear reduction to 38	V_B
±50	linear reduction to 25	V_C, V_H
±25	linear reduction to 12.5	V_D, V_L

V_B = Speed for maximum gust intensity
V_C, (V_H) = Design cruise speed (max. level flight speed at max. continuous power for military aircraft)
V_D = Design dive speed (1.25 V_C or 1.25 V_L)

Source: MIL-SPEC Mil-F-8861.

$$W_0 = \left(\frac{W_c + W_{pl}}{1 - \left(\frac{W_e}{W_0}\right) - \left(\frac{W_f}{W_0}\right)} \right) \text{(lb)}$$

where (W_e/W_0) is the empty weight fraction and (W_f/W_0) is the fuel fraction.

The empty weight fraction in conventional design is a targeted allowance for those items listed in Figure 13.15.

The likely value for this fraction depends on the function of the aircraft (see Figure 13.16). In conceptual design, empty weight component weights are often based on aircraft with missions similar to your conceptual design (Table 13.4). This

Structure weight
- Wing weight
- Empennage weight
- Fuselage weight (including furnishings)

Propulsion weight

Systems weight
- Controls
- Environmental
- Hydraulic
- Pneumatic
- Fuel
- Electric
- Avionics
- Weapons

Landing gear weight

FIGURE 13.15 Components included in empty weight.

$$\frac{W_e}{W_0} = C \times (W_0)^D$$

W_e = empty weight (lb) W_0 = gross weight (lb)

Aircraft type	C	D
Fighters	0.6954	−0.0275
Business jets	1.0595	−0.0598
Wide-body jets	0.8342	−0.0375

FIGURE 13.16 Typical empty weight fractions (*Aerospace Source Book* 1999).

TABLE 13.4 Some Typical Group Weight Values

Group weight (lb)	F/A-18A	Gulfstream II	Airbus A300
Wing group	3,798	6,372	44,131
Empennage group	945	1,965	5,941
Fuselage group	4,685	5,944	35,820
Engine section	143	1,239	7,039
Landing gear group	1,992	2,011	13,611
Structure total	**11,563**	**17,531**	**106,542**
Engines & systems	6,277	6,886	22,897
Fixed equipment	5,135	11,203	35,053
Total empty weight	**22,974**	**35,620**	**164,492**
Flight design gross weight	**32,357**	**64,800**	**302,000**

Source: Roskam 1980.

approach assumes that the structural efficiency and resulting weight of the systems for the design being developed are at least as good as those of existing aircraft.

The fuel fraction depends entirely on the mission, the anticipated L/D ratio of the aircraft, and the specific fuel consumption of the engine(s) to be used. The use of the Breguet (or similar) range and endurance equations will permit fuel fractions to be determined. For example, using the mission profile of Figure 13.17, we can postulate that the overall mission weight fraction W_{end}/W_0 is given by

$$\left(\frac{W_7}{W_0}\right) = \left(\frac{W_1}{W_0}\right)\left(\frac{W_2}{W_1}\right)\left(\frac{W_3}{W_2}\right)\left(\frac{W_4}{W_3}\right)\left(\frac{W_5}{W_4}\right)\left(\frac{W_6}{W_5}\right)\left(\frac{W_7}{W_6}\right) = \left(\frac{W_{\text{end}}}{W_{\text{start}}}\right)$$

for a seven-element mission profile, where (W_3/W_2) represents an application of Breguet's range and (W_5/W_4) the endurance equation. All other fractions represent what can be termed standard weight fractions for design purposes. These are a

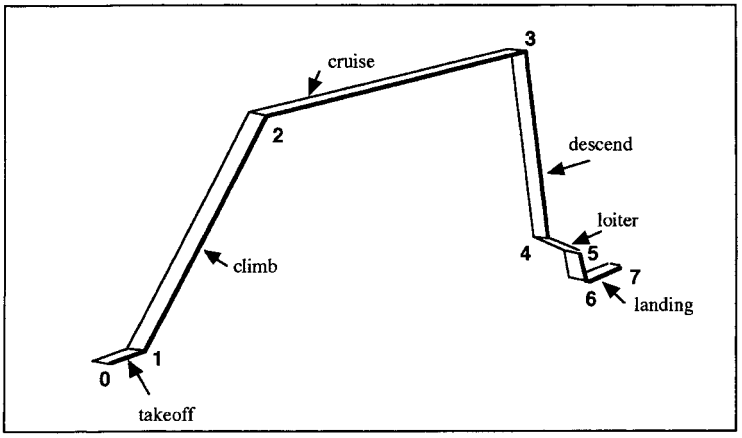


FIGURE 13.17 Typical mission profile.

function of aircraft type, i.e., fighter or commercial transport. An excellent summary (Roskam 1980) lists various values for various concepts, and it can be seen that for conceptual design purposes each of these can be assumed to be about 0.99.

$$1 - (W_7/W_0) = \text{Fuel fraction} = (W_f/W_0)$$

and if a reserve percentage of fuel is required, e.g., 6% fuel reserve, then the total mission fuel fraction is given by 1.06 (W_f/W_0), which is called the total fuel fraction and represents the fuel required for the mission.

With knowledge of the crew weight and payload weight, with the total fuel fraction (W_f/W_0), and with (W_e/W_0) magnitudes, it is now possible to solve iteratively for W_0. This represents the design value for the takeoff weight, which now permits the estimation of total fuel weight and target empty weight through

$$\text{Total fuel weight} = (W_f/W_0) \cdot W_0 \text{ (lb)}$$

$$\text{Target empty weight} = (W_e/W_0) \cdot W_0 \text{ (lb)}$$

Likely values of certain design parameters necessary for use in the Breguet range and endurance equations are given in Table 13.5.

Estimation of Thrust to Weight (T/W), Wing Loading (W/S), and definition of the High Lift System Requirements

These design parameters are established by evaluations of the performance and maneuver requirements of the mission. Whereas the T/W and W/S values are generally established by in-flight performance requirements, the high-lift systems are usually dictated by airfield performance demands in terms of field length or takeoff and landing distances.

The relationship between T/W and W/S is given as follows,

$$\frac{T}{W} = \left(\frac{C_{D0}q}{W/S}\right) + \left(\frac{n^2(W/S)}{q \cdot \pi \cdot A \cdot e}\right)$$

where C_{D_0} = profile drag coefficient,
q = dynamic pressure (psf)
W/S = wing loading (psf)
n = maneuver load factor
T/W = thrust to weight

TABLE 13.5 Likely Values of Design Parameters

Aircraft type	Wing aspect ratio	$(L/D)_{max}$	Wing taper ratio
Fighters • Subsonic	4.00–6.00	11.0	0.28–0.57
• Supersonic	2.00–3.00	9.00	0.07–0.25
Jet transports (military and civil)	7.00–9.00	17.00	0.28–0.42

For subsonic flight, e = 0.80 to 0.85 (clean wing).

all at the flight conditions being evaluated.

By choosing appropriate values of aspect ratio (A), and induced drag factor, e, the above equation can be evaluated to show the relationship between T/W and W/S for the particular mission segment. Mission segments can be combinations of those shown in Figure 13.18. Likely values of A and e are shown in Table 13.5.

It is not unusual to assume at the conceptual design phase that the drag polar for the aircraft is of parabolic form, resulting in the generation of algebraic relationships for many design-related parameters (Table 13.6). These clean aircraft relationships can be modified when external stores or other drag-producing devices are used during parts of the mission.

The maximum value of L/D is given by

$$\frac{L}{D} = 0.5\sqrt{\frac{\pi \cdot A \cdot e}{C_{D_0}}}$$

Choosing the design value for maximum value of L/D, A and e will determine a target value of C_{D_0}.

Presenting the resulting T/W and W/S relationships in graphical format (Figure 13.19) will identify the T/W and W/S combination that will satisfy all of the mission requirements.

Figure 13.19 shows that a single combination of T/W and W/S, called the *design point,* will satisfy or exceed all mission requirements. It identifies which mission requirements define the configuration and which do not. The design point shown here is the lowest thrust-to-weight ratio and the highest wing loading necessary to meet all the requirements. The $(T/W)_{\min}$ and $(W/S)_{\max}$ chosen lead to the lowest

Maximum speed

Cruise speed

Maneuver requirements

Climb requirements – Single engine

Climb requirements – Multiengine

– All engines operating

– One engine failure

– One engine go-around

Supercruise speed

Absolute ceiling – Rate of climb = 0

Service ceiling – Rate of climb = 100 feet per minute (Mil.@ max. power)

= 500 feet per minute (Commercial jet)

FIGURE 13.18 Mission segments defining T/W and W/S.

TABLE 13.6 Some Useful Design Relationships

Maximum range	Maximum endurance
$C_D = 2\,C_{D_o}$	$C_D = 4C_{D_o}$
$C_L = \sqrt{C_{D_o}\,\pi Ae}$	$C_L = \sqrt{3C_{D_o}\pi Ae}$
$V = \sqrt{\frac{W}{S}\frac{2}{\rho}\frac{1}{\sqrt{C_{D_o}\pi Ae}}}$ (fps)	$V = 0.76\sqrt{\frac{W}{S}\frac{2}{\rho}\frac{1}{\sqrt{C_{D_o}\pi Ae}}}$ (fps)
$\left(\frac{L}{D}\right)_{\max} = 0.500\sqrt{\frac{\pi Ae}{C_{D_o}}}$	$\left(\frac{L}{D}\right) = 0.433\sqrt{\frac{\pi Ae}{C_{D_o}}}$

For parabolic drag polar: $C_D = C_{D_o} + \frac{C_L^2}{\pi Ae}$ (Propeller-powered aircraft)

Maximum range	Maximum endurance
$C_D = \frac{4}{3}\,C_{D_o}$	$C_D = 2\,C_{D_o}$
$C_L = \sqrt{\frac{C_{D_o}\pi Ae}{3}}$	$C_L = \sqrt{C_{D_o}\pi Ae}$
$V = 1.316\sqrt{\frac{W}{S}\frac{2}{\rho}\frac{1}{\sqrt{C_{D_o}\pi Ae}}}$ (fps)	$V = \sqrt{\frac{W}{S}\frac{2}{\rho}\frac{1}{\sqrt{C_{D_o}\pi Ae}}} \sim$ F.P.S.
$\left(\frac{L}{D}\right) = 0.433\sqrt{\frac{\pi Ae}{C_{D_o}}}$	$\left(\frac{L}{D}\right)_{\max} = 0.500\sqrt{\frac{\pi Ae}{C_{D_o}}}$

For parabolic drag polar: $C_D = C_{D_o} + \frac{C_L^2}{\pi Ae} \sim$ Jet powered aircraft.

installed thrust and the lightest wing weight. Any point above the two defining mission requirement lines will also satisfy the mission requirements, but probably not with the lowest cost. The area above these lines is sometimes called the design space.

The resulting airplane geometric parameters and engine requirements can then be established as follows:

$$S = (W)/(W/S)_{\max}\ \text{ft}^2 = \text{wing area}$$

and

$$T = (T/W)_{\min}(W)\ \text{lb} = \text{total performance thrust}$$

Use of the geometric and aerodynamic relationships of Figures 13.20–13.22 permits the aircraft configuration to develop once a wing taper ratio has been chosen.

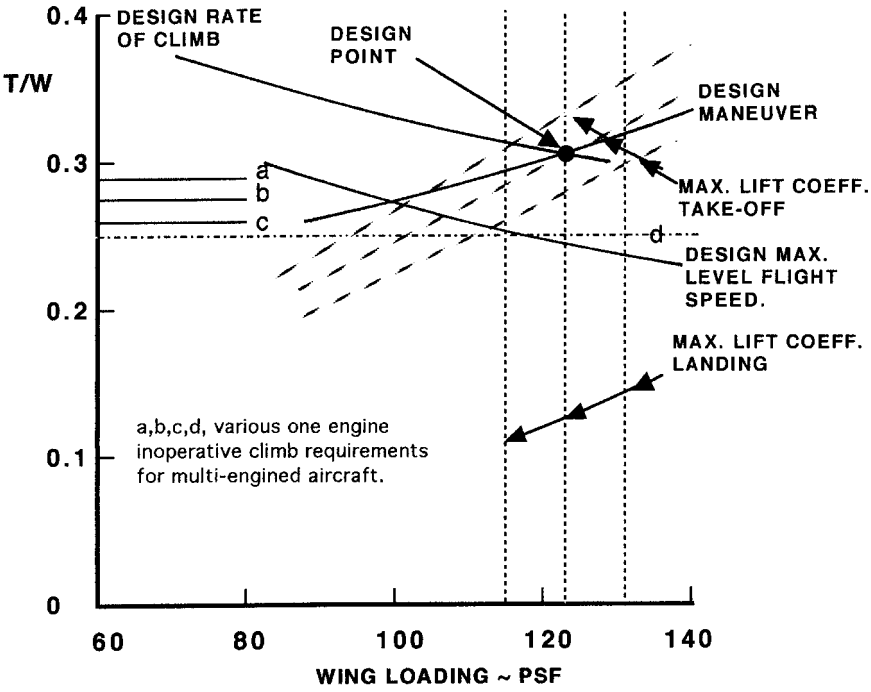


FIGURE 13.19 Typical design chart showing various mission requirements.

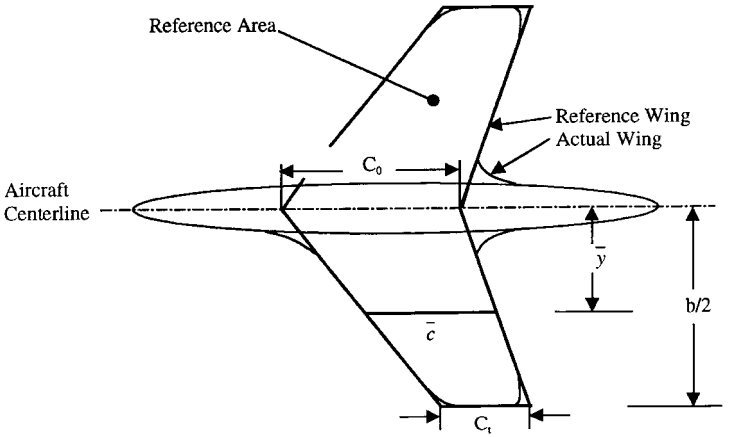
Trade studies are performed to assess the influences on the design (primarily on W_0) of variations in certain mission parameters and in certain assumed design parameters. Included in these are range, payload, aspect ratio, specific fuel consumption, and C_{D_0}, in which the effects of systematic individual variations are assessed.

Influence of Takeoff Distance Requirements on Aircraft Design

Traditionally, takeoff and landing distances are defined as being composed of a ground distance and an air distance. In each instance the air distance is measured relative to an imaginary vertical obstacle situated at the end of the runway (Figure 13.23). For civil aircraft this obstacle is 50 ft for FAR 23 category aircraft on takeoff and 50 ft for landing. For FAR 25 category aircraft this obstacle is 35 ft on takeoff and 50 ft for landing. For landing the approach speed, V_A, is specified as $1.3(V_{SL})$, which becomes $1.2(V_{SL})$ for military land based aircraft and $1.15(V_{SL})$ for carrier-based aircraft. The design factors that are influenced are wing loading, thrust or power loading, aerodynamic drag, and lift for an RFP-dictated takeoff distance and runway surface type. The design process is centered on establishing the appropriate combinations of W/S, T/W, or W/P and C_L necessary to meet the design takeoff distance. These relationships have been extensively explored (Loftin 1980, as shown in Raymer 1989).

For FAR 23 rules, the following can be used for takeoff distance estimation (Loftin 1980, as shown in Raymer 1989).

Wing / Horizontal Tail Geometries



$$\bar{c} = \left(\frac{2}{3}\right) c_0 \frac{(1+\lambda+\lambda^2)}{(1+\lambda)} \qquad \bar{y} = \left(\frac{b}{6}\right)\frac{(1+2\lambda)}{(1+\lambda)}$$

Typically, Wing/Horizontal Tail Aerodynamic Center $\begin{cases} = 0.25\bar{c} & (Subsonic) \\ = 0.40\bar{c} & (Supersonic) \end{cases}$

S = Reference Wing Area = $\left\{\frac{(c_0 + c_t)}{2}\right\} * b$ AR = Aspect Ratio = $\left\{\frac{b^2}{S}\right\}$

λ = Taper Ratio = $[c_t / c_0]$ b = Span

Given W/S, AR, and λ we can find the following:

$$b = \sqrt{AR * S} \qquad c_0 = \frac{2 * S}{b * (1+\lambda)} \qquad c_t = \lambda * c_0$$

FIGURE 13.20 Some geometric relationships for the wing and empennage.

$$S_{\text{TOG}} = 4.9(\text{TOP}_{23}) + 0.009(\text{TOP})_{23})^2 \text{ (ft)}$$

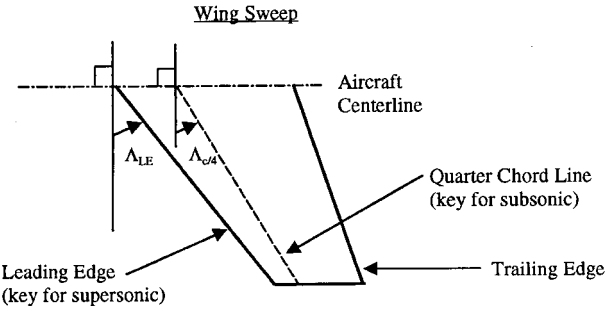
where $$(\text{TOP}_{23}) = \frac{(W/S)_{TO}(W/P)_{TO}}{\sigma \cdot C_{L_{\text{MAX}\,TO}}} \text{ (lb}^2/\text{ft}^2 \cdot \text{hp)}$$

and $$S_{tO} = 1.66(S_{\text{TOG}})$$

For FAR 25 takeoff estimates the following is applicable:

$$S_{TOFL} = \frac{37.5(W/S)_{TO}}{\{\sigma \cdot C_{L_{\max TO}}(T/W)_{TO}\}} \text{ (ft)}$$

In each instance solution of the above equations will give a combination of W/S,



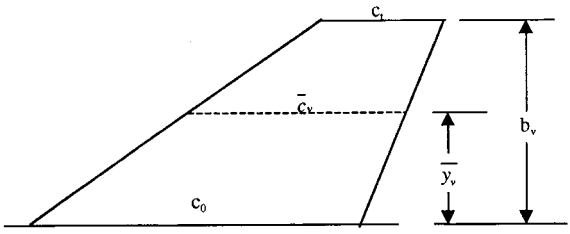
Wing and Horizontal Tail

$$TAN(\Lambda_{LE}) = TAN(\Lambda_{C/4}) + \left(\frac{[1-\lambda]}{AR[1+\lambda]}\right)$$

Vertical Tail

$$TAN(\Lambda_{LE}) = TAN(\Lambda_{C/4}) + \frac{(1+\lambda)}{2AR(1+\lambda)}$$

Vertical Tail Geometric Properties



$$\lambda = c_t / c_0 \qquad S_v = \left[(c_t + c_0)/2\right] * b_v$$

$$\bar{c}_v = \frac{2}{3}\left\{\frac{c_0(1+\lambda+\lambda^2)}{(1+\lambda)}\right\} \qquad \bar{y}_v = \frac{b_v}{3}\left\{\frac{(1+2\lambda)}{(1+\lambda)}\right\}$$

Geometric Mean Chord, $\bar{c}_{geom} = \frac{c_0}{2}(1+\lambda)$ Aspect Ratio = $\frac{b_v^2}{S_v}$

FIGURE 13.21 Some geometric relationships for the wing and empennage.

T/W (or W/P), and $C_{L\max}$ that will generate the necessary takeoff distance required. (Here P = power, σ = relative density, and subscript TO refers to takeoff conditions.)

In military aircraft design takeoff evaluations, it is usual to consider only the takeoff ground run, S_{TOG}, and (from Roskam 1980) an effective equation for this is

$$S_{TOG} = \frac{k_1 \cdot (W/S)_{TO}}{\rho \cdot \{C_{L_{\max TO}} \cdot ((T/W) \cdot k_2 - \mu_G)_{TO} - 0.72 \cdot C_{D_0}} \text{ (ft)}$$

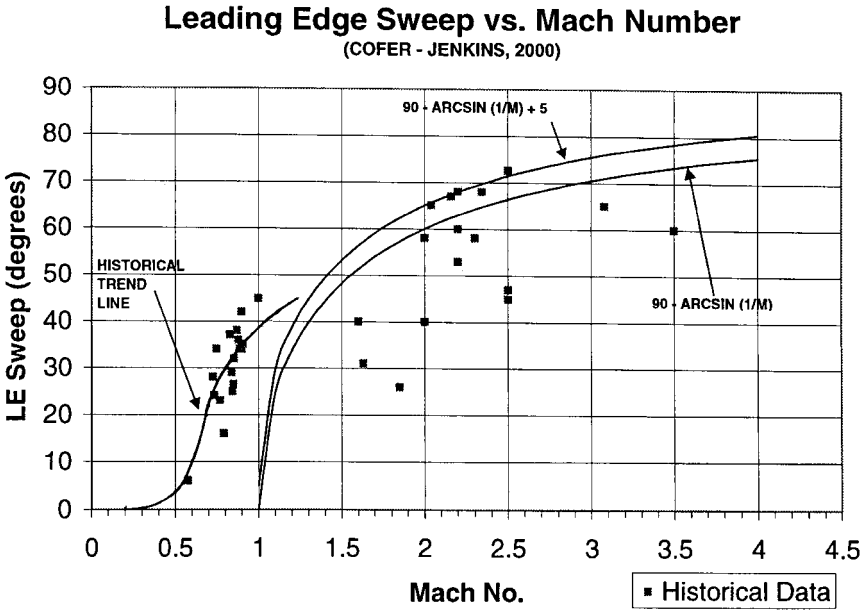


FIGURE 13.22 Historical leading edge sweep versus Mach number.

where k_1 = 0.0447 for jet aircraft,
= 0.0376 for propeller aircraft
k_2 = a factor to allow for the reduction thrust with forward velocity on the takeoff ground run
ρ = air density, slugs/ft^3
C_{D_o} = profile drag coefficient including the effects of landing gear and flaps
μ_G = ground rolling coefficient of friction
$C_{L_{\max}}$ TO = lift coefficient at the takeoff configuration

This equation will also permit the determination of combinations of W/S, T/W, and $C_{L_{\max}}$ that will generate the design value of takeoff ground run. Typical results for takeoff and landing are shown in Figure 13.19.

Typical ground rolling coefficients of friction, μ_G, are given in Table 13.7.

Influence of Landing Distance Requirements on Aircraft Design

Landing wing loading, approach speed, deceleration techniques, and pilot techniques all influence the resulting landing distance. There is usually a design landing weight specified for the type of aircraft (Table 13.8), as well as an obstacle height and a set of government requirements. Typically, for CTOL aircraft the total landing distance consists of an air distance and a ground distance. The design aim is to establish the appropriate combination of wing loading and $C_{L_{\max L}}$ necessary to meet the stated landing distance required. It can be shown that for civil/commercial design the landing distance is a function of the square of the approach speed, V_A^2,

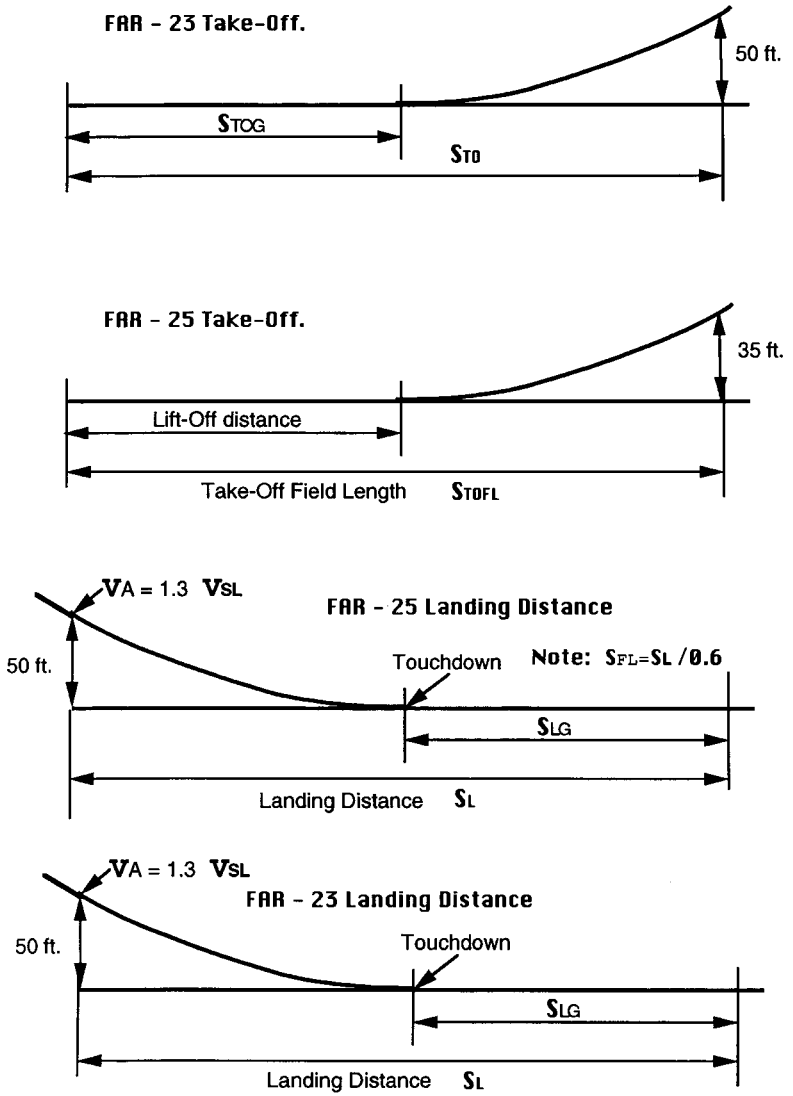


FIGURE 13.23 Takeoff and Landing Profiles (from Roskam 1980).

where $V_A = 1.3\ V_s$; for land-based military aircraft $V_A = 1.2\ V_s$; and for carrier-based aircraft $VA = 1.15\ V_s$, where V_s = the appropriate stall speed in the landing configuration.

For landing design evaluations the following equations permit the generation of a range of equivalent takeoff wing loadings necessary to meet a given landing distance requirement. For FAR 23 rules, for landing,

TABLE 13.7 Typical Ground Rolling coefficients of Friction

Surface type	μ_G
Concrete	0.02–0.03
Asphalt	0.02–0.03
Hard turf	0.05
Short grass	0.05
Long grass	0.10
Soft ground	0.10–0.30

Source: Roskam 1980.

TABLE 13.8 Maximum Landing Weight Design Fractions

Airplane type		Average (W_L/W_0)
Propeller	• Single engine	1.00
	• Twin engine	0.99
Business jets		0.88
Transport jets		0.84
Military trainers		0.99
Military patrol, bombers, and transports	• Jets	0.76
	• Turboprop	0.84

Source: Roskam 1980.

$$(W/S)_{TO} = \frac{S_L \cdot \rho \cdot C_{L_{\max}}}{0.3606\ (W_L/W_0)}\ \text{(psf)},\ \text{where}\ V_A = 1.3\ V_{SL}$$

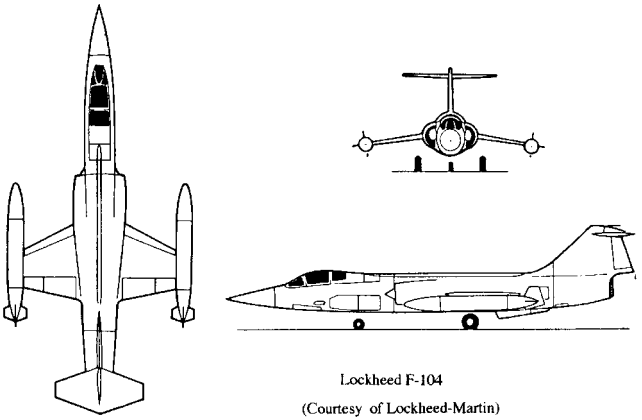
For FAR 25 rules, for landing,

$$(W/S)_{TO} = \frac{S_{FL} \cdot \rho \cdot C_{L_{\max}}}{0.3559\ (W_L/W_0)}\ \text{(psf)},\ \text{where}\ V_A = 1.3\ V_{SL}$$

and for military aircraft design, for landing,

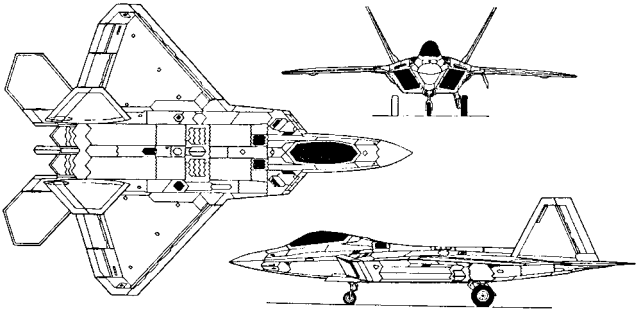
$$(W/S)_{TO} = \frac{S_L \cdot \rho \cdot C_{L_{\max}}}{0.3033\ (W_L/W_0)}\ \text{(psf)},\ \text{where}\ V_A = 1.2\ V_{SL}$$

but using the FAR 25 ground rules as a good approximation (S = distance, ρ = air density). In all landing design evaluations, V_{SL} is the stall speed with the aircraft in the appropriate configuration such as landing gear down, flaps down and the engine at the appropriate setting. S_L is defined as the desired landing distance in feet and (W_L/W_0) is the design value of landing weight to gross weight ratio (Table 13.8). No allowance is made in these distance estimates for the reduction in landing ground run due to the effects of reverse thrust, drag parachute, or spoiler application.



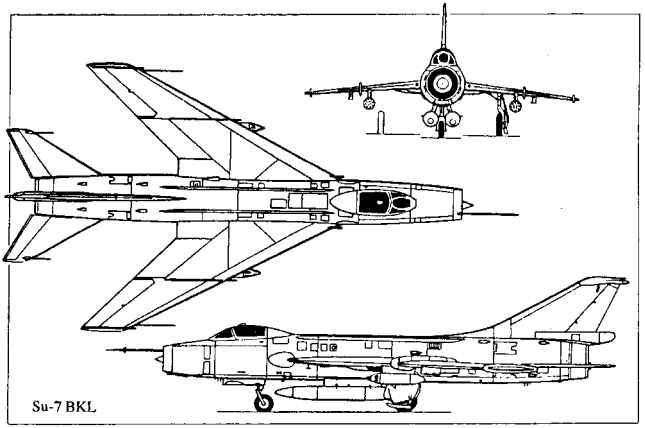
Lockheed F-104

(Courtesy of Lockheed-Martin)



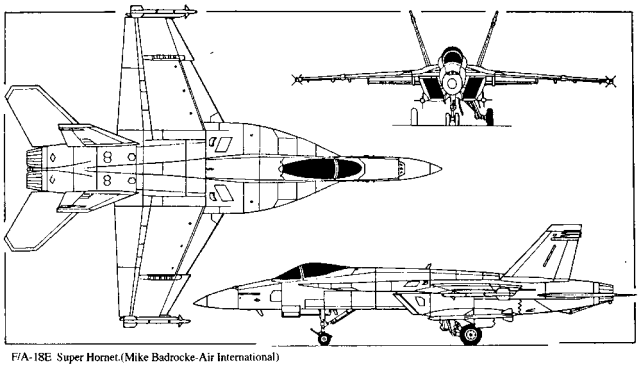
Lockheed Martin F-22A

(Courtesy of Lockheed-Martin)



(Air International. April 1994)

FIGURE 13.24 Typical fighter aircraft configurations.



F/A-18E Super Hornet.(Mike Badrocke-Air International)

FIGURE 13.24 (*Continued*) Typical fighter aircraft configurations.

Geometric Layout of the Wing

Experience has shown that the magnitudes of the taper ratio (λ), aspect ratio (A), and thickness-to-chord ratio (t/c) are driven by the operational function of the aircraft. The taper ratio chosen is a compromise between structural efficiency (low values) and wing stall (high values). For example, it can be shown that the spanwise stall point on the semispan of the wing from the centerline of the aircraft is approximately equal to the following:

$$\text{Stall point} = (1 - \lambda)(b/2), \text{ where } (b/2) \text{ is the wing semispan}$$

Wing planform layout and fuselage shape are determined primarily by the performance demands of the concept, in terms of Mach number, maximum dynamic pressure, and function or purpose of the concept.

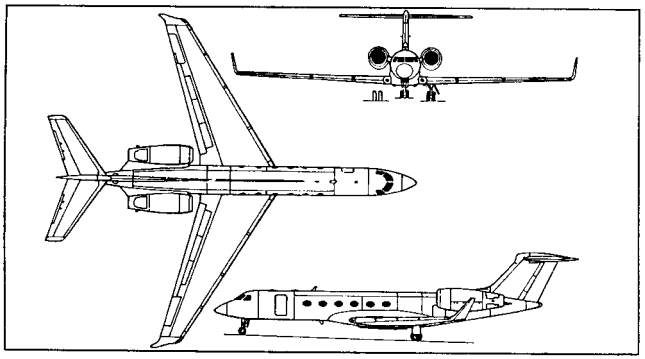
Examples of these planforms are shown in Figures 13.24–13.27.

The wing and sometimes the fuselage will be required to carry fuel. An approximate estimate of the fuel volume available in the wing is given by

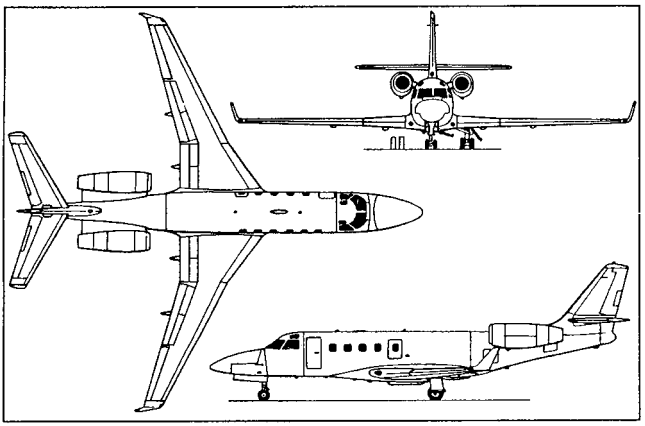
$$\text{Fuel volume} = 1.07\,(\Delta c)\left(\frac{S^2}{b}\right)\left(\frac{t}{c}\right)_o \left\{\frac{1 + \lambda_w(\tau_w)^{0.5} + (\lambda_w)^{2.0}\,\tau_w}{(1 + \lambda_w)^{2.0}}\right\} (\text{ft}^3)$$

where $(t/c)_o$ = thickness to chord ratio of the wing at the wing root
S = wing area (ft^2)
b = wing span (ft)
λ = wing taper ratio
$\tau_w = \left[\dfrac{(t/c)_{\text{tip}}}{(t/c)_o}\right]$, for the wing

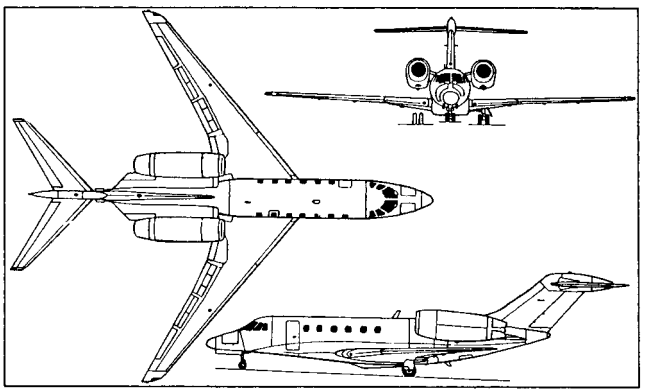
The magnitudes of S and b are those values covering the extent of the wing assumed to carry the fuel. Thus subscript o defines the inboard extent of the wing fuel tank, and Δc is the distance between the front spar and the rear spar of the wing which encloses the wing fuel. For example, $\Delta c = (0.75 - 0.25) = 0.50$ for a wing with the front spar at 25% chord and the rear spar at 75% chord. As can be seen, this volume is a function of the placement of the spars. Even if multiple spars are used, this equation is still valid as a first pass estimate.



Above: ***Gulfstream V. (Mike Badrocke/AI)***

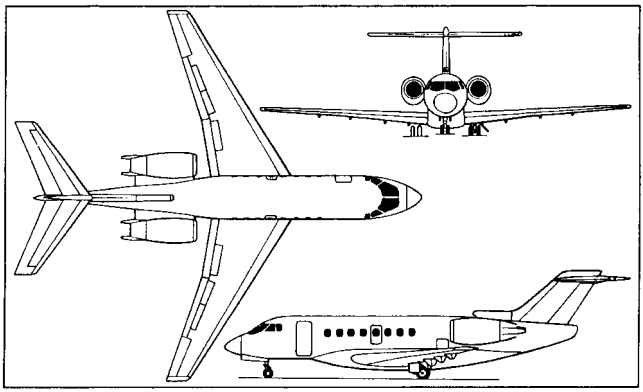


Above: ***Galaxy Aerospace 1125A Astra SPX. (Mike Badrocke/AI)***



Above: ***Cessna 750 Citation X. (Mike Badrocke/AI)***

FIGURE 13.25 Typical business jet configurations (Mike Badrocke, *Air International,* April 1994).



Above: ***Raytheon Hawker 4000 Horizon. (Mike Badrocke/AI)***

FIGURE 13.25 (*Continued*) Typical business jet configurations (Mike Badrocke, *Air International*).

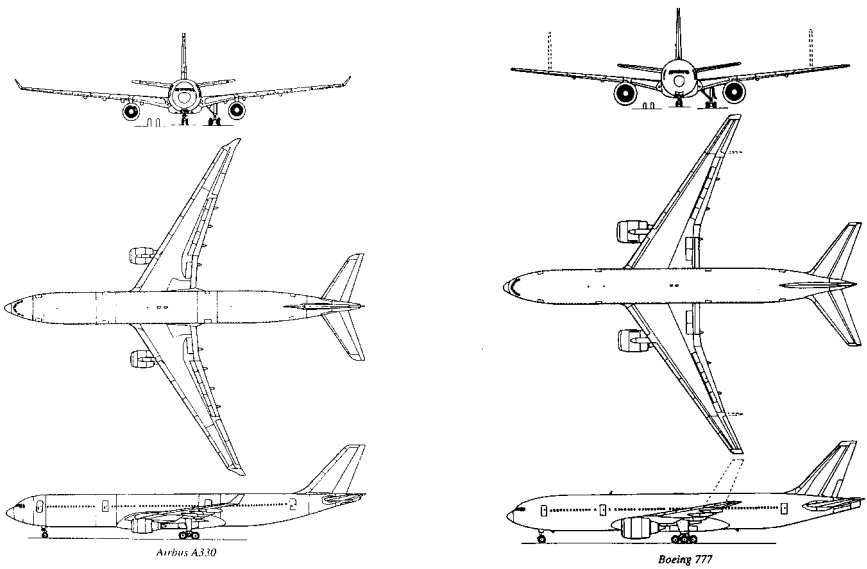
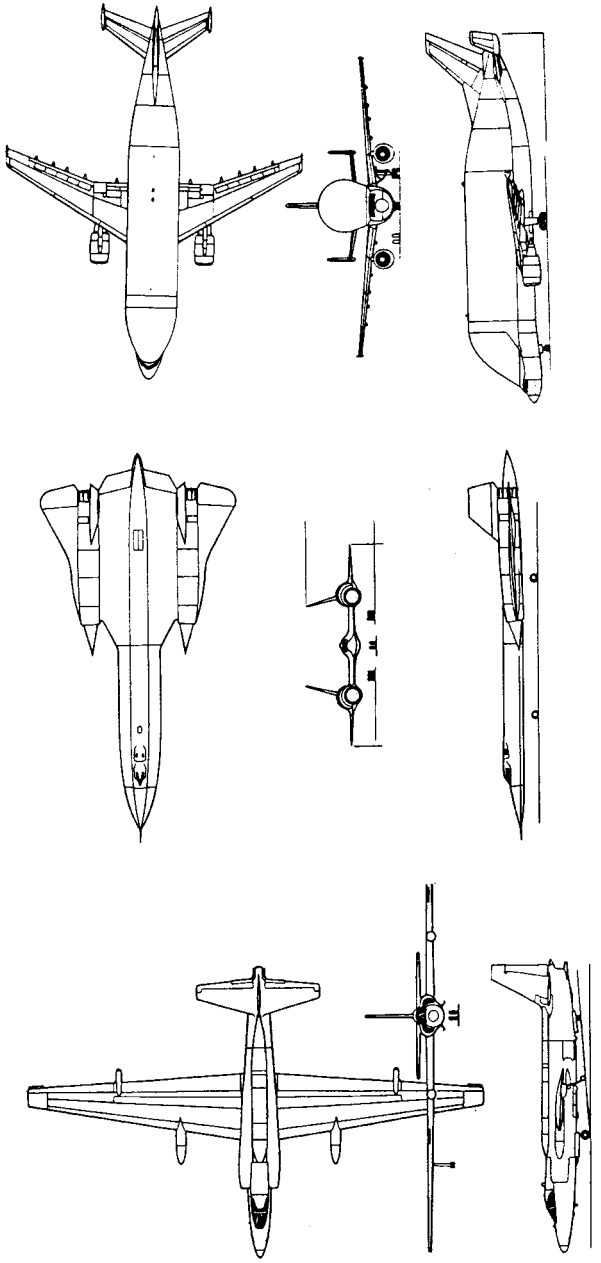


FIGURE 13.26 Typical large commercial transports (Mike Badrocke, *Air International,* April 1994).



SATIC A300-608ST "Beluga"(Mike Badrocke)
(Air International, November 1997)

Lockheed SR-71
(Courtesy of Lockheed-Martin)

Lockheed TR-1
(Courtesy of Lockheed-Martin)

FIGURE 13.27 Some special-purpose configurations (Lockheed-Martin and *Air International*, November 1997).

Fuselage

Designed to house the crew, fuel, and cargo, the fuselage design must also make allowance for loading and unloading of cargo, and for single-engine aircraft the fuselage usually also houses the engine. Drag produced by the fuselage is almost always exclusively parasitic and so must be minimized. The advantage of flying wings is evident in this respect.

Historically the length and diameter of the fuselage have been shown to be a function of the mission of the aircraft, which is usually tied to the cargo or payload required. For civil and commercial aircraft the dimensions of the fuselage are closely tied to the proposed seating layout in terms of numbers of seat rows, called seat pitch, and the number of seats abreast (Figure 13.28).

These requirements reflect market forecasts and passenger acceptance and vary with the class of accommodations. FAA/JAR regulations determine the numbers of aisles required in the cabin. The maximum number of seats on either side of an aisle must not exceed three. For military cargo aircraft, loading and unloading of cargo is a paramount design parameter and usually dictates the aircraft configurational layout (Figure 13.6). For fighter/bomber type aircraft with weapons bays the fuselage requires in-flight activated fuselage doors. For low-observable aircraft these doors must open and close with great rapidity.

Lacking explicit length needs, an approximation can be obtained for a range of concepts from Table 13.9.

Empennage

Described as the combined vertical and horizontal tail, the empennage function is to provide trim, maneuver, and stability for a range of aircraft centers of gravity. The two parameters which encompass these functions are the horizontal tail volume ($\overline{V}_H$) and the vertical tail volume ($\overline{V}_V$). These tail volumes are a function of the aircraft type and aircraft engine layout (Table 13.10) and depend on the required center of gravity range to be used on the aircraft (Table 13.11).

$$\overline{V}_H = \left[\frac{S_T l_T}{S_w \overline{c}}\right]$$

where S_T = horizontal tail area
l_T = distance from aircraft center of gravity to aerodynamic center of the horizontal tail
S_W = wing reference area
$\overline{c}$ = mean aerodynamic chord of the wing

and

$$\overline{V}_V = \left[\frac{S_V l_V}{S_w b}\right]$$

where S_V = vertical tail area
l_V = distance from aircraft center of gravity to aerodynamic center of the vertical tail
b = wingspan

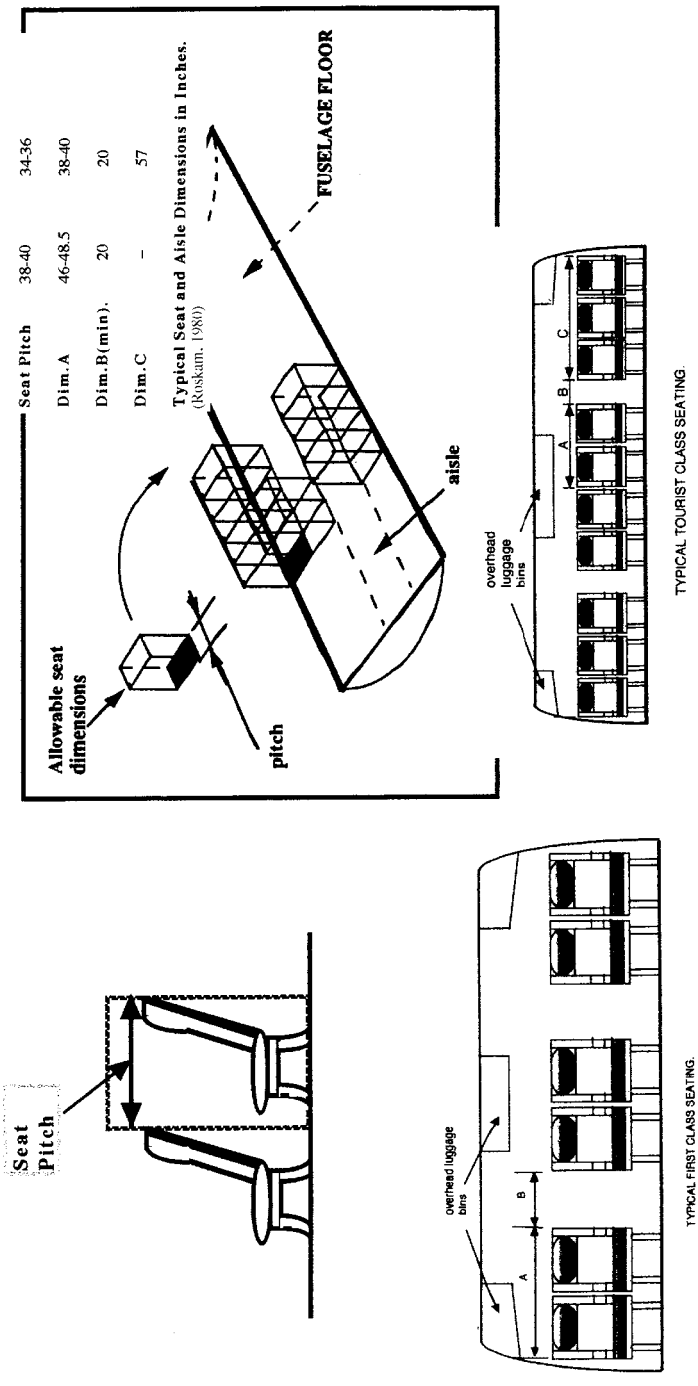


FIGURE 13.28 Passenger seat arrangement.

TABLE 13.9 Fuselage Length Relationships

Fuselage length $= A \cdot W_0^B$ (ft)		
Aircraft type	A	B
Fighter	1.1106	0.3700
Business jets	1.3286	0.3738
Wide-body jets	1.9311	0.3522

Source: K. Wagoboff.

TABLE 13.10 Examples of Tail Volume Coefficients

	Typical tail volume coefficients	
Aircraft Type	Vertical	Horizontal
General Aviation		
Single engine	0.04	0.70
Twin engine	0.07	0.80
Jet fighters	0.07	0.40
Military cargo/bomber	0.08	1.00
Jet transports	0.09	1.00

Source: Raymer 1989.

TABLE 13.11 Some Typical Center of Gravity Ranges as Fractions of Wing Mean Aerodynamic Chord, $\bar{c}$

Aircraft type	Center of gravity range (fr $\bar{c}$)	Aircraft type	Center of gravity range (fr $\bar{c}$)
Single-engine Prop.	0.06–0.27	Military trainers	0.10
Twin-engine Prop.	0.12–0.22	Fighters	0.20
Business jets	0.10–0.21	Military patrol, bombers transports	0.30
Jet transports	0.12–0.32	Flying boats, amphibious float	0.25

Source: Roskam 1980.

The geometric relationships of Figures 13.20 and 13.21 are used to generate the remaining empennage parameters.

Landing Gear

Certain aerospace companies now specialize in the design and manufacture of landing gear units. In modern aircraft the total weight of the gear is approximately 3% to 7% of the gross weight of the aircraft, or as high as 38,000 lb (Table 13.12). Correct interface with the airframe (and the gear design itself) is necessary for a

TABLE 13.12 Total Landing Gear Weight

Aircraft type	Total landing gear weight, (lb) (typical)	% (Gear wt/W_0) (typical)
Propeller • Single engine	191	5.6
• Twin engine	263	5.5
Business jets	659	2.8
	2,011	3.1
Transport jets	25,761	4.6
Military trainers	457	3.9
Military fighters	1,992	6.5 (U.S. Navy)
	1,848	5.9 (USAF)
Military transports • Jets	38,353	5.0
• Turboprops	5,309	3.4

Source: Roskam 1980.

successful aircraft design. Interface entails the appropriate ground clearance in the pitch and roll axes and the necessity to avoid lateral tip-over (Figures 13.29–13.31), as well as the ability to provide effective stowage within the airframe if necessary.

A general ground rule is that if the cruise speed of the aircraft is greater than 150 knots, then retraction is recommended to reduce cruise drag.

In the gear design itself the following factors of the gear strut and tire combination need to be addressed: gear vertical, longitudinal, and lateral loads; and gear dynamics, including vertical damping and spring constants and directional stability (sometimes called shimmy). These dynamic requirements are a strong function of the design-to aircraft rate of sink. Longitudinal placement of the gear for a nosewheel aircraft is a function of the actual most aft center of gravity of the aircraft as well as by the loads to be carried by the nose gear to ensure effective nosewheel steering (Figure 13.32). Lateral placement of the main gear is determined primarily by the most forward center of gravity of the aircraft and by the desire to avoid lateral-longitudinal tip-over (Figure 13.31).

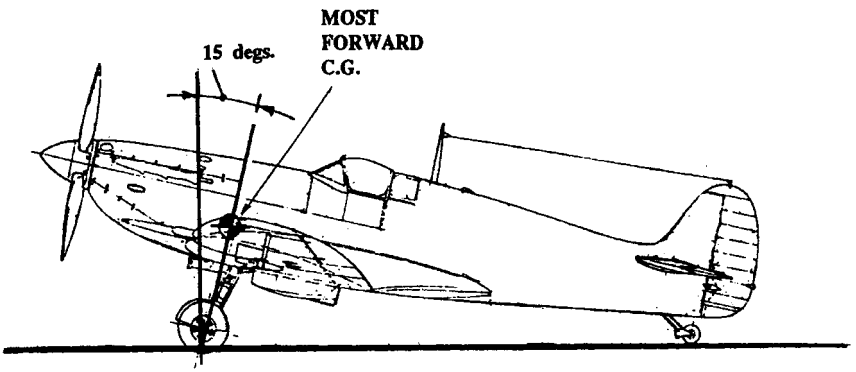


FIGURE 13.29 Longitudinal tip-over guideline—tail dragger.

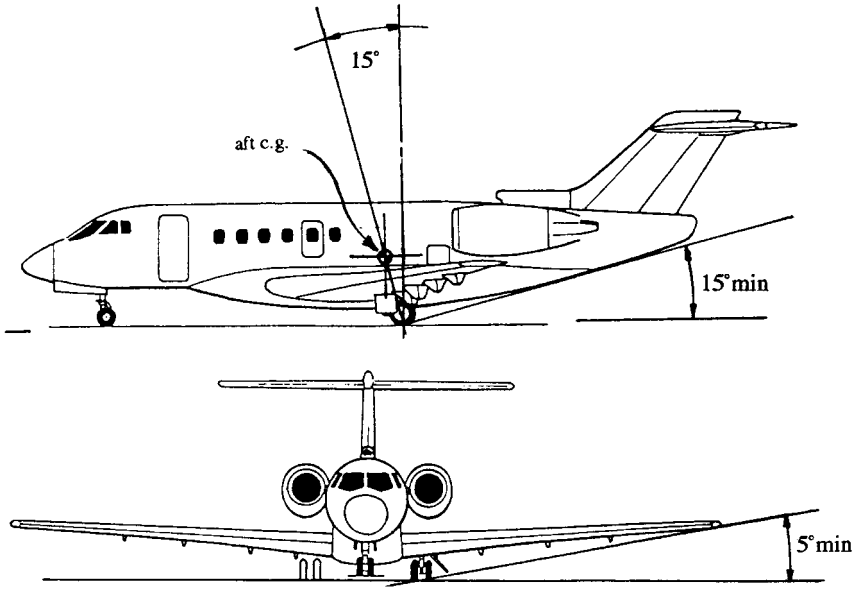


FIGURE 13.30 Lateral clearance and tip-back criteria (aircraft by Mike Badrocke, *Air International*).

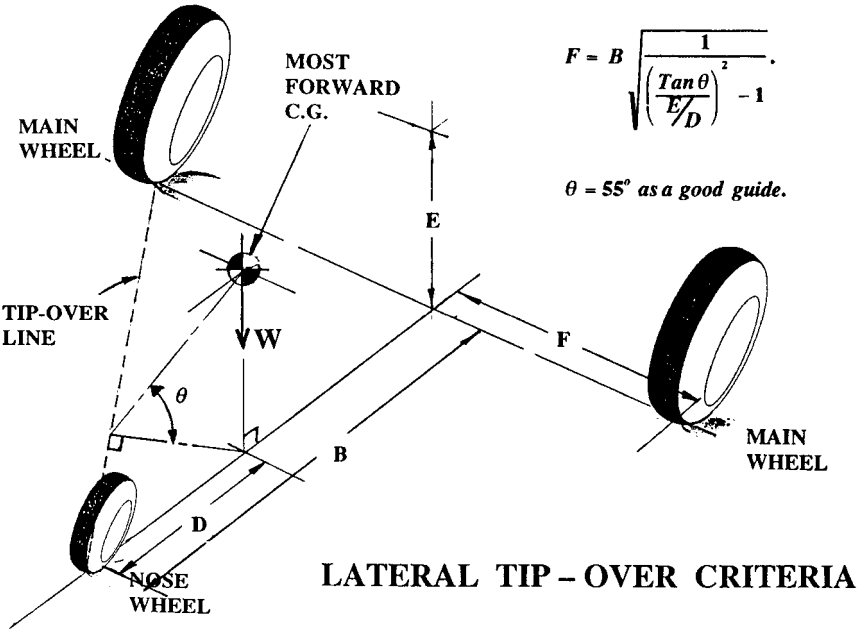


FIGURE 13.31 Lateral-directional tip-over criteria (from Roskan 1980).

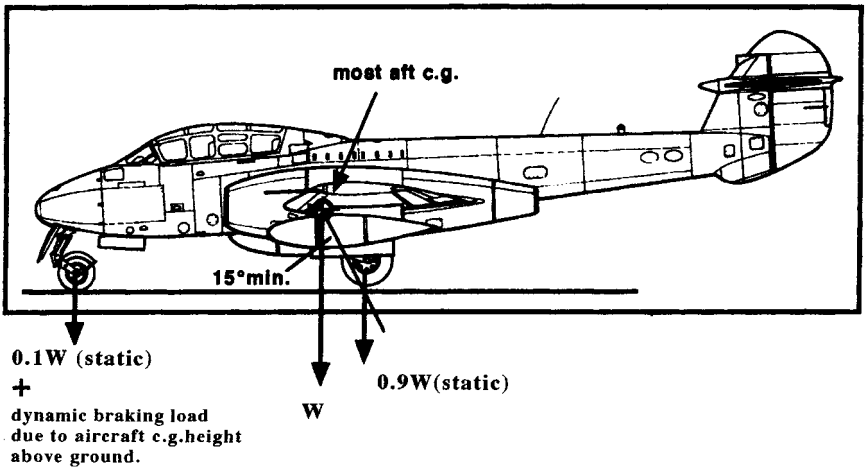


FIGURE 13.32 Longitudinal tip-back guidelines and wheel load guideline (after Raymer 1989).

Sizing of the tires is a function of the tire loads and the tire operating speeds. The number and configuration of the tires on the strut has become a function of the gross weight of the aircraft, runway surface pressure (Figure 13.33), as well as the stowage location in the airframe. Gears are primarily retracted into the fuselage and/or the wings, but occasionally retraction is made into specialized fuselage pods to accommodate special requirements, usually military, and into wing pods for other needs (Figure 13.34 and 13.35).

Weight/Center of Gravity Influence on Design Layout

Every element comprising an air vehicle contributes to the total weight and resulting center of gravity (see Figure 13.36). For all air vehicles fuel burn-off will influence weight and the center of gravity; for military aircraft, weapons delivery will also influence these items. Figure 13.37 illustrates a method for estimating the longitudinal location of the center of gravity, and the following equation locates this center of gravity on the mean aerodynamic chord of the wing from the leading edge.

$$\frac{\text{c.g. location}}{\bar{c}} = \left(\frac{1}{\bar{c}}\right)\frac{\sum_{n=1}^{n=x} l_n w_n}{\sum_{n=1}^{n=x} w_n} - \left(\frac{l_A}{\bar{c}}\right) - \left(\frac{b}{4c_0}\right)\left(\frac{1+2\lambda}{1+\lambda+\lambda^2}\right)\text{Tan}\,\Lambda_{L.E.}$$

where c_o = centerline chord of the wing
λ = wing taper ratio

Vertical c.g. location can be determined by similar methods by rotating the datum line to a horizontal location beneath the aircraft.

Figure 13.38 shows a typical variation of center of gravity throughout a simple mission. The required c.g. range is shown to be $0.30\bar{c}$ as an example, and the

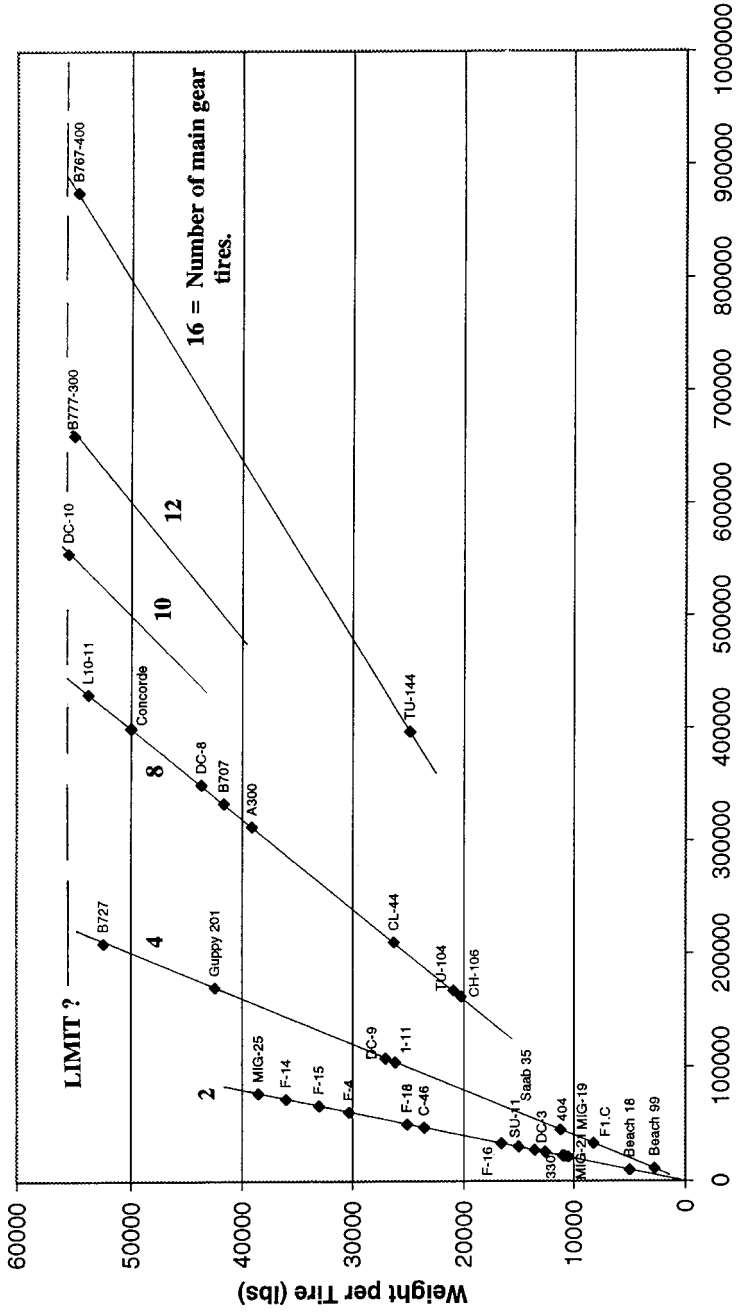


FIGURE 13.33 Weight per tire versus gross weight (T. Cofer).

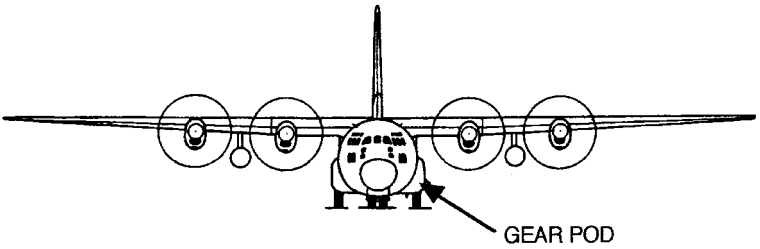


FIGURE 13.34 Special landing gear pods—C 130 (Lockheed-Martin).

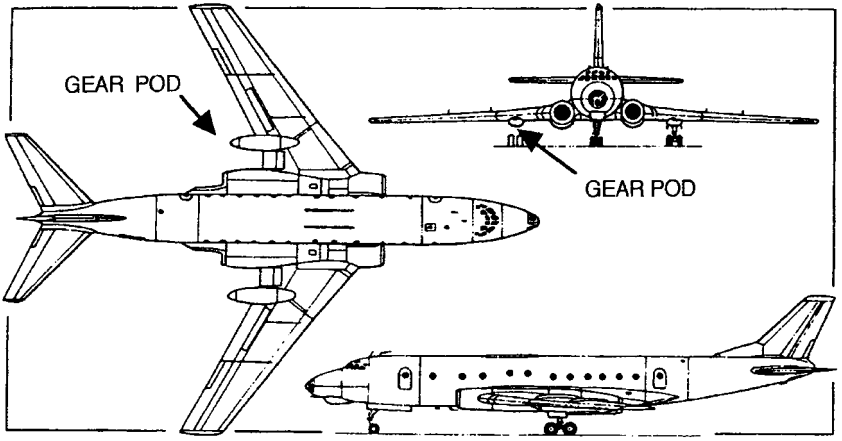


FIGURE 13.35 Special landing gear pods—Tu 154 (Mike Badrocke, *Air International*).

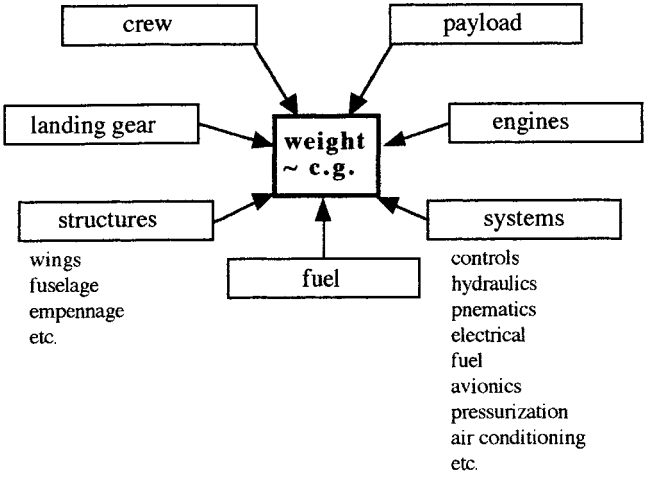


FIGURE 13.36 Contributors to center of gravity estimation.

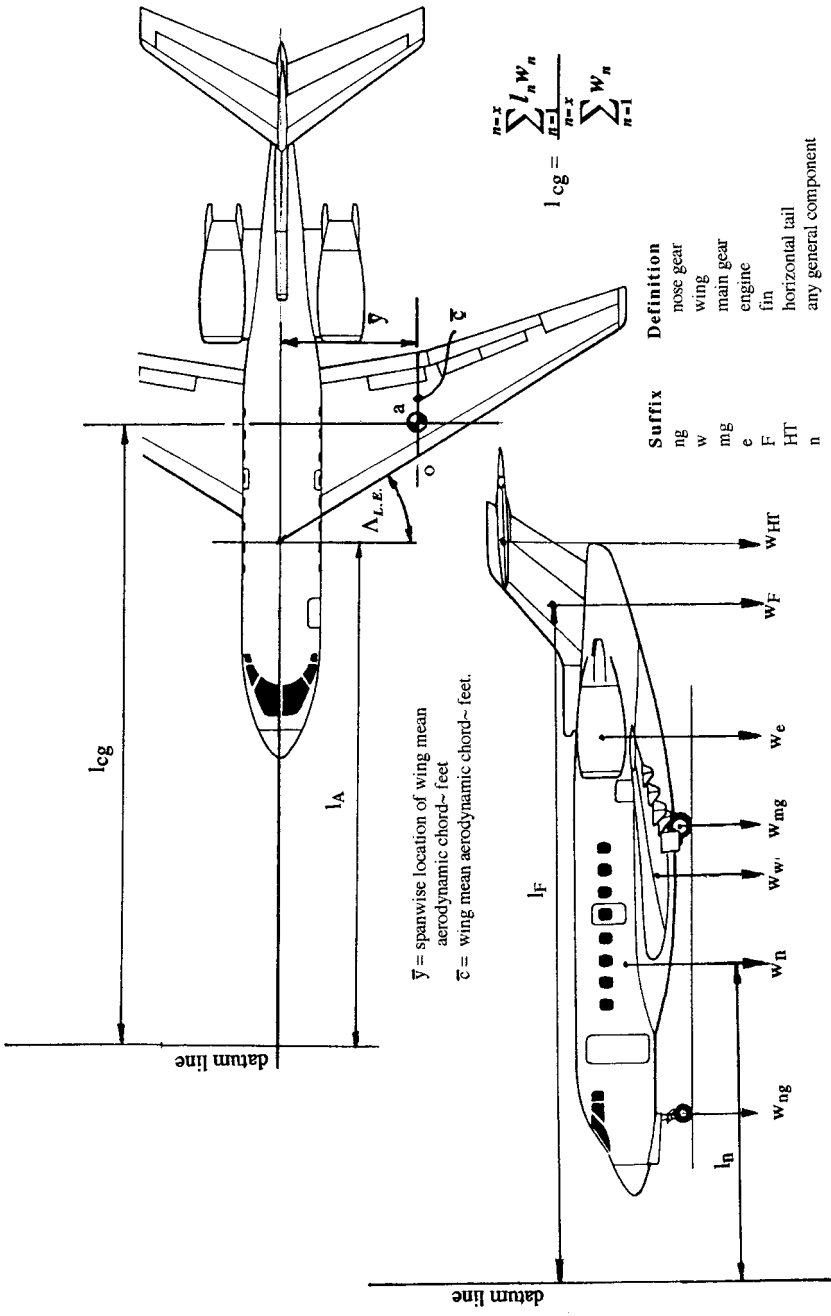


FIGURE 13.37 Center of gravity estimation (aircraft drawing from Mike Badrocke, *Air International*).

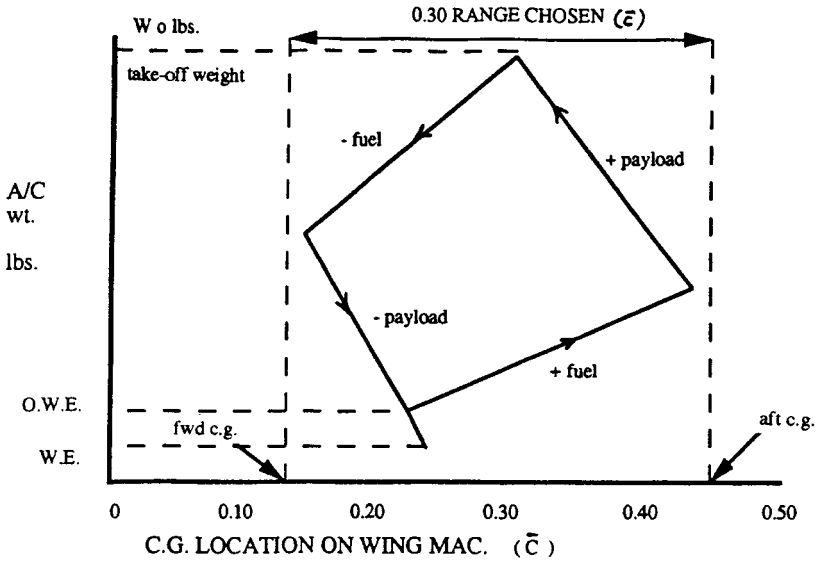


FIGURE 13.38 Typical center of gravity variation during mission.

forward (fwd) and aft c.g. limits must be sustainable by the aerodynamics generated by the horizontal tail or its equivalent.

Ground Clearance

Landing gear design also must ensure that lateral and longitudinal (Figure 13.30) ground clearances of 5° minimum and 15° minimum are observed. The latter promotes safe aircraft rotation on takeoff as well as structural clearance on landing. By observing a 15° aft c.g./main gear line relationship, inadvertent tail sinking can be avoided on rotation during takeoff.

Propulsion

The propulsive method can be propellers, with reciprocating or gas turbine power sources, or fans with gas turbines or just pure jets. The physical characteristics of typical modern gas turbines are given in Figure 13.39 for maximum thrust values up to 110,000 lb.

Typical cruise values of specific fuel consumption, sfc, in pounds of fuel per pound of thrust per hour are 0.35 to 0.38 for thrusts up to about 70,000 lb, and approximately 0.55 to 0.58 for engines with maximum thrusts from 70,000 lb to 110,000 lb. Figure 13.39 also shows the corresponding physical properties for shaft turbines up to 6,000 shaft horse power (shp). For these engines cruise sfc is between 0.45 and 0.55 lb/shp/hour. Typical fan and shaft gas turbine engines are shown in Figure 13.40, which also illustrates the techniques used to extract power to drive the various aircraft systems. Engines are usually mounted to the aircraft structure with a fixed forward mount and a flexible aft mount (Figure 13.41).

	GAS TURBINE	SHAFT TURBINE
overall length (in.)	$2.2941(T)^{0.3849}$	$35.459e^{0.0001(SHP)}$
maximum diameter (in.)	$0.5649(T)^{0.4775}$	$5.0679(SHP)^{0.2310}$
weight (lb)	$0.7191(T)^{0.8642}$	$1.0287(SHP)^{0.8044}$
	Up to 110,000 lb thrust	Up to 6,000 shp

FIGURE 13.39 Typical physical characteristics of propulsion types (*Aerospace Source Book* 1999).

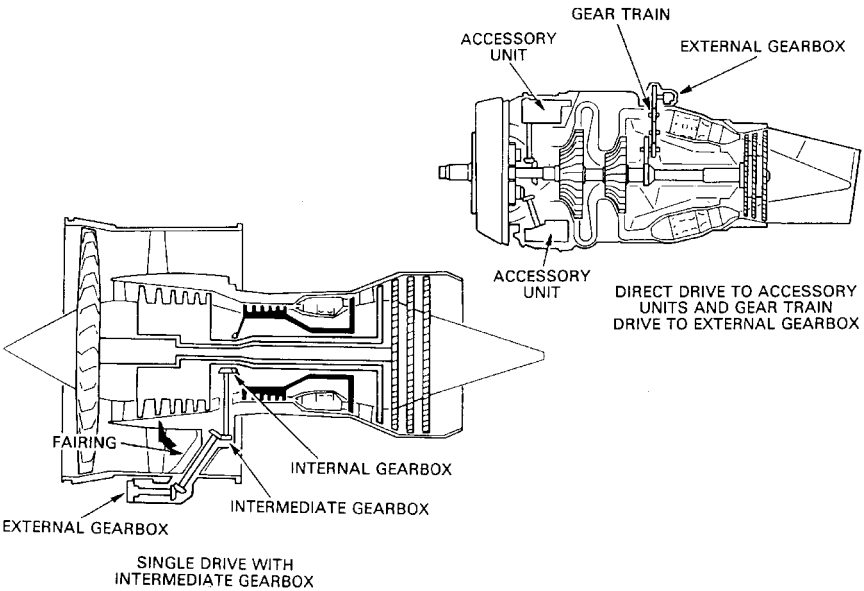


FIGURE 13.40 Typical power plants (reproduced with kind permission of Rolls-Royce plc 1986).

In addition to providing power (thrust) for performance needs, the power plant must also generate sufficient power to drive all the other aircraft systems necessary for the concept to perform the design mission. Some typical systems are listed in Table 13.13. Table 13.14 lists the resulting typical system power requirements. Typically, JP-4 fuel weighs 6.5 lb/gallon with a density 49 lb/cubic foot.

13.6 MILITARY AIRCRAFT DESIGN

The basic design principles are the same as given in Subsection 13.1, with the additional need to account for the effects of specific military requirements. In general, these requirements influence the design by changes in the following:

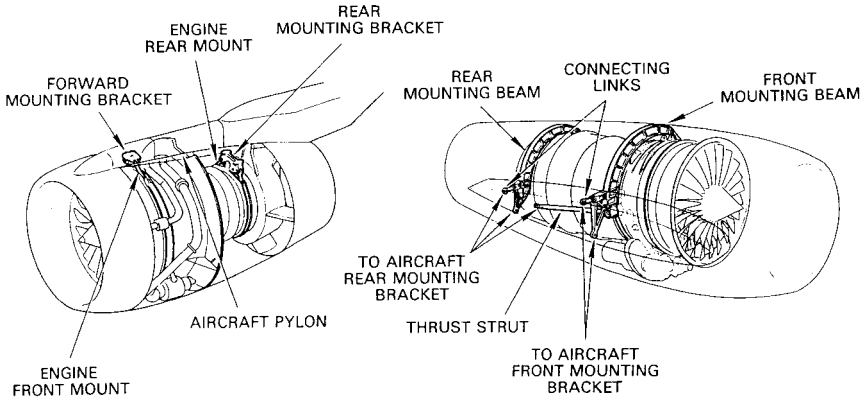


FIGURE 13.41 Engine mounting techniques (reproduced with kind permission of Rolls-Royce plc 1986).

TABLE 13.13 Typical Onboard Systems

Fuel pumps	Deicing
Hydraulic pumps	Antiicing
Cooling fans	Engine starting system
Heating system	Pressurization of fuel tanks
Cooling system	Pressurization system
Radar	Electronic warfare systems

TABLE 13.14 Onboard System Typical Power Requirements

Aircraft type	Electrical, shp	Mechanical, shp	Pneumatic bleed slugs/s
Single engine, light a/c	2–4	3–5	0.01 $m_{\dot{a}}$
Twin-engine turbojets/fans	8–10	9–11	0.025 $m_{\dot{a}}$
Jet fighters—attack	100–200	100–200	0.04 $m_{\dot{a}}$
Jet transports—civil	0.0007 W_0	0.0006 W_0	0.03 $m_{\dot{a}}$

W_0 = maximum design weight (lb)
$m_{\dot{a}}$ = Engine air mass flow rate (slugs/s)

Source: Roskam 1980.

- Engine thrust/installation/shielding
- Fuselage volume/dimensions
- External shaping of aircraft (radar reflectivity)
- External additions to the aircraft:
  - Arresting hook
  - Launch attachments (carrier)
  - Probe for in-flight refueling

- Underwing/fuselage pylons for external stores
  - Internal or external weapons
  - Folding wings
  - External surface treatment such as radar-absorbing material or finishes

All of these induce weight changes, cost increases, and performance changes, and the resulting designs are usually accomplished with the aim of minimizing these effects.

Radar cross-section values are a function of measurement azimuth and shape of the aircraft with many contributors (Figure 13.42). The rapid development of aircraft external shape/radar reflectivity technology is shown by comparing the early faceted approach to design (Figure 13.43) and the more modern approach (the F-22 in Figure 13.24), each aircraft design coming from the same design organization.

An indication of radar reflectivity, in terms of meters squared, and the published approximate values for a range of aircraft are shown in Table 13.15. These values are meant to be relative only and will vary with azimuth and radar frequency (*Air International*).

Minimizing radar reflectivity can be achieved by a combination of (a) absorbing incoming radar energy and (b) redirecting the reflected wave away from receiving sensors. Design centers on the development and use of radar-absorbing materials and coatings and also the appropriate external shaping and surface integrity of the aircraft. Surface integrity effects primarily involve the avoidance of long surface gaps such as weapons doors and control surface shroud lines because these reflect surface currents unidirectionally, causing reflection spikes. Aircraft external shaping is aimed at minimizing reflection spikes or on concentrating a spike in a given direction.

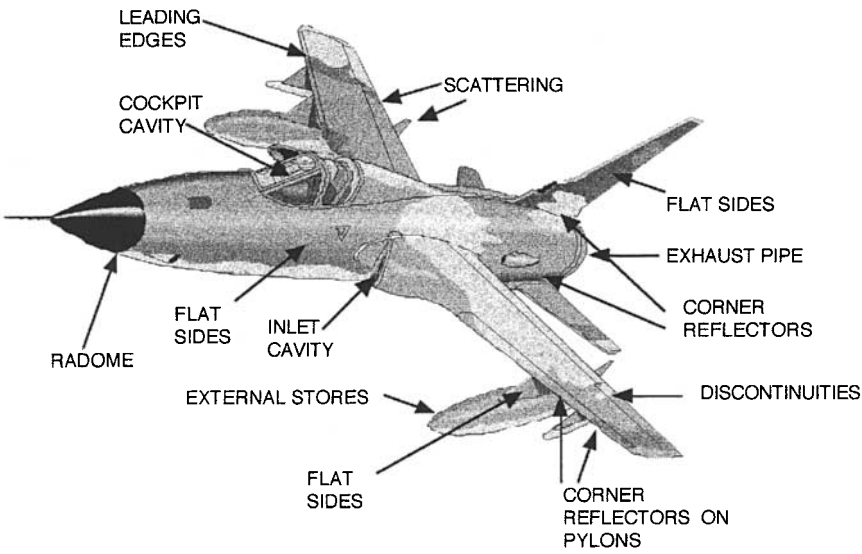


FIGURE 13.42 Major RCS contributors.

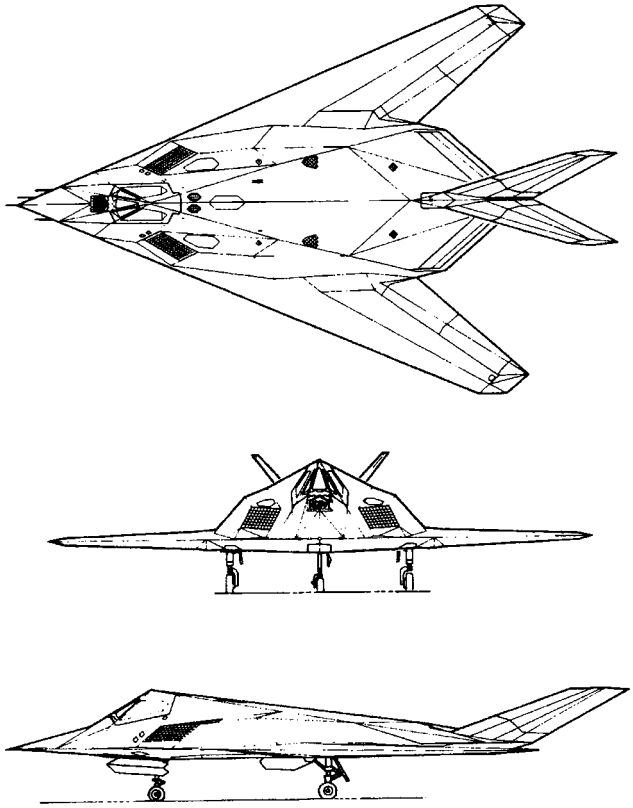


FIGURE 13.43 Lockheed-Martin F-117A (Lockheed-Martin).

TABLE 13.15 Comparison of Typical Radar Cross-Sections

Aircraft	Radar cross section (m^2)
McDonnell Douglas F-15	40.5
Boeing B-52	99.5
Rockwell B-1A	10
Rockwell B1-B	1.02
Lockheed SR-71	0.014
Lockheed F-22A	0.0065
Lockheed F-117A	0.003
Northrop B-2	0.0014

Source: *Air International Magazine* 1995.

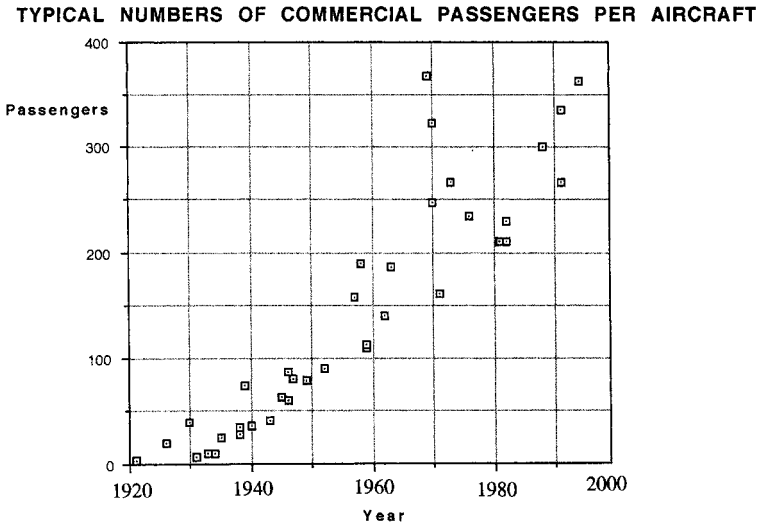


FIGURE 13.44 Passengers per aircraft trends.

13.7 COMMERCIAL AND CIVIL AIRCRAFT DESIGN

Aircraft that are nonmilitary and that carry passengers and or freight make up this classification. The basic design principles outlined earlier in this section apply, starting from a commercial or civil mission requirement. Since their inception, the number of passengers carried by typical commercial airliners has increased significantly (Figure 13.44). Aircraft cruise speed and cruise altitude have also increased (Figure 13.45).

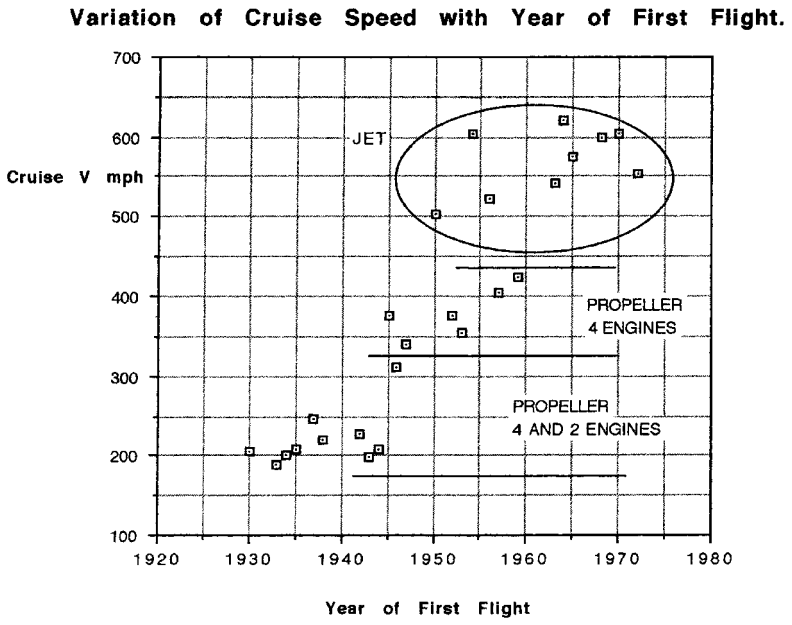
The difference between military and commercial/civil aircraft empty weight is primarily due to the necessary comfort and conveniences required for the passengers in commercial/civil aircraft. These include galleys, seats, multiple toilets, overhead luggage racks, baggage storage, in-flight entertainment, and cabin pressurization. Figure 13.46 shows a typical cabin versus outside pressure comparison for a typical airliner.

Passenger seating space allowances are a function of the class of travel chosen (Figure 13.28), and are a significant factor in the passenger acceptability of a particular design. Seating space is characterized by seat pitch and seat width.

13.8 LIFE CYCLE COST (LCC)

Life cycle cost is considered to be the total cost of a concept from cradle to grave (Figure 13.47). It is known that 95% of LCC is usually locked in or committed at the end of the preliminary design phase. Extensive methodology exists to estimate LCC for aircraft (see Roskam 1980).

Four elements are considered part of LCC: costs for RDT&E, acquisition and manufacturing, operations, and disposal.



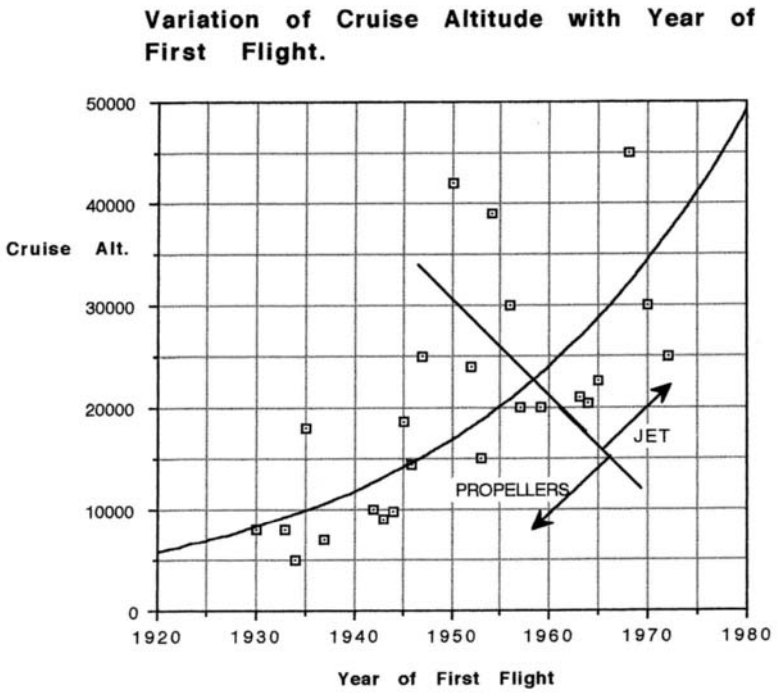


FIGURE 13.45 Cruise speed and cruise altitude trends.

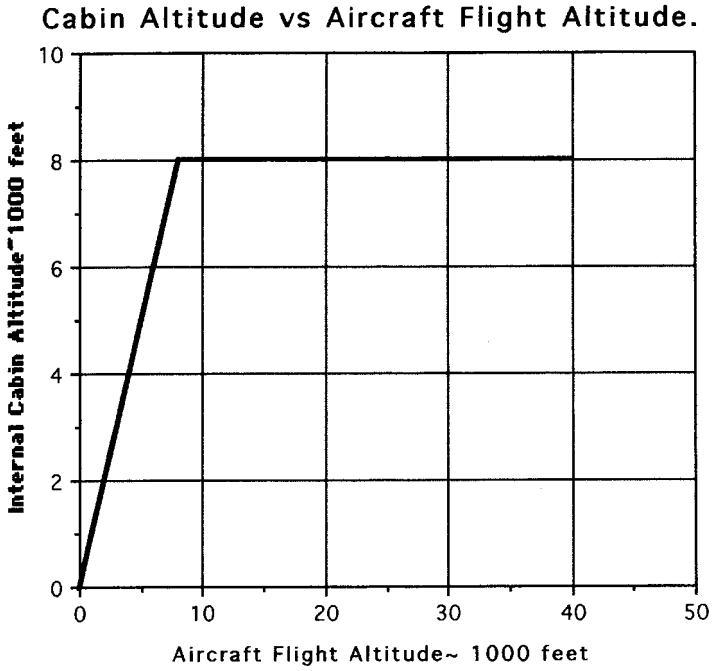


FIGURE 13.46 Typical cabin altitude variation.

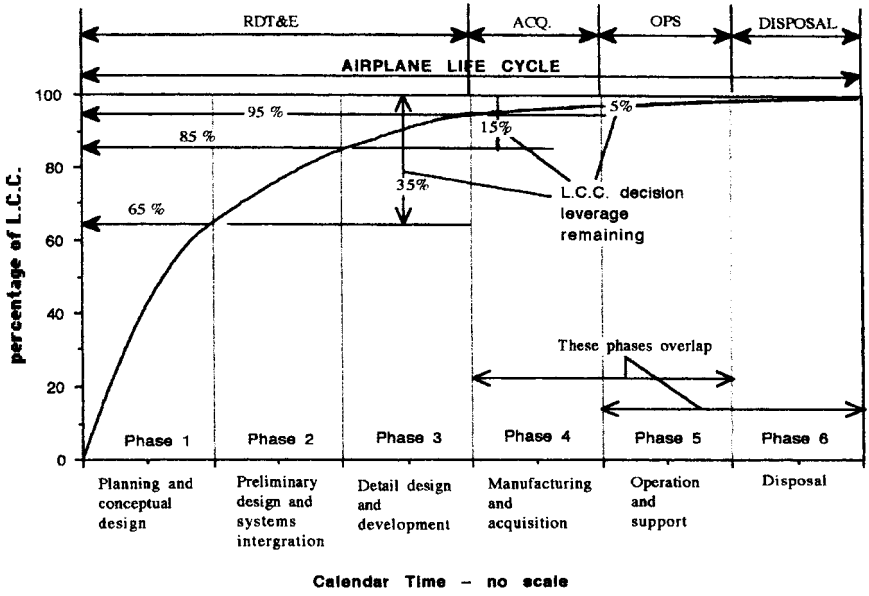


FIGURE 13.47 Overview of life cycle cost for aircraft (from Roskam 1980).

In general, RDT&E costs and acquisition and manufacturing costs are composed of many contributions, each possessing the following format:

$$\Delta \text{ Costs} = f(V)^a \cdot (W^b\text{ampr}) \cdot (\overline{n})^c \cdot (\$/\text{hr})$$

where V = maximum airspeed (knots e.a.s.)
W ampr = ampere weight (lb)
$\overline{n}$ = number of aircraft built
$/hr = specific labor rates
exponents a, b, c, represent experience based numbers
f = numerical values based upon allowances for such things as the use of CADAM/CATIA, new materials, special observable requirements, and program security needs

The total RDT&E, manufacturing, and acquisition costs are then

$$\sum_1^n \Delta \text{ Costs}$$

where n = number of cost elements involved. Figure 13.48 summarizes some of these cost element sources and shows one derivation of airplane estimated price.

Other factors, such as prototype and production aircraft manufacturing rates, avionics cost, inflation, profit, cost of money, and number of aircraft produced, all influence the individual aircraft price. Figure 13.49 shows a typical variation of aircraft price with number of aircraft manufactured.

13.9 COMMERCIAL AIRCRAFT OPERATING COSTS

Total operating costs (TOC) are composed of the sum of the indirect operating costs (IOC) and the direct operating costs (DOC). These costs can be expressed as either $ per nautical mile or $ per block time. (Sometimes ¢ or £ are substituted for $). Block time is defined as flight time plus 25 minutes for long haul or flight time plus 15 minutes for short haul (Jenkinson et al. 1999).

The magnitude of DOC is influenced in the design phase of an aircraft, whereas in general the magnitude of IOC is not.

Elements of each cost component are summarized below (see Figure 13.50).

- Direct operating costs:
  - Flight costs:
    - Crew costs (not including flight attendants)
    - Petroleum, oil, and lubricants
    - Insurance costs:
      - Airframe ground and flight risk
      - Passenger liability
      - Third-party liability
      - Cargo damage

$C_{RDTE} = C_{aedr} + C_{dstr} + C_{ftar} + C_{ftor} + C_{tsfr} + C_{pror} + C_{finr}$

= Research, development, test and evaluation cost

where C_{aedr} = Airframe engineering and design cost

C_{dstr} = Development support and testing cost

C_{ftar} = Flight test airplanes cost

C_{ftor} = Flight test operations cost

C_{tsfr} = Test and simulation facilities cost

C_{pror} = RDTE profit

C_{finr} = Cost to finance the RDTE phases

$C_{MAN} = C_{aedm} + C_{apcm} + C_{ftom} + C_{finm}$

= Manufacturing and acquisition cost

where C_{aedm} = Airframe engineering and design cost

C_{apcm} = Airplane program production cost

C_{ftom} = Production flight test operations cost

C_{finm} = Cost to finance the manufacturing phase

Estimated airplane price $= (C_{RDTE} + C_{MAN} + C_{MANP})/(N_M)$

where C_{MANP} = Manufacturing profit

N_M = Number of aircraft made to the production standard

each C_{XXXX} above = $\Delta \text{cost}_{XXXX}$ = a cost element

and $\sum \Delta \text{ cost}_{XXXX}$ = Total cost

Each appropriate C_{XXXX} will also establish appropriate man hour requirements necessary for the completion of that particular task. A summation of these man hour requirements will establish the foundation for the development of program planning, manpower distribution, and headcount needs for the program.

FIGURE 13.48 Typical Cost Elements (from Roskam 1980).

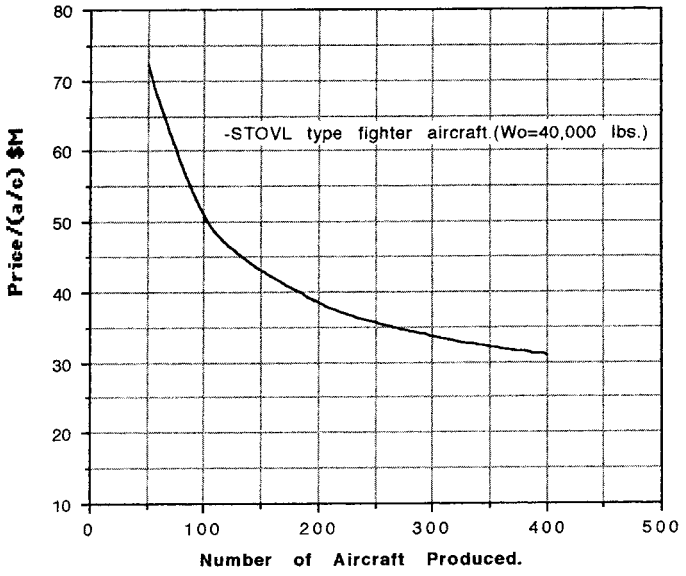


FIGURE 13.49 Typical influence of production run on aircraft cost.

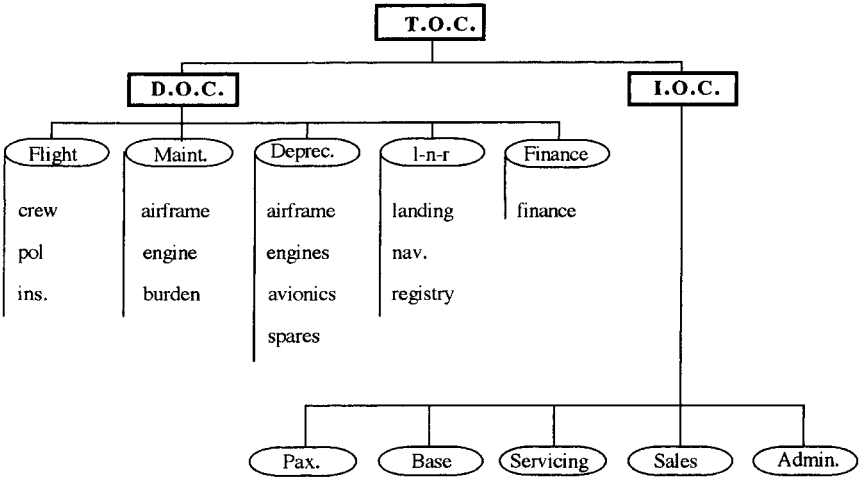


FIGURE 13.50 Elements of total operating costs (TOC) for commercial aircraft.

- Maintenance costs:
  - Airframe:
    - Labor costs
    - Material costs
  - Engine:
    - Labor costs
    - Material costs
  - Maintenance burden
- Depreciation costs:
  - Airframe depreciation costs, less engine, avionics, and spares
  - Engine depreciation costs
  - Avionics depreciation costs
  - Depreciation of spare parts:
    - Airframe
    - Engines
- Landing fees, navigation and registry taxes:
  - Landing fees
  - Navigation fee per aircraft
  - Registry taxes: function of aircraft size and particular state or country
- Finance DOC costs:
  - Function of operators' philosophy in financing fleet
- Indirect operating costs:
  - Passenger services (PAX): meals, cabin attendants, passenger insurance, sales and reservations, baggage handling, security
  - Maintenance (and depreciation) of ground equipment and base facilities
  - Airplane freight and traffic servicing
  - Sales: travel agent commissions; publicity; entertainment
  - Administrative: administrative and accounting and requirements for corporate staffers and their facilities

Equations for the estimation of the contributions of the above components to the DOC and IOC are extensively covered in Roskam (1980). The calculations rely upon historical data for the many classes of aircraft available in the commercial field and are shown to be fairly accurate in comparison with published TOCs.

All estimations of TOC other than those performed by the actual aircraft designers and operators can only be approximate and comparative between concepts. This is because actual inputs and actual mathematical models are by necessity commercially most sensitive and not available in the public domain.

13.10 UNMANNED AIR VEHICLES

Generically, UAVs are flying vehicles that rely upon wing, rotor, or gas bag methods to generate lift and that do not carry a human. Historically, UAVs have acquired

changing acronyms starting as remotely piloted vehicles (RPVs) through UAVs to the currently popular uninhabited aerial vehicles. All have a variety of control or navigation techniques such as direct control from the ground, relying upon line-of-sight control, through preprogrammed systems to semiautonomous and autonomous navigation systems. Launch and recovery techniques are ground based or air based, with the former using runways similar to manned aircraft or zero-zero assisted launch and recovery. Catapults, launch rails, and rocket assists are methods used for assisted takeoff, and nets or arresting wires for landing. In-flight parachute recovery and capture by specially modified manned aircraft have also been used. Propellers and jets form standard propulsion methods, with solar cells being constantly evaluated.

Air vehicle design of RPVs follows the principles outlined for manned aircraft, with appropriate (empty weight to gross weight) ratios, with suitable C_{D_o}'s, and with correct specific fuel consumption appropriate to the type of propulsion used. Table 13.16 shows characteristics of a range of typical RPVs.

UAV missions are many and varied and are constantly under development (see Table 13.17.

The takeoff gross weight of RPVs can range from 90 lb up to 26,000 lb. For weights up to 1,000 lbs, a good design value of $(W_e/W_0) = 0.4095\ (W_0)^{0.0648}$, and for weights greater than 1,000 lb lightweight aircraft design guidelines can be used initially. Propulsion types vary according to mission. For propeller-driven UAVs, the required shaft horsepower can be correlated with W_0, and this shows shp = $0.4109\ (W_0)^{0.7513}$, and this relationship can be used for conceptual design for W_0 up to 2000 lb.

For jet-driven vehicles propulsion can be sized following the manned aircraft design approach.

13.11 LIGHTER-THAN-AIR VEHICLES (LTA)

Generically called Aerostats, there are three classes of LTA: nonrigid, semirigid, and rigid airships. All rely on buoyancy for lift, and modern airships use inert helium as the gas. (Early airships used inflammable hydrogen as gas: $\rho = 0.00532$ lb/ft^3 for hydrogen and $\rho = 0.01056$ lb/ft^3 for helium). Some nonrigid aerostats use hot air for buoyancy; these are called hot air balloons. All airships use ballast weight and gas release for vertical ascent and descent (Figure 13.51). Horizontal thrust from jet engines or propellers controls the forward motion of the airship, and tailfins provide stability and airship angular control. More modern airship designs use vectored thrust from swiveling engines to assist in angular orientation and vertical ascent and descent. All modern airships enclose the gas in sealed gas bags with a controllable vent valve. Ascent speed, operating altitude, and payload weight depend on the percentage of available gas volume in the gas bag to the volume available.

$$\text{Available lift} = F \cdot V(\rho_a - \rho_g) - W_e - W_f - W_c \text{ (lb)}$$

where F = fraction of gas used to volume available
V = volume available in gas bag (ft^3)
ρ_a = density of air at local altitude (lb/ft^3)
ρ_g = density of gas at same condition as ρ_a (lb/ft^3)

TABLE 13.16 Typical Unmanned Aerial Vehicles, Drones, and Rotary Wing Aircraft

Unmanned serial Vehicles and Drones

Manufacturer	Aerovironment Inc.	BAI Aerosystems Inc.	BTA Automatic Piloting International	Bombardier Inc.	CAC Systemes	Finmeccanica	Flight Refueling Ltd.
Designation	HILINE	Exodrone	Sheddon Mk3	CL 327 Guardian	Fox TX	Mirach 20	Falconet
Country	USA	USA	Israel	Canada	France	Italy	England
Mission	High alt Long-end recon.	Multipurpose	Recon / surv / target acq.	Recon / surv / target acq.	EW	Surv. Target, acq.	Target
Overall Length (ft)	19	6.1	8.6	6	9.1	11.8	12
Maximum Overall Span (ft)	50	8.1	13.5	13.3	12	13.6	10
Launch Weight (Less Booster) (lbs)	770	9.1	90	770	250	374	450
Power Plant	(1) Akkeman eng. recip.	(1) 8 hp 2-stroke	(1) 100 cc, 10 hp 2-stroke	(1) Williams WTS 117-5	(1) Limbach 22 hp. recip.	(1) 26 hp.	(1) Microturbo TJA24
Payload type	Autop. data link, nav. comp.	TV, IR, EW	TV	Various	Autop., nav., GPS	TV or FLIR / ELINT	—
Speed (MPH or Mach)	120	100	45–80	98	60–120	138	M 0.72
Endurance	1–2 days	2.5 hr.	6 hr	6.5 hr.	5 hr.	4 hr. +	1 hr.
Maximum Altitude	40,000	—	15,000	18,000	10,000	11,500	—

Manufacturer	General Atomics	Israel Aircraft Industries	Meggitt Defence Systems Ltd.	Oerlikon Contraves AG	Lockheed Martin Skunk Works	A.S. Yakoviev Design Bureau	Teledyne Ryan Aeronautical
Designation	I-GNAT	Searcher	Spectre	Ranger	DeskStar	Yak-061 (Shimel)	Global Hawk
Country	USA	Israel	England	Switzerland	USA	Russia	USA
Mission	Recon. / surv.	Recon. / surv / target acq.	Surveillance, EW	Recon. /surv. / target acq.	Acq. / recon / surv.	Recon.	Recon.
Overall Length (ft)	20.8	16.7	8.9	15.1	15	9.1	44.4
Maximum Overall Span (ft)	42.2	23.7	10.7	18.7	65	10.6	116.2
Launch Weight (Less Booster) (lbs)	1,400	818	—	605	8600	285	25,600
Power Plant	(1) Rotax 914 gasoline	(1) Limbach 550 47 hp.	(1) TTL WAE 342	(1) Goebler-Hirt 42 hp. 2-stroke	(1) Williams Rolls FJ44 tf	(1) Kuznetsov P-032 re	(1) All. AE3007H tf
Payload	EO, IR, SAR	TV & FLIR	CCD camera, FLIR, EW	EO / IO sensor	SAR	TV or IR	SAR / MTI, EO, IR
Speed (MPH or Mach)	161	121	77-150	60–137	288+	87	454
Endurance	40 hr. +	12 hr.	3–6 hr.	5 hr.	8 hr +	2 hr.	42 hr.
Maximum Altitude	23,000	15,000	23,000	14,765	45,000 +	9,840	65,000

TABLE 13.16 Typical Unmanned Aerial Vehicles, Drones, and Rotary Wing Aircraft (*Continued*)

Rotary-Wing Aircraft

Manufacturer	Eurocopter	Kamov	Bell Helicopter Textron	The Boeing Co.	Finmeccanica	MIL	Bell Boeing
Name	AS365 N4	Ka-29	OH-58D	AH-64D	A-129	MI-28N	V-22
Gross Weight (lbs)	10,580	24,230	5,500	17,650	11,243	25,300	57,000
Empty Weight (lbs)	5,564	—	3,300	11,150	6,989	—	33,140
Rotor Diameter (ft)	41.3	52.1	35	48	39	56.4	38
Max. Length Unfolder Blades (ft)	41.7	52.1	40.7	58.1	47.7	67.3	62.3
Plower Plant	(2) Tur. Amiel 2C ts	(2) TV3-117VK	(1) All. 250-C3OR	(2) GE T700-GE-701	(2) LHTEC T800 ts	(2) TV3-117VMA	(2) All. T406-AD-400
Maximum Speed (Mach or MPH)	190.00	160	150	226	185	200	345
Hover Ceiling	12,480	14,100	10,000	14,300	13,600	4,593	18,500
Still-Air Range (nm)	—	255	256	290	305	270	2,400
No. of Passengers	13	16	—	—	—	—	24

Manufacturer	The Boeing Co.	Eurocopter	MIL	EH Industries Ltd.	Hindustan Aeronautics Ltd.	Mitsubishi Heavy Industries	United Technologies Corp.
Name	CH-47D	AS332L1	MI-26TC	EH101	ALH	MH-2000	UH-6OL
Gross Weight (lbs)	50,000	18,960	123,000	32,188	9,900	9,920	16,994
Empty Weight (lbs)	23,455	9,920	62,040	19,861	5,500	5,510	11,284
Rotor Diameter (ft)	60	51.2	105	61	43.3	40	53.7
Max. Length Unfolder Blades (ft)	99	61.4	110.7	74.8	52.1	45.9	64.8
Plower Plant	(2) ASE T55-L-712 ts	(2) Tur. Makila 1A1	(2) Progress D-136	(3) GE CT7-6A	(2) Tur TM 333 of LHTEC T-800s	(2) MG5-100 ts	(2) GE T700-Ge-700 / 701C ts
Maximum Speed (Mach or MPH)	187	173	155	173	180	170	184
Hover Ceiling	10,100	10,663	4,922	12,500	9,850	7,700	9,500
Still-Air Range (nm)	1,450	516	416	864	460	435	373
No. of Passengers	33	24	82	30	12	8	11

Source: *Aerospace Source Book* 1999.

TABLE 13.17 Some Typical Potential Missions for UAVs.

Small-area surveillance	Reconnaissance
Large-area surveillance	Command and control
Strike	Deception
Aerial spraying	Electronic intelligence
Communications relay	Jamming
Atmospheric sampling	Air-to-air combat
Monitoring ground sensors	Electronic relay
Aircraft research	Systems research

W = weight (lb)
$_e$ = empty weight (lb)
$_f$ = fuel weight (lb)
$_c$ = crew weight (lb)

Empty weight is composed of weight for the necessary structure, operating systems, ballast, and propulsion systems. The gas and air in the buoyancy system behave according to Boyle's law and Charles' law. For airships with partially inflated gas bags, ascent causes gas expansion at constant lift until the gas totally fills the available volume. That altitude is called pressure altitude, and further ascent will require venting of the excess gas volume caused by the increased altitude (see Figure 13.52). This venting will reduce the available lift. Operating altitude is reached when the available lift is equal to the total weight of the airship.

Nonrigid Airship. A nonrigid airship is primarily a balloon whose shape is maintained by internal gas pressure. Elongated shapes are used to minimize air resistance in forward flight, and fins, if used, provide directional and longitudinal stability and control. Ventral gondolas are supported by a system of gas bag internal wires which distribute the gondola weight along the dorsal part of the gas bag. Fore and aft air-filled ballonets on the lower side of the hull are pressurized to retain external shape at sea level. With increasing altitude, gas expansion causes the air-filled ballonets to be compressed and to vent, and pressure altitude is reached when all ballonets are empty of air. Longitudinal stiffeners are placed on the nose of nonrigid airships to retain nose shape at forward velocities and provide some rigidity for mooring.

Semirigid Airship. A girder-like ventral keel forms a payload and gas bag mounting structure on semirigid airships. The shape of the gas bag is maintained by its pressure. Aft-mounted fins provide directional and longitudinal stability and angular control.

Rigid Airship. The external shape consists of a series of circular frames connected by longitudinal structural members extending from nose to tail of the hull (Figure 13.53). The hull is covered by a layer of fabric material. This shape, designed to minimize drag, is filled with individual (up to 15) gas bags. The structure supports payload compartments, propulsive system, control systems, and crew areas. Internal to the structure are walkways, passages, and ladders permitting access to all critical flight components and to mooring areas.

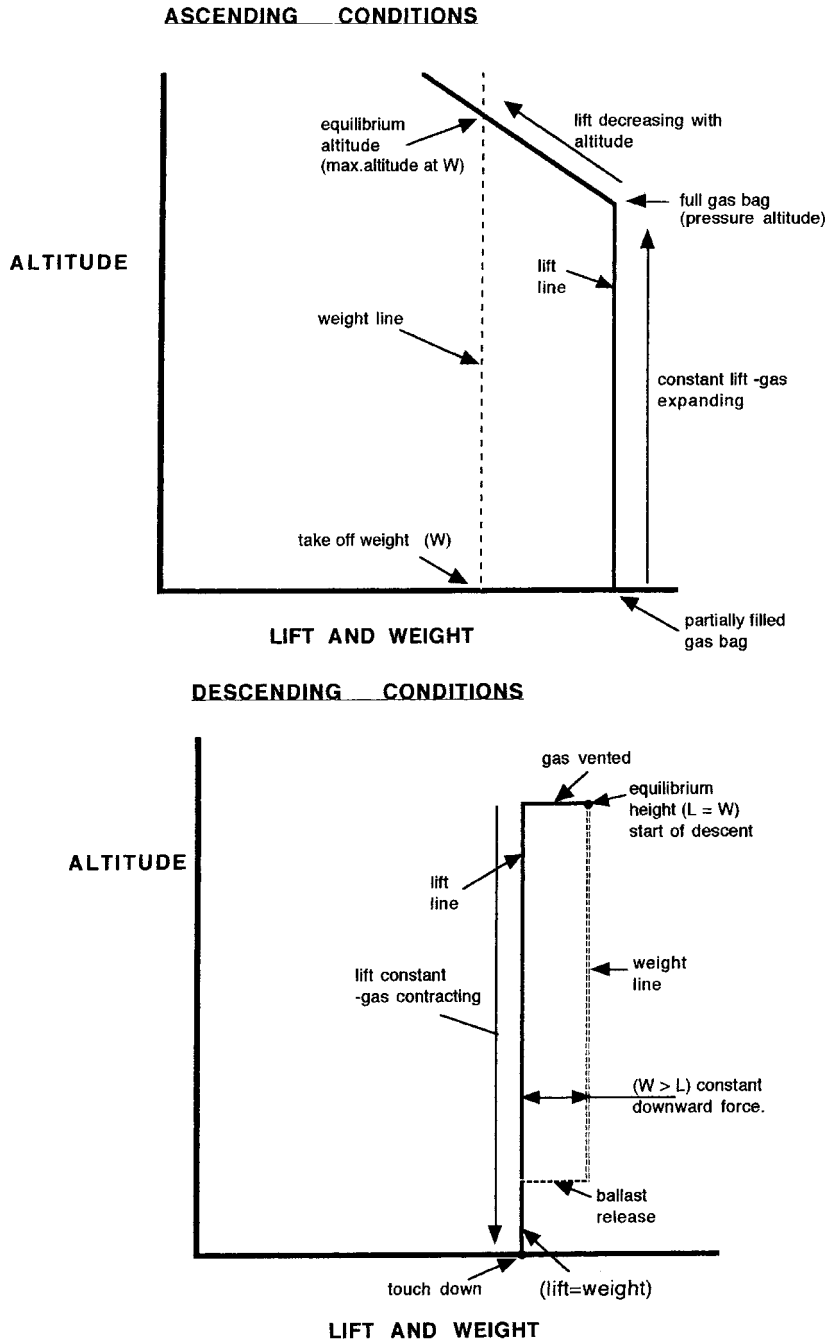


FIGURE 13.51 Ascending and descending conditions for LTA vehicles.

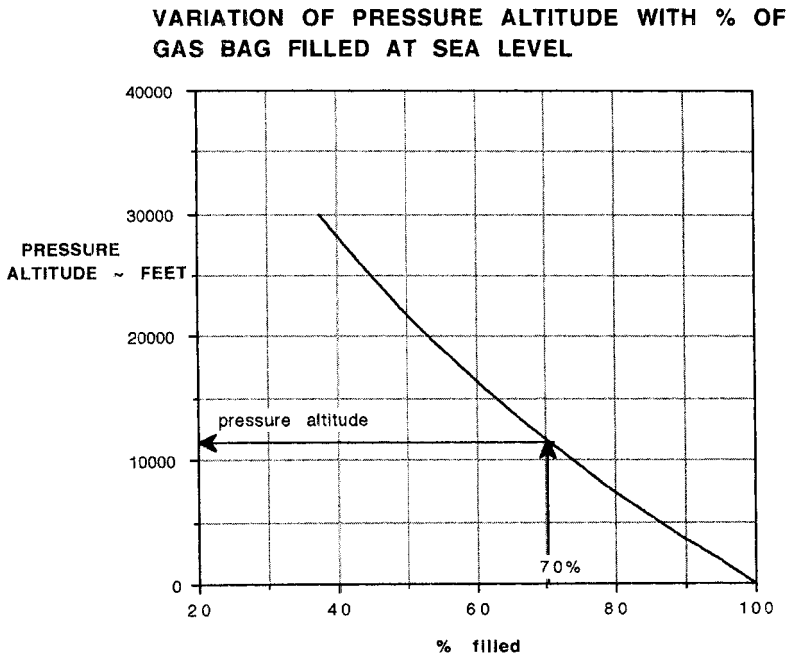


FIGURE 13.52 Pressure altitudes for LTA vehicles.

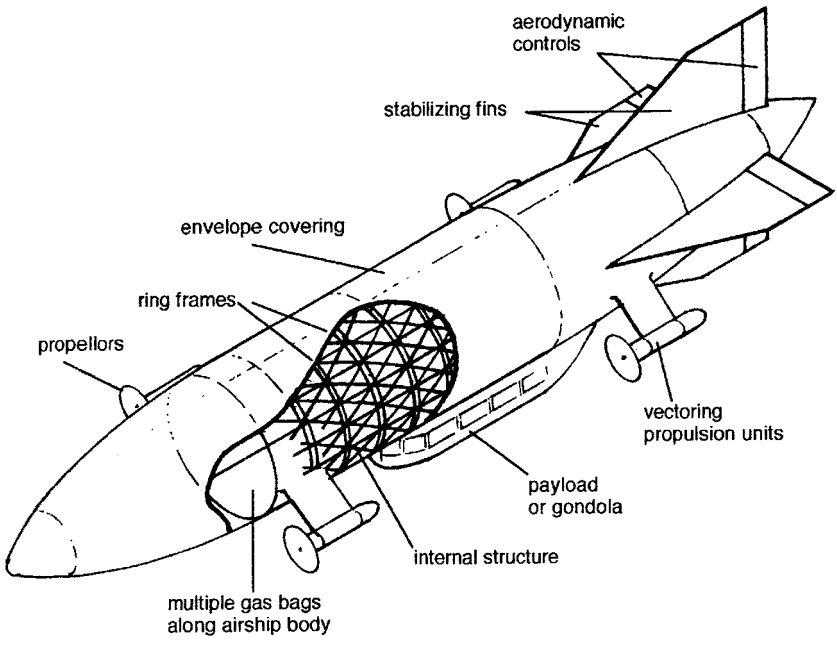


FIGURE 13.53 Cutaway of typical rigid airship.

13.12 V/STOL AIR VEHICLES

This covers a vast array of configurations, with the principle discriminators being shown in Figure 13.54. Power assist is defined as the integration of thermodynamic units with the lift system. These units can be either the main forward propulsive engines or lift-dedicated engines. Aerodynamic STOL entails the application of leading-edge and trailing-edge devices and the general use of low wing loading. Power-assist STOL and STOVL concepts require the discrete airframe impingement of or vectoring of the main propulsive jet efflux (vectored thrust). Fixed-wing VTOL concepts use either vectored thrust or a combination of vectored thrust and direct-lift thrust generators. The latter can be either direct-lift fans or dedicated direct lift-engines. Rotary-wing air vehicles are designated as either helicopters or autogyros. The former use engines to drive the main lift rotor and tail rotor (Figure 13.55). The latter use a forward-thrusting engine to establish a forward velocity and generate lift by autorotation of the main lift rotor.

Too many combinations of configuration variables exist in the design of V/STOL air vehicles to cover here. The general guidelines for fixed-wing aircraft, however, still apply for their design. These are primarily minimum weight, maximum availability, and minimum acquisition and operating cost. In wing-supported conventional flight of V/STOL concepts, fixed-wing performance and design guidelines apply. In hover and in transition from hover to wing-supported flight, other performance and design requirements need to be met for V/STOL air vehicles.

To provide for vertical linear excursions in hover, it is necessary that the vertical component of thrust, T, be greater than the air vehicle operating weight, W, with T/W on the order of 1.1 as a typical value. Linear and angular control about all three axes is required in hover, with angular control requirements being quoted in terms of angular accelerations at time t = zero. These control requirements are mission dependent and are usually specified by the procuring agency.

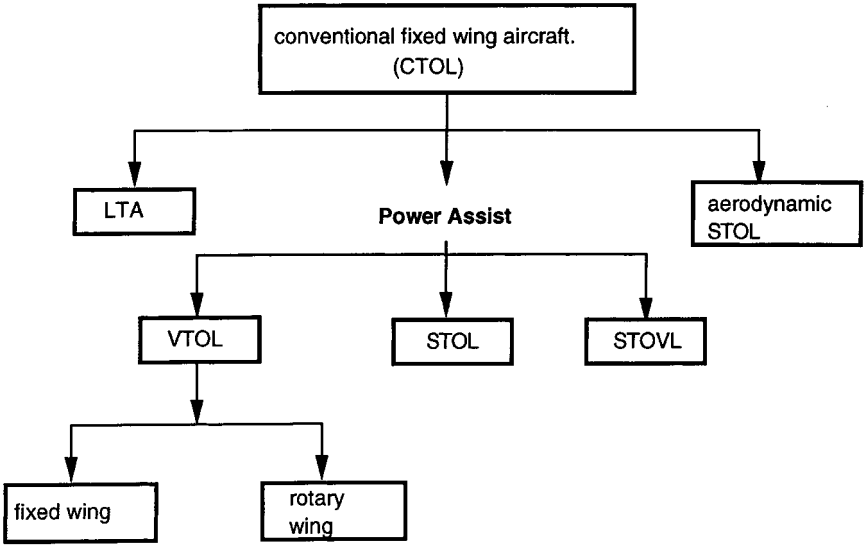


FIGURE 13.54 CTOL–STOL–VTOL relationships.

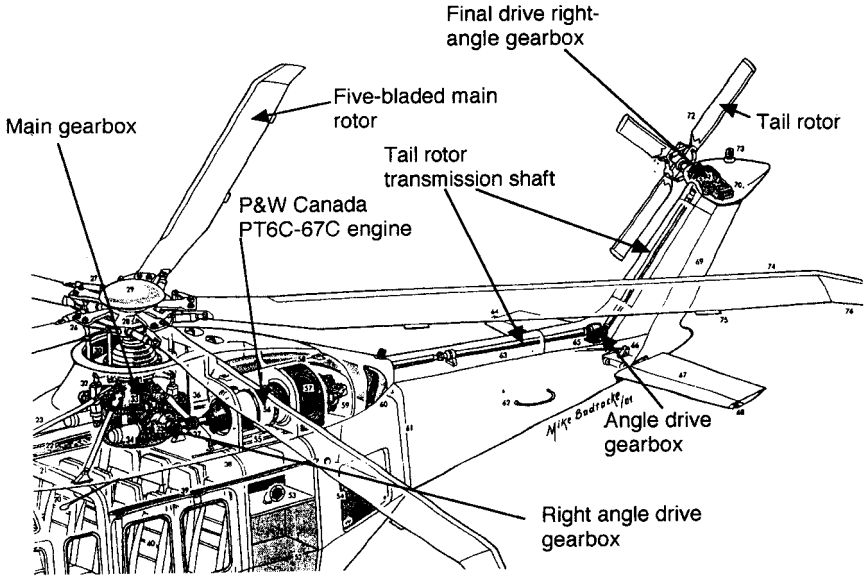


FIGURE 13.55 Typical helicopter power train (after Mike Badrocke, *Air International*).

Transition is the stage of flight during which the primary lift is gradually transferred from the hover type to the wing-supported type. (Control methods must also be similarly transitioned.) Transition can occur in a climbing, descending, or level flight trajectory. In a level flight transition a primary design goal is to ensure that lift equals weight throughout. This can be achieved by controlled or programmed variations in angle of attack and in pitch angle throughout the transition. Such variations, coupled with corresponding propulsive lift changes, result in what is known as a transition corridor. There are many likely transition corridors for each air vehicle. It is usual to minimize transition time.

Helicopters use the main rotor to generate lift, forward velocity, pitch, lateral, and roll control. These quantities are varied through the use of cyclic pitch control of the rotor and the use of flapping hinges and lag mechanisms on the rotor blades. The tail rotor generates directional or yaw control (Figure 13.56). Multiple rotors are possible, and multiple blades are possible on each rotor (Figure 13.57), with blades in excess of four per rotor being generally reserved for heavy-lift vehicles. Performance capability depends upon the relative magnitudes of drag and power (or thrust) available from the main rotor (Figure 13.58).

The total power required is the sum of the required induced, profile, and parasite powers. Induced power is referred to as the power required to produce lift; profile power is that required to drag the lift-producing rotor blades through the air; parasite power is that necessary to drag the fuselage through the air in forward flight. In hover, at zero forward velocity, induced and profile power requirements predominate, and, as shown in Figure 13.58, in forward flight parasite power increases. The maximum horizontal speed is defined where the power available equals the total power required.

At lower speeds the excess power can be used to establish horizontal acceleration, establish a rate of climb, or set up a combination of both. Figure 13.59 shows

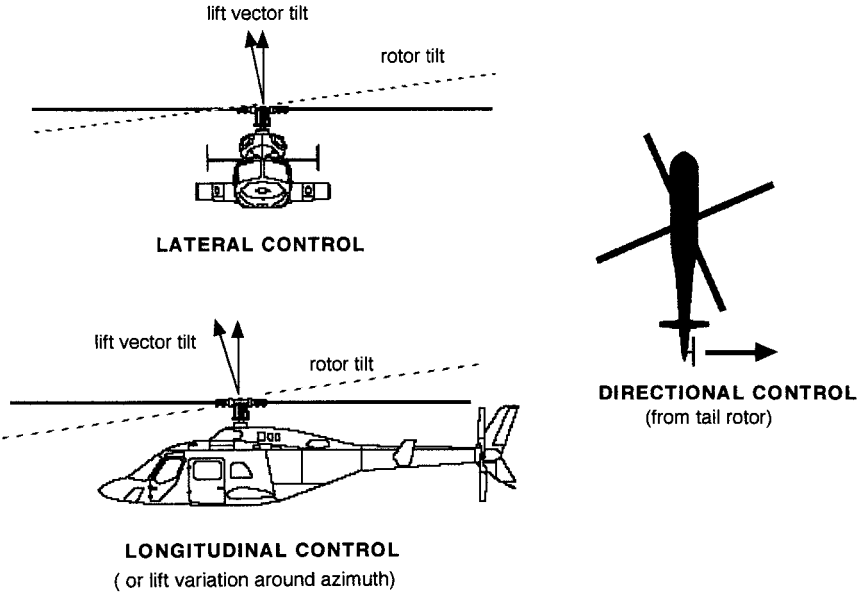


FIGURE 13.56 Typical helicopter control methods.

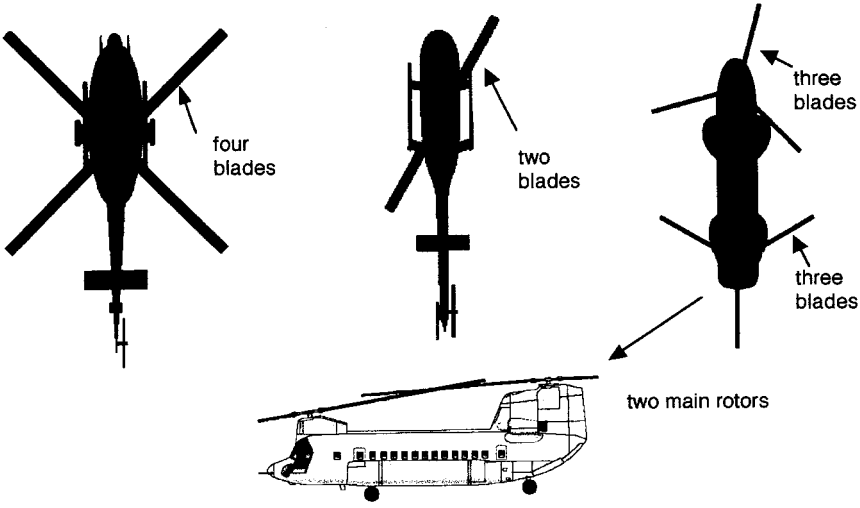


FIGURE 13.57 Typical rotor blade configurations.

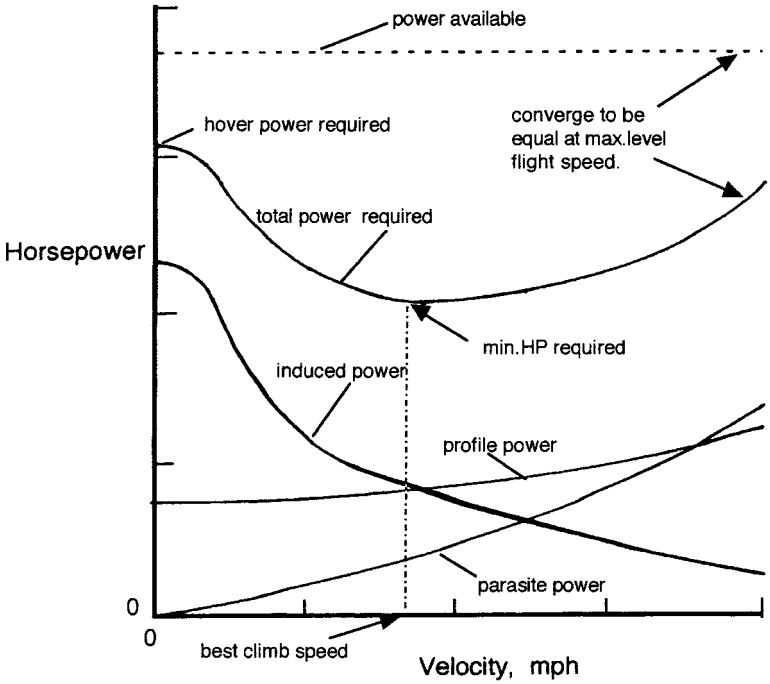


FIGURE 13.58 Helicopter power requirements.

a typical rate-of-climb variation with speed for a helicopter. Speeds for maximum endurance and maximum range can be established through the identification of minimum power-required and the tangential point on the power required curve, respectively, as also shown in Figure 13.59.

A criterion for establishing rotor hovering efficiency can be established through the use of the rotor figure of merit, M. Dimensionally,

$$M = \left[\frac{1}{\sqrt{2}}\right] \cdot \left\{\frac{T}{P}\right\} \cdot \sqrt{\frac{T}{\rho\ \pi\ R^2}}$$

where T = hover thrust (lb)
P = horsepower
R = rotor diameter, (ft)
ρ = air density (slugs/ft^3)

For sea-level conditions, this reduces to

$$P.L. = 37.93 \left[\frac{M}{\sqrt{D.L.}}\right]$$

where $P.L.$ = power loading (lb/horsepower)
$D.L.$ = disk loading (psf)

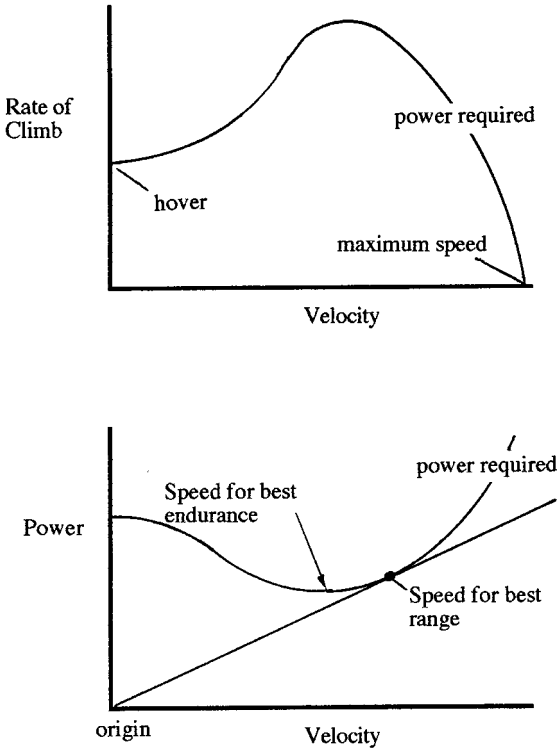


FIGURE 13.59 Some performance aspects of helicopters.

An ideal rotor will have $M = 1.00$. This is a rotor with zero profile drag power, minimum induced power (constant inflow), and zero rotational and tip losses. Such a rotor represents an upper limit to the hovering performance of any helicopter. Figure 13.60. shows the variation of *D.L.* and *P.L.* for a range of *M*.

The gross takeoff weight, W_0, of helicopters can be defined as follows:

$$W_0 = W_c + W_{pl} + W_F + W_S$$

where subscript c = crew weight
subscript *p.l.* = payload weight
subscript F = fuel weight
subscript S = structure weight

$$\text{Power required, } HP, = HP_o + HP_i + HP_p + HP_{acc} + HP_{tr}$$

where subscript o = profile power
subscript i = induced power
subscript p = parasite power
subscript *acc* = accessory power

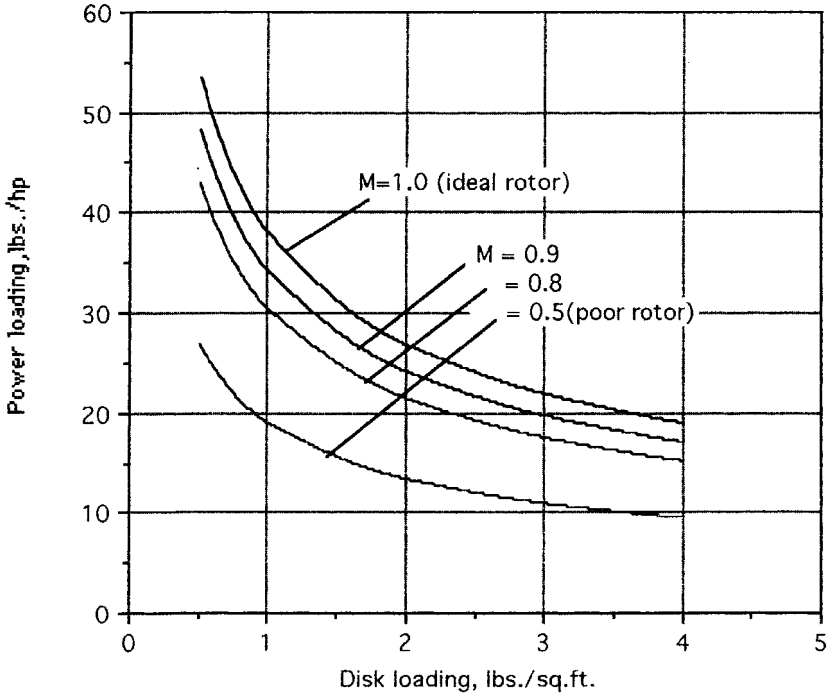


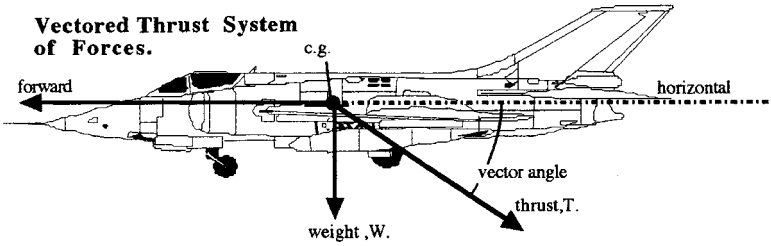
FIGURE 13.60 Power loading and disk loading relationships.

subscript tr = tail rotor power

If special on-board mission oriented payloads are required, then an additional power requirement can be added.

Helicopter design then centers on the establishment of suitable inputs to these two equations to reflect current technology. The main driver in the design of the helicopter is the rotor aerodynamics, structure, mechanics and hub. Typical helicopter design data are shown in Table 13.16.

For pure vectored thrust V/STOL concepts, without exhaust impingement on the airframe, it is possible to shorten the takeoff ground roll considerably by distributing the total thrust between a lift component and a forward air vehicle accelerating component. Figure 13.61 shows the assumed system of forces acting on an aircraft with thrust deflected through an angle measured from the horizontal. If ground rolling friction and aerodynamic drag are ignored, a good approximation can be obtained of the effectiveness of vectored thrust in reducing ground roll in a rolling takeoff. The equations of Figure 13.61, developed from the force diagram, are graphed in Figure 13.62 which illustrates the effectiveness of vectoring with a constant vector angle. The optimum relationship between T/W and thrust vector angle is also shown.



$$\left(\frac{S_{g\theta}}{S_g}\right) = \left(\frac{1}{\cos\theta_T}\right) - \left(\frac{T}{W}\right) Tan\,\theta_T$$

$$\left(\frac{S_{g\theta}}{S_g}\right)_{opt.} = \sqrt{1-\left(T/W\right)^2} = Cos\,\theta_T$$

where θ_T = **thrust deflection angle measured from the horizontal,**

$\left(\frac{T}{W}\right)$ = **thrust to weight ratio,**

S = **ground roll dist., subscript,** θ **implies dist. with thrust deflected through angle** θ_T

S_g = **dist. with thrust undeflected**

FIGURE 13.61 Vectored thrust system of forces.

13.13 PERFORMANCE

Axes System

For the purposes of analysis and design, a system of orthogonal axes is assumed to exist in the airframe (Figure 13.63). Aircraft rotational motions are then measured relative to these axes and are termed *longitudinal,* about the *y*-axis; *directional,* about the *z*-axis; and *rolling,* about the *x*-axis. Performance measures are usually established in the longitudinal plane, which involves motion fore and aft, along the *x*-axis, and in the pitching plane, that is, nose-up and nose-down rotation about the *y*-axis. Under these rules it is usual to establish a force diagram (Figure 13.64) from which a series of equations can be set up to establish performance capability.

For equilibrium in straight flight the following equations can be established from Figure 13.64:

$F_x = T\cos\alpha - D - W\sin\theta$, force parallel to the flight path

$F_z = W\cos\theta - L - T\sin\alpha$, force perpendicular to the flight path

which for steady, level, unaccelerated flight, when a small angle of attack is also assumed, becomes

AERONAUTICAL DESIGN

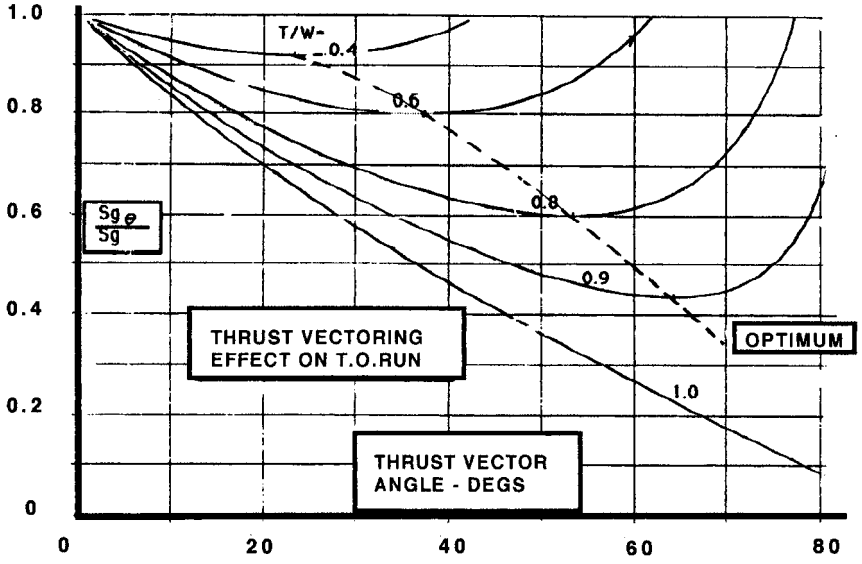
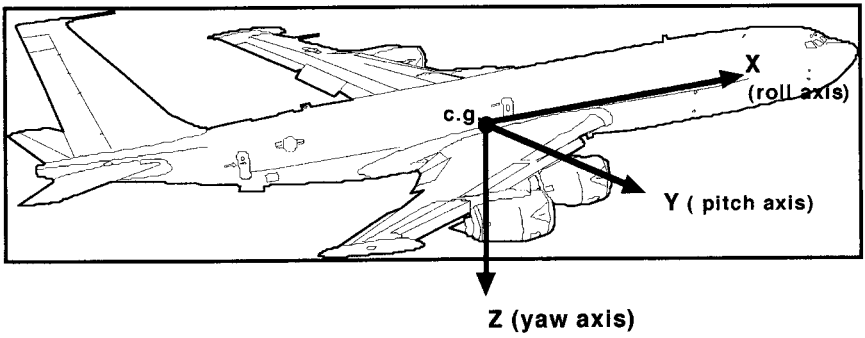


FIGURE 13.62 Takeoff ground run with vectored thrust.



Positive linear directions shown

Positive angular directions are
- **nose up**
- **nose right**
- **right wing down**

FIGURE 13.63 Axes system.

$$T = D \quad \text{and} \quad L = W$$

where n = load factor
θ = climb angle
α = angle of attack

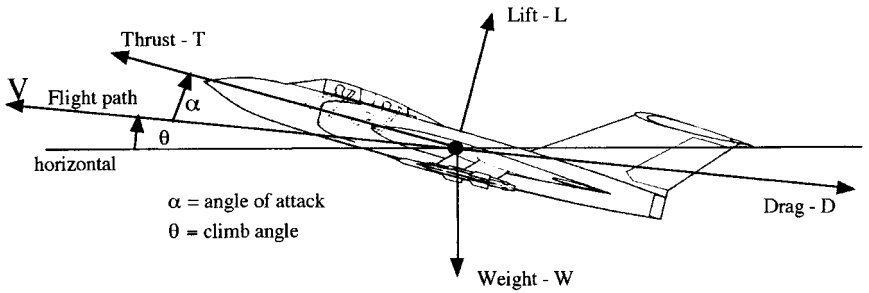


FIGURE 13.64 Force diagram.

Drag

For the complete aircraft, drag is composed of three major components: profile, induced, and wave drag. Profile is made up of skin friction drag and form or pressure drag. Induced drag is generated by the backward inclination of the lifting surface vector, caused by tip vortices. Lifting surface definition includes wing and empennage surfaces; the latter drag is sometimes referred to as trim drag. Wave drag is caused by the formation of shock waves and is sometimes referred to as compressibility drag.

For subsonic aircraft, the total drag coefficient, C_D, is given by

$$C_D = \frac{\text{Profile drag}}{q \cdot s} + \frac{\text{Induced drag}}{q \cdot s}$$

or

$$C_D = C_{Do} + \frac{C_L^2}{\pi A \cdot e}$$

where C_{Do} is the profile drag coefficient and C_{Di} is the induced drag coefficient. Figure 13.65 shows a typical variation of total drag with airspeed. Minimizing profile drag is a key factor in establishing maximum performance. Figure 13.66 shows typical values of C_{Do} for a range of aircraft, and these remain essentially invariant with Mach number until the formation of shock waves (near mach number = 1.0), at which point C_{Do} begins to increase. Delaying this increase to the highest possible Mach number is a good design aim. The onset of this increase in drag is a function of the cross-sectional area distribution of the complete aircraft. The addition of external stores to military aircraft will increase C_{Do}.

Figure 13.67 shows how the total drag will vary with altitude, and, as can be seen, minimum drag magnitude does not vary with altitude, but the speed for minimum drag increases with increasing altitude.

Thrust

Thrust can come from a jet engine or from a propeller. In each case thrust will vary with altitude and airspeed. For jet engines the aircraft basic performance capabilities can be assessed by assuming thrust is invariant with velocity and varies

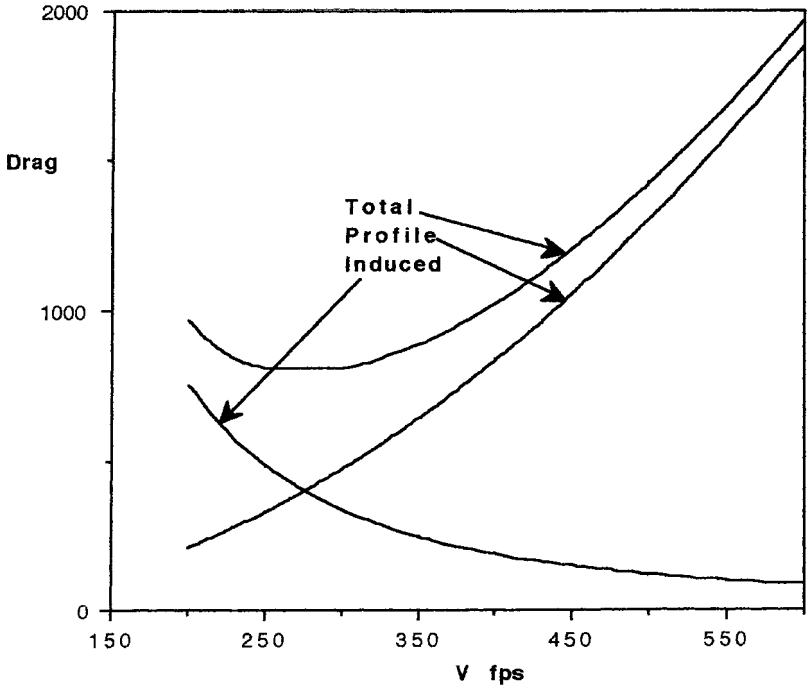


FIGURE 13.65 Aircraft drag elements.

only with altitude. (Graphical methods can be used to evaluate performance for engines with significant thrust–speed dependency). Figure 13.68 shows a typical variation of thrust with altitude for a high bypass ratio engine.

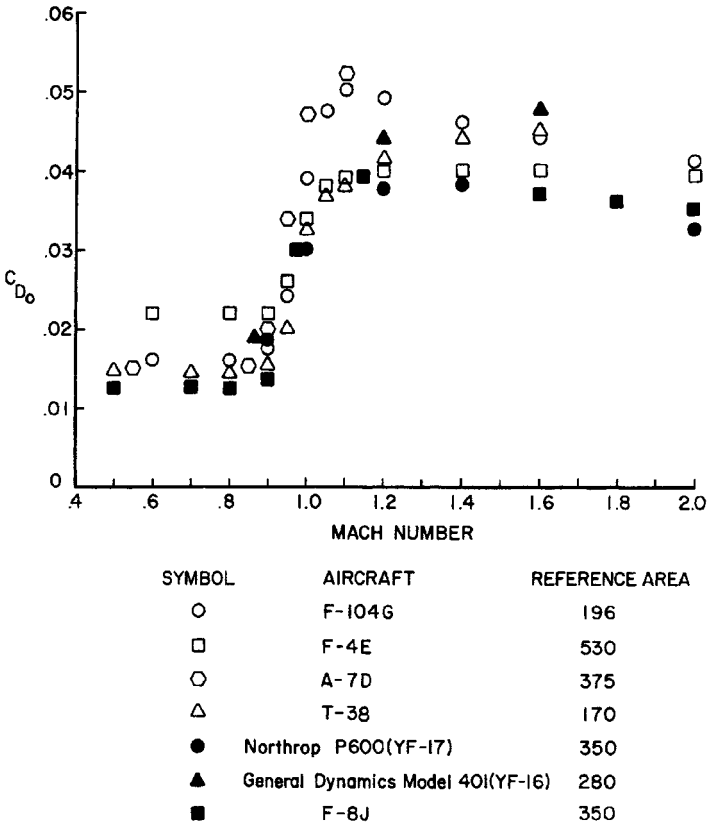
For propeller-driven aircraft, power can also be assumed to be independent of forward velocity with the same caveat for power–speed dependency as for jet engines.

Airplane Performance

The relative interplay between thrust available and drag constitutes the available performance of an aircraft. Figure 13.69 illustrates a performance estimation sequence showing the two major elements, air performance and field performance. It is usual to estimate the performance of aircraft using a set of standard atmosphere tables to describe the environmental conditions at the appropriate mission segment. This permits performance comparisons between concepts to be made effectively.

Steady flight or unaccelerated level flight conditions are usually evaluated, and under these conditions the performance is said to be static performance. For some high-performance aircraft, such as modern-day fighters, accelerated flight conditions are taken into account in estimating the maneuverability capability of the concept.

Both the thrust available and drag vary with altitude and Mach number. Therefore, the performance capability of a concept becomes a function of speed and



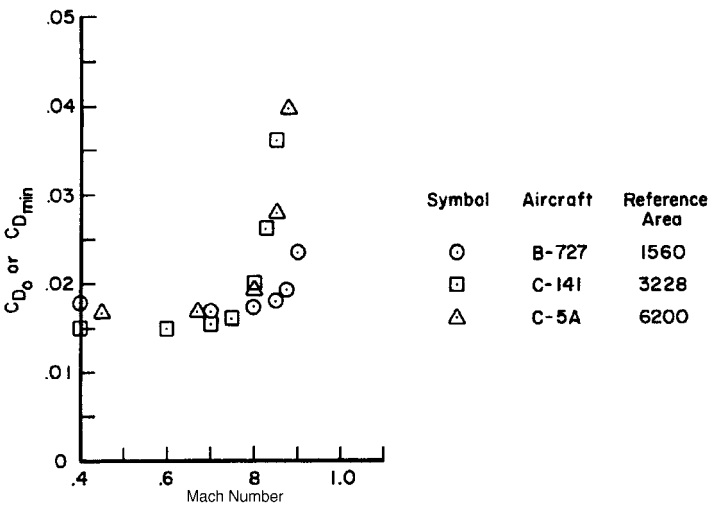


FIGURE 13.66 Measured profile drag coefficients. (from Nicolai 1984).

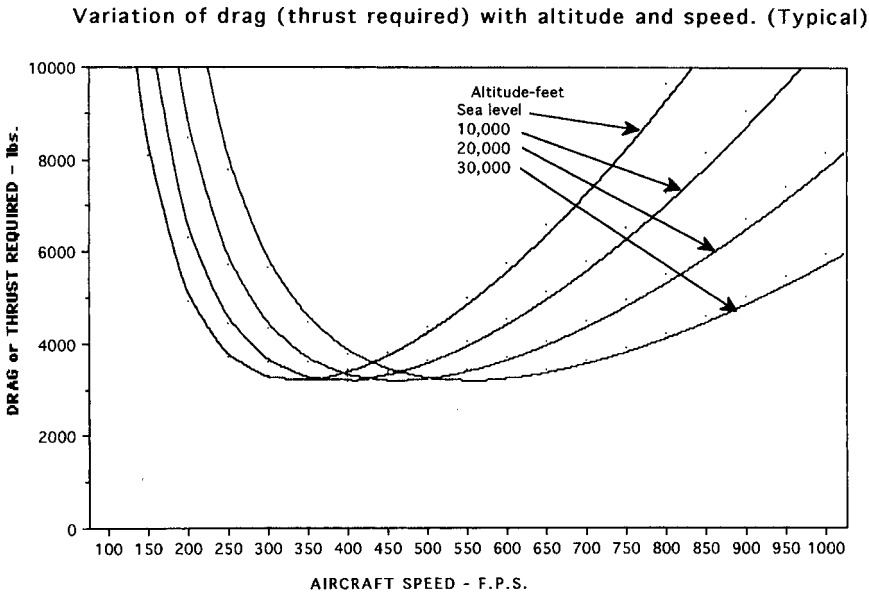


FIGURE 13.67 Typical drag variation with altitude.

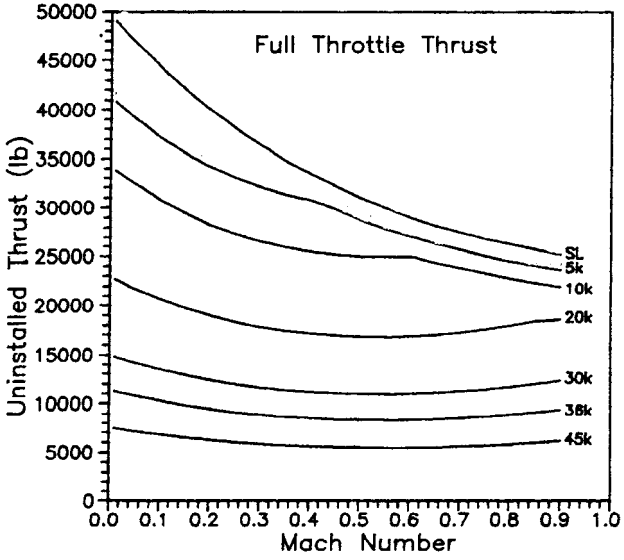


FIGURE 13.68 Typical thrust variation with altitude (from Raymer 1989).

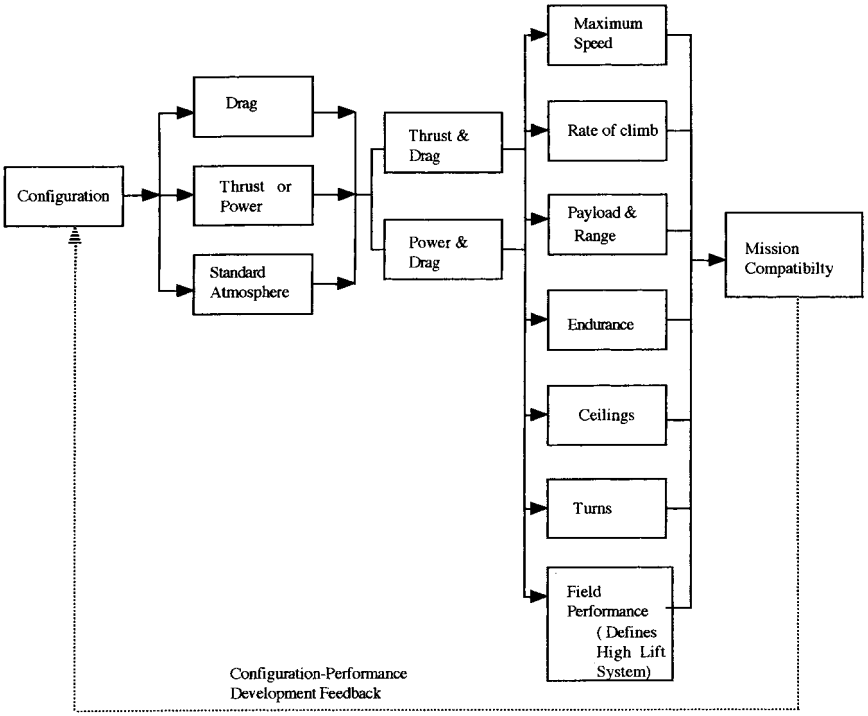


FIGURE 13.69 Overview of performance estimation sequence.

altitude. The thrust variation is a function of the engine type and depends on engine bypass ratio, whereas the drag variation is a function of the aircraft shape.

An overview of performance capability can be obtained from a graphical representation of thrust and drag at a given altitude, (Figure 13.70). Here maximum and minimum speeds occur at $(T - D) = 0$, minimum drag speed is identified, and at speeds where $(T - D) > 0$, excess power is available for use in climb or in level flight acceleration. As altitude increases overall, $T - D$ over the speed range decreases (Figure 13.71), eventually reaching zero or less than zero for all speeds. Absolute ceiling is reached when $(T - D) = 0$ and theoretically the aircraft can fly at one speed only, which is that speed where the thrust available curve is tangential to the drag curve. Service ceilings are a more practical measure of operating altitudes, and the rate-of-climb conditions for these are tabulated in this design section for commercial and military aircraft.

Much simplification in performance estimations can be achieved by assuming that the drag polar is parabolic. This results in the development of many performance equations which eloquently display the influence of many design parameters. The parabolic assumption is made for the performance methods presented herein.

***Rate of-Climb* (*R*/*C*).** Defined as $V(T - D)/W$, the rate of climb will vary with altitude (see Figure 13.72), with a maximum value occurring for each altitude at a speed given by

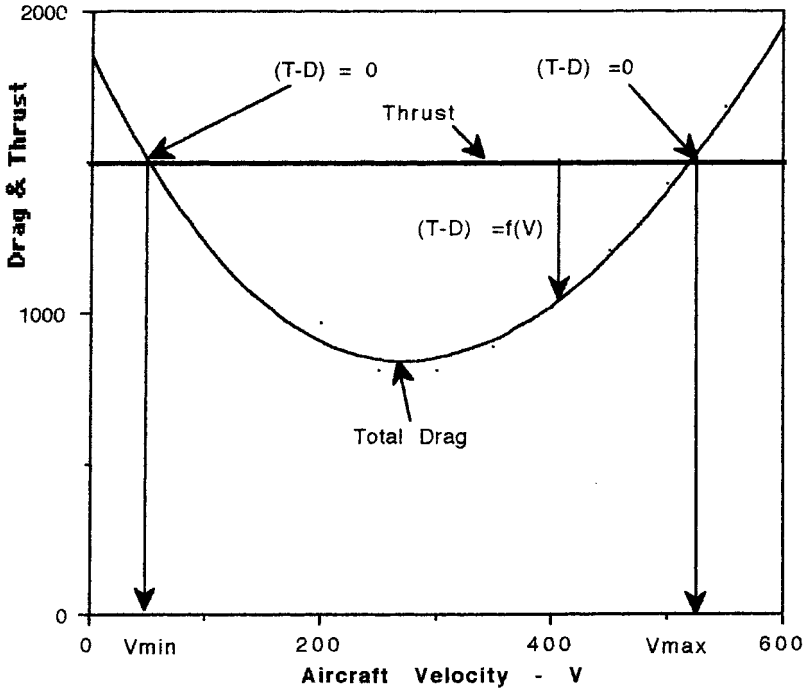


FIGURE 13.70 Typical variation of thrust and drag with velocity.

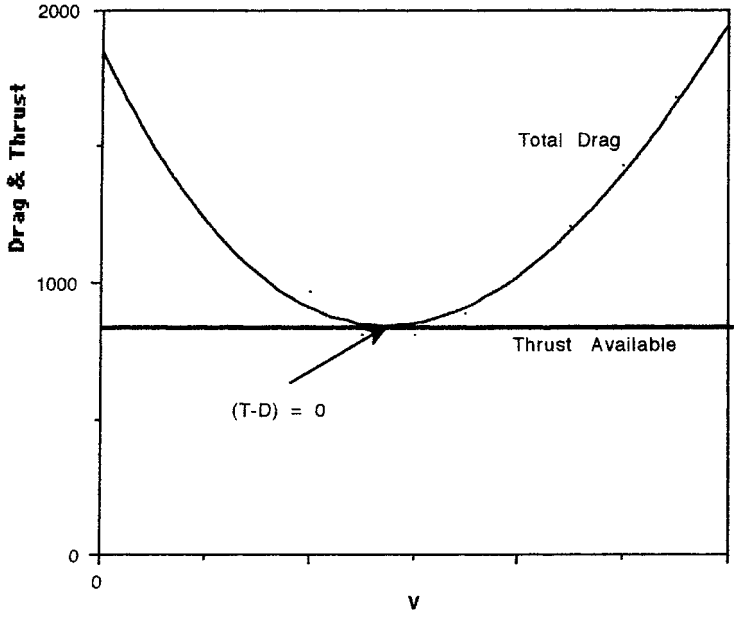


FIGURE 13.71 Typical variation of thrust and drag relationships at ceiling.

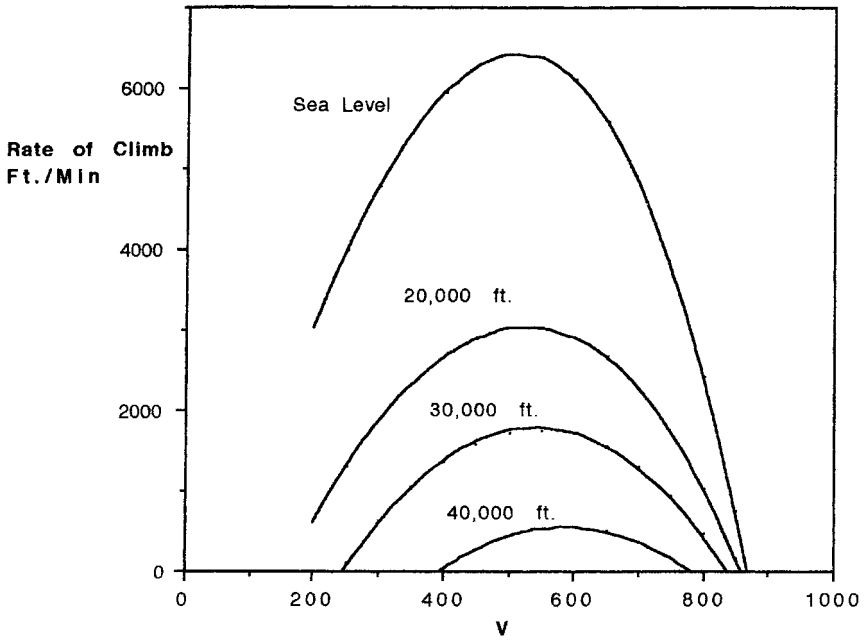


FIGURE 13.72 Typical rate of climb versus altitude.

$$V = \left\{ \frac{(T/W)\left(\frac{W}{S}\right) \pm \sqrt{\left(\frac{T}{W}\right)^2 \left(\frac{W}{S}\right)^2 + 12\left(\frac{W}{S}\right)^2 \left(\frac{C_{Do}}{\pi Ae}\right)}}{3C_{Do}\rho} \right\}^{0.5} \text{(fps)}$$

The climb angle corresponding to maximum rate of climb, $(R/C)_{\max}$, is given by

$$\Theta = \sin^{-1}\left(\frac{(R/C)_{\max}}{V}\right) \text{(degrees)}$$

The maximum climb angle is given by

$$\Theta_{\max} = \sin^{-1}\left(\frac{T}{W} - \frac{1}{L/D}\right) \text{(degrees)}$$

at a speed of

$$V = \sqrt{\frac{2}{\rho}\left(\frac{1}{C_{Do}\pi Ae}\right)^{0.5} \frac{W}{S} \cos \Theta_{\max}} \text{ (fps)}$$

All of the above are summarized in the climb hodograph of Figure 13.73.

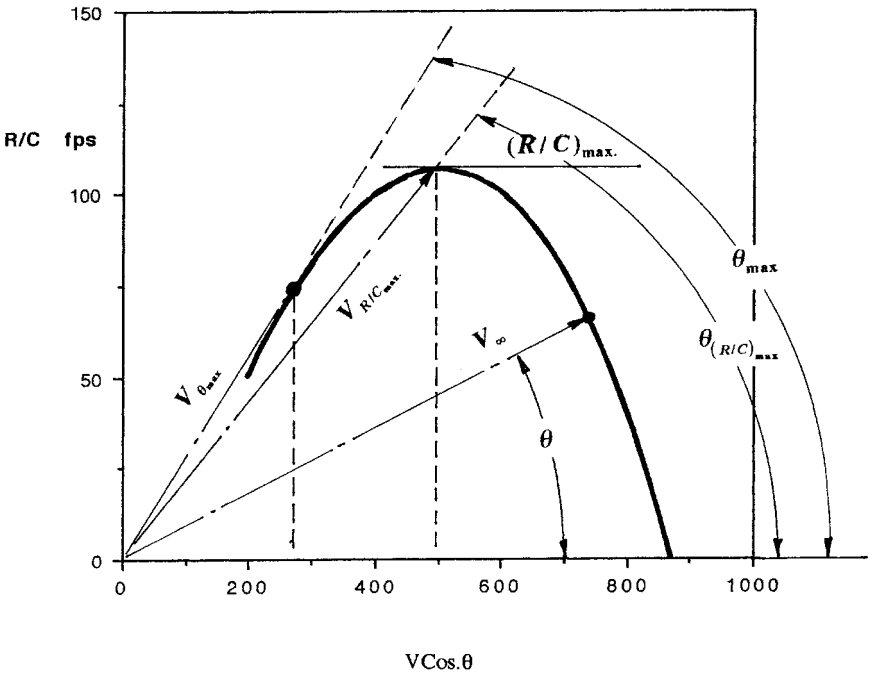


FIGURE 13.73 Climb hodograph.

Gliding Flight. In power-off flight the aircraft will glide and descend (Figure 13.74). The maximum straight-line ground range in a glide from height h is given by

$$\text{Max. horizontal distance, } l, = h\left(\frac{L}{D}\right)_{\max}$$

and it will occur at an aircraft speed given by

$$V = \sqrt{\frac{W}{S}\left(\frac{2}{\rho}\right)\frac{1}{\sqrt{C_{Do}\ \pi Ae}}}\ \text{(fps)}$$

and at a descent angle given by

$$\Theta_{\min} = \tan^{-1}\left(\frac{1}{(L/D)_{\max}}\right)\ \text{(degrees)}$$

Minimum sink speed (vertical velocity) in a glide is given by

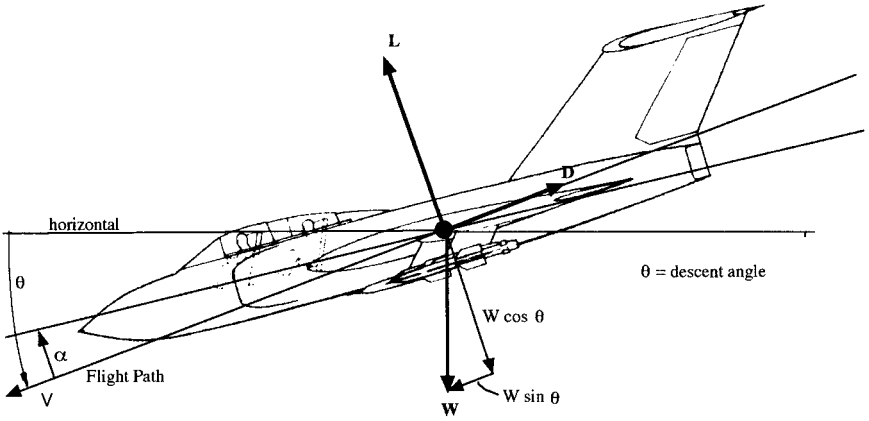


FIGURE 13.74 Forces in power-off flight.

$$V_v = \sqrt{\frac{2}{\rho((C_L)^3/(C_D)^2)}\left(\frac{W}{S}\right)} \text{ (fps)}$$

Figure 13.75 summarizes these glide equations on a glide hodograph.

Maximum Range and Maximum Endurance. The flight conditions for these performance parameters are specified by aircraft speeds; that is, a speed for maximum endurance and a speed for maximum range. Maximum endurance occurs at a flight condition appropriate to the longest period of *time* spent in the air for a given

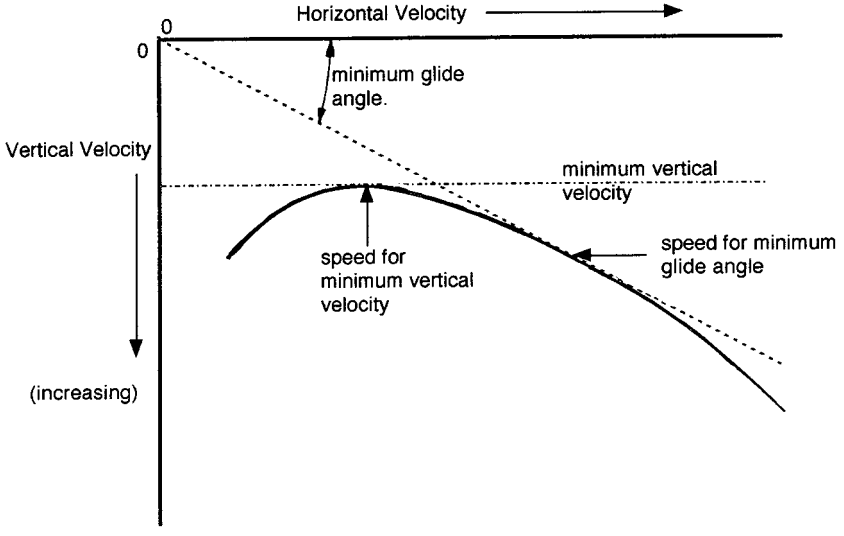


FIGURE 13.75 Glide hodograph.

quantity of fuel, and maximum range corresponds to the maximum *distance* traveled on a given quantity of fuel. Table 13.18 summarizes the flight conditions for propeller and jet aircraft, and Figure 13.76 compares the maximum range and endurance speeds for two identical aircraft, one powered by a jet and the other by a reciprocating engine propeller combination. The maxima on each curve of the various aerodynamic parameters indicate "best" or maximum conditions. Table 13.19 summarizes the Breguet range and endurance formulae for propeller and jet aircraft.

Table 13.6 summarizes the analytical expressions for the maximum range and maximum endurance speeds for both jet-propelled and propeller-driven aircraft based on an assumption that the drag polar can be assumed parabolic in nature. Also shown are the corresponding C_L, C_D, and L/D values.

Figure 13.77 summarizes a typical variation of the above speeds with altitude for a jet-propelled aircraft and includes the stall speed based upon the following equation:

$$V_{\text{stall}} = \sqrt{\left(\frac{W}{S}\right)\left(\frac{2}{\rho}\right)\left(\frac{1}{C_{L\max}}\right)} \text{ (fps)}$$

where ρ = air density (slugs/ft^3)
$C_{L\max}$ = maximum lift coefficient
W/S = wing loading (psf)

TABLE 13.18 Speed, Maximum Range, Maximum ENdurance Relationships

Propulsion type	Maximum range	Maximum endurance
Propeller	V such that C_L/C_D is a maximum	V such that $C_L^{3/2}/C_D$ is a maximum
Jet	V such that $C_L^{1/2}/C_D$ is a maximum	V such that C_L/C_D is a maximum

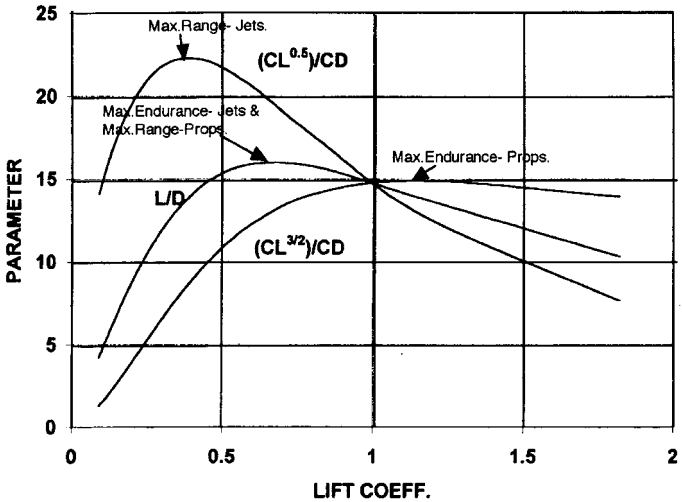


FIGURE 13.76 Performance comparisons—jet and propeller.

TABLE 13.19 Maximum Range, Maximum Endurance Relationships

Propulsion type	Maximum range (ft)	Maximum endurance (hrs)
Propeller	$R = \frac{\eta}{c}\frac{C_L}{C_D}\ln\frac{W_0}{W_1}$	$E = \frac{\eta}{c}\frac{C_L^{3/2}}{C_D}(2\rho S)^{1/2}\{W_1^{1/2} - W_0^{-1/2}\}\left(\frac{1}{3600}\right)$
Jet	$R = 2\sqrt{\frac{2}{\rho S}}\cdot\left(\frac{1}{c}\right)\cdot\left(\frac{C_L^{1/2}}{C_D}\right)\cdot\{W_0^{1/2} - W_1^{1/2}\}$ at constant altitude	$E = \left(\frac{1}{3600\ c}\right)\frac{C_L}{C_D}\ln\frac{W_0}{W_1}$

where c = specific fuel consumption is lb/lb thrust/s for jets, lb/ft lb/s/s for reciprocating engines
ρ = air density (slugs/ft^3)
S = wing area (ft^2)
W_0 = weight at start of range (endurance) segment (lb)
W_1 = weight at end of range (endurance) segment (lb)

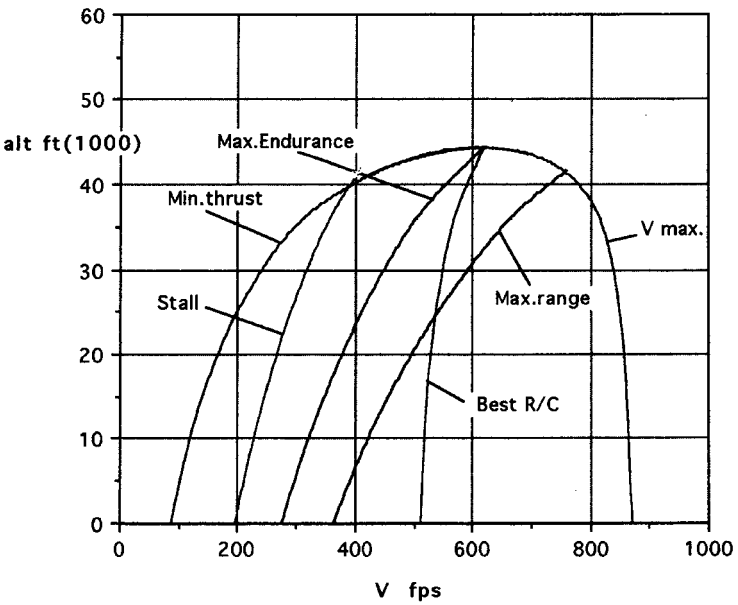


FIGURE 13.77 Typical speed-altitude limits.

Pull-up, Push-over, and Horizontal Turns. The magnitude of $T - D$ can also be used for maneuver in the form of horizontal and vertical flight path changes. In the horizontal level flight turn the airspeed and altitude are assumed to remain constant. The corresponding load factor and bank angle are defined as

$$n = \left(\frac{L}{W}\right)$$

$$\text{Bank angle, } \phi = \sec^{-1}\left(\frac{1}{n}\right)$$

All of these maneuvers are initiated from trimmed, straight, and level flight. In a pull-up the forces acting on the aircraft are shown in Figure 13.78 and the pull-up is evaluated at the instant of maneuver initiation, with the resulting performance being strictly applicable to that instant. In a pull-down the aircraft is rolled inverted and a pull-down maneuver is initiated, with the resulting performance measures being applicable to the instant that pull-down is initiated. In this case a similar system of forces acts on the aircraft with, in this instance, lift and weight acting together in a downward direction to oppose centrifugal force, which is acting upwards.

The system of forces acting in a horizontal turn is shown in Figure 13.79, and in this case these forces can be assumed to be valid throughout the turn provided sufficient additional thrust is added. In this maneuver the bank angle, φ, and V are assumed constant throughout the turn.

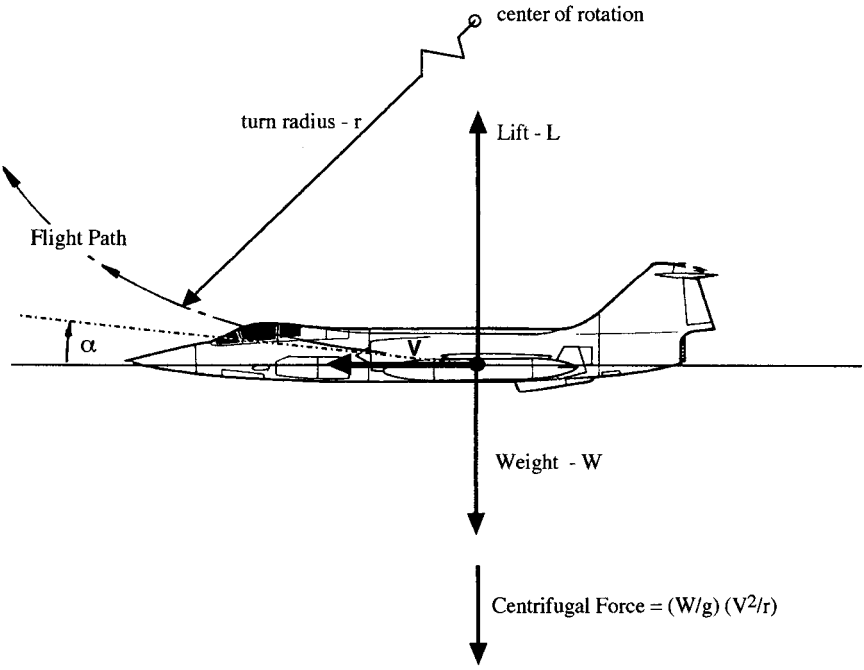


FIGURE 13.78 Forces in a pull-up.

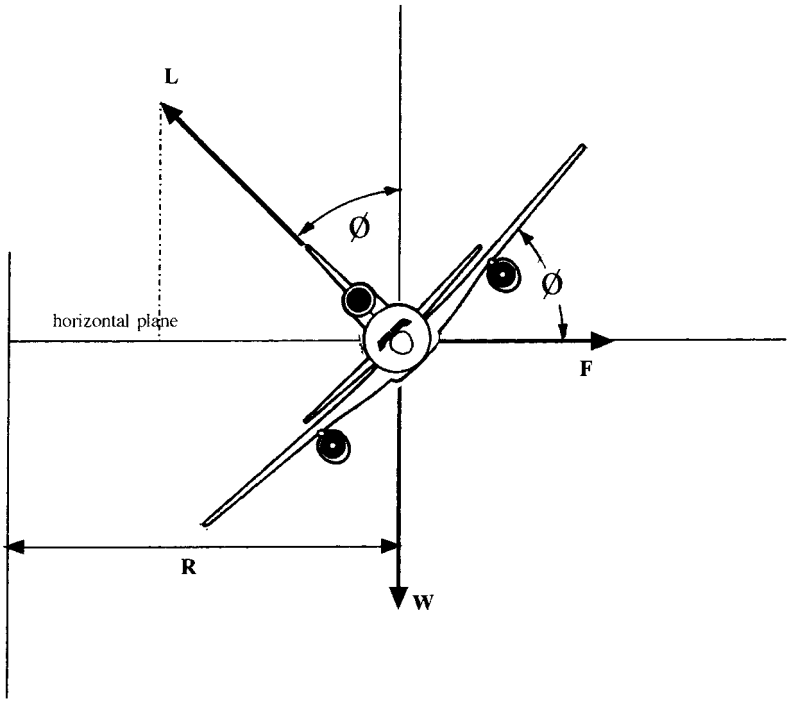


FIGURE 13.79 Forces in a horizontal turn.

All of these maneuvers depend on the application of normal load factor, n, primarily through an increase in angle of attack. The resulting turn performance measures are the available radius of turn, R ft, and resulting rate of turn, w, radians per second. The equations to evaluate pull-up, pull-down, and horizontal turns are given in Table 13.20, where the applicable load factor, n, and corresponding speed, V, are derived from the design maneuver flight envelope.

Takeoff and Landing. This aspect of performance is concerned with estimating the distance required to clear a given obstacle height (Figures 13.80 and 13.81). This distance consists of ground, transition, and air distance elements for takeoff, and an air distance and ground roll for landing (Anderson 1999).

In general, the ground roll distance, s_g, for takeoff can be shown to be approximated by the following equation, which serves to illustrate the influence of certain aircraft parameters:

$$s_g = \frac{1.21(W/S)}{g \cdot \rho \cdot C_{L\max}\left(\dfrac{T}{W} - \dfrac{D}{W}\right)} \text{ (ft)}$$

TABLE 13.20 Horizontal Turns, Pull-up and Pull-down Relationships

	Pull down	Pull up	Horizontal turn
R (ft)	$\dfrac{V^2}{g(n+1)}$	$\dfrac{V^2}{g(n-1)}$	$\dfrac{V^2}{g\sqrt{n^2-1}}$
ω (rads/s)	$\dfrac{g(n+1)}{V}$	$\dfrac{g(n-1)}{V}$	$\dfrac{g\sqrt{n^2-1}}{V}$
t (s for 180° turn)	—	—	$\dfrac{\pi}{\omega}$

$$\sec \varphi = \left(\frac{L}{W}\right) = n \text{ (load factor)}$$

φ = bank angle (deg s)

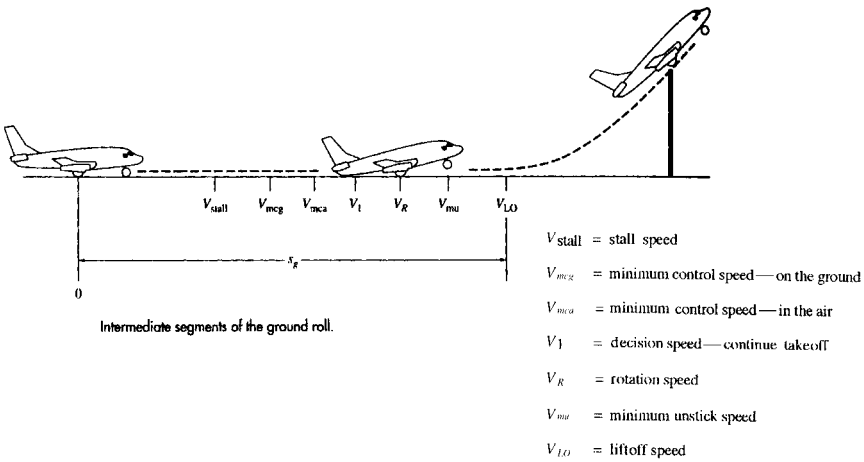


FIGURE 13.80 Interim segments of ground roll—takeoff (from Anderson 1999).

where T/W = aircraft thrust-to-weight ratio
D = aircraft total drag

and where 1.21 corresponds to the assumption that liftoff speed, V_{LO}, occurs at $(1.1\ V_s)^2$. To achieve short takeoff distances, it is necessary that high values of $C_{L\max}$ and $T - D$ be present, coupled with low values of W/S.

The following equation can be used to estimate the landing ground roll distance, s_L, for modern jet transports (Anderson 1999):

$$s_L = \frac{1.69\ W^2}{g \cdot \rho \cdot S \cdot C_{L\max}[T_R + (D + \mu_G(W - L))]} \text{ (ft)}$$

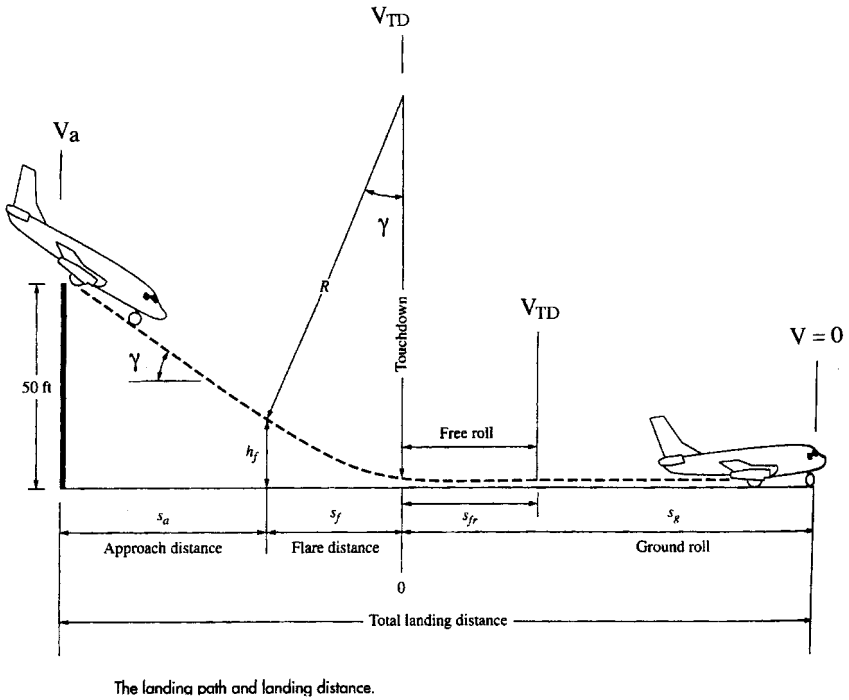


FIGURE 13.81 Interim segments of ground roll—Landing (from Anderson 1999).

where V_T is the touchdown velocity = $1.3(V_{stall})$ (fps)
g = 32.2 ft/s^2
ρ = air density (slugs/ft^3)
S = wing reference area (ft^2)
$C_{L\,max}$ = maximum lift coefficient on landing
T_R = reverse thrust (lb)
D = total drag (lb)
μ_B = braking coefficient of friction (usually = 0.40)
W = aircraft landing weight (lb)
L = residual lift on the wing (lb) with a proviso that all the terms in the [ ] brackets are evaluated at a speed equal to 0.70 (V_T)

To minimize s_L, it is necessary to increase the design terms in the denominator and reduce the landing weight. The residual lift on the wing can be minimized or made equal to zero by deploying wing-mounted upper surface spoilers at the start of the ground run on landing. This in turn reduces the magnitude of s_L.

The remaining elements of the takeoff and landing distance estimates can be established through methods developed in Anderson (1999).

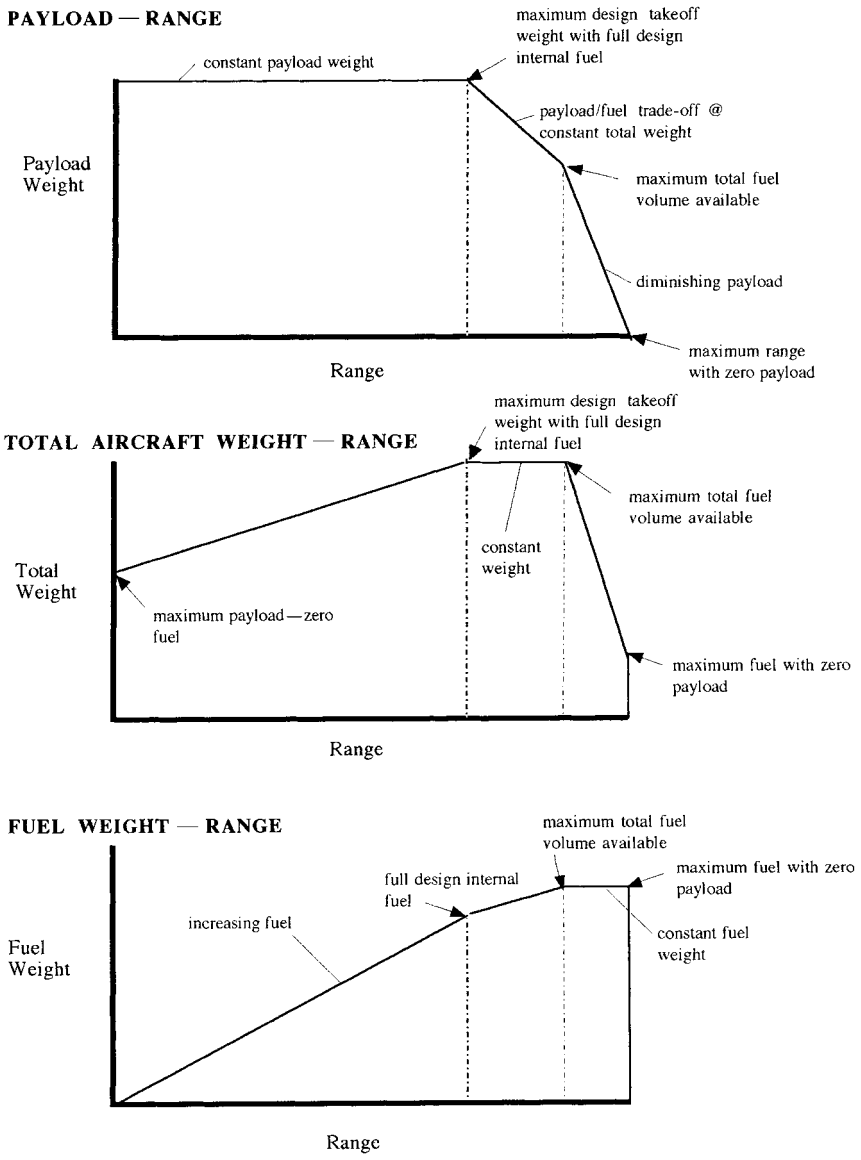


FIGURE 13.82 Definition of payload range and corresponding fuel and total weight.

Payload Range. A measure of the mission effectiveness of a cargo or passenger carrying aircraft is displayed by the payload range diagram, which shows what payload weight can be carried for what range. Figure 13.82 shows a typical payload range diagram and how an increase in range, beyond that for maximum payload, is possible by trading additional fuel for payload. The corresponding total mission weight variation with range is also shown, and this assumes that only enough fuel is loaded to achieve the chosen range. The fuel weight required to establish the corresponding range concludes the figure. In all instances additional fuel above the design full internal fuel can be accommodated by using temporary fuselage fuel cells which can be located inside the cargo space in place of the normal cargo.

REFERENCES

Aerospace Source Book. 1999. *Aviation Week & Space Technology,* McGraw-Hill, New York.

Air International. 1989–2001, Vol. 36–60, Key Publishing LTD, Stamford, Lincolnshire, UK.

Anderson, J. D., Jr. 1999. *Aircraft Performance and Design,* WCB/McGraw-Hill, Boston.

Jenkinson, D., Simpson, P., and Rhodes, D. *Civil Jet Aircraft Design,* AIAA Education Series, Washington, DC.

Loftin, L. K., Jr. 1980. *Subsonic Aircraft: Evolution and the Matching of Size to Performance,* NASA Reference Publication 1060.

Nicolai, L. M. 1984. *Fundamentals of Aircraft Design,* 1st ed., METS, San José, CA.

Raymer, D. P. 1989. *Aircraft Design—A Conceptual Approach,* AIAA Education Series, Washington, DC.

Rolls-Royce plc. 1986. *The Jet Engine,* Rolls-Royce plc, Derby.

Roskam, J. 1980. *Aircraft Design,* Parts I through VIII, DAR Corporation, Lawrence, KS.

FURTHER READING

Avallone, E. A. and Baumeister, T., III, eds., *Marks' Standard Handbook for Mechanical Engineers,* 10th ed., McGraw-Hill, New York (1996).

Gessow, A. and Myers, G. C., Jr., *Aerodynamics of the Helicopter,* Macmillan, New York (1952).

Huenecke, K., *Modern Combat Aircraft Design,* Naval Institute Press, Annapolis, MD (1987).

Khoury, G. A. and Gillett, J. D., *Airship Technology,* Cambridge Aerospace Series 10, Cambridge University Press, Cambridge (1999).

Mair, W. A. and Birdsall, D. L., *Aircraft Performance,* Cambridge Aerospace Series 5, Cambridge University Press, Cambridge (1992).

McCormick, B. W., Jr., *Aerodynamics of V/STOL Flight,* Academic Press, New York (1967).

Schaufele, R. D., *The Elements of Aircraft Preliminary Design,* Aires, Santa Ana, CA (2000).

Shapiro, J., *Principles of Helicopter Engineering,* McGraw-Hill, New York (1955).

Smetana, F. O., *Flight Vehicle Performance and Aerodynamic Control,* AIAA Education Series, Washington, DC (2001)

Stinton, D., *The Anatomy of the Airplane-Second Edition,* Blackwell Science, Oxford, and AIAA, Washington, DC (1998).

Stinton, D., *The Design of the Airplane,* Van Nostrand Reinhold, New York (1983).

Stinton, D., *Flying Qualitites and Flight Testing of the Airplane,* AIAA Education Series, Washington DC (1996).

SECTION 14
ASTRODYNAMICS

Roy Y. Myose

NOTATION

a semimajor axis length (defines the orbit size)
E eccentric anomaly angle (used for time of flight in elliptic orbit)
F hyperbolic eccentric anomaly angle (used for time of flight in hyperbolic orbit)
H angular momentum per unit mass
p semilatus rectum (defines the orbit size)
r radial distance (measured from center of gravitational body)
T orbital period (for circular and elliptic orbits)
t time since periapsis passage
U mechanical energy per unit mass (sum of kinetic and potential energies per unit mass)
v velocity
γ flight path angle (direction of velocity vector measured above horizon line)
ε eccentricity (defines the orbit shape)
θ true anomaly angle (measured in direction of travel from periapsis)
μ gravitational parameter

Subscripts

a apoapsis (location in elliptic orbit farthest from gravitational body)
c condition corresponding to circular orbit
p periapsis (location in orbit closest to gravitational body)
∞ condition where $r = \infty$ (for hyperbolic and parabolic orbits)

14.1 ORBITAL MECHANICS

Introduction

Over the past 50 years, there have been significant achievements in the field of astronautics—astronauts have visited the moon, cosmonauts have stayed up in space for more than a year, and robotic probes have visited most of the planets in the solar system. Closer to home, there are hundreds of communication satellites circling the globe today. These satellites provide communication links between distance countries and help to support the Internet-driven global economy of the 21st century. The ability to accomplish even the simplest of these space missions depends upon a fundamental understanding of orbital mechanics. The foundation for the modern-day study of orbital mechanics was laid by Johannes Kepler and Isaac Newton some four centuries ago. Many of the terms used in orbital mechanics, however, originate from ancient Greek studies of celestial mechanics and analytical geometry. A list of nomenclature is given above as a convenient reference to aid readers who are new to the study of orbital mechanics. Although there are many standard textbooks covering the topic of orbital mechanics, some of the notation used differs from reference to reference. Thus, the nomenclature list also helps to clarify the notation used in this section of the handbook.

Terminology and Orbit Types

The orbital mechanics part of the space mission begins with the delivery of a spacecraft into orbit around a gravitational body such as the Earth. Depending upon the amount of energy (and angular momentum) provided by the launch vehicle, the spacecraft will follow a specific flight path. The shape of this flight path can be circular, elliptic, parabolic, or hyperbolic. Closed orbits around the Earth follow circular or elliptic flight paths, while parabolic or hyperbolic flight paths are followed by spacecraft leaving the Earth's gravitational field, as shown in Figure 14.1. With the exception of a circular orbit, there is one location in the orbit which is closest to the gravitational body. For the generic gravitational body, this location is called periapsis and has radius r_p, as shown in Figure 14.2. For specific gravitational bodies, the suffix is changed; the point of closest approach to the Earth is called perigee, while the closest approach to the sun is called perihelion. The periapsis location is used as the reference starting point for measuring the true anomaly angle θ so that $\theta = 0$ at the periapsis. Another important location in the orbit is the semilatus rectum p, which is defined as the location where $\theta = 90°$, as shown in Figure 14.2. For elliptic orbits, there is one location in the orbit which is farthest from the gravitational body. This location, called the apoapsis, corresponds to $\theta = 180°$ and has radius r_a. The periapsis and apoapsis are used as two points which define the so-called line of apsides. In addition, the periapsis and apoapsis are the end points of the major axis of the ellipse. Thus, for an elliptic orbit the semimajor axis length a is given by:

$$a = (r_p + r_a)/2 \tag{14.1}$$

For circular orbits, the radius r_c is constant everywhere which means that for this special case:

$$r_c = r_p = p = r_a = a \tag{14.2}$$

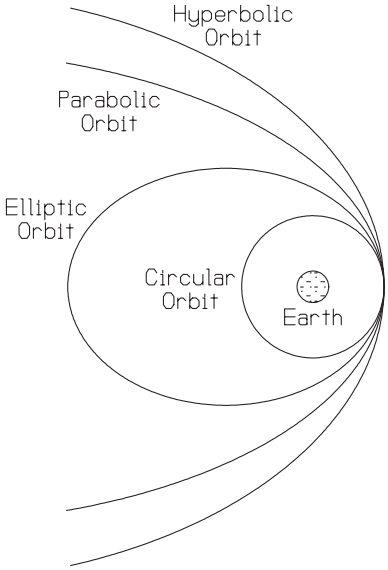


FIGURE 14.1 Different orbit types (adapted from Myose 2001).

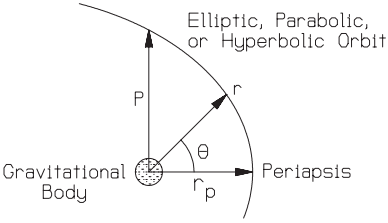


FIGURE 14.2 Definition of periapsis, semilatus rectum, and true anomaly angle (adapted from Myose 2001).

Scalar Relationships for Position, Energy, and Velocity

The four different orbit shapes (circular, elliptic, parabolic, and hyperbolic) are all geometric shapes which can be formed by sectioning off a cone, as shown in Figure 14.3. Consequently, the motion of a spacecraft is governed by the well-known equation for a conic section, which is given by:

$$r = p/(1 + \varepsilon \cos \theta) \tag{14.3}$$

where r is the radial distance measured from the center of the gravitational body, p is the semilatus rectum, ε is the eccentricity, which is a measure of the orbit's shape, and θ is the true anomaly angle. Equation (14.3), called the orbit equation, describes the functional relationship between r, the radial distance from the gravitational body, and θ, the true anomaly angle. Since cosine is an even function,

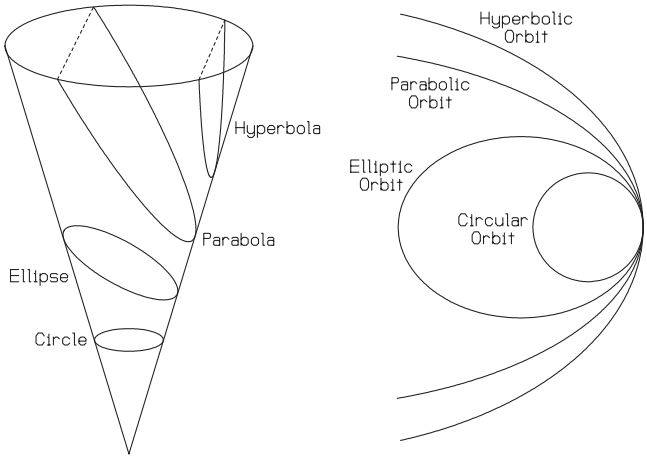


FIGURE 14.3 Conic section orbital shapes (adapted from Myose 2001).

equation (14.3) shows that the radius at $+\theta$ is the same as the radius at $-\theta$. The value of the eccentricity is $\varepsilon = 0$ for a circular orbit, $0 < \varepsilon < 1$ for an elliptic orbit, $\varepsilon = 1$ for a parabolic orbit, and $\varepsilon > 1$ for a hyperbolic orbit.

One unique parameter which is often of interest in hyperbolic orbits is the asymptote angle θ_∞, which is shown in Figure 14.4. This angle is measured between the asymptote line and the line of apsides. The asymptote line is simply the extension of the spacecraft's direction of travel when it escapes the planet's gravity well. A second asymptote line can be drawn from a mirror image along the line of apsides. This second asymptote line is an extension of the direction of travel for a spacecraft which starts outside the planet's gravity well and is captured by the planet's gravity well. Since the spacecraft must start off or end up outside the planet's gravity well, the radial distance is ∞ in this case. Substituting $r = \infty$ into the orbit equation results in an asymptote angle of:

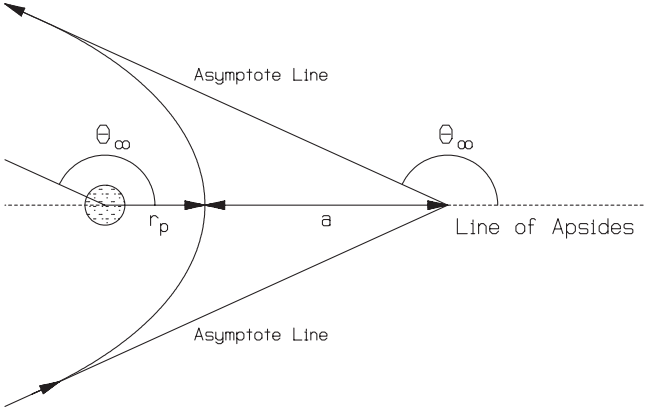


FIGURE 14.4 Hyperbolic orbit (adapted from Myose 2001).

$$\theta_\infty = \cos^{-1}(-1/\varepsilon) \tag{14.4}$$

The orbit equation given in equation (14.3) only describes the flight path taken by the spacecraft and not the energy or angular momentum required to place the spacecraft into such an orbit. For the nonmaneuvering case associated with the basic orbital mechanics problem, the mechanical energy and the angular momentum of the spacecraft must remain constant. The conservation laws for energy and angular momentum are derived from Newton's laws. If **r** represents the vector measured from the gravitational body and pointed towards the spacecraft, then application of Newton's second law together with Newton's law of gravitation results in (see Myose 2001; Bate et al. 1971; Hale 1994; Wiesel 1997):

$$d^2\mathbf{r}/dt^2 = -[G(m_1 + m_2)/r^3]\mathbf{r} \tag{14.5}$$

where G is the universal gravitational constant. For typical astrodynamics problems, the mass m_2 of the spacecraft is significantly less than the mass m_1 of the gravitational body. In this case $G(m_1 + m_2) \approx Gm_1 = \mu$, where μ is called the gravitation parameter. Typical values for the gravitational parameter as well as other data (based on Hale 1994; Brown 1992; Beatty and Chaikin 1990) are given in Tables 14.1 and 14.2. Using the gravitational parameter μ, the modified form of Newton's law given in equation (14.5) can be written as:

$$d^2\mathbf{r}/dt^2 = -(\mu/r^3)\mathbf{r} \tag{14.6}$$

It should be noted that the spacecraft mass is not a part of equation (14.6). Thus, the conservation laws are related to the energy per unit mass and the angular momentum per unit mass.

The conservation of energy can be derived by taking a dot product between the radius vector **r** and the modified form of Newton's law given by equation (14.6) (see Myose 2001; Bate et al. 1971; Hale 1994; Wiesel 1997). The result is a scalar equation involving the spacecraft's velocity v and is given by:

$$U = \tfrac{1}{2}v^2 - (\mu/r) = \text{constant} = -\mu/(2a) \tag{14.7}$$

where U is the total mechanical energy per unit mass, $\frac{1}{2}v^2$ is the kinetic energy per unit mass, and $-(\mu/r)$ is the potential energy per unit mass. It should be noted that $r = \infty$ is the reference location for zero potential energy in astrodynamics

TABLE 14.1 Physical Data

Celestial body	μ (km^3/s^2)	Equatorial radius of surface (km)
Sun	1.3271×10^{11}	6.960×10^5
Mercury	2.2032×10^4	2.439×10^3
Venus	3.2486×10^5	6.051×10^3
Earth	3.9860×10^5	6.378×10^3
Moon	4.9028×10^3	1.783×10^3
Mars	4.2828×10^4	3.393×10^3
Jupiter	1.2671×10^8	7.1492×10^4
Saturn	3.7940×10^7	6.0268×10^4
Uranus	5.7816×10^6	2.5559×10^4
Neptune	6.8713×10^6	2.4764×10^4
Pluto	1.0209×10^3	1.500×10^3

TABLE 14.2 Orbital Data

Celestial body	Semimajor axis length of orbit (km)	Eccentricity	Inclination (deg)
Mercury	5.7910×10^7	0.2056	7.004
Venus	1.0820×10^8	0.0068	3.394
Earth	1.4960×10^8	0.0167	0.000
Moon	3.8440×10^5	0.05	18.3–28.6
Mars	2.2794×10^8	0.0934	1.850
Jupiter	7.7833×10^8	0.0483	1.308
Saturn	1.4270×10^9	0.0560	2.488
Uranus	2.8710×10^9	0.0461	0.774
Neptune	4.4971×10^9	0.0097	1.774
Pluto	5.9135×10^9	0.2482	17.148

rather than the surface of the Earth, which is commonly used in physics textbooks. Thus, a spacecraft has negative potential energy inside the Earth's gravitational well. It requires a significant amount of kinetic energy to climb out of the gravity well and leave the Earth's gravitational field. The total energy for circular and elliptic orbits are negative since the semimajor axis lengths are finite positive values. Thus, spacecraft in circular and elliptic orbits remain trapped inside the gravity well. A spacecraft in parabolic orbit has just enough energy to climb out of the gravity well with zero residual velocity at $r = \infty$. Thus, the total energy for a parabolic orbit is zero and the semimajor axis length is ∞. Finally, a spacecraft in hyperbolic orbit has positive residual velocity at $r = \infty$. This means that the total energy for a hyperbolic orbit is greater than zero and the semimajor axis length is negative. The orbital characteristics for the different orbit types are summarized in Table 14.3.

The conservation of energy given by equation (14.7) can be rearranged to obtain a general equation for velocity as follows:

$$v = \sqrt{\frac{2\mu}{r} - \frac{\mu}{a}} \tag{14.8}$$

In the special case of a circular orbit which has a constant radius, the semimajor axis length is given by $a = r_C$. Substituting this result into equation (14.8) results in a circular orbit velocity of

$$v_c = \sqrt{\frac{\mu}{r_c}} \tag{14.9}$$

Another velocity of interest is the escape velocity, which is the special case where

TABLE 14.3 Orbital Characteristics

Orbit type	Total energy	Semimajor axis length	Eccentricity
Circular	$U < 0$	$a > 0$	$\varepsilon = 0$
Elliptic	$U < 0$	$a > 0$	$0 < \varepsilon < 1$
Parabolic	$U = 0$	$a = \infty$	$\varepsilon = 1$
Hyperbolic	$U > 0$	$a < 0$	$\varepsilon > 1$

a spacecraft has just enough energy to escape the gravity well. In this case, the spacecraft has zero excess energy when it reaches $r = \infty$. Since energy must be conserved, this means that the total energy of this orbit is zero, which corresponds to the case of a parabolic orbit. Substituting $a = \infty$ into equation (14.8) results in an escape velocity of

$$v_{\text{escape}} = \sqrt{\frac{2\mu}{r}} \tag{14.10}$$

When equation (14.10) is applied to determine the escape velocity from the surface of the Earth using μ and r from Table 14.1, the well established result of $v_{\text{escape, Earth}} = 11.2$ km/s is obtained.

Vector Relationships for Angular Momentum and Orientation of the Orbit

The conservation of angular momentum can be derived by taking a cross-product between the radius vector **r** and the modified form of Newton's law given by equation (14.6) (see Myose 2001; Bate et al. 1971; Hale 1994; Wiesel 1997). The result is a vector equation given by:

$$\mathbf{H} = \mathbf{r} \times \mathbf{v} = \text{constant} \tag{14.11}$$

where H is the angular momentum per unit mass. The radius and velocity vectors are the two lines which define the orbital plane while the angular momentum is perpendicular to the orbital plane as shown in Figure 14.5. If no maneuvers are performed, the spacecraft remains in orbit within this plane and the magnitude as well as the direction of the angular momentum vector remains constant according to equation (14.11).

For the circular orbit shown in Figure 14.5, the velocity vector is always perpendicular to the radius vector. In general, the velocity does not have to be perpendicular to the radius. Instead, the velocity must always be tangent to the flight path. For noncircular orbits, this means that the velocity vector is typically oriented

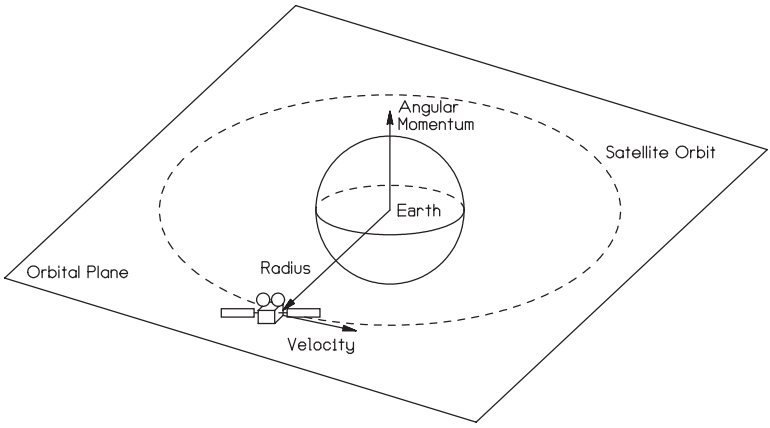


FIGURE 14.5 Orbital plane and angular momentum vector (from Myose 2001).

above or below the horizon line as shown in Figure 14.6. The angle between the velocity vector and the horizon line is called the flight path angle γ. When the spacecraft is moving away from the periapsis, $\gamma > 0$, while $\gamma < 0$ when the spacecraft is moving towards the periapsis. At the periapsis and apoapsis, the velocity and radius are perpendicular to each other which means that the flight path angle is equal to zero. When the idea of flight path angle is applied to equation (14.11), a scalar form of the conservation of angular momentum results:

$$H = rv\cos\gamma \tag{14.12}$$

For the special case when the spacecraft is located at the periapsis and apoapsis, equation (14.12) becomes:

$$H = r_p v_p = r_a v_a \tag{14.13}$$

Since the angular momentum vector is perpendicular to the orbital plane, the vector **H** can be used as part of a three-axis reference system which defines the orientation of an orbit in three dimensions. Within the orbital plane, a two-axis system can be constructed using the periapsis and semilatus rectum directions. Determining the periapsis direction requires a quantity called the eccentricity vector $\boldsymbol{\varepsilon}$ which is pointed in the periapsis direction and has a magnitude equal to the eccentricity. Starting with the cross-product between the modified form of Newton's law given by equation (14.6) and the angular momentum, it can be shown that (see Myose 2001; Bate et al. 1971; Hale 1994; Wiesel 1997):

$$\mathbf{v} \times \mathbf{H}/\mu - \mathbf{r}/r = \boldsymbol{\varepsilon} \tag{14.14}$$

If a dot product is taken between **r** and the eccentricity vector equation given by equation (14.14), a slightly different form of the orbit equation is obtained (see Myose 2001; Bate et al. 1971; Hale 1994; Wiesel 1994):

$$r = (H^2/\mu)/(1 + \varepsilon\cos\theta) \tag{14.15}$$

One consequence of this result is that the orbits of all celestial objects must be conic sections in shape. This was deduced by Johannes Kepler in the 16th century, who stated in his first law that all planetary orbits are elliptic orbits with the sun at one focus.

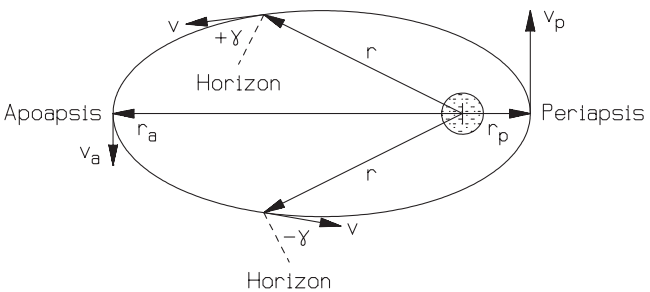


FIGURE 14.6 Flight path angle (from Myose 2001).

Conic Section Relationship and Orbital Elements

A one-to-one relationship between total energy and semimajor axis length was given in equation (14.7). A similar one-to-one relationship between angular momentum (magnitude) and semilatus rectum is obtained by comparing the two different forms of the orbit equation given by equations (14.3) and (14.15). If the orbit equation is evaluated at the periapsis, then a relationship between semilatus rectum and periapsis radius is also obtained. Furthermore, the definition of semimajor axis length given by equation (14.1) can be used to relate the semilatus rectum to the semimajor axis length. The result is an equation relating the angular momentum to the various conic section parameters (see Myose 2001; Bate et al. 1971; Hale 1994; Wiesel 1997):

$$H^2/\mu = p = r_p(1 + \varepsilon) = a(1 - \varepsilon^2) \tag{14.16}$$

Equation (14.16) shows that any pair of (independent) conic section parameters is sufficient to define a spacecraft's angular momentum and total energy. For example, the semimajor axis length and eccentricity can uniquely determine a spacecraft's angular momentum and total energy except in the case of a parabolic orbit.

The angular momentum and total energy are two dynamical quantities which can be used to define a spacecraft's orbital characteristics, also known as orbital elements. Since these two dynamical quantities are directly related to conic section parameters, two geometric quantities can also be used to define the characteristics of an orbit. Oftentimes the semimajor axis length and eccentricity are chosen as the two orbital elements which define the in-plane orbital characteristics. The current position of the spacecraft can be defined by the true anomaly angle (or time since periapsis passage). Three other angles are used to define the orientation of the orbit in three dimensions. The complete set of orbital elements can then be formed from these six different parameters.

Time of Flight Relationships

One consequence of the conservation laws (for energy and for angular momentum) is that a spacecraft moves quickly at periapsis and slowly at apoapsis. This means that the circumferential distance traveled by the spacecraft is large at periapsis and small at apoapsis. Suppose the area swept out by the radius vector in a given amount of time is considered. Due to the conservation laws, the area swept out near the periapsis would be comparable to the area swept out near the apoapsis as shown in Figure 14.7. In fact, it can be shown from first principles that the area A swept out by the radius vector in a given amount of time is related to the angular momentum according to (see Myose 2001; Bate et al. 1971; Hale 1994; Wiesel 1997):

$$dA/dt = H/2 \tag{14.17}$$

Since the angular momentum is constant, equation (14.17) proves Kepler's second law, which states that the radius vector connecting the spacecraft and the gravitational body sweeps out equal areas in equal time.

The orbital period T, which is the time it takes to complete one full orbit, can be determined if the area for the closed orbit is known. For elliptic orbits, the area is given by πab and the semiminor axis length b follows the relationship $b^2 =$

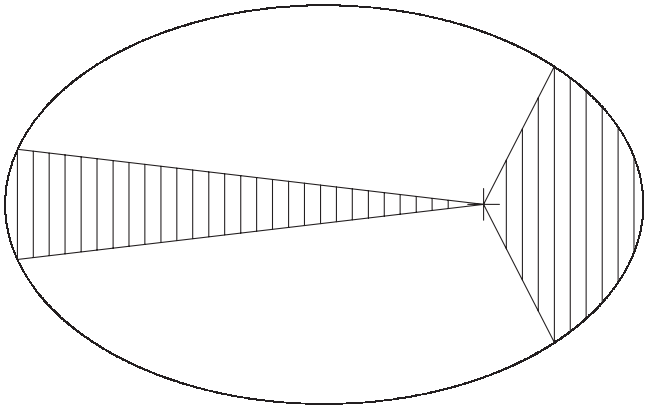


FIGURE 14.7 Area swept out by radius vector near apoapsis (left) and periapsis (right).

$a^2(1 - \varepsilon^2)$. This means that the orbital period is given by (see Myose 2001; Bate et al. 1971; Hale 1994; Wiesel 1997):

$$T = 2\pi\sqrt{a^3/\mu} \tag{14.18}$$

Equation (14.18) is Kepler's third law, which states that the orbital period squared is proportional to the semimajor axis length cubed.

Many problems of interest require determining the time of flight between the periapsis and a location on the orbit where the true anomaly angle θ is known. In order to determine this time since periapsis passage in an elliptic orbit, a concept called the auxiliary circle must be used. The auxiliary circle is a circle with a radius equal to the semimajor axis length of the ellipse which just touches the periapsis and apoapsis as shown in Figure 14.8. If a line perpendicular to the line of apsides

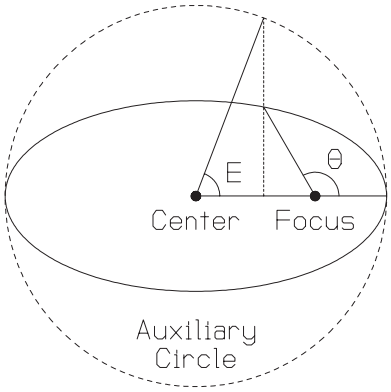


FIGURE 14.8 Auxiliary circle (adapted from Myose 2001).

is extended from the location on the orbit to the auxiliary circle, the resulting angle on the auxiliary circle is given by the eccentric anomaly angle E as shown in Figure 14.8. Using relationships from analytical geometry, it can be shown that the true anomaly and eccentric anomaly angles are related as follows (see Myose 2001; Bate et al. 1971; Hale 1994; Wiesel 1997):

$$\cos E = \frac{\varepsilon + \cos \theta}{1 + \varepsilon \cos \theta} \quad \text{and} \quad \cos \theta = \frac{\cos E - \varepsilon}{1 - \varepsilon \cos E} \tag{14.19}$$

The time since periapsis passage t is given by (see Myose 2001; Bate et al. 1971; Hale 1994; Wiesel 1997):

$$t = \sqrt{\frac{a^3}{\mu}} (E - \varepsilon \sin E) \tag{14.20}$$

It should be noted that the eccentric anomaly angle E used in equation (14.20) must be in units of radians. One interesting property of all orbits is that the orbital behavior is mirror imaged along the line of apsides. This means that the time since periapsis passage for a location corresponding to a true anomaly angle of θ is exactly the same as the time it takes to travel from a location corresponding to a true anomaly angle of $-\theta$ to the periapsis.

Although the time since periapsis passage given by equation (14.20) is valid only for an elliptic orbit, a similar equation is available for a hyperbolic orbit. In this case, the hyperbolic eccentric anomaly angle must be defined. The hyperbolic eccentric anomaly angle F is related to the true anomaly angle as follows (see Myose 2001; Bate et al. 1971; Hale 1994; Wiesel 1997):

$$\cosh F = \frac{\varepsilon + \cos \theta}{1 + \varepsilon \cos \theta} \quad \text{and} \quad \cos \theta = \frac{\cosh F - \varepsilon}{1 - \varepsilon \cosh F} \tag{14.21}$$

The time since periapsis passage for a hyperbolic orbit is given by (see Myose 2001; Bate et al. 1971; Hale 1994; Wiesel 1997):

$$t = \sqrt{\frac{-a^3}{\mu}} (\varepsilon \sinh F - F) \tag{14.22}$$

Example of Geotransfer Orbit

A very useful and often crowded orbital location is the geosynchronous orbit. Many geosynchronous orbit-bound satellites are first brought to a low Earth parking orbit by a rocket launch vehicle before final delivery to geosynchronous orbit. The transfer orbit between the low Earth orbit (LEO) and the final geosynchronous Earth orbit (GEO) is an elliptic orbit whose perigee is equal to the LEO radius and whose apogee is equal to the GEO radius. As an example, we will determine the characteristics of this transfer ellipse if LEO corresponds to 100 nautical mile altitude (i.e., 185 km above the Earth's surface) and GEO corresponds to a radius of 42,164 km.

Problem statement: given h_p = 185 km and r_a = 42,164 km, determine the orbital characteristics of this transfer ellipse and the time of flight from perigee to apogee.

Problem solution: the radius of the Earth's surface, from Table 14.1, is $r_{\text{Earth}} = 6378$ km, which means that $r_p = r_{\text{Earth}} + h_p = 6{,}563$ km. From equation (14.1), the semimajor axis length of the transfer ellipse would be $a = (r_p + r_a)/2 = 24{,}363.5$ km. Based on the conservation of energy, the total energy would be $U = -\mu/(2a) = -8.18$ km^2/s^2. Using the conic section relationship, equation (14.16), the eccentricity of this transfer ellipse would be given by $\varepsilon = 1 - (r_p/a) = 0.7306$. Also from the conic section relationship, the semilatus rectum and angular momentum would be $p = r_p(1 + \varepsilon) = 11{,}358$ km and $H = \sqrt{\mu p} = 67{,}285$ km^2/s. The orbital period is given by equation (14.18), $T = 2\pi\sqrt{a^3/\mu} = 37{,}846$ s. Since the orbit is a mirror image along the line of apsides, the time of flight from perigee to apogee is the same as the time of flight from apogee to perigee. This means that the time of flight from perigee to apogee is half the orbital period, or $T/2 = 18{,}923$ s. An alternative approach would be to determine the eccentric anomaly angle at $\theta = 180°$ (the answer is $E = \pi$) and then substitute it into equation (14.20).

14.2 ORBITAL MANEUVERS

Introduction

An important orbit mentioned in the example of the previous subsection is the geosynchronous Earth orbit (or GEO). A large number of communication satellites as well as some of the Earth observation weather satellites are located in GEO. The idea of using GEO for communication relays was first proposed in 1945 by Arthur C. Clarke, who was trained as an engineer before becoming a science fiction writer. The Early Bird satellite made by Hughes (which is now a part of the Boeing Company) was the first commercial communication satellite to be put into GEO. Twenty years later, in 1984, the number of commercial communication satellites in GEO had grown to 75. Today (in 2001) that number has increased threefold to 223 commercial communication satellites in GEO. To put this into perspective, a large communication satellite relays as many as 100,000 simultaneous telephone or internet connections. Thus, there has been a tremendous increase in global communication traffic over the last 35 years.

Three conditions must be met in order for a satellite to be located in GEO. First, the satellite must have an orbital period equal to the rotation rate of the Earth. Contrary to popular belief, the rotation rate of the Earth is not one revolution in 24 hours. A mean solar day, which corresponds to 24 hours, is the time it takes for the sun to be located at the same position when viewed by a ground observer. During this 24-hour period, the Earth not only rotates about its own axis, but it also moves in its orbit around the sun. This motion around the sun contributes a 360° rotation in 365¼ days, i.e., an additional degree-per-day rotation. Thus, when this additional contribution is taken into account, the rotation rate of the Earth is determined to be one revolution every 23 hours 56 minutes and 4 seconds, which is called a sidereal day. A satellite with an orbital period equal to a sidereal day would have a semimajor axis length of 42,164 km based on equation (14.18). The second requirement for a GEO satellite is that it must be in a circular orbit, i.e., an eccentricity of $\varepsilon = 0$. Since the satellite travels at a constant rate in a circular orbit, there is no eastward or westward shift of the satellite from the perspective of a ground observer. The third requirement for a GEO satellite is that the orbit must

lie in the equatorial plane. This prevents any north-south drift of the satellite during its orbit. Since the satellite motion is synchronized with the rotation of the Earth, a GEO satellite will appear to be located at a fixed location above the equator for a ground observer. One consequent advantage of GEO is that satellite tracking is not required, which makes it easy to use fixed ground-based antennas. Another advantage for GEO satellites is continuous observation of a fixed area of the Earth, which is a desirable quality for an Earth observation weather satellite.

Regardless of whether the spacecraft is intended to be a GEO communication satellite or a scientific probe to another planet, a typical spacecraft is delivered to a low Earth parking orbit by a rocket launch vehicle. This subsection focuses on the issue of how to maneuver a spacecraft from one orbit to another, in particular the velocity changes required to accomplish the orbit change. This may entail not only an increase (or decrease) in the spacecraft's velocity within the plane of the orbit, but also a change in the spacecraft's orbital plane.

Hohmann Transfer

The most basic problem in orbital maneuvers is a transfer from one circular orbit to another circular orbit, e.g., from low Earth orbit (LEO) to GEO. For this situation, the most fuel efficient transfer orbit is half of an ellipse which is tangent to the two circular orbits—in other words, an ellipse whose perigee corresponds to the low circular orbit radius and apogee corresponding to the high circular orbit radius as shown in the left-hand illustration of Figure 14.9. Such an elliptic transfer orbit is called the Hohmann transfer, which was named after Walter Hohmann, who discovered this maneuver in 1925. In order to go from low to high orbit, the velocity must be increased at perigee and at apogee. These two engine burns are often called the perigee and apogee kicks. The velocity increment associated with each of these engine burns is simply the difference in the final and initial velocities. The initial and final velocities are the circular orbit velocity given by equation (14.9) and the perigee or apogee velocity on the transfer ellipse which is given by equation (14.8). At perigee, the initial velocity is the low circular orbit velocity (v_{CL}) while the final velocity is the perigee velocity (v_p) for the Hohmann elliptic transfer orbit. Thus, the first velocity increment (Δv_1) is given by:

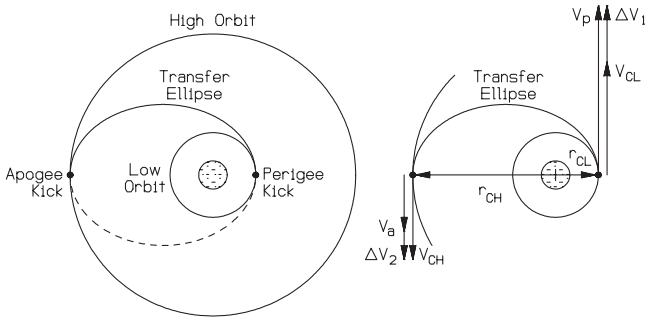


FIGURE 14.9 Hohmann transfer from low to high orbit (from Myose 2001).

$$\Delta v_1 = v_p - v_{CL} \tag{14.23a}$$

$$v_{CL} = \sqrt{\frac{\mu}{r_{CL}}} \tag{14.23b}$$

$$v_p = \sqrt{\frac{2\mu}{r_{CL}} - \frac{\mu}{a}} \tag{14.23c}$$

$$a = \frac{r_{CL} + r_{CH}}{2} \tag{14.24}$$

At apogee, the final velocity is the high circular orbit velocity (v_{CH}) while the initial velocity is the apogee velocity (v_a) for the Hohmann elliptic transfer orbit. Thus, the second velocity increment (Δv_2) is given by:

$$\Delta v_2 = v_{CH} - v_a \tag{14.25a}$$

$$v_{CH} = \sqrt{\frac{\mu}{r_{CH}}} \tag{14.25b}$$

$$v_a = \sqrt{\frac{2\mu}{r_{CH}} - \frac{\mu}{a}} \tag{14.25c}$$

where equation (14.24) gives the semimajor axis length a for the Hohmann transfer orbit.

As an example, consider a Hohmann transfer from 6,563 km radius circular LEO to GEO which corresponds to a circular orbit radius of 42,164 km. Based on equation (14.24), the semimajor axis length of the Hohmann transfer is $a = 24{,}363.5$ km. An Earth orbit means that the gravitational parameter is $\mu_{\text{Earth}} = 3.986 \times 10^5$ km^3/s^2 based on Table 14.1. The circular orbit, perigee, and apogee velocities determined using equations (14.23b), (14.23c), (14.25b), and (14.25c) are $v_p = 10.252$ km/s, $v_{CL} = 7.793$ km/s, $v_{CH} = 3.075$ km/s, and $v_a = 1.596$ km/s. This means that the required velocity increments, based on equations (14.23a) and (14.25a), are $\Delta v_1 = 2.459$ km/s and $\Delta v_2 = 1.479$ km/s. The total velocity increment (Δv_{total}), which relates to the fuel expenditure associated with the perigee and apogee kicks, is given by the sum of the two velocity increment magnitudes, i.e., $\Delta v_{\text{total}} = 3.938$ km/s. The time of flight (TOF) was determined in the example of the previous subsection to be TOF $= T/2 = 18{,}923$ s, where the orbital period T was given by equation (14.18).

Fast Transfers

The time of flight in a Hohmann transfer is equal to half the period of an elliptic orbit since a Hohmann transfer is formed from half of an ellipse. If a reduction in the time of flight is desired, then the transfer orbit must be changed from an ellipse, which is tangent to the low and high orbits, to another conic section with a larger eccentricity. Such a transfer orbit is called a fast transfer. The advantage of a fast

transfer is its reduced time of flight, which comes at the cost of significantly increased fuel expenditure.

The conservation laws discussed in the previous subsection showed that a spacecraft moves quickly at perigee and slowly at apogee. This means that truncating the ellipse to exclude the apogee area can decrease the time of flight. The type I fast transfer shown in Figure 14.10 is such a transfer orbit. A type I fast transfer must have its perigee at the low orbit radius while the high orbit is intercepted at a true anomaly angle of $\theta_2 < 180°$. A type II fast transfer, shown in Figure 14.11, must start from the low orbit radius with a true anomaly angle past perigee ($\theta_1 > 0$), while its apogee is located at the high orbit radius. A type III fast transfer, shown in Figure 14.12, is the general case with $\theta_1 > 0$ and $\theta_2 < 180°$.

The relevant velocity vectors are aligned with each other if the transfer orbit intercepts the circular orbit at either the perigee or apogee. For the type I fast transfer, the first velocity increment will then be given by equation (14.23) as indicated by Figure 14.10. For the type II fast transfer, the second velocity increment is given by equation (14.25) as indicated by Figure 14.11. At locations other than perigee and apogee, the transfer orbit's velocity is oriented at a nonzero flight path angle. This means that the velocity increment is not aligned with the other velocities in the case of Δv_2 for type I (Figure 14.10), Δv_1 for type II (Figure 14.11),

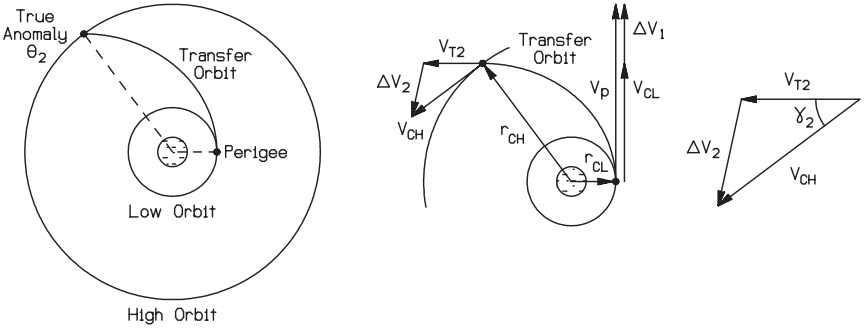


FIGURE 14.10 Fast transfer type I: starting from perigee (adapted from Myose 2001).

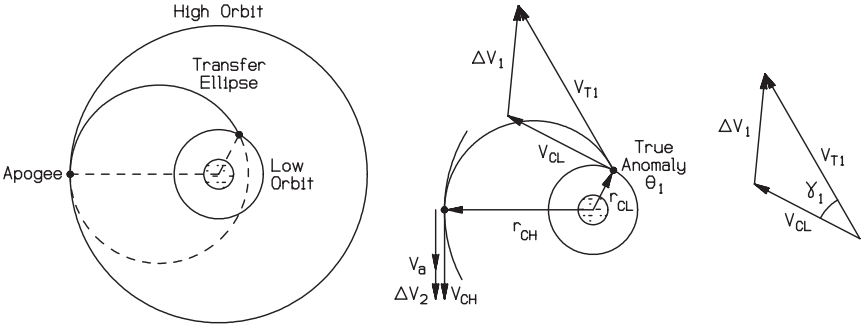


FIGURE 14.11 Fast transfer type II: starting from apogee (adapted from Myose 2001).

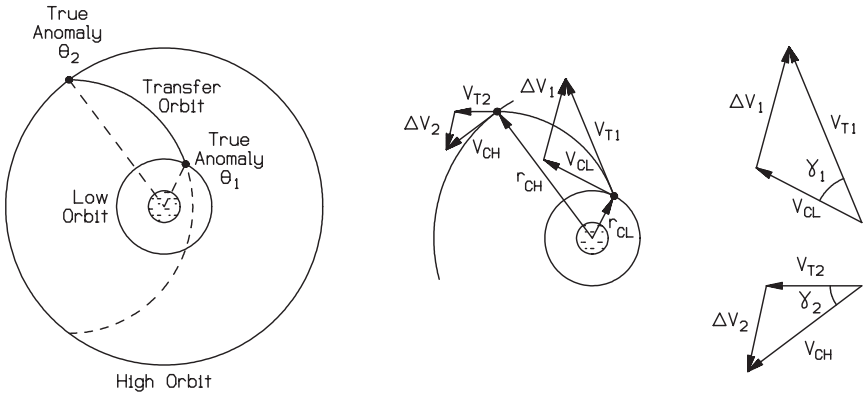


FIGURE 14.12 Fast transfer type III: general case (adapted from Myose 2001).

and both Δv_1 and Δv_2 for type III (Figure 14.12). In such a situation, the cosine law must be used to determine the velocity increment as follows:

$$\Delta v_1 = \sqrt{v_{T1}^2 + v_{CL}^2 - 2v_{T1}v_{CL}\cos\gamma_1} \tag{14.26}$$

$$\Delta v_2 = \sqrt{v_{T2}^2 + v_{CH}^2 - 2v_{T2}v_{CH}\cos\gamma_2} \tag{14.27}$$

As an example, consider transfer from 6,563 km radius circular LEO to GEO which corresponds to a circular orbit radius of 42,164 km. Suppose a type I fast transfer with a high orbit intercept at a true anomaly angle of $\theta_2 = 120°$ is used. The characteristics of this type I fast transfer can be determined from the orbit equation, which is given by either equation (14.3) or (14.15). Since the semilatus rectum p and angular momentum H are fixed for the fast transfer, the orbit equation can be rearranged to determine the eccentricity as follows:

$$r = \frac{H^2/\mu}{1 + \varepsilon\cos\theta} \Rightarrow r_{CL}(1 + \varepsilon\cos\theta_1) = \frac{H^2}{\mu} = r_{CH}(1 + \varepsilon\cos\theta_2) \Rightarrow$$

$$\varepsilon = \frac{r_{CH} - r_{CL}}{r_{CL}\cos\theta_1 - r_{CH}\cos\theta_2} \tag{14.28}$$

Substituting the relevant values result in an eccentricity of $\varepsilon = 1.2878$ for this fast transfer, which means that this transfer orbit is hyperbolic. Since $r_{CL} = r_p$ for a type I fast transfer, the semimajor axis length of $a = -22{,}805$ km can be determined from the conic section relationship given by equation (14.16). An Earth orbit means that the gravitational parameter is $\mu_{\text{Earth}} = 3.986 \times 10^5$ km^3/s^2 based on Table 14.1. Substituting these values into equation (14.23a, b, c) results in $v_p = 11.788$ km/s, $v_{CL} = 7.793$ km/s, and $\Delta v_1 = 3.995$ km/s. This first velocity increment is 1.6 times the first velocity increment for the Hohmann transfer found in the previous example.

The transfer orbit's velocity at high orbit, v_{T2}, can be found from the general velocity equation given by equation (14.8). Substituting the relevant values of $\mu_{\text{Earth}} = 3.986 \times 10^5$ km^3/s^2, $r_{CL} = 42{,}164$ km, and $a = -22{,}805$ km results in $v_{T2} = 6.032$ km/s. Since the perigee conditions are known, the angular momentum can be found using equation (14.13) to be $H = r_p v_p = 77{,}362$ km^2/s. The scalar form of the conservation of angular momentum given by equation (14.12) can then

be used to determine the flight path angle at the time of high orbit intercept. That is, $\gamma_2 = \cos^{-1}(H/r_{CH}v_{T2}) = 72.29°$. Based on equation (14.25b), the required final velocity in GEO is $v_{CH} = 3.075$ km/s. Using the cosine law relationship given in equation (14.27) results in a second velocity increment of $\Delta v_2 = 5.878$ km/s. This second velocity increment is four times the second velocity increment for the Hohmann transfer found in the previous example. Finally, the total velocity increment for this fast transfer is $\Delta v_{\text{Total}} = 9.873$ km/s, which is 2½ times the total velocity increment for the Hohmann transfer.

Substituting $\theta_2 = 120°$ and $\varepsilon = 1.2878$ into equation (14.21) results in a hyperbolic eccentric anomaly angle of $F = 1.4317$ radians. This value along with $\mu_{\text{Earth}} = 3.986 \times 10^5$ km^3/s^2 and $a = -22{,}805$ km gives a time of flight of 6,053 s based on equation (14.22). In comparison, the Hohmann transfer time of flight found in the previous example was about three times longer at 18,923 s.

Orbital Inclination and Launch Site Latitude

The geosynchronous orbit is one important orbit mentioned earlier in the introduction. One requirement for GEO is that the orbit must lie in the equatorial plane. The initial low Earth parking orbit, on the other hand, oftentimes does not lie in the equatorial plane. Figure 14.13 shows such a case where the inclination angle i is defined to be the angle between the angular momentum vector and the Earth's north pole direction. Orbits with an inclination angle of $i = 0$ are called equatorial orbits. Orbits with an inclination angle in the range $0 < i < 90°$ are called prograde orbits because its direction of motion has an eastward component which is the rotation direction of the Earth. Orbits with an inclination angle of $i = 90°$ are called polar orbits because they fly over both poles. Orbits with an inclination angle in the range $90° < i \leq 180°$ are called retrograde orbits because their direction of motion is towards the west, which is counter to the rotation direction of the Earth.

The smallest inclination angle for the low Earth parking orbit is obtained when the rocket is launched toward the east. In this case, the resulting inclination angle is exactly equal to the latitude angle of the launch site since the cross-product

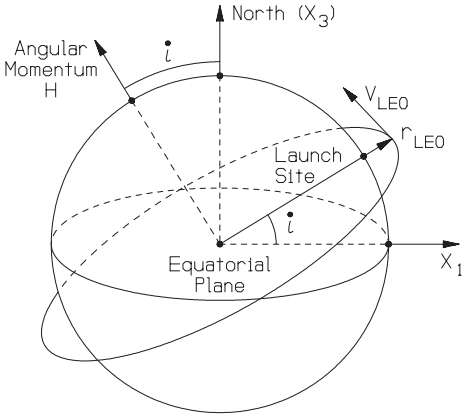


FIGURE 14.13 Launch latitude and orbital inclination (from Myose 2001).

between the launch site's radius vector and the due east velocity vector results in an inclined angular momentum vector, as shown in Figure 14.13. If the rocket is launched in any other direction, then the resulting inclination angle of the low Earth parking orbit is greater than the latitude angle of the launch site. This is because northeastwardly launches would mean that the rocket would continue to travel farther north while southeastwardly launches would indicate that the rocket originates from a higher more northerly latitude.

A launch site located on or near the equator has two advantages. First, the inclination angle of the low Earth parking orbit would be relatively small if the rocket were launched towards the east. This means that very little inclination change is required to transfer from the parking orbit to GEO. Second, a launch site located on the equator would obtain a free boost of 0.464 km/s due to the rotation of the earth. However, this boost in velocity diminishes with increasing latitude. Table 14.4 lists the latitudes of some of the major launch sites (from Isakowitz 1995; Brown 1992). Most launch sites are located within the territorial limits of the national entity launching the rocket. Consequently, there are no ground-based rocket launch facilities located on the equator. One notable exception is the multinational Sea Launch rocket developed by the Boeing Company, which uses a floating base located at the equator in the Pacific Ocean to launch the Russian Zenit rocket launch vehicle.

Pure Inclination Change

Table 14.4 shows that most launch sites are not located on the equator. This means that the low Earth parking orbit is typically inclined. If the desired final destination is located at a different inclination angle, then an inclination change maneuver must be performed. Figure 14.14 depicts such a situation with an inclined initial orbit and an equatorial orbit such as GEO for the desired final orbit. The inclination can only be changed at the intersection points of the two orbits. There are two possible intersection points called nodes—the ascending node, where the spacecraft is headed northward, and the descending node, where the spacecraft is headed southward. A line connecting the two nodes is called the line of nodes.

If only the inclination of the orbit is to be changed without affecting any other orbital characteristic, then the total energy as well as the magnitude of the angular momentum must remain the same. Such a maneuver, shown in Figure 14.14, is called a pure inclination change. In this case, the radius and velocity magnitudes

TABLE 14.4 Launch Site Latitudes

Country	Launch site	Location	Latitude
China	Xichang	China	28.2° N
Europe	Kourou	French Guiana	5.2° N
India	Sriharikota	India	13.9° N
Japan	Tanegashima	Japan	30.2° N
Russia	Baikonur	Kazakhstan	45.6° N
Russia	Plesetsk	Russia	62.8° N
United States	Cape Canaveral (KSC)	Florida, United States	28.5° N
United States	Vandenberg AFB	California, United States	34.7° N

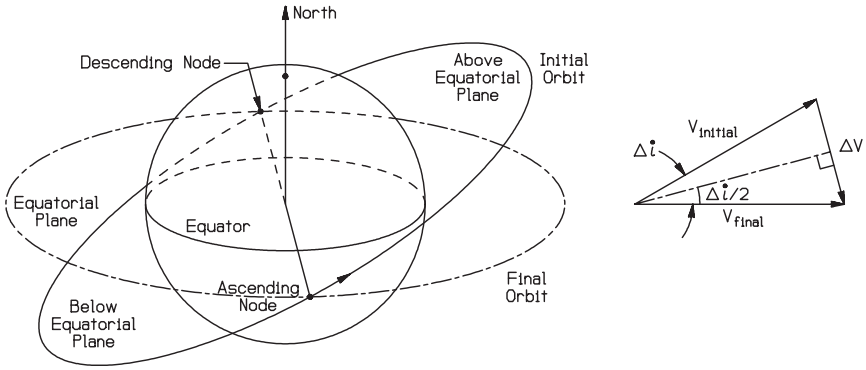


FIGURE 14.14 Intersecting nodes and pure inclination change (from Myose 2001).

are the same before and after the inclination change. Based on the bisection of a right triangle, the velocity increment associated with a pure inclination change is given by:

$$\Delta v_{\text{pure inclination}} = 2v \sin(\Delta i/2) \tag{14.29}$$

Pure inclination changes are extremely fuel-expensive maneuvers to perform. As an example, consider a spacecraft which is initially in a 28.5° inclination 6,563 km radius circular low Earth parking orbit. If an equatorial orbit with the same circular orbit radius is desired, then $v = v_{CL} = 7.793$ km/s and $\Delta i = 28.5°$ would give a pure inclination change velocity increment of $\Delta v_{\text{pure inclination}} = 3.837$ km/s. In comparison, a Hohmann transfer which delivers a spacecraft from 6,563 km radius to GEO at 42,164 km requires a velocity increment of $\Delta v_{\text{Hohmann}} = 3.938$ km/s. If, on the other hand, the pure inclination change is performed at the high orbit, a much smaller fuel expenditure is involved. In this case, $v = v_{CH} = 3.075$ km/s, so that the pure inclination change velocity increment is $\Delta v_{\text{pure inclination}} = 1.514$ km/s.

Combined Hohmann Transfer and Inclination Change Maneuver

The example cited above involved a three-engine burn maneuver. Perigee and apogee kicks for the Hohmann transfer within the plane of the transfer orbit and a separate engine burn to change the inclination of the orbit. In an actual LEO to GEO transfer, the out-of-plane maneuver is combined with the two in-plane maneuvers. Figure 14.15 depicts such a combined Hohmann transfer and inclination change maneuver. Suppose the change in inclination angle during the first engine burn is Δi while the total change in inclination is i. Based on the cosine law, the velocity increments at perigee and apogee are given respectively by:

$$\Delta v_1 = \sqrt{v_p^2 + v_{CL}^2 - 2v_p v_{CL} \cos(\Delta i)} \tag{14.30}$$

$$\Delta v_2 = \sqrt{v_a^2 + v_{CH}^2 - 2v_a v_{CH} \cos(i - \Delta i)} \tag{14.31}$$

The total velocity increment is the sum of these two velocity increments. That is,

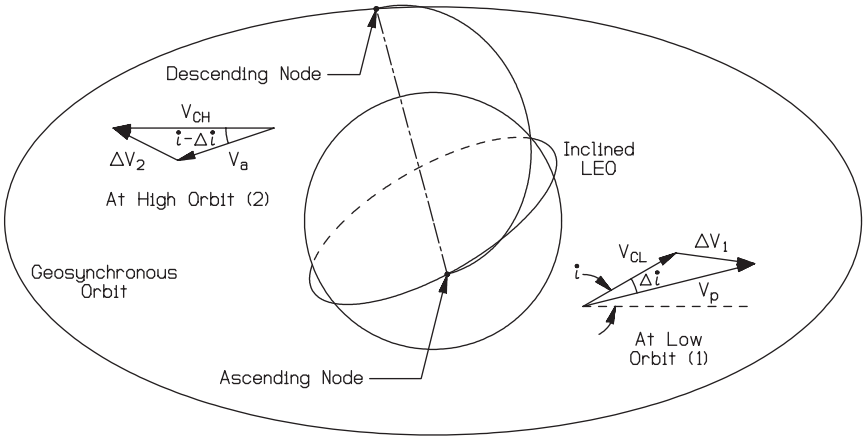


FIGURE 14.15 Combined Hohmann transfer and inclination change maneuver (from Myose 2001).

$$\Delta v_{\text{total}} = \sqrt{v_p^2 + v_{CL}^2 - 2v_p v_{CL}\cos(\Delta i)} + \sqrt{v_a^2 + v_{CH}^2 - 2v_a v_{CH}\cos(i - \Delta i)} \tag{14.32}$$

Based on the results of the previous example, one might conclude that all of the inclination change should be performed at the high orbit since the velocity involved is smallest at that location. However, this is not the case. The optimum solution is the minimization of the total velocity increment, which is given by equation (14.32). Since v_p, v_{CL}, v_a, v_{CH}, and i are fixed constants, the only variable in equation (14.32) is the amount of inclination change Δi during the first engine burn. This means that the minimum total velocity increment is obtained if $\partial(\Delta v_{\text{total}})/\partial(\Delta i) = 0$. Taking a partial derivative of equation (14.32) results in:

$$\frac{v_p v_{CL}\sin(\Delta i)}{\sqrt{v_p^2 + v_{CL}^2 - 2v_p v_{CL}\cos(\Delta i)}} - \frac{v_a v_{CH}\sin(i - \Delta i)}{\sqrt{v_a^2 + v_{CH}^2 - 2v_a v_{CH}\cos(i - \Delta i)}} = 0 \tag{14.33}$$

The optimum solution is dependent upon the specific values for v_p, v_{CL}, v_a, v_{CH}, and i. This means that equation (14.33) must be solved numerically to determine the optimum inclination change and total velocity increment. The optimum solution is obtained at small nonzero values of Δi because the first and second terms involve the sine (and cosine) of two different angle values.

For the 28.5° inclination 6563 km radius circular low Earth parking orbit to GEO example cited earlier, the optimum inclination change angle turns out to be $\Delta i = 2.2°$. This results in a total velocity increment of $\Delta v_{\text{combined}} = 4.273$ km/s. This compares to an in-plane Hohmann followed by a pure inclination change at high orbit requiring a total velocity increment of $\Delta v_{\text{Hohmann}} + \Delta v_{\text{pure inclination}} = 3.938$ km/s + 1.514 km/s = 5.452 km/s.

14.3 EARTH ORBITING SATELLITES

Introduction

The previous subsection covered the basic aspects of orbital maneuvers, e.g., transferring a satellite from an inclined low Earth orbit (LEO) to geosynchronous Earth orbit (GEO). This subsection focuses on additional issues which are relevant to Earth orbiting satellites, such as the field of view, ground tracks, orbital rendezvous, orbit determination, the effect of Earth oblateness, and special types of Earth orbits.

Field of View

One critical piece of information with regard to mission specification is the issue of the satellite's field of view (FOV). The FOV determines the amount of area on the Earth's surface which is visible from the satellite. This information is critical for determining the number of satellites required to provide continuous global coverage in the case of a satellite constellation such as the NAVSTAR Global Positioning System located in medium Earth orbit (MEO). From the perspective of a ground observer, the FOV determines the north and south latitude limits from which the observer can sight a satellite such as the INTELSAT communication satellite located in GEO.

The upper and lower limits of a satellite's FOV are defined by lines which are tangent to the Earth's surface. If a satellite is located infinitely far away, then the lines are parallel and the FOV is 50% of the Earth's surface area as shown in the left-hand illustration of Figure 14.16. If a satellite is a finite distance away, then the FOV is less than 50% of the Earth's surface area as shown to the right of Figure 14.16. The northern and southern latitude limits of the FOV (i.e., the angle L in Figure 14.16) is given by the relationship (see Myose 2001; Brown 1992; Hale 1994):

$$\cos L = r_{\text{Earth}}/r \tag{14.34}$$

where r is the orbital radius of the satellite. The FOV fraction, which is the fraction of the Earth's surface area visible by the satellite, is given by (see Myose 2001; Brown 1992; Hale 1994):

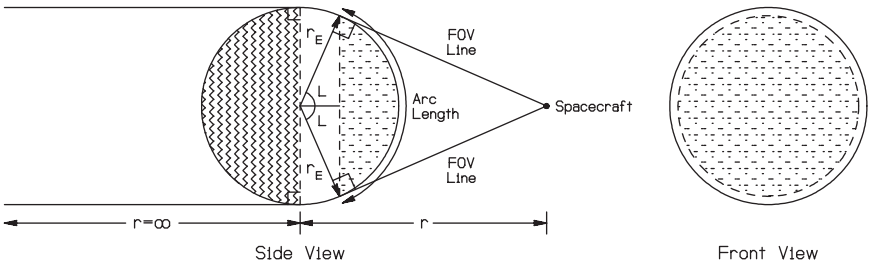


FIGURE 14.16 Field of view (adapted from Myose 2001).

$$\text{FOV fraction} = \frac{1}{2}\left(1 - \frac{r_{\text{Earth}}}{r}\right) \tag{14.35}$$

The circumferential arc length visible by the satellite is given by (see Myose 2001; Brown 1992):

$$\text{FOV arc length} = 2Lr_{\text{Earth}} \tag{14.36}$$

The FOV for a GEO satellite can easily be determined from equations (14.34)–(14.36). Since $r_{\text{Earth}} = 6378$ km and $r = r_{\text{GEO}} = 42{,}164$ km, the FOV latitude limits are $L = \pm 81.3°$ (i.e., 81.3° north to 81.3° south latitude), FOV fraction = 0.424 or 42.4%, and FOV arc length is 18,100 km. It should be noted that a minimum of three GEO satellites is required to obtain global coverage since the FOV fraction for each GEO satellite is 42.4%.

A theoretical latitude limit of 81.3° was obtained for the FOV of GEO satellites in the example above. In practice, however, the latitude limit of the FOV is less than this theoretical value because even a small vertical obstruction would get in the way of the line of sight for a ground observer located at 81.3° north or south latitude. This further reduction in the FOV is not a significant problem for most communication satellite applications since the major population centers are not located in the far north or south. For countries such as Russia, however, this is a problem since a large part of their land mass is located in the far north, especially the undeveloped areas, which have poor communication lines. Russian engineers devised an ingenious solution called the Molniya orbit, which involves a large eccentricity ($\varepsilon = 0.69$) and a high inclination angle ($i = 63.435°$). The Molniya orbit, shown in Figure 14.17, results in a very large hang time near its apogee. This allows the Molniya satellite to be visible in the northern hemisphere during 8 hours of its 12-hour orbital period. A constellation of three Molniya satellites therefore provides continuous coverage. A major drawback of the Molniya system, however, is that satellite tracking is required.

Ground Tracks

One distinct character of a GEO satellite is that it appears to be stationary to a ground observer, hovering over a fixed point above the equator. The main reason is because the orbital period is the same as the Earth's rotation rate, and consequently the satellite moves in sync with the Earth. If the satellite is located in the equatorial plane but at a radius which is less than (or greater than) 42,164 km, then it will no longer move in sync with the Earth. In this case, the satellite would appear to fly over the equator. The flight path taken by the satellite traced onto the Earth's surface is called the ground track.

When a satellite is located in an inclined orbit, it follows a flight path as shown in Figure 14.18. The satellite starts at the ascending node labeled A1 for the first orbital pass, reaches its northernmost location, labeled C, crosses the equator again, but this time at the descending node, and so on. The resulting ground track for a circular LEO is a sinusoidal path as shown in Figure 14.19. Three interesting pieces of information can be obtained by studying the satellite's ground track. First, the northernmost and southernmost latitudes correspond to the inclination of the orbit. Thus, the inclination of the orbit for the ground track shown in Figure 14.19 would be 40°. Second, the symmetry of the sinusoidal ground track pattern indicates that the orbit type is circular. If the orbit is noncircular, the resulting ground track lacks

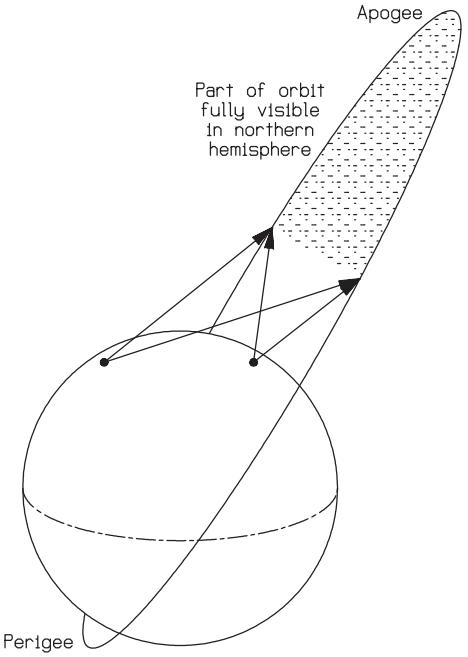


FIGURE 14.17 Molniya orbit (adapted from Myose 2001).

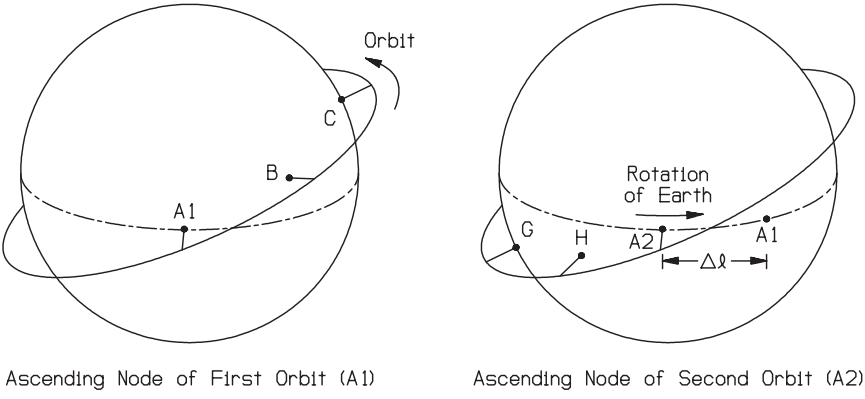


FIGURE 14.18 Ground track of a satellite in a prograde orbit (adapted from Myose 2001).

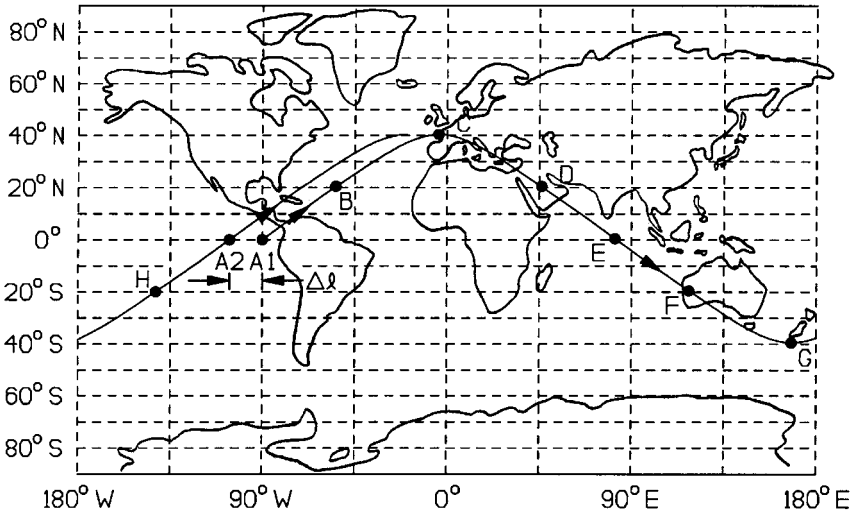


FIGURE 14.19 Ground track of a satellite in 40° inclination circular orbit (adapted from Myose 2001).

some form of symmetry in the sinusoidal pattern due to the difference in speed at perigee and apogee. See Myose (2001) and Hale (1994) for additional details on this issue. The third interesting piece of information which can be obtained from the ground tracks is the orbital period. Suppose the satellite's ground track is traced out starting from the beginning of the first orbital pass, which is labeled A1 in Figures 14.18 and 14.19. During the time it takes for the satellite to complete one orbit, the Earth does not stand still but moves to the east. Since the satellite is left behind by the Earth, the ground track appears to shift west, from the perspective of a ground observer, to the location labeled A2 in Figures 14.18 and 14.19. Subsequent orbits result in further westward shifts in the ground track. Since the Earth's rotation rate is 360° in one sidereal day (which is 86,164 s), each longitudinal shift $\Delta\ell$ in the ground tracks is given by:

$$\Delta\ell = t\ (360°/86{,}164) \text{ or } \Delta\ell = t/239.34 \tag{14.37}$$

where the time t which is relevant to the present situation is the orbital period T. As an example, the longitudinal shift in the ground track for 6,563 km radius circular LEO is 22.11° (to the west) since its orbital period is 5,291.3 s based on equation (14.18).

Orbital Rendezvous

Satellites in geosynchronous orbits are assigned to specific orbital slots corresponding to particular longitude locations on the equator. Longitudes are defined towards the east or west starting from 0° longitude, corresponding to the meridian location of Greenwich, England. Longitude values range from 0° to 180° in both the eastward and westward direction, as shown in Figure 14.19. In order to distinguish east longitudes from west longitudes, we will define east longitudes as positive angle

values and west longitudes as negative angle values. For example, Wichita State University in Kansas is located at 97.3° west longitude, so its longitude will correspond to −97.3°.

The geosynchronous satellite placement problem is illustrated in Figure 14.20. The goal is to place a satellite at a particular orbital slot, for example a final destination corresponding to a longitude of −97.3°. Suppose the transfer orbit involves a true anomaly angle difference of $\theta_2 - \theta_1$ between high orbit intercept (at θ_2) and transfer initiation (at θ_1). For a Hohmann transfer which involves a perigee to apogee half ellipse, the true anomaly angle difference is $\theta_2 - \theta_1 = 180°$, whereas it would be less than 180° for a fast transfer. An initial glance at the problem might suggest that the Hohmann transfer should be initiated (i.e., its perigee should be located) 180° west of the final destination longitude. However, this ignores the rotation of the Earth during the time it takes to complete the Hohmann transfer. Based on equation (14.37), the rotation of the Earth $\Delta\ell$ during the time of flight (*TOF*) is given by:

$$\Delta\ell = TOF/239.34 \tag{14.38}$$

This means that the final destination longitude is situated differently depending on the phase of the transfer, i.e., the start or end of the transfer orbit as shown in the right-hand illustration in Figure 14.20. The lead angle κ_ℓ is the angle difference between the transfer initiation longitude and the final destination longitude at the time the Hohmann transfer is initiated at perigee. Based on Figure 14.20, the lead angle is given by:

$$\kappa_\ell = (\theta_2 - \theta_1) - \Delta\ell \tag{14.39}$$

where $(\theta_2 - \theta_1) = 180°$ for a Hohmann transfer. Since the required transfer initiation longitude is west of the desired final destination longitude by an angle κ_ℓ, the lead angle value must be subtracted from the final destination longitude:

$$\ell_{\text{transfer initiation}} = \ell_{\text{final destination}} - \kappa_\ell \tag{14.40}$$

As an example, consider placing a satellite in GEO (with 42,164 km radius) at

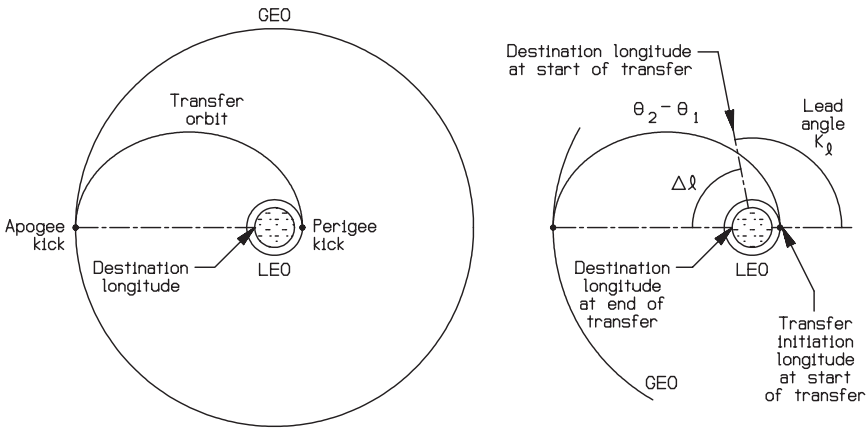


FIGURE 14.20 Transfer initiation for a geosychronous satellite (adapted from Myose 2001).

an orbital slot of 97.3° west longitude starting from a 6563 km radius 28.5° inclination circular LEO. The semimajor axis length of the Hohmann transfer is a = 24,363.5 km based on equation (14.1) or (14.24). Using the orbital period relationship given by equation (14.18), TOF = $T/2$ = 18,923 s. Based on equation (14.38), the Earth rotates $\Delta\ell$ = 79.06° during the Hohmann transfer. Since the true anomaly angle difference is 180° for a Hohmann transfer, the lead angle is given by equation (14.39) to be κ_ℓ = 180° − 79.06° = 100.94°. Based on equation (14.40), the required transfer initiation longitude is $\ell_{\text{transfer initiation}} = -97.3° - 100.94° = -198.24°$. Since longitudes are only defined between ±180°, the required transfer initiation longitude is $\ell_{\text{transfer initiation}} = -198.24° + 360° = 161.76°$ east longitude.

The discussion in the orbital maneuvers subsection showed that the combined Hohmann transfer and inclination change maneuver must be performed at the node locations. The nodes for a low Earth parking orbit depend upon the orbital radius, and information about their locations is usually given by the organization providing the launch service. However, a rough idea about the process involved in determining the node locations can be illustrated by a simple example. Suppose a due east launch is performed from Cape Canaveral, which is located at 80.5° west longitude. The resulting orbit will have a 28.5° inclination, and the first crossing of the equator, which is the descending node, occurs one quarter of an orbit later, or 90° east of the launch site. However, there is a 22.11° westward shift per orbit due to Earth rotation for a 6,563 km radius LEO. This means that there is a westward shift of about 22.11°/4 = 5.53° during this quarter orbit. Thus, the first descending node occurs at roughly −80.5° + 90° − 5.53° = 3.97° east longitude. The first ascending node will occur 180° to the east, less the shift due to Earth rotation, which is 22.11°/2 = 11.06° for this half of the orbit. This means that the first ascending node occurs at 3.97° + 180° − 11.06° = 172.91° east longitude. Neither the descending nor the ascending node for this first orbit corresponds to the required launch initiation longitude. The ascending nodes for subsequent orbits are shifted 22.11° to the west. Thus, the ascending nodes for the second and subsequent orbits are 150.8°, 128.69°, 106.58°, 84.47°, 62.36°, 40.25°, 18.14°, −3.97°, and so on. A further refinement of this process requires knowledge of the actual rocket launch trajectory to determine the initial node location and the inclusion of Earth oblateness effect, which is discussed in the last part of this subsection.

None of the orbits in the example above have ascending nodes corresponding to the required transfer initiation longitude. The closest longitude is the ascending node for the first orbit, i.e., 172.91°. This is 11.15° ahead (i.e., farther east than) the required transfer initiation longitude of 161.76°. If the transfer is initiated at the ascending node of the first orbit, the satellite will arrive 11.15° ahead of the desired final longitude. Thus, some type of longitude correction must be performed once the satellite arrives at the high orbit radius. This type of problem, shown in the left-hand illustration in Figure 14.21, is called the secondary intercept problem. Since the satellite is currently ahead of the desired longitude, an initial thought would be to slow down and allow the Earth to catch up. However, this braking maneuver would cause the satellite to be placed into a secondary intercept ellipse whose orbital period is less than the GEO period of 86,164 s, which is synchronized with the Earth's rotation. This will cause the satellite to increase its longitudinal error angle after one full orbit is completed, as shown in the middle illustration of Figure 14.21. Therefore, the correct secondary intercept maneuver is initially to increase the speed, which would place the satellite into an ellipse whose orbital period is greater than the GEO period. The required secondary intercept period and semimajor axis length are given by the following relationship:

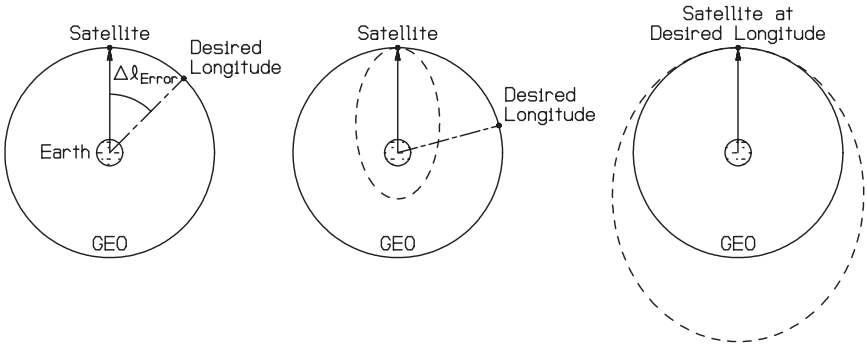


FIGURE 14.21 Secondary intercept problem (from Myose 2001).

$$2\pi\sqrt{\frac{a^3}{\mu_{\text{Earth}}}} = T = 86{,}164\left(1 + \frac{\Delta\ell_{\text{error}}}{360^\circ}\right) \tag{14.41}$$

where $\Delta\ell_{\text{error}}$ is positive if the satellite is ahead of the desired final longitude and negative if it is behind the desired final longitude. The velocity increment required to inject the satellite into this secondary intercept ellipse is given by:

$$\Delta v_1 = v_{\text{secondary intercept}} - v_{CH} \tag{14.42a}$$

$$v_{\text{secondary intercept}} = \sqrt{\frac{2\mu_{\text{Earth}}}{r_{CH}} - \frac{\mu_{\text{Earth}}}{a}} \tag{14.42b}$$

$$v_{CH} = \sqrt{\frac{\mu_{\text{Earth}}}{r_{CH}}} \tag{14.42c}$$

where r_{CH}, the high circular orbit radius, is the 42,164 km for GEO. After one full orbit on the secondary intercept ellipse is completed, a second velocity increment which slows the speed down and places the satellite back onto GEO is required. This second maneuver is exactly the opposite of the first secondary intercept maneuver, so that

$$\Delta v_2 = -\Delta v_1 \tag{14.43}$$

Fuel expenditure is associated with any type of engine burn, whether it increases the velocity or decreases it. Thus, the total velocity increment associated with the initiation and completion of the secondary intercept is given by the sum of the absolute values of these velocity increments. That is,

$$\Delta v_{\text{total}} = |\Delta v_1| + |\Delta v_2| \tag{14.44}$$

As an example, consider the situation described above where the satellite arrives at the GEO radius 11.15° ahead of the desired longitude location. In this case, $\Delta\ell_{\text{error}} = +11.15^\circ$ so the required secondary intercept orbital period and semimajor axis length based on equation (14.41) are $T = 88{,}833$ s and $a = 43{,}030.3$ km,

respectively. Using equations (14.42)–(14.44), the required velocities are found to be $v_{SI} = 3.106$ km/s, $v_{CH} = 3.075$ km/s, $\Delta v_1 = 0.031$ km/s, $\Delta v_2 = -0.031$ km/s, and Δv_{total} 0.062 km/s.

In practice, the longitudinal error between the actual and required transfer initiation longitude is dealt with in a slightly different manner. A GEO-bound satellite is placed in a Hohmann transfer orbit with an apogee slightly larger (or smaller) than the GEO radius (Blevis and Stoolman 1981). The subsequent apogee kick places the satellite into a high orbit with an orbital period which is slightly different than the GEO period. The satellite is then allowed to drift until the final destination longitude is reached, at which point small velocity adjustments are made.

Orbit Determination and Orbital Elements

Delivering a GEO satellite to the assigned orbital slot requires a matchup between the required transfer initiation longitude and the actual node location for the low Earth parking orbit. Otherwise, some sort of secondary intercept maneuver is required to correct the longitude error. Either method requires a precise determination of the satellite's orbital characteristics in the low Earth parking orbit and in the high orbit before any type of maneuver is performed. A total of six parameters, called orbital elements, is required to define the characteristics of an orbit in three dimensions. The reason is because the governing equations of motion, the modified form of Newton's law given in vector form by equation (14.6), are three separate second order differential equations. Since two constants of integration are required for each second order differential equation, a total of six constants of integration or orbital elements is involved in three-dimensional orbits.

In the basic orbital mechanics discussion, two parameters were mentioned as the required elements for determining the characteristics of an orbit within the plane of the orbit. They can come in different forms (dynamic or geometric), but the most commonly used parameters are the semimajor axis length a and the eccentricity ε. One additional in-plane parameter determines the current position of the satellite. Typically, the time since perigee passage is used. However, the true anomaly angle θ will be used for the purposes of the discussion which follows. There is indeed a one-to-one correspondence between the true anomaly angle and the time since perigee passage, as can be seen from equation pairs (14.19) and (14.20) for elliptic orbits and (14.21) and (14.22) for hyperbolic orbits. However, equations (14.20) and (14.22) are transcendental, and no analytical solutions are available to determine the eccentric anomaly angle. Determining the anomaly angles from the time information therefore requires an iterative numerical technique.

Three other orbital elements besides the semimajor axis length, eccentricity, and true anomaly angle are required to define the characteristics of an orbit. The remaining three orbital elements define the orientation of the orbit in three dimensions measured with respect to a three-axis geocentric (or Earth-based) coordinate system. This Earth-based coordinate system $\mathbf{X}$ illustrated in Figure 14.22 has its origin at the center of the Earth. The X_3 axis is pointed in the north pole direction and the X_1 axis is pointed towards the vernal equinox direction. The vernal equinox is the location where the Sun crosses the equator on the first day of spring. Although not shown, the X_2 axis lying in the equatorial plane completes the geocentric coordinate system, and it is simply orthogonal to the other two axes according to the usual right-hand rule involving cross-products where $\mathbf{X}_2 = \mathbf{X}_3 \times \mathbf{X}_1$. The orbit-based coordinate system with its origin at the center of the Earth has already been intro-

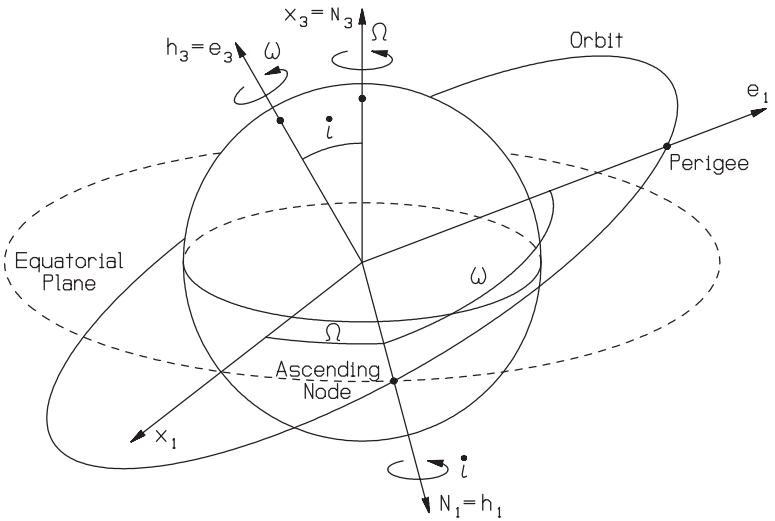


FIGURE 14.22 Orbital elements (from Myose 2001).

duced in the orbital mechanics subsection. The e_1 axis corresponds to the perigee direction, the e_2 axis to the semilatus rectum direction, and the e_3 axis to the angular momentum direction.

A set of three angles called Euler rotation angles separates the geocentric and orbit-based axes. Euler rotation angles are involved whenever a coordinate system undergoes a series of rotations to obtain another coordinate system. Starting from the geocentric coordinate system **X**, the three rotations are: (1) the right ascension of ascending node Ω, (2) the inclination angle i, and (3) the argument of perigee ω. The first rotation is made about the north pole direction X_3, towards the east by an angle Ω, until the one axis is pointed in the ascending node direction. This results in the formation of a new set of axes called the node crossing coordinate system **N** with N_1 pointed toward the ascending node, N_3 in the north pole direction, which makes it coincident with X_3, and N_2 completing the three-axis orthogonal coordinate system. The second rotation is made about the ascending node direction N_1, by the inclination angle i, until the three axis is pointed in the orbit's angular momentum direction. This results in the formation of another set of axes **h**, where h_1 points toward the ascending node and is therefore coincident with N_1, h_3 is pointed in the angular momentum direction, and h_2 completes the three-axis orthogonal coordinate system. The third and final rotation is made about the angular momentum direction h_3, in the direction of travel in the orbit by an angle ω, until the one axis is pointed in the perigee direction. The result of this final rotation is the orbit-based coordinate system **e**.

During a typical satellite mission, multiple sightings of the satellite are obtained through ground-based visual observations or radar measurements. These geocentric measurements must be converted into orbit-based parameters such as a, ε, θ, Ω, i, and ω. That is, a set of dynamical information such as the radius and velocity vectors at one point on the orbit must be converted into the geometric orbital elements. Given **r** and **v**, the orbital elements can be determined from the relation-

ships which follow (see Myose 2001; Hale 1994; Bate et al. 1971). The semimajor axis length a can be determined from equation (14.7), the conservation of energy, which after some rearrangement is given by:

$$a = \frac{1}{(2/r) - (v^2/\mu)} \tag{14.45}$$

The eccentricity can be determined from the magnitude of the eccentricity vector $\boldsymbol{\varepsilon}$. Substituting equation (14.11) into equation (14.14) and then applying the vector triple product rule, the eccentricity vector is given by:

$$\boldsymbol{\varepsilon} = \frac{1}{\mu}\left[\left(v^2 - \frac{\mu}{r}\right)\mathbf{r} - (\mathbf{r}\cdot\mathbf{v})\mathbf{v}\right] \tag{14.46}$$

The true anomaly angle is the angle between the perigee and the radius vector. Since the eccentricity vector $\boldsymbol{\varepsilon}$ is pointed in the perigee direction $\mathbf{e}_1$, the true anomaly angle can be found from:

$$\theta = \cos^{-1}\left(\frac{\boldsymbol{\varepsilon}\cdot\mathbf{r}}{\varepsilon\, r}\right) \tag{14.47}$$

Since cosine is an even function, two solutions are possible. One possible technique for determining the correct true anomaly angle is to observe the sign of the dot product $\mathbf{r}\cdot\mathbf{v}$, which is the second parenthetical term in equation (14.46). If $\mathbf{r}\cdot\mathbf{v} > 0$, then the true anomaly angle is in the range $0 < \theta < 180°$. If $\mathbf{r}\cdot\mathbf{v} < 0$, then the true anomaly angle is in the range $180° < \theta < 360°$. The inclination angle lies between the angular momentum vector and the north pole direction. Thus,

$$i = \cos^{-1}\left(\frac{\mathbf{H}\cdot\mathbf{X}_3}{H}\right), \text{ where } \mathbf{H} = \mathbf{r}\times\mathbf{v} \tag{14.48}$$

Quadrant correction is not required in the case of the inclination angle, since it is only defined between 0 and 180°. Before the right ascension of ascending node and argument of perigee angles can be determined, the node crossing direction N_1 must first be determined. The node crossing direction must be perpendicular to both the angular momentum direction h_3 and the north pole direction X_3 as indicated in Figure 14.22. Thus,

$$\mathbf{N}_1 = \left(\frac{\mathbf{X}_3\times\mathbf{H}}{H}\right)\Bigg/\left\|\frac{\mathbf{X}_3\times\mathbf{H}}{H}\right\| \tag{14.49}$$

It should be noted that $(\mathbf{X}_3\times\mathbf{H})/H$ is not a unit vector, so that this resultant vector must be normalized by its magnitude in order to make the $\mathbf{N}_1$ axis direction given in equation (14.49) a unit vector. Based on the definition for the right ascension of ascending node and the argument of perigee,

$$\Omega = \cos^{-1}(\mathbf{X}_1\cdot\mathbf{N}_1) \tag{14.50}$$

$$\omega = \cos^{-1}\left(\frac{\mathbf{N}_1\cdot\boldsymbol{\varepsilon}}{\varepsilon}\right) \tag{14.51}$$

Quadrant correction may be required since cosine is an even function, the details of which are given in Myose (2001) and Bate et al. (1971).

Since large numerical values are involved in astrodynamics problems, measurement data used to determine the orbital elements are oftentimes normalized by units of measure called canonical units. For Earth orbiting satellites, the normalization for distance is made with respect to the radius of the Earth's surface, where 1 distance unit (DU) is 6378 km. Also, the gravitational parameter for Earth is used as a reference, i.e., 1 distance unit cubed divided by a time unit squared (DU^3/TU^2) is 3.986×10^5 km^3/s^2 which means that a time unit (TU) is equivalent to 806.786 s. For interplanetary missions a different set of references is used to define the canonical units. In this case, the Earth's orbital radius around the Sun ($1\ DU = 1.496 \times 10^8$ km) and the gravitational parameter for the Sun ($1\ DU^3/TU^2 = 1.327 \times 10^{11}\ km^3/s^2$) are used as the reference normalization values. Canonical units are used simply as a matter of tradition and convenience, just as the problems discussed thus far used kilometers rather than the SI standard of meters.

To illustrate the use of the relationships discussed above, consider the following example from Myose (2001). Problem statement: determine the six orbital elements given that the radius and velocity vectors written in terms of geocentric coordinates are:

$$\mathbf{r} = (-0.73993\mathbf{X}_1 + 0.73993\mathbf{X}_2 + 0.60415\mathbf{X}_3)\ \text{DU}$$

$$\mathbf{v} = (-0.77264\mathbf{X}_1 - 0.64637\mathbf{X}_2 + 0.05155\mathbf{X}3)\ \text{DU/TU}$$

Solution: the radius and velocity magnitudes are $r = 1.2083$ DU (i.e., $r = 1.2083 \times 6378 = 7706.5$ km) and $v = 1.0087$ DU/TU (i.e., $v = 1.0087 \times 6378/806.786 = 7.974$ km/s). Based on equation (14.45), the semimajor axis length is $a = 1.5679$ DU. Keeping in mind that $\mu = 1$ in canonical units, the first term in equation (14.46) is $(v^2 - \mu/r)\mathbf{r} = -0.14044\mathbf{X}_1 + 0.14045\mathbf{X}_2 + 0.11467\mathbf{X}_3$. The parenthetical part of the second term is $\mathbf{r} \cdot \mathbf{v} = +0.12458$, so that $(\mathbf{r} \cdot \mathbf{v})\ \mathbf{v} = -0.09625\mathbf{X}_1 - 0.08052\mathbf{X}_2 + 0.00642\mathbf{X}_3$. Finally, the eccentricity vector is $\boldsymbol{\varepsilon} = -0.04419\mathbf{X}_1 + 0.22097\mathbf{X}_2 + 0.10825\mathbf{X}_3$, which means that its magnitude is $\varepsilon = 0.25$. Since $\mathbf{r} \cdot \mathbf{v} > 0$, the true anomaly angle based on equation (14.47) is $\theta = 30°$ rather than 330°. From equation (14.11), $\mathbf{H} = \mathbf{r} \times \mathbf{v} = (0.42865\mathbf{X}_1 - 0.42865\mathbf{X}_2 + 1.04996\mathbf{X}_3)$ DU^2/TU, which means that its magnitude is $H = 1.21239\ DU^2/TU$. The inclination angle $i = 30°$ based on equation (14.48). The node crossing direction, which is a unit vector, is given by $\mathbf{N}_1 = 0.70711\mathbf{X}_1 + 0.70711\mathbf{X}_2 + 0.0\mathbf{X}_3$ based on equation (14.49). Finally, the right ascension of ascending node and the argument of perigee from equations (14.50) and (14.51) are $\Omega = 45°$ and $\omega = 60°$, respectively.

In the discussion above, the radial position and velocity vectors in terms of geocentric coordinates were used as the starting basis for determining the orbital elements. Determining the orbital elements based on more rudimentary measurements such as position and time is beyond the scope of this discussion. In addition, there are reference coordinate systems besides the geocentric and orbit-based systems described above. The interested reader is referred to Bate et al. (1971); Chobotov (1991); Escobal (1976); and Vallado (2001).

Orbital Precession due to Earth Oblateness

The westward shift of the ground track discussed above is only an apparent shift in the ground track viewed from the perspective of the ground observer. On the other hand, the oblateness of the Earth results in an actual precession or wobble in the orbit. The reason for this effect is illustrated in Figure 14.23 where the Earth

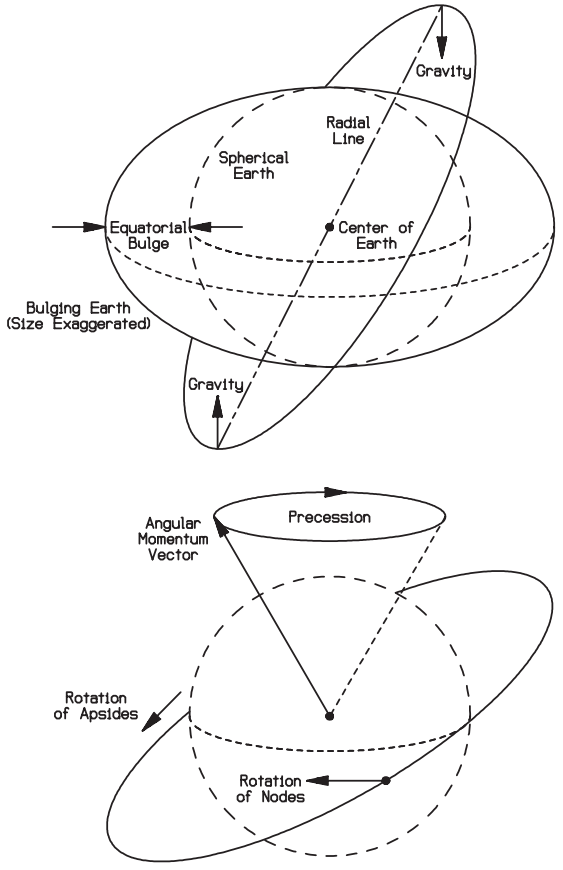


FIGURE 14.23 Earth oblateness effect (adapted from Myose 2001).

bulges out at the equator by about 20 km, radially. Because of this bulging effect, the satellite experiences a nonradial gravitational force at the maximum and minimum latitude locations in its orbit. This nonradial force causes the angular momentum vector of the orbit to precess in a manner similar to the wobbling of an Earth-based gyroscope. This precession motion causes the ascending node to shift physically westward in prograde orbits and eastward in retrograde orbits.

As discussed earlier, the right ascension of ascending node Ω is the angle measured from a reference axis on the equator (called the vernal equinox) to the ascending node. The rate $\dot{\Omega}$ at which the node regresses westward is given by the relationship (see Prussing and Conway 1993; Griffin and French 1991; Kaplan 1976):

$$\dot{\Omega} = \left[\frac{3}{2} J_2 \sqrt{\frac{\mu}{a^3}} \left(\frac{r_{\text{equator}}}{p}\right)^2 \cos i\right]\left(\frac{180}{\pi}\right) \tag{14.52}$$

where $\dot{\Omega}$ is given in units of degrees per second after the $180/\pi$ conversion factor,

$r_{equator}$ is the equatorial radius of the planet given in Table 14.1, and J_2, the zonal harmonic, for the Earth is 1.0826×10^{-3}. Brown (1992) provides the zonal harmonic coefficients for many of the other planets in the solar system. Contrary to some references, a sign convention of positive for westward shift and negative for eastward shift is used in equation (14.52) for the sake of consistency with the earlier discussion on longitudinal shifts in the ground track.

As an example, consider a 6,563 km radius 28.5° inclination circular LEO. Since this is a circular Earth orbit, the semimajor axis length and semilatus rectum are given by $a = p = 6563$ km and the gravitational parameter from Table 14.1 is $\mu_{Earth} = 3.986 \times 10^5$ km^3/s^2. Based on equation (14.52), the regression of nodes would then be $\dot{\Omega} = 9.170 \times 10^{-5}$ deg/s. Since the orbital period is 5291 s based on equation (14.18), the westward shift in the ascending node after one orbit would be $\Delta\Omega = \dot{\Omega}T = 0.49°$. This compares to the 22.11° westward shift in the ground tracks due to Earth rotation, discussed earlier.

The effect of precession due to Earth oblateness can be used as an advantage in a certain type of orbital application called Sun synchronous orbits. Earth resource observation satellites such as the American Landsat or French Spot require comparable lighting conditions from one day to the next, and this requirement can be met using a Sun synchronous orbit. The idea is to match the precessional rotation of the orbit with the Earth's orbital motion around the Sun. This means that the precession rate must be 360° in 365¼ days, which corresponds to the Earth completing one orbit around the Sun in one year. One important issue, however, is that this precession must be oriented towards the east, i.e., a negative regression of nodes. Thus, $\dot{\Omega} = -360°/365¼$ days $= -0.9856$ deg/day or a regression of nodes of about 1° per day towards the east. This value for the regression of nodes can be obtained from a multitude of orbital radius and inclination angle pairs. The Landsat 4, for example, is in a 7083 km radius 98.2° inclination circular LEO (Damon 1995).

For noncircular orbits, there is an additional Earth oblateness effect called the rotation of apsides, where the perigee is shifted due to a physical rotation of the orbit within the orbital plane. As discussed earlier, the argument of perigee ω is the angle measured from the ascending node to the perigee. The rate $\dot{\omega}$ which the argument of perigee is rotated is given by the relationship (see Prussing and Conway 1993; Griffin and French 1991; Kaplan 1976):

$$\dot{\omega} = \left[\frac{3}{2} J_2 \sqrt{\frac{\mu}{a^3}} \left(\frac{r_{equator}}{p}\right)^2 (2 - 2.5 \sin^2 i)\right]\left(\frac{180}{\pi}\right) \tag{14.53}$$

Here, positive values indicate a rotation of apsides in the direction of travel (i.e., same rotation direction as the orbit's angular momentum vector) while negative values indicate a rotation of apsides opposite to the direction of travel. The highly eccentric Molniya orbit described earlier is located at a unique inclination angle (i.e., $i = 63.435°$), where the parenthetical term $(2 - 2.5 \sin^2 i) = 0$. This means that the Molniya orbit does not entail any rotation of apsides. This ensures that the apogee is always located in the northern hemisphere, where the largest hang time (or coverage) is required.

14.4 INTERPLANETARY MISSIONS

Introduction

The basic interplanetary mission is a truly complex problem involving multiple gravitational bodies. There is the Earth where the spacecraft originates from, the

Sun during transit across the solar system, and the destination planet. The simple two-body problem consisting of a spacecraft and a gravitational body such as the Earth is easy to solve, and its governing equations were discussed in subsection 14.1. However, there is no solution available for the general *N*-body problem, which is associated with interplanetary missions. Instead of an exact solution, an approximate technique called the method of patched conics will be used. In this technique, the interplanetary mission is broken up into phases according to the primary gravitational body that influences the spacecraft. The first phase (I) is the hyperbolic Earth departure, the second phase (II) is the heliocentric (i.e., Sun-centered) transit across the solar system, and the third phase (III) is the hyperbolic arrival at the destination planet. These orbits or conic sections around different primary gravitational bodies are then patched together to complete the full interplanetary mission. Another complex problem associated with interplanetary missions is the issue of proper launch timing in order to intercept the destination planet, which is a moving target. This issue will be addressed in the final part of this subsection. Keeping these complexities in mind, it is indeed an accomplishment to visit other planets that are trillions of kilometers away.

For the purposes of this discussion, a number of simplifying assumptions will be made. First, a Hohmann transfer will be used for the heliocentric transit from the Earth to the destination planet. Interested readers should see Myose (2001) and Hale (1994), which discuss the issue of fast heliocentric transfers. The second and third assumptions are as follows: planetary orbits around the Sun are circular and planetary orbits are in the same orbital plane as the Earth. With the exception of Mercury and Pluto, these are reasonable assumptions to make according to Table 14.2. Brown (1992) discusses the issue of non-coplanar interplanetary transfers as well as planetary ephemeris, which specify the planetary positions within their elliptic orbits around the Sun.

Sphere of Influence and Method of Patched Conics

As discussed earlier, the interplanetary mission must be broken up into separate phases according to the primary gravitational bodies that influence the spacecraft. This means that the boundary, or, in the case of three dimensions, the sphere of influence for each gravitational body, must be determined. Since gravitational force is inversely proportional to the distance squared, spacecraft located far away from a planet would primarily be under the influence of the Sun. On the other hand, spacecraft located close to a planet are influenced primarily by the planet. The sphere of influence (SOI) radius which defines this demarcation was determined by Laplace to be:

$$r_{\text{SOI}} = R_{\text{planet}}(\mu_{\text{planet}}/\mu_{\text{Sun}})^{0.4} \tag{14.54}$$

where R_{planet} is the planet's orbital radius around the Sun, which should not be confused with r_{planet}, the equatorial radius for the surface of the planet. For planets in circular orbit around the Sun, as was assumed in this subsection, R_{planet} would correspond to the orbit's semimajor axis length. Table 14.5 lists the SOI radius for the various planets based upon this assumption. It should be noted that the SOI radii for Mercury and Pluto deviate from the average value much more than the other planets since their orbits are relatively eccentric.

A spacecraft located inside the SOI boundary, i.e., $r < r_{\text{SOI}}$, is viewed to be in orbit around the planet, and all values such as the gravitational parameter, radius, and velocity must be referenced with respect to the planet. On the other hand, a

TABLE 14.5 Sphere of Influence Radius for Planets in the Solar System

Planet	Semimajor axis length (a) for heliocentric orbit (km)	μ (km^3/s^2)	r_{SOI} (km)	r_{planet}, equatorial radius of surface (km)	r_{SOI}/r_{planet}
Mercury	5.7910×10^7	2.2032×10^4	1.124×10^5	2.439×10^3	46.1
Venus	1.0820×10^8	3.2486×10^5	6.162×10^5	6.051×10^3	101.8
Earth	1.4960×10^8	3.9860×10^5	9.247×10^5	6.378×10^3	145.0
Mars	2.2794×10^8	4.2828×10^4	5.772×10^5	3.393×10^3	170.1
Jupiter	7.7833×10^8	1.2671×10^8	4.821×10^7	7.1492×10^4	674.3
Saturn	1.4270×10^9	3.7940×10^7	5.456×10^7	6.0268×10^4	905.3
Uranus	2.8710×10^9	5.7816×10^6	5.172×10^7	2.5559×10^4	2023.7
Neptune	4.4971×10^9	6.8713×10^6	8.681×10^7	2.4764×10^4	3505.6
Pluto	5.9135×10^9	1.0209×10^3	3.359×10^6	1.500×10^3	2239.7

spacecraft located outside the SOI is viewed to be in orbit around the Sun. In this case, the gravitational parameter, radius, and velocity must be referenced with respect to the Sun. In order to distinguish between the different phases of an interplanetary mission, upper and lower case lettering will be used for radius and velocity. During phase II, radius and velocity values are measured with respect to the Sun, and they will be denoted with upper case lettering. For phases I and III, lower case lettering will be used for radius and velocity measured with respect to the Earth or the destination planet.

Whenever the SOI boundary is crossed, the reference system must be changed from the Earth to the Sun or from the Sun to the destination planet. In each case, the gravitational parameter corresponding to the new gravitational body must be used. Since the SOI radius is several orders of magnitude less than the heliocentric orbital radius as shown in Table 14.5, it is reasonable to assume that the radius at the start and end of phase II are given by R_{Earth} and R_{planet}, respectively. Finally, the velocity after crossing over the SOI boundary can be determined from the following relationships:

$$\mathbf{V}_{T1} = \mathbf{v}_{\infty,\text{Earth}} + \mathbf{V}_{\text{Earth}} \tag{14.55}$$

$$\mathbf{v}_{\infty,\text{ln}} = \mathbf{V}_{T2} - \mathbf{V}_{\text{planet}} \tag{14.56}$$

In equations (14.55) and (14.56), the subscripts $T1$ and $T2$ refer to the start and end of the phase II heliocentric transfer orbit, respectively. Subscripts ∞, Earth and ∞, In refer to the hyperbolic excess velocity measured with respect to the Earth and the destination planet in phases I and III, respectively. This is the spacecraft's velocity when it is located infinitely far away from the Earth or the planet. In this subsection, planetary orbits around the Sun were assumed to be circular. Thus, V_{Earth} and V_{planet} can be determined from equation (14.9), which in the present situation can be written as:

$$V_{\text{Earth}} = \sqrt{\mu_{\text{Sun}}/R_{\text{Earth}}} \tag{14.57a}$$

$$V_{\text{planet}} = \sqrt{\mu_{\text{Sun}}/R_{\text{planet}}} \tag{14.57b}$$

The phase II heliocentric transfer orbit velocities can be determined from the conservation of energy, equation (14.7) or (14.8), to be:

$$V_{T1} = \sqrt{\frac{2\mu_{\text{Sun}}}{R_{\text{Earth}}} - \frac{\mu_{\text{Sun}}}{a}} \tag{14.58a}$$

$$V_{T2} = \sqrt{\frac{2\mu_{\text{Sun}}}{R_{\text{planet}}} - \frac{\mu_{\text{Sun}}}{a}} \tag{14.58b}$$

Since the phase II heliocentric transfer orbit is a Hohmann transfer, the semimajor axis length is given by equation (14.1) or (14.24), which in the present case can be written as:

$$a = (R_{\text{Earth}} + R_{\text{planet}})/2 \tag{14.59}$$

Since R_{Earth} corresponds to the perihelion in the present Hohmann transfer case, the eccentricity for the phase II heliocentric transfer orbit can be determined from the conic section relationship, equation (14.16), to be:

$$R_{\text{Earth}}(1 + \varepsilon) = a(1 - \varepsilon^2) \quad \text{or} \quad R_{\text{Earth}} = a(1 - \varepsilon) \quad \text{or} \quad \varepsilon = 1 - \frac{R_{\text{Earth}}}{a} \tag{14.60}$$

Figure 14.24 graphically illustrates the change in reference frame for the case of a Hohmann transfer to an outer planet. In equations (14.55) and (14.56), all the velocities (V_{T1}, $v_{\infty,\text{Earth}}$, V_{Earth}, V_{T2}, $v_{\infty,\text{In}}$, and V_{planet}) are given in vector form. For the present case of a Hohmann transfer, however, the equations can be written in scalar form since the velocities are all co-linear and aligned with each other as illustrated in Figure 14.24. Although all three phases are shown in the figure, the focus is the phase II heliocentric portion. Thus, the phase II perihelion is located to the right of the origin (i.e., the Sun) in the middle illustration just as the $+X$

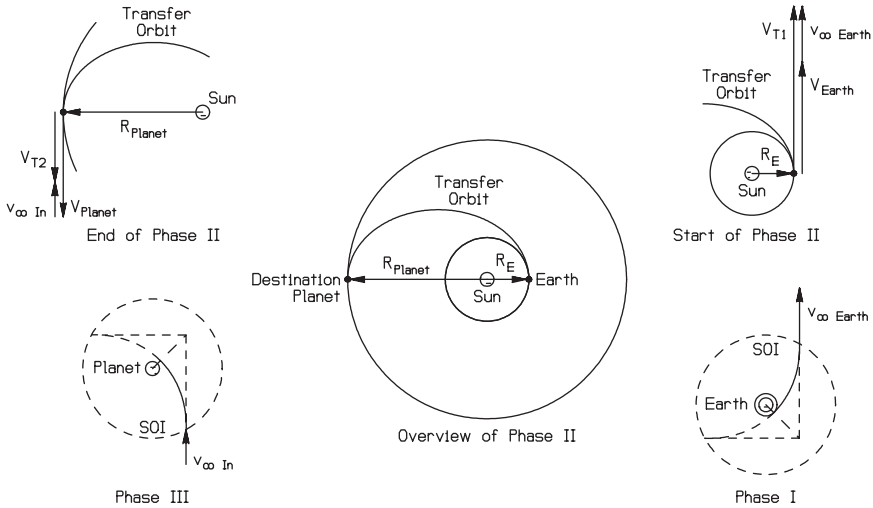


FIGURE 14.24 Interplanetary mission using the method of patched conics for a Hohmann transfer to an outer planet (from Myose 2001).

axis would normally be oriented towards the right. Consequently, the perigee and periapsis for phases I and III are skewed at an angle in the lower right and lower left SOI illustrations. Although the hyperbolic orbits within the SOI of the Earth and the destination planet are shown in the figure, they will be discussed later in more detail. The focus of the following example, just as in the figure, will be on the phase II heliocentric transfer orbit, which is the first design step associated with an interplanetary mission.

As an example, consider sending a spacecraft from Earth to Jupiter using a Hohmann transfer. In this case, the perihelion would correspond to $R_{\text{Earth}} = 1.496 \times 10^8$ km and the aphelion would correspond to $R_{\text{Jupiter}} = 7.7833 \times 10^8$ km according to Table 14.5. This means that the semimajor axis length for the phase II heliocentric transfer orbit is $a = 4.6397 \times 10^8$ km based on equation (14.59). Using equation (14.60), the eccentricity of the phase II heliocentric transfer orbit is determined to be $\varepsilon = 0.6776$, which is an ellipse as expected for a Hohmann transfer. The gravitational parameter for the Sun is $\mu_{\text{Sun}} = 1.3271 \times 10^{11}$ km^3/s^2 according to Table 14.1. The planetary velocities, based on equation (14.57), are $V_{\text{Earth}} = 29.784$ km/s and $V_{\text{Jupiter}} = 13.058$ km/s. Based on equation (14.58), the heliocentric velocities at the start and end of phase II are $V_{T1} = 38.577$ km/s and $V_{T2} = 7.415$ km/s, respectively. The hyperbolic excess velocity required at the end of phase I, based on equation (14.55), is $v_{\infty,\text{Earth}} = 8.793$ km/s. Finally, the hyperbolic excess velocity of the spacecraft upon arrival at Jupiter's SOI is $v_{\infty,\text{In}} = -5.643$ km/s based on equation (14.56). Note that the minus sign simply indicates that the incoming velocity in phase III (i.e., $v_{\infty,\text{In}}$) is oriented in a direction opposite to V_{Jupiter} as shown in Figure 14.24. For a Hohmann transfer to an outer planet such as Jupiter, the hyperbolic excess velocity $v_{\infty,\text{Earth}}$ at Earth departure is oriented in the same direction as V_{Earth} while the hyperbolic excess velocity $v_{\infty,\text{In}}$ upon arrival at the destination planet is oriented opposite to V_{planet}. The reverse is true for a Hohmann transfer to an inner planet such as Mercury or Venus. That is, $v_{\infty,\text{Earth}}$ is opposite to V_{Earth} while $v_{\infty,\text{In}}$ is in the same direction as V_{planet} upon arrival at an inner planet.

Earth Departure

When an interplanetary mission is planned, the phase II part of the mission is analyzed first in order to determine the required phase I hyperbolic excess velocity. Once this is found, the Earth departure condition can be determined from basic orbital mechanics relationships. In a typical mission, the spacecraft is initially placed in a circular low Earth parking orbit by a rocket launch vehicle. Then, the spacecraft is injected into the Earth departure hyperbolic orbit at the perigee of the hyperbola as illustrated in Figure 14.25. The orientation chosen for the figure is such that the perigee of the hyperbola is towards the right of the origin, which is the center of the Earth. The radius and velocity at this injection point will be denoted with a subscript i rather than with a subscript p in order to avoid confusion with the perihelion of the phase II transfer orbit or with the planet's heliocentric orbital radius and velocity. An assumption will be made that the spacecraft's radial distance at the SOI boundary is large enough so that its contribution to the potential energy is negligible. Assuming that $\mu_{\text{Earth}}/r_{\text{SOI}} << v_\infty^2/2$ contributes only a small error, on the order of about 1%. The required injection velocity v_i at the perigee of the Earth departure hyperbola is then a much simpler relationship that can be determined from the conservation of energy, equation (14.7), to be:

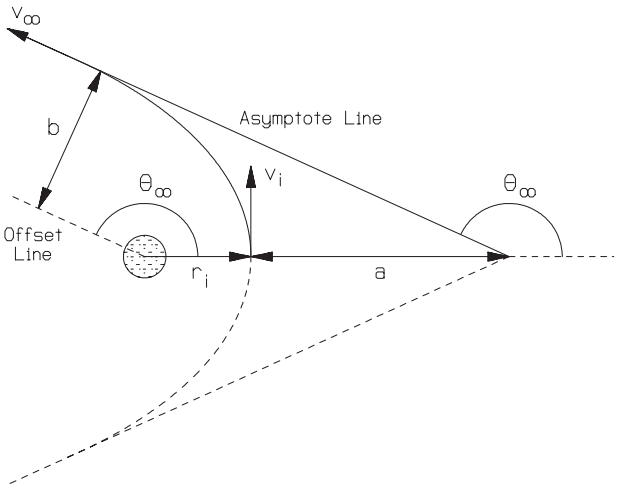


FIGURE 14.25 Phase I hyperbolic Earth departure condition (adapted from Myose 2001).

$$\frac{1}{2}v_i^2 - \frac{\mu_{\text{Earth}}}{r_i} = U = \frac{1}{2}v_\infty^2 - \frac{\mu_{\text{Earth}}}{r_{\text{SOI}}} \approx \frac{1}{2}v_\infty^2 \tag{14.61a}$$

$$v_i = \sqrt{\frac{2\mu_{\text{Earth}}}{r_i} + v_\infty^2} \tag{14.61b}$$

For a circular orbit, the velocity vector is perpendicular to the radius vector. Since the spacecraft is injected into the Earth departure hyperbola starting from a circular parking orbit, the injection velocity v_i is in the same direction as the initial circular parking orbit velocity v_C. This means that the velocity increment Δv would simply be the difference between these two velocities. That is,

$$\Delta v = v_i - v_C \tag{14.62a}$$

$$v_C = \sqrt{\frac{\mu_{\text{Earth}}}{r_i}} \tag{14.62b}$$

Since the total energy is directly related to the semimajor axis length, the Earth departure hyperbola's semimajor axis length is found to be:

$$\frac{1}{2}v_\infty^2 = U = -\frac{\mu_{\text{Earth}}}{2a} \quad \text{or} \quad a = -\frac{\mu_{\text{Earth}}}{v_\infty^2} \tag{14.63}$$

The remaining characteristic of the Earth departure hyperbola, the eccentricity ε, can be found from the conic section relationship, equation (14.16), to be:

$$r_i(1 + \varepsilon) = a(1 - \varepsilon^2) \quad \text{or} \quad r_i = a(1 - \varepsilon) \quad \text{or} \quad \varepsilon = 1 - \frac{r_i}{a} \tag{14.64}$$

The equation for the asymptote angle, equation (14.4), was given earlier in the basic orbital mechanics subsection, and is presented here once again:

$$\theta_\infty = \cos^{-1}(-1/\varepsilon) \tag{14.65}$$

The offset distance b, which is also the semiminor axis length, is given by $b^2 = a^2(1 - \varepsilon^2)$. Substituting equations (14.64) and (14.65) into this semiminor axis length relationship results in three additional relationships for the offset distance (see Myose 2001; Brown 1992):

$$b = \sqrt{a^2(\varepsilon^2 - 1)} = r_i \sqrt{\frac{\varepsilon + 1}{\varepsilon - 1}} = r_i \sqrt{\frac{-2a}{r_i} + 1} = r_i \sqrt{\frac{2\mu_{\text{Earth}}}{r_i v_\infty{}^2} + 1} \tag{14.66}$$

For a Hohmann transfer, the hyperbolic excess velocity $v_{\infty,\text{Earth}}$ must be oriented in the same direction as V_{Earth} in the case of transfer to an outer planet and opposite to V_{Earth} in the case of transfer to an inner planet. Thus, V_{Earth} must be oriented along the offset line that is shown as a dashed line in Figure 14.25 in the case of a Hohmann transfer to an outer planet.

As an example of the Earth departure condition, consider the Earth-to-Jupiter Hohmann transfer discussed earlier. The required hyperbolic excess velocity at the end of phase I was found to be $v_{\infty,\text{Earth}} = 8.793$ km/s. Suppose a spacecraft is injected into this departure hyperbola starting from a circular low Earth parking orbit of radius $r_i = 6563$ km. Based on equation (14.61), the required injection velocity is $v_i = 14.099$ km/s. Since the circular parking orbit velocity is $v_C = 7.793$ km/s from equation (14.62b), the velocity increment required to inject the spacecraft into the Earth departure hyperbola, based on equation (14.62a), is $\Delta v = 6.306$ km/s. This is a substantial Δv requirement in light of the fact that the original rocket booster raised the velocity from zero on the ground to a value of $v_C = 7.793$ km/s for the parking orbit. Based on equations (14.63) and (14.64), the characteristics of the Earth departure hyperbola are semimajor axis length of $a = -5155.4$ km and eccentricity of $\varepsilon = 2.273$, i.e., $a < 0$ and $\varepsilon > 1$, as expected for a hyperbolic orbit. It should be noted that the semimajor axis length and eccentricity for phase I are quite different from those for phase II that were found in the previous example. Care should therefore be taken to avoid confusing the two sets of values in light of the fact that the same notation (a and ε) is used. The asymptote angle of $\theta_\infty = 116.1°$, found using equation (14.65), specifies the injection location measured relative to the Earth's velocity vector V_{Earth}. Based on equation (14.66), the offset distance is $b = 10{,}523$ km. This value is several orders of magnitude smaller than the Earth's orbital radius around the Sun (i.e., $R_{\text{Earth}} = 1.496 \times 10^8$ km). Since $b << R_{\text{Earth}}$, the heliocentric radius at the start of phase II being equal to R_{Earth} is a reasonable assumption to make for the method of patched conics.

Planetary Arrival

When the spacecraft reaches the SOI boundary of the destination planet, the spacecraft can possibly arrive at four different quadrants. For the Hohmann transfer case

to an outer planet, this is illustrated in Figure 14.26. The orientation chosen for the figure is consistent with Figure 14.24, where the phase II perihelion is located to the right of the origin, which is the Sun. If the phase II aphelion radius is less than R_{planet}, then the spacecraft arrives at the sunny side of the planet's SOI as shown in the two lower middle illustrations. If the phase II aphelion radius is larger than R_{planet}, then the spacecraft arrives at the dark side of the planet's SOI (i.e., the side farther away from the Sun) as shown in the lower left and lower right illustrations. Furthermore, the spacecraft's arrival at the planet's SOI can be timed so that it is ahead of the planet, as shown in the two lower left illustrations, or behind the planet, as shown in the two lower right illustrations. It is evident from the two lower-right illustrations that it is not possible for the spacecraft to penetrate the planet's SOI from behind, since the spacecraft moves away from the planet. Therefore, arrival in front of the planet is required for SOI penetration in the case of a Hohmann transfer to an outer planet. The situation is reversed in the case of a Hohmann transfer to an inner planet whereby arrival behind the planet is required for SOI penetration. The interested reader should see Myose (2001) for the penetration and nonpenetration quadrants of a fast transfer, which are more complicated.

Another issue that must be considered is whether the spacecraft impacts the planet or successfully enters an orbit around the planet. If the spacecraft's phase II aphelion radius is exactly equal to R_{planet}, then the spacecraft will be headed directly into the planet and it will obviously impact the planet. Due to the geometry associated with hyperbolic orbits, offset distances much larger than the planet's surface radius will still result in planetary impact. Figure 14.27 illustrates the situation where the spacecraft's periapsis just touches the planet's surface. The orientation chosen for the figure is such that the periapsis of the arrival hyperbola is towards the right of the origin, which is the center of the destination planet. Using the

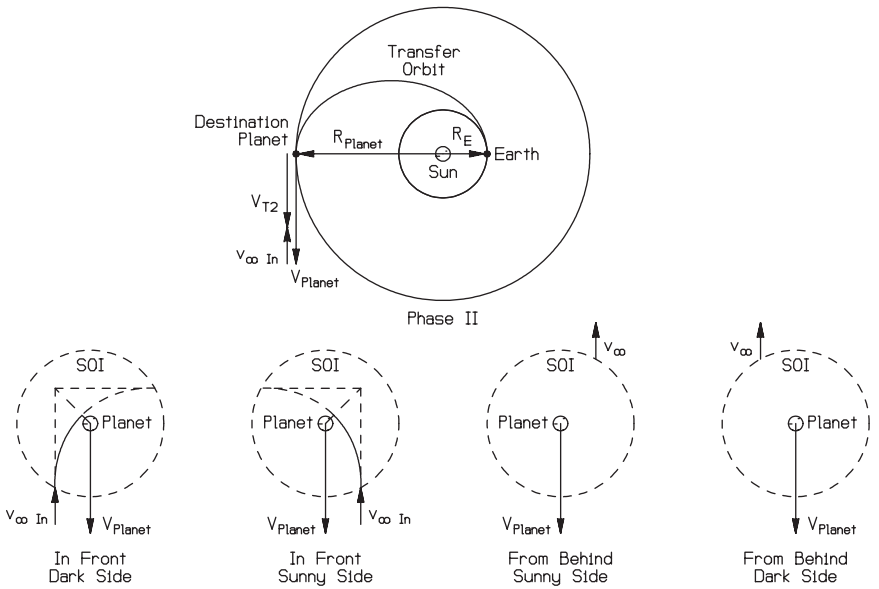


FIGURE 14.26 Encounter with an outer planet after a Hohmann transfer (from Myose 2001).

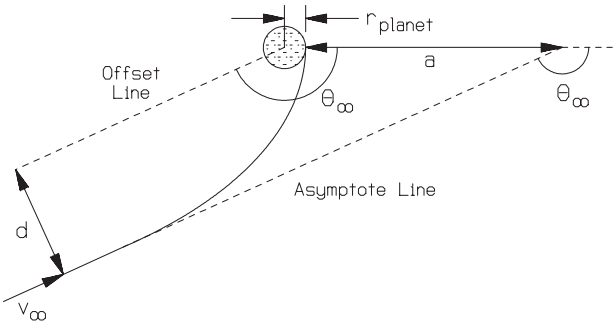


FIGURE 14.27 Impact distance (adapted from Myose 2001).

appropriate gravitation parameter and substituting the planet's surface radius r_{planet} for the periapsis radius r_i into equation (14.66) results in an impact distance d of:

$$d = r_{\text{planet}} \sqrt{\frac{2\mu_{\text{planet}}}{r_{\text{planet}}\, v_\infty^2} + 1} \tag{14.67}$$

If the offset distance b for the arrival hyperbola is less than or equal to the impact distance d (i.e., $b \leq d$), then the spacecraft will impact the planet.

Finally, if a closed orbit such as a circular parking orbit around the planet is to be established, then a braking maneuver must be performed at the periapsis of the arrival hyperbola. The process is essentially the reverse of the Earth departure procedure outlined in equations (14.61)–(14.66), with the exception that the appropriate gravitational parameter for the planet must be substituted for μ_{Earth}. It should be noted, however, that the arrival hyperbola's characteristics, such as the periapsis radius r_i and the offset distance b, are the desired target values. If the actual offset distance b is different from the desired value, then a midcourse correction must be performed (see Brown 1992).

Gravity Assist Maneuver

As discussed earlier, it is possible for a spacecraft to penetrate the SOI after a Hohmann transfer to an outer planet by arriving in front of the planet on either the sunny side or the dark side. The resulting flyby of the planet is illustrated in Figure 14.28 for the sunny-side case and Figure 14.29 for the dark-side case. The orientation chosen for Figure 14.28 is such that the periapsis of the arrival hyperbola is towards the right of the origin, which is the center of the destination planet. On the other hand, the orientation chosen for Figure 14.29 is such that V_{planet} is oriented in the same direction as V_{planet} in Figure 14.28. The orientation chosen for both of these figures is different from the orientation of the same situations shown in the two lower-left illustrations of Figure 14.26. Notice that the hyperbolic orbit resulting from a sunny-side penetration shown in Figure 14.28 results in a counterclockwise flyby. On the other hand, the hyperbolic orbit resulting from a dark-side penetration shown in Figure 14.29 results in a clockwise flyby. Both of these orbits, however, are trailing-side flybys since the orbital path never cuts across V_{planet}. In general,

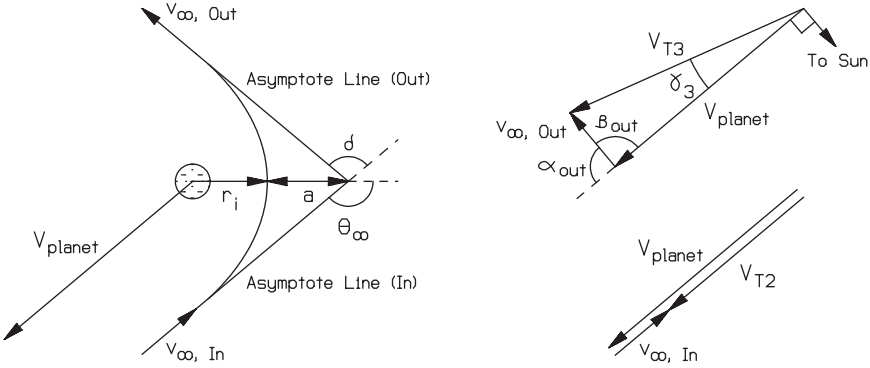


FIGURE 14.28 Counterclockwise planetary flyby after a Hohmann transfer to an outer planet (adapted from Myose 2001).

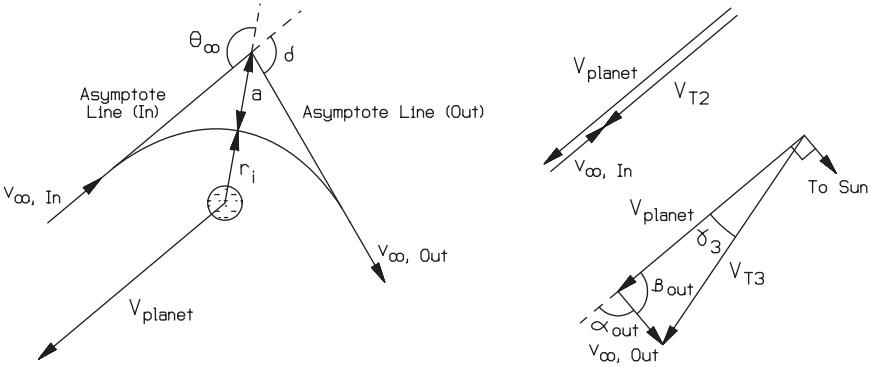


FIGURE 14.29 Clockwise planetary flyby after a Hohmann transfer to an outer planet (adapted from Myose 2001).

trailing-side flybys result in an increase in the spacecraft's heliocentric velocity, i.e., its velocity after exiting the planet's SOI increases. This behavior of velocity increase can be explained from two different physical perspectives. One perspective is the gravity assist view, where gravity acts to "pull" the spacecraft as it flybys the rear of the planet. This "pulling" motion is in the same direction as the spacecraft's heliocentric travel direction, and therefore increases the spacecraft's heliocentric velocity. Another perspective is the "slingshot" view, where the hyperbolic excess velocity's direction is changed. In this view, $v_{\infty,\mathrm{In}}$ opposes the heliocentric travel direction (i.e., opposite to V_{planet}) when it penetrates the SOI, but it is turned by the time it departs so that $v_{\infty,\mathrm{Out}}$ is less opposed. In comparison to trailing-edge flybys, described above, leading-side flybys cut across V_{planet}, resulting in a reduction in the spacecraft's heliocentric velocity. Such a situation can occur from a planetary flyby after a Hohmann transfer to an inner planet or after a fast transfer. The interested reader should see Myose (2001) and Hale (1994), which address the Hohmann transfer to an inner planet case and the fast transfer case.

A number of different angles associated with the hyperbolic flyby and the heliocentric velocities need to be defined in order to explain the change in velocity resulting from a planetary flyby. From a "slingshot" perspective, the important parameter that must be determined is the amount of turning accomplished by the hyperbolic flyby. As can be seen from Figures 14.28 and 14.29, this turning angle δ is related to the asymptote angle according to:

$$2\theta_\infty = \delta + 180° \tag{14.68}$$

Substituting equation (14.65) and then using a trigonometric identity results in:

$$\delta = 2 \sin^{-1}(1/\varepsilon) \tag{14.69}$$

The flight path angle γ is defined to be the angle between the velocity of interest and the horizon line, where the horizon line is defined to be perpendicular to the radius vector. Normally, the radius vector is drawn from the center of the gravitational body of interest. In the case of the heliocentric velocity V_T, it is evident that flight path angle γ should be measured with respect to the Sun. In the case of the hyperbolic excess velocity v_∞, the appropriate gravitational body is not so clear. The normal perspective would be that v_∞ is the hyperbolic excess velocity measured with respect to the destination planet (or the Earth when appropriate). Another perspective, however, is that v_∞ is the difference between the heliocentric velocity V_T and the planet's orbital velocity around the Sun V_{planet}. In this case, the appropriate gravitational body would be the Sun, and the horizon line should be drawn with respect to the Sun. In order to avoid confusion, α will be used to denote the heliocentric flight path angle for the hyperbolic excess velocity as shown in Figures 14.28 and 14.29. Finally, β will be defined as the supplementary angle to α, so that

$$\alpha + \beta = 180° \tag{14.70}$$

In the case of a Hohmann transfer to an outer planet, $v_{\infty,\text{In}}$ opposes V_{Earth} so that $\alpha_{\text{In}} = 180°$, which means that $\beta_{\text{In}} = 0°$. The heliocentric flight path angle α_{Out} associated with the outgoing hyperbolic excess velocity $v_{\infty,\text{Out}}$ will depend upon the type of flyby. For a counterclockwise flyby,

$$\alpha_{\text{Out}} = \alpha_{\text{In}} - \delta \tag{14.71}$$

For a clockwise flyby,

$$\alpha_{\text{Out}} = \alpha_{\text{In}} + \delta \tag{14.72}$$

where flight path angles greater than 180° correspond to the situation where the velocity lies below the horizon line. Thus, the range of values for α (as well as β) should be limited to $\pm 180°$. Once the outgoing heliocentric flight path angle α_{Out} is determined within this range, the supplement angle β_{Out} can be found using equation (14.70). Since no engine burns are performed during the flyby, the outgoing hyperbolic excess velocity $v_{\infty,\text{Out}}$ is exactly the same magnitude as the incoming hyperbolic excess velocity $v_{\infty,\text{In}}$. Only the direction has been changed. Once the spacecraft completes the flyby and exits the planet's SOI, the spacecraft is in a new conic section patch, i.e., phase IV, which is a heliocentric orbit. The new heliocentric velocity V_{T3} and its flight path angle γ_{T3} corresponding to the start of phase IV can be determined from the cosine and sine laws as follows:

$$V_{T3} = \sqrt{v_\infty^2 + V_{planet}^2 - 2v_\infty V_{\text{planet}} \cos \beta_{\text{Out}}} \tag{14.73}$$

$$\gamma_{T3} = \sin^{-1}\left(\frac{v_\infty}{V_{T3}} \sin \beta_{\text{Out}}\right) \tag{14.74}$$

As an example, consider a Jupiter flyby where the periapsis of the hyperbolic orbit is r_i = 72,000 km. Note that this periapsis radius is greater than Jupiter's surface radius of 71,492 km listed in Table 14.5. In an earlier example, the hyperbolic excess velocity at Jupiter arrival and Jupiter's orbital velocity were found to be $v_{\infty,\text{In}}$ = 5.643 km/s and V_{Jupiter} = 13.058 km/s, respectively. Although the equations outlined in the Earth departure condition will be used, the appropriate gravitational parameter, from Table 14.5, is $\mu_{\text{Jupiter}} = 1.2671 \times 10^8$ km^3/s^2 in this case rather than μ_{Earth}, which is written in the equations. Substituting μ_{Jupiter}, $v_{\infty,\text{In}}$, and r_i into equations (14.63)–(14.65) results in $a = -3.979 \times 10^6$ km, ε = 1.018, θ_∞ = 169.2°, and $b = 7.6 \times 10^5$ km. These values are the characteristics required for the arrival hyperbola in order to obtain a periapsis of 72,000 km. It should be noted that the offset distance of $b = 7.6 \times 10^5$ km is well within Jupiter's SOI radius of 4.821×10^7 km and is substantially smaller than Jupiter's orbital radius around the Sun of $R_{\text{Jupiter}} = 7.7833 \times 10^8$ km given in Table 14.5. Substituting ε = 1.018 into equation (14.69) or θ_∞ = 169.2° into equation (14.68) results in a turning angle of δ = 158.4°. This is a very large amount of turning, which leads to a substantial increase in the heliocentric velocity, as will be shown below. As discussed earlier, α_{In} = 180° and β_{In} = 0 in the case of a Hohmann transfer to an outer planet. If the SOI penetration occurs in front on the sunny side (i.e., $R_{T2} = R_{\text{Jupiter}} - b$) as shown in the lower middle-left illustration of Figure 14.26, then a counterclockwise flyby will result, and α_{Out} = 21.6° according to equation (14.71). In this case, β_{Out} would then be 158.4° based on equation (14.70). If the SOI penetration occurs in front on the dark side (i.e., $R_{T2} = R_{\text{Jupiter}} + b$) as shown in the lower-far left illustration in Figure 14.26, then a clockwise flyby will result, and α_{Out} = 338.4° = −21.6° according to equation (14.71). In this case, the outgoing supplement angle would be β_{Out} = 201.6° = −158.4° based on equation (14.70). Since cosine is an even function, the heliocentric velocity V_{T3} at the start of phase IV does not depend upon the sign of the outgoing supplement angle β_{Out}. This means that the heliocentric velocity V_{T3} = 18.422 km/s for both counterclockwise and clockwise flybys based on equation (14.73). This compares to a heliocentric velocity of V_{T2} = 7.415 km/s that the spacecraft had before entering Jupiter's SOI. Based on equation (14.74), the heliocentric flight path angle at the start of phase IV is γ_{T3} = 6.475° for counterclockwise flyby and γ_{T3} = −6.475° for clockwise flyby. A positive flight path angle indicates that the velocity vector is oriented above the horizon line and the spacecraft therefore moves farther away from the Sun. On the other hand, a negative flight path angle indicates that the velocity vector is oriented below the horizon line and the spacecraft moves closer to the Sun.

Once the heliocentric velocity and flight path angle at the start of phase IV are found, the characteristics of this phase IV heliocentric orbit can be determined. The appropriate gravitational parameter is $\mu_{\text{Sun}} = 1.3271 \times 10^{11}$ km^3/s^2, and the heliocentric radius at the start of phase IV, i.e., upon exiting Jupiter's SOI, is $R_{\text{Jupiter}} = 7.7833 \times 10^8$ km. The angular momentum and total energy based on equations (14.12) and (14.7) are:

$$U = \tfrac{1}{2}V_{T3}^2 - (\mu_{\text{Sun}}/R_{\text{Jupiter}}) = -0.8210 \text{ km}^2/\text{s}^2 = -\mu_{\text{Sun}}/(2a) \quad \text{or}$$

$$a = 8.082 \times 10^{10} \text{ km} \quad H = R_{\text{Jupiter}} V_{T3} \cos \gamma_{T3} = 1.425 \times 10^{10} \text{ km}^2/\text{s}$$

Using the conic section relationship given by equation (14.16), the semilatus rec-

tum, eccentricity, and perihelion for the phase IV heliocentric orbit are found to be $p = 1.529 \times 10^9$ km, $\varepsilon = 0.9905$, and $R_{\text{perihelion}} = 7.68 \times 10^8$ km. It should be noted that the Jupiter flyby has increased the heliocentric eccentricity from a moderately eccentric one at $\varepsilon = 0.6776$ for phase II to a highly eccentric one with $\varepsilon = 0.9905$ for phase IV. Based on equation (14.1), the aphelion for the phase IV heliocentric orbit is $R_{\text{aphelion}} = 1.609 \times 10^{11}$ km, which is well beyond the average radial position of Pluto since its semimajor axis length is 5.9135×10^9 km according to Table 14.5. Both the counterclockwise flyby and the clockwise flyby result in the same overall orbital characteristics for phase IV. The only difference between the two is that they are located at different points on the orbit. Using the orbit equation, equation (14.3) or (14.15), the true anomaly angle upon exiting Jupiter's SOI is found to be $\theta_{T3} = 13°$ for the counterclockwise flyby, whereas $\theta_{T3} = -13° = 347°$ for the clockwise flyby. That is, the spacecraft is headed towards the aphelion after a counterclockwise flyby, whereas it would be headed towards the perihelion after a clockwise flyby.

Intercept Problem

In the previous example, the heliocentric orbit was substantially energized by a Jupiter flyby to the point that the phase IV heliocentric orbit is nearly parabolic with $\varepsilon = 0.9905$ and its aphelion is beyond the outer edges of the solar system. It would not be possible to obtain a significantly more energetic phase IV heliocentric orbit from a single Jupiter flyby alone, since the hyperbolic flyby's periapsis is already close to Jupiter's surface. If multiple flybys are used, however, it is possible to increase the orbital energy and then completely escape the solar system. This was accomplished by the Voyager spacecraft through a series of gravity assists from Jupiter, Saturn, Uranus, and Neptune. The planetary alignment required to obtain such a sequence of planetary flybys is very rare, and the technical issue of timing the spacecraft launch is extremely complex. Although only the most basic case will be considered, this timing issue is the subject of the final part of this subsection.

The key issue associated with the intercept problem is the difference in orbital period between the spacecraft and the planets. Kepler's third law, given by equation (14.18), shows that the orbital period is proportional to the semimajor axis length (raised to the $\frac{3}{2}$ power). This means that the smaller the orbital radius, the faster an object travels in its orbit. If the spacecraft is delivered to an outer planet using a Hohmann transfer, the Earth is located in the lowest orbit since $R_{\text{Earth}} < a_{\text{Hohmann transfer}} < R_{\text{planet}}$. This means that it takes less time for the Earth to complete one lap in its orbit, i.e., traveling 360° of true anomaly compared to the spacecraft in the Hohmann transfer or the destination planet. To view this from a different perspective, suppose the time of flight is fixed to the time it takes for the spacecraft to travel from the Earth to the outer planet. During this time, the spacecraft travels 180° of true anomaly in the Hohmann transfer. In comparison, the destination planet travels less than 180°, say $\theta_{\text{planet}} = 180° - K_1$, since it moves at the slowest rate while the Earth travels more than 180°, say $\theta_{\text{Earth}} = 180° + K_2$, since it moves at the fastest rate. K_1, called the lead angle, is the angle difference between the Earth and the destination planet at the time the spacecraft is launched from the Earth. K_2, called the arrival difference angle, is the angle difference between the Earth and the destination planet at the time the spacecraft arrives at the destination planet. These angle differences are shown graphically in Figure 14.30.

The equations associated with the intercept problem are simply an extension of the time of flight relationships discussed in subsection 14.1. If the semimajor axis

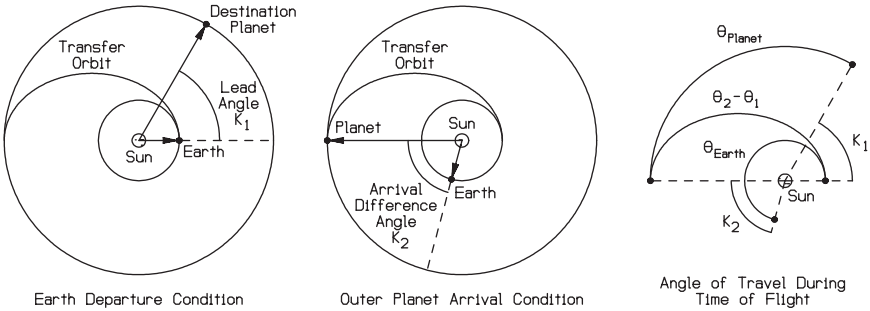


FIGURE 14.30 Planetary alignment for an outbound Hohmann transfer to an outer planet (adapted from Myose 2001).

length of the Hohmann transfer orbit is known, then its time of flight TOF can be determined from equation (14.18) since the time of flight for a Hohmann transfer is one half its orbital period. For the Earth and the destination planet, which are in circular orbits, the angular velocity is a constant and is given by the relationship

$$n_{\text{Earth}} = \sqrt{\frac{\mu_{\text{Sun}}}{(R_{\text{Earth}})^3}} \tag{14.75a}$$

$$n_{\text{planet}} = \sqrt{\frac{\mu_{\text{Sun}}}{(R_{\text{planet}})^3}} \tag{14.75b}$$

The angle traveled by the Earth and the destination planet, given in units of radians, would then be the product of the angular velocity n and the time of flight TOF. Consequently, the lead angle K_1 and the arrival difference angle K_2 would be given by:

$$K_1 = (\theta_2 - \theta_1) - \theta_{\text{planet}} = (\theta_2 - \theta_1) - n_{\text{planet}} \cdot \text{TOF} \tag{14.76a}$$

$$K_2 = \theta_{\text{Earth}} - (\theta_2 - \theta_1) = n_{\text{Earth}} \cdot \text{TOF} - (\theta_2 - \theta_1) \tag{14.76b}$$

where $\theta_2 - \theta_1 = 180° = \pi$ in the case of a Hohmann transfer orbit.

The fact that the Earth has traveled ahead of the destination planet poses a problem for a return-type mission. If a Hohmann transfer is to be used in the return mission, then the Earth must be located behind the outer planet. Thus, proper timing associated with a return mission is a more complicated problem involving some waiting time. The interested reader should see Myose (2001), Hale (1994), Brown (1992), and Wiesel (1997) for further details.

REFERENCES

Bate, R. R., Mueller, D. D., and White, J. E. 1971. *Fundamentals of Astrodynamics*, Dover, New York.

Beatty, J. K. and Chaikin, A. 1990. *The New Solar System,* Cambridge University Press, Cambridge.

Blevis, Z. O. and Stoolman, L., eds. 1981. *Hughes Aircraft Company Geosynchronous Spacecraft Case Histories,* Hughes Aircraft Company, El Segundo, CA.

Brown, C. D. 1992. *Spacecraft Mission Design,* American Institute of Aeronautics and Astronautics, Washington, DC.

Chobotov, V. A., ed. 1991. *Orbital Mechanics,* American Institute of Aeronautics and Astronautics, Washington, DC.

Damon, T. D. 1995. *Introduction to Space,* Krieger Publishing, Malabar, FL.

Escobal, P. R. 1976. *Methods of Orbit Determination,* Krieger Publishing, Malabar, FL.

Griffin, M. D. and French, J. R. 1991. *Space Vehicle Design,* American Institute of Aeronautics and Astronautics, Washington, DC.

Hale, F. J. 1994. *Introduction to Space Flight,* Prentice-Hall, Englewood Cliffs, NJ.

Isakowitz, S. J. 1995. *International Reference Guide to Space Launch Systems,* American Institute of Aeronautics and Astronautics, Washington, DC.

Kaplan, M. H. 1976. *Modern Spacecraft Dynamics and Control,* John Wiley & Sons, New York.

Myose, R. Y. 2001. *Fundamentals of Astronautical Engineering,* Aerospace Engineering 415 Class Notes Version 0.6, Copyright 2001, Department of Aerospace Engineering, Wichita State University, Wichita, KS.

Prussing, J. E. and Conway, B. A. 1993. *Orbital Mechanics,* Oxford University Press, Oxford.

Vallado, D. A. 2001. *Fundamentals of Astrodynamics and Applications,* Microcosm Press, Malabar, FL.

Wiesel, W. E. 1997. *Spaceflight Dynamics,* McGraw-Hill, New York.

SECTION 15
SPACECRAFT

Section Editor: Nick Larter

PART 1

INTRODUCTION

Nick Larter

INTRODUCTION

Man has only been able to design and build successful spacecraft since 1957, when the Russian Sputnik spacecraft was launched. Since then space technology and activity have grown enormously. Commercial interests have led to the large array of communications satellites that now orbit the earth, the need for military intelligence has resulted in very high spatial resolution imaging of the earth's surface that has both changed the nature of war and resulted in far safer transport systems, and the thirst for scientific understanding has resulted in the continuous monitoring of earth and space by orbiting craft and the launch of mission craft that have landed on the Moon, Venus, and Mars.

Because of the high costs and risks involved in space technology, it can be characterized as being relatively conservative. A new spacecraft design will often be made up of many elements of well-established technology. This does not, however, lead to much repetition in craft, as is found in the aircraft industry. Spacecraft tend to be tailored for a given payload and mission. It is consequently more difficult to generalize about spacecraft than aircraft. This Section therefore describes many spacecraft failures through the use of examples.

Information relevant to spacecraft is also detailed in other Sections of the book; the space environment in Section 16; spacecraft astrodynamics and mission analysis in Section 14; propulsion systems in Section 8; and telecommunications and telemetry in Section 11.

To give an idea of expenditure on space, Tables 15.1 and 15.2 give the space budgets of the Department of Defense and NASA, respectively.

TABLE 15.1 Department of Defense Space Programs Procurement (Millions of Dollars)

Agency and Program	1998		1999E		2000E	
			Air Force			
Defense Support Program	$85.8	$17.6	$88.7	$11.8	$11.6	$7.5
EELV	—	23.3	—	259.1	70.8	324.8
Medium launch vehicles	195.5	2.0	175.1	7.3	64.9	1.2
Milstar	—	609.7	—	546.5	—	361.3
NAVSTAR C PS	162.6	96.4	93.6	94.4	170.9	98.9
SBIRS—low	—	213.5	—	192.2	—	229.0
SBIRS—high	—	337.9	—	539.4	—	328.7
Spaced-based laser	118.3	—	125.0	33.8	75.0	63.8
Titan launch vehicles	453.3	62.4	583.8	77.2	431.3	45.4
			Army			
DSCS	$88.1	$13.8	$110.2	$16.1	$80.3	$9.0

E Estimate. Latest year reflects Administration's budget proposal.
Key: DSCS = Defense Satellite Communications System
EELV = Evolved Expendable Launch Vehicle
GPS = Global Positioning System
SBIRS = Space-Based lnfrared System
Source: Department of Defense, *Program Acquisition Costs by Weapon System.*

TABLE 15.2 NASA Budget (millions of inflation-adjusted FY 1999 dollars)

FY	Inflation factors	NASA total	NASA space	DoD	Other	Energy	DoC	Ag	NSF	DoT	EPA	Total space
1985	1.440	10,91	9,976	18,393	836	49	609	22	153	0	0	29,204
1986	1.395	10,89	9,997	19,710	660	49	431	32	145	0	0	30,367
1987	1.362	14,89	13,368	22,196	630	65	279	26	147	1	0	36,193
1988	1.327	12,03	11,049	23,472	978	320	467	24	147	1	0	35,499
1989	1.286	14,18	12,987	23,031	720	125	387	27	149	4	6	36,738
1990	1.238	15,26	14,191	19,337	634	98	301	31	155	5	6	34,162
1991	1.193	16,72	15,563	16,917	831	299	299	31	156	5	6	33,312
1992	1.15	16,46	15,176	17,273	884	256	376	33	167	5	8	33,333
1993	1.125	16,09	14,690	15,862	785	186	364	28	156	4	9	31,366
1994	1.097	15,98	14,281	14,439	659	81	342	34	154	5	9	29,379
1995	1.072	14,85	13,446	11,411	674	64	377	34	151	6	8	25,531
1996	1.05	14,58	13,195	12,088	787	48	496	39	154	6	6	26,071
1997	1.03	14,12	12,831	12,079	750	36	461	40	157	6	6	25,659
1998	1.013	13,82	12,478	12,517	778	64	462	43	154	6	6	25,773
1999	1	0	0	0	0	0	0	0	0	0	0	0

Source: National Aeronautics and Space Administration.

PART 2

PLATFORM TECHNOLOGIES

Tara Dalton

15.1 SPACECRAFT STRUCTURE

Much of the discussion of structures in the Section 9 will apply equally well to spacecraft structures—particularly those issues relating to weight control and the lifting surfaces used on the Space Shuttle.

The primary mechanical functions of the spacecraft structure, where no lifting surfaces are used, are to support the various subsystems, particularly during the launch, where stresses tend to be highest; to allow for whatever deployments maybe used on the mission, to provide overall rigidity; and to allow the craft to be handled on the ground during assembly and transportation.

The materials requirements are that the structure be both light and rigid. The forces along the main axis of the spacecraft are generally the greatest, so the structure is often built around a more rigid core. Typical materials used are aluminum alloys, honeycomb panels, and carbon fibers. Table 15.3 gives properties for some commonly used materials.

Before any material may be deployed in space, it must meet rigorous criteria, both in determining its material properties on earth and in space and in the way the material is deployed in the design. Both NASA and the ESA publish standards and guidelines covering all significant aspects of this. As an example, NASA's Materials and Processing Policy is reproduced below.

Policies

- Materials and Process selection and control requirements shall be developed covering all phases of NASA space flight programs based on functional, test, transportation, storage, mission and safety requirements.
- A materials and processes selection, control and verification plan shall be established and used to implement materials and processes requirements. The implementation approach to be used for materials and processes selection, control and verification shall be documented between the hardware developer and the responsible NASA center materials organization. NASA center implementation procedures shall be used for Government in-house programs.
- The content of the materials and processes program shall be limited to the minimum that will result in a cost effective approach commensurate with acceptable risk.
- Control of fabrication, test, qualification and verification of hardware shall be based on documented materials and processes standards.
- Existing standards and requirements shall be used, wherever possible from the following sources: international, national, government and performing organization.

Processes

The materials and processes selection, control and verification plan should define the following:

- The management process to be used for implementing requirements, and the interfaces among the organizations including subcontractors and suppliers.

TABLE 15.3 Materials Properties

	Density (kg/m^3)	Young's modulus E (Gpa)	Yield strength f (Mpa)	Thermal expansion ($\mu m/mK^{-1}$)	Fracture toughness (MPa m)	Fatigue strength (Mpa)	Comment
Aluminum alloy							
6061 . T6	2,800	68	276	23.6	186	97	Good CR
7075 . T6	2,700	71	503	23.4	24	159	Prone to SCR
Magnesium alloy							
A2 31 B	1,700	45	220	26			Prone to SCR
ZK 60 A.T5							
Extrn	1,700	45	234	26		124	
Titanium alloys							
T1-6A1-4V	4,400	110	825	9	75	500	
Beryllium alloys							
S 65 A	2,000	304	207	11.5			Hot pressed
S R 200 E			345				sheet
Ferrous alloys							Low
INVAR		150	275/415	1.66			expansion Ferromagnetic
Stainless steel							
AM 350 (SCTY850)	7,700	200	1034	11.9	40/60	550	Austentic
304L Ann	7,800	193	170	17.2			
Composites							
Kevlar 49 0°	1,380	76	1379′	−4			Structural
(Aramid fiber)							Members
90°	1,380	5.5	9.6	57			Pressure
Graphite epoxy							Vessels
sheets	1,620	282	586	−11.7 (Longitudinal)			Rocket casings
(undirectional)				29.7 (Transverse)			
GY70/934							Sheet

CR = Corrosion resistance
SCR = Stress corrosion cracking

- The hardware selection, control and verification process for approval, prior to release, of all materials and processes identified in each engineering drawing and engineering change. The detailed technical requirements for performance, safety and durability, including tailoring as necessary.
- The methods for documenting the acceptability of materials and processes usage in specific applications where materials test data indicate a potential performance limitation.
- Inter-center Agreements. NASA Inter-center agreements defining the detailed certification process or implementation of this policy are encouraged to mutually accept hardware, facilitate communication, and preclude redundant materials and processes review activities.

Materials and Process Selection

Technical factors to be addressed in materials and process selection shall include, but are not limited to the following:

- Material and process selection should consider factors such as controllability, reproducibility, inspectability, sensitivity to environmental factors, verification of properties, and special requirements for personnel training and certification.

- Materials Data Base. The NASA Materials and Processes Technical Information System (MAPTIS) database (available from MSFC electronically or in hard copy as MSFC-HDBK-527/JSC 09604, MSFC-HDBK-527/JSC-09604) shall be consulted for materials selection.
- Design allowables—Materials shall be selected for use only in the range of properties documented in accepted sources.
- Environment Compatibility—Materials shall be evaluated for compatibility with the anticipated ground as well as the natural and induced space environment to which they are exposed. Examples of materials evaluation rating/results can be found in MSFC-HDBK-527/JSC-09604 and NHB 8060.1
- Fluid Compatibility—Materials used in fluid systems shall be evaluated for compatibility with the fluid in the worst-case environment to which they are exposed. Standard evaluation methods and criteria for oxygen and propellant fuels and oxidizers are contained in NHB 8060.1.
- Flammability—Materials shall be evaluated for flammability in the worst-case environment to which they are exposed. Standard evaluation methods and criteria are contained in reference NHB 8060.1.
- Toxic Offgassing—Materials used in crew-habitable areas shall be evaluated for offgassing of potentially toxic trace gas contaminants. Standard evaluation methods and criteria are contained in reference NHB 8060.1.
- Thermal Vacuum Stability—Materials used in a space vacuum environment shall be evaluated for contamination effects of vacuum outgassing on critical surfaces. A standard screening test is contained in reference ASTM E595.
- Corrosion—Metallic materials shall be evaluated for general corrosion and stress corrosion cracking. Stress corrosion characteristics and evaluation criteria are contained in MSFC-SPEC-522.

The key material characteristics addressed are the specific strength, which is a measure of the materials strength per unit weight at launch, the stiffness, which determines the vibrational modes present and the relative position of subelements of the craft, stress corrosion resistance, fracture and fatigue resistance, sublimination erosion, and thermal expansion coefficient. The last of these is particularly important in space, where some materials are exposed to widely varying temperatures whilst being required to accurately maintain their relative position.

Just as in the aeronautics industry, composite materials are finding interesting uses in space. Composites are tailored to give the maximum strength in the major load-bearing direction by fibers being aligned that way. Titanium end fittings are often used for securing composite struts. By careful laying up of the composite, the thermal expansion coefficient can also be tailored to the application. Panels are often produced as honeycombs. These consist of an aluminum honeycomb sandwiched between an upper and lower skin. They are very stiff for their mass. Material properties for these cannot be simply stated because they depend upon the design, and designs tend to be application specific. Particular care must be given to verifying the honeycomb trapped volume of air and in the design of the attachment points to the panels.

Each new design of craft must be verified before it is launched. The approach is to load a prototype to greater strength and operational loads than are expected in service. Flight acceptance may require a load of 110% of design, for instance. A range of different vibration inputs is used to verify the vibrational modes of the craft and determine its vibrational stiffness. Part of this test will be to shock test the system. The design and test will be in part determined by the individual design.

15.2 THERMAL CONTROL

Thermal control is necessary in spacecraft to ensure the comfort of any people on board, the survival of life forms being used for scientific experiments, and the reliability of the constituent engineering subsystems.

The three important sources and sinks for the overall spacecraft are the sun at of order 6,000 K surface temperature, space at 4 K, and the heat generated within the craft. Consideration may also be given to radiation reflected and emitted by nearby planets. Most of the craft must be kept at temperatures around those found on earth. Typical operating ranges are given below.

Antenna	−150°C to +80°C
Operating electronics	−5°C to +40°C
Hydrazine fuel	9°C to 40°C
Bearings	−45°C to 65°C
Solar cells	−60°C to 55°C
Operating batteries	0°C to 10°C

Typical values of solar radiation intensity and albedo values are given in Table 15.4.

The Sun acts as a blackbody with an effective temperature of 5,800 K. The radiation intensity from the sun is given by:

$$\dot{q} = \sigma\, 5{,}800^4 \quad W/m^2$$

where σ is the Stefan–Boltzmann constant, 5.6×10^{-8} W/m^2K^4. This may be factored by the distance squared from the sun to give the local intensity.

The fraction that is reflected by a planet is called the albedo. For the Earth, albedo values range from 0.8 for cloud cover to 0.3 for vegetation. Typical albedo values are also given in Table 15.4.

Planets also emit radiation. The spectral distribution is described by the variation of emissive power with wavelength. For the Earth this variation is far from a blackbody curve. It is usual to assign a single mean value, therefore, of 237 W/m^2. A visibility or shape factor, relating the fraction of emitted energy that falls on the

TABLE 15.4 Solar Radiation Intensity as a Multiple of That Found in Earth's Orbit (at 1 AU), 1371 W/m^2, and Albedo Values

Planet	Solar radiation intensity	Albedo
Mercury	6.67	0.06
Venus	1.91	0.61
Earth	1	0.34
Mars	0.43	0.15
Jupiter	0.037	0.41
Saturn	0.011	0.42
Uranus	0.002	0.45
Neptune	0.001	0.52
Pluto	0.00064	0.16

body of interest, is used to calculate the radiation heat balance. Details of this are given in Section 16.

Thermal control of spacecraft is achieved by ensuring that the absorbed and generated thermal energy is balanced by that emitted. A common target temperature for the core of a spacecraft is 25°C. Care must be taken because many of the materials commonly used have values of absorbance that differ greatly from their value of emittance. Spacecraft thermal design handbooks describe those effects and the effects of wavelength.

Thermal design may be broken down into its passive and active elements. The former uses material and fixed geometry design to achieve control, while the later relies upon sensors and activators to respond to given conditions. The primary passive elements rely upon choosing the correct ratio of the absorbtivity and emmisivity, $\alpha/\in$, of the exterior surface. Solar reflectors reflect the solar input while still radiating, and thus have low $\alpha/\in$. Solar absorbers achieve the opposite with a high $\alpha/\in$ ratio.

There are many active devices either in use or proposed as design possibilities. Some examples are actuated external shutters to vary the radiation properties and geometry of a surface, controlled heaters, variably finned surfaces, thermodynamic cooling, and a fluid transport system within the craft.

The final thermal test is when the complete spacecraft is tested in a vacuum tank which simulates solar, albedo, and Earth radiation. The increasing size of spacecraft causes the greatest design challenge here.

15.3 SOLAR CELLS

The electric power supply system for a satellite is designed to provide a continuous, low weight and reliable sources of electricity. It can be broken into three parts: the primary source of energy, which is now almost always solar, the secondary source, which is the battery backup to the primary, and the conditioning and protection circuits.

The normalized solar flux at one astronautical unit, the earth's mean position, from the Sun is of order 1353 W/m^2. The Sun can be modeled as a blackbody source at 5,800 K.

Solar cells operate by generating a voltage at a p–n junction when it is subjected to a solar flux. Solar cells generally produce a constant current variable voltage source over their designed range of operation. Their output tends to drop with increasing temperature and with increasing life in operation. Their beginning of life efficiency is of order 1.5 times that found at end of life.

The basic solar cell equations are as follows:
The delivered electrical power is:

$$P_c = \Phi e A(1 - l) \quad \text{(W)}$$

where Φ = solar flux (W/m^2)
e = cell efficiency ($\approx$15% for silicon)
A = area
l = parasetic losses ($\approx$ 10%)

The solar flux is given by:

$$\Phi = W\left(\frac{a}{d}\right)^2 \cos\theta$$

where θ = panel inclination
a = mean Sun–Earth distance
d = distance to Sun
W = nominal solar flux

The efficiency is a function of time, sometimes modeled for silicon cells as a relationship between the end-of-life (EOL) and beginning-of-life (BOL) efficiency:

$$e_{EOL} = e_{BOL}e^{-0.043T}$$

where T is the time in orbit in years.

The necessary surface area for a solar panel is given by:

$$A = ns/f$$

where s = area of each cell
n = number of cells
f = filling efficiency (as high as 90%)

A typical silicon cell specification for a monocrystalline chip of 200μm thickness is; e_{BOL} = 12.6%, e_{EOL} = 8.7% (0.17 and 0.12 kw/m^2 respectively), weight 0.46kg/m^3, the cell cost 1,000 \$/m^2.

Cells are connected in series and in parallel with the aim of producing the required voltages, currents, and reliability. Diodes are used in series to isolate a line of cells should they become defective.

On spin-stabilized satellites the cells are arranged on the skin of the satellite. Because of the curved surface and the time in shadow, the theoretically required surface area is π times that of a planar panel. In practice this is of order two. For three axis-stabilized satellites flexible, semirigid, or rigid panels are used. Once deployed, these generator wings must be maneuvered to face the Sun continuously. This is achieved by means of solar sensors and drive motors.

The orbit a satellite is designed for also determines the solar panel measurements. Low earth orbits are characterized by a relatively long period spent in eclipse. The panels therefore need to be relatively large and must be able to meet the necessary battery charging requirements. This is in contract to geostationary orbit, where the time in eclipse is at most 70 minutes and the consequent battery requirement is very much shorter.

15.4 ATTITUDE CONTROL

The attitude control system (ACS) is designed to orient the spacecraft and its subsystems in the right direction and to the design accuracy. Solar panels need to face the Sun, cooling thermal radiators need to face away from the Sun, antennas need to point at the receiving station, and thrusters need to face the opposite direction to the required force. Many different methods are deployed. Here the subject is only touched upon; for details, see Hughes (1995).

The basic requirement arises when a craft in orbit changes its orientation relative to a given location due to its orbit. For example, if the orbit is circular at a rotational frequency f, then the spacecraft must rotate at f around its normal to maintain its attitude. For a mission where a time-varying torque is required, a momentum storage device is often used. This consists of a motor driving a spinning wheel. The spinning wheel can give up momentum to the craft that can be restored to the wheel when the maneuver is complete, thereby saving on the use of fuel. Momentum bias is used to ensure that one axis of the spacecraft is highly resistant to change. For example, if the craft is to maintain an axis perpendicular to its orbital plane, then momentum bias will be designed in that direction.

The attitude control system will consist of sensors that measure the attitude relative to given reference points. The output from these is fed to the on-board processor and to ground control. Signals are then sent to the torque controllers, when necessary, to maintain the required attitude.

The following are the major sources of torque on an orbiting spacecraft:

- Thrusters mounted externally for attitude control
- Magnetic torquers that use interaction between the craft and the earth's magnetic field to generate torque
- Solar radiation pressure generated by the impact of photons on a surface, significant only on solar panels set a significant distance from the rotational axis.
- Aerodynamic torques, not significant above 600 km altitude
- Mass movements; moving a mass within a spacecraft to create a torque

The most accurate attitude sensors use one or more stars as reference points. An accuracy of 1 arcsecond is achievable. This may be compared with only an accuracy of only 5–10 arcmeters for an Earth sensor system. Star sensors are, however, larger and more expensive.

REFERENCES

Hughes, P. C. 1995. *Spacecraft Attitude Dynamics,* John Wiley & Sons, New York.

PART 3

LAUNCH AND REENTRY VEHICLES

Nick Larter

Table 15.5 shows the activity of the leading centers in launching commercial satellites into space. Table 15.6 lists some details on the satellites put into orbit on US launch vehicles in 1998–1999; it gives a good overview of the variety and number of missions mounted.

15.5 SEMIREUSABLE LAUNCH VEHICLES

There are no completely reusable launch vehicles in the world today, and there is only one substantially that is reusable (*refurbishable* may be the most appropriate

TABLE 15.5 Worldwide Commercial Space Launches

	1990	1991	1992	1993	1994	1995	1996	1997	1998	Total 1990–98
					United States					
Athena	0	0	0	0	0	1	0	0	0	1
Atlas	0	2	3	1	3	5	6	6	3	29
Conestoga	0	0	0	0	0	1	0	0	0	1
Delta	4	4	3	1	1	1	2	7	11	34
Pegasus	0	0	0	1	0	0	0	3	3	7
Titan	3	0	0	0	0	0	0	0	0	3
Total	7	6	6	3	4	8	8	16	17	75
					Europe					
Ariane 4	5	6	6	6	8	8	9	11	9	68
Total	5	6	6	6	8	8	9	11	9	68
					Russia					
Proton	0	0	0	0	0	0	2	6	4	12
Shtil	0	0	0	0	0	0	0	0	1	1
Start	0	0	0	0	0	0	0	1	0	1
Total	0	0	0	0	0	0	2	7	5	14
					Ukraine					
Zenit	0	0	0	0	0	0	0	0	1	1
Total	0	0	0	0	0	0	0	0	1	1
					China					
Long March 2C	0	0	0	0	0	0	0	1	4	5
Long March 2E	0	0	2	0	1	3	0	0	0	6
Long March 3	1	0	0	0	1	0	1	0	0	3
Long March 3B	0	0	0	0	0	0	1	2	0	3
Total	1	0	2	0	2	3	2	3	4	17
Total space launches	13	12	14	9	14	19	21	37	36	175

Source: U.S. Department of Transportation.

TABLE 15.6 Successful Launches to Orbit—on U.S. Launch Vehicles: October 1, 1998–September 30, 1999

Launch date Spacecraft name COSPAR designation Launch vehicle	Mission objectives	Apogee and perigee (km) period (min) inclination to equator (°)	Remarks
Oct. 3, 1998 USA 140 55A Taurus	Military satellite	Orbital parameters unavailable	
Oct. 5, 1998 USA 141 55C Taurus	Military reconnaissance satellite	Orbital parameters unavailable	
Oct. 9, 1998 Hotbird 5 57A Atlas 11A	Communications satellite	Geosynchronous	Eutelsat consortium spacecraft
Oct. 20, 1998 UHF F/O F9 58A Atlas 11A	Military communications satellite	Geosynchronous	
Oct. 23, 1998 SCD2 60A Pegasus	Environmental data relaying mini-spacecraft	769 km 743 km 99.9 min 25.0	Brazilian spacecraft
Oct. 24, 1998 Deep Space 1 61A Delta 11	Experimental spacecraft with ion propulsion engine	Orbital parameters unavailable	
Oct. 24, 1998 SEDSAT1 61B Delta 11	Earth imaging student spacecraft	1,079 km 547 km 101 min 31.4	Students for the exploration and development of space
Oct. 29, 1998 STS-95 64A Space shuttle	Carried experiments on microgravity science and ageing	561 km 551 km 95.8 min 31.4	Second space flight of John Glenn
Oct. 30, 1998 Pansat 64B Space shuttle	Communications satellite	Orbital parameters unavailable	Amateur student minisatellite
Nov. 1, 1998 Spartan 201-05 64C Space shuttle	Solar observatory	Similar to spacecraft on STS-95	
Nov. 6, 1998 Iridium 2,83-86 66A-E Delta 11	Communications satellite	536 Km 517 km 95 min 86	

TABLE 15.6 Successful Launches to Orbit—on U.S. Launch Vehicles: October 1, 1998–September 30, 1999 (*Continued*)

Launch date Spacecraft name COSPAR designation Launch vehicle	Mission objectives	Apogee and perigee (km) period (min) inclination to equator (°)	Remarks
Nov. 22, 1998 Bonum 1 68A Delta 11	Communications satellite	Geosynchronous	Russian television satellite
Dec. 4, 1998 STS-88 69A Space shuttle		401 km 388 km 92.4 min 51.6	
Dec. 14, 1998 SAC-A 69B Space shuttle	Carried a GPS receiver, magnetometer, and CCD camera.	Similar to STS-88	Argentine minisatellite
Dec. 15, 1998 Mightysat 1 69C Space shuttle	Advanced technology demonstrator experiments	Similar to STS-88	Minisatellite
Space shuttle Dec. 18, 1998 ISS-Unity 69F Space shuttle	U.S. module of the International Space Station	410 km 390 km 93 min 51.6	
Dec. 6, 1998 SWAS 71A Pegasus-XL	Submillimeter wave astronomy satellite	651 km 638 km 97.6 min 69.9	
Dec. 11, 1998 MCO 73A Delta 11	Mars Climate Orbiter	Interplanetary spacecraft	
Jan. 3, 1999 MPL 1A Delta 11	Mars Polar Lander	Interplanetary spacecraft	
Jan. 27, 1999 Rocsat 1 2A Athena 1	Earth resources monitoring satellite	601 km 589 km 96.6 min 35.0	Taiwanese satellite
Feb. 7, 1999 Stardust 3A Delta 11	Spacecraft to collect interstellar dust using aerogel technology	Interplanetary spacecraft	

TABLE 15.6 Successful Launches to Orbit—on U.S. Launch Vehicles: October 1, 1998–September 30, 1999 (*Continued*)

Launch date Spacecraft name COSPAR designation Launch vehicle	Mission objectives	Apogee and perigee (km) period (min) inclination to equator (°)	Remarks
Feb. 16, 1999 JCSAT6 6A Atlas IIAS	Communications Satellite	Geosynchronous	Japanese satellite
Feb. 23, 1999 ARGOS 8A Delta 11	Advanced research global observation satellite	842 km 822 km 102 min 98.7	Technology demonstrator that should permit some observations
Feb. 23, 1999 Oersted 8B Delta 11	Ionospheric science spacecraft	857 km 644 km 100 min 96.5	
Feb. 23, 1999 Sunsat 8C Delta 11	Research and education satellite	857 km 644 km 100 min 96.5	South African satellite
Mar. 5, 1999 WIRE 11A Pegasus-XL	Astronomical research spacecraft	593 km 539 km 96 min 97.5	Payload became inoperable due to malfunction after launch
Apr. 9, 1999 USA 142 (DSP19) 17A Titan IVB	Military (missile warning) spacecraft	Highly elliptical and useless orbit	Planned to be geosynchronous but now in useless orbit
Apr. 12, 1999 Eutelsat W3 18A Atlas IIas	Communications satellite	Geosynchronous	European consortium satellite
Apr. 15, 1999 Landsat 7 20 A Delta 11	Remote-sensing satellite	698 km 669 km 98.4 min 98.2	
Apr. 30, 1999 USA 143 (Milstar 2) 23A Titan IVB	Military communications satellite	4,997 km 740 km 147 min	In useless low Earth orbit

Source: National Aeronautics and Space Administration.

term) system in regular use—the National Space Transportation System (NSTS), known as the Space Shuttle. One other substantially reusable system has flown on occasion, the Russian Energiya/Buran system. All other operational launch vehicles are known as expendable launch vehicles (ELVs), with no recoverable or reusable elements.

A completely reusable vehicle, that is to say, one with no expendable system elements such as propulsion stages and with a maintenance regime more akin to that of an aircraft than the multimillion-dollar refurbishment that takes place after each NSTS Shuttle Orbiter flight, is a central long-term aim of the launch vehicle industry, simply for the reduction in the per-kilogram cost of putting a payload into orbit that such a development would deliver.

Much work is being carried out on reusable launch vehicles with medium to large payload capacities, many of which are single- or two-stage-to-orbit winged designs, with horizontal or vertical takeoff and horizontal landing. Some unwinged vertical takeoff and landing designs have also been studied, notably giving rise to the DC-X launch vehicle demonstrator. An interesting design in recent years has been the Roton project of the Rotary Rocket Company, with its ''sycamore seed'' return system.

All future launcher concepts currently at the design stage require considerable technological development, whether it be in propulsion systems such as air-breathing engines, or new materials to enable lighter structures to be developed or to provide the thermal protection required for sharp-edged aerodynamic surfaces during reentry.

15.6 EXPENDABLE LAUNCH VEHICLES (ELVs)

Currently operational ELVs can be divided into two broad classes: those that make up what could be called the world's current commercial ELV fleet, and the rest. The former class consists of comparatively few basic vehicles, such as Europe's Ariane family, Russia's Soyuz and Proton vehicles, America's Titan, Atlas, and Delta vehicles and China's Long March family. The Indian Satellite Launch Vehicle family is a good example of a typical ELV series.

Table 15.7 lists the Ariane family. Other launchers on the fringe of this group include Russia's Zenit and Tsyklon vehicles and the American Pegasus air-launched vehicle. These vehicles are mostly of large size. They exist in a wide range of

TABLE 15.7 Ariane 4: Configurations with Different Strap-on Boosters in Ascending Order of Payload Launch Mass

Denomination	Number of boosters	Solid	Liquid
A4	0	0	0
A4-2P	2	2	0
A4-2L	2	0	2
A4-4P	4	4	0
A4-4LP	4	2	2
A4-4L	4	0	4

variants, reflecting either incremental improvements in design over time or else small differences in configuration, depending on mission type. Between them and the NSTS they are responsible for the launch of all the major military and civil satellite systems in orbit at the present time. The latter class is more varied and can be said to be composed of the following:

- A few small scientific ELVs with infrequent launch schedules
- Military missiles which have been taken out of service and are looking for a commercial market
- A wide variety of would-be commercial ELVs being developed by entrepreneurs, generally small in size, based on a mixture of tried and tested and novel technologies
- Vehicles from small or developing countries, developed under programs for political, military, or technology development purposes

15.7 SOUNDING ROCKETS

Not all launch vehicles are orbital. Sounding rockets are basically ground-to-air missiles that have been modified for scientific purposes, originally for upper atmosphere physics applications but now also for microgravity science. Flights are of short duration and describe a high parabola. The greatest scientific contribution of sounding rockets to date is probably the discovery of the Van Allen radiation belt in the 1950s. The most interesting-sounding rockets are probably those associated with microgravity science, first because of the long duration of the flight (delivering of order 15 minutes of microgravity at the free fall phase at the top of the parabola), secondly because the payload is recovered, and thirdly because recent examples are multistaged purpose-built designs incorporating components from larger civil launchers.

An example of a sounding rocket is Black Brant, a solid propellant rocket in single- and multistage versions that can carry payloads of 70–850 kg to altitudes from 150 km to more than 1,500 km. It provides up to 20 minutes of useful time for microgravity experiments, auroral studies, deep space observations, and other extraterrestrial research. Black Brants are launched from conventional boom rails or towers. Since 1962, more than 800 Black Brants have been launched, with a vehicle success rate of 98%.

Black Brant XII characteristics are as follows:

Stages:	4
Length:	18.5m
Talos diameter:	0.76m
Taurus diameter:	0.58m
Black Brant V diameter:	0.44m
Nihka diameter:	0.44m
Payload weight:	136–522 kg
Altitude:	500–1,500 km

There are two main future trends in the development of sounding rockets. The first is towards more and more powerful rockets, such as the Castor-based MAXUS, that can fly higher parabolas and thereby offer longer microgravity times. The second is number of private sounding rocket companies emerging, such as the Wickman Spacecraft and Propulsion Company, offering flight services on their own sounding rockets. Given the poor success rate to date of the commercial small launcher developers offering services to orbit, this is an interesting development.

15.8 GUNS, MASS DRIVERS, SLED, ENERGY BEAM-ASSISTED VEHICLES, PARABOLIC FLIGHT AIRCRAFT, AND DROP TOWERS

The basic driver of all of the concepts in the following paragraphs is similar to that of the air-breathing winged launchers—the aim of reducing launch vehicle mass through reducing the fuel mass and providing some of the propulsive energy input from off-vehicle sources.

The launch of space vehicles as projectiles fired from a gun follows the trend begun with the HARP program in the 1960s. The concept is viable for the launch of small unmanned satellites, the acceleration and g-forces developed being impracticable for a manned ascent, even assuming that a gun capable of launching a large enough projectile could be developed. Even so, the projectiles would have to include upper-stage rocket motors for orbital insertion to be achieved. The most recent research on guns as delivery systems has focused on novel breech technologies to increase the chamber pressure that could be developed. In practice, the concept of the gun launch has probably been discredited owing to the strong military connotations of such a satellite delivery system.

The mass driver (or rail gun) has been studied as a concept to accelerate winged launch vehicles during takeoff. The assistance gained would likely not be sufficient to offset the cost of building such a device nor the energy inputs required to run it.

Much more promising is the use of mass drivers as systems to launch lunar material from the Moon's surface into orbit as part of a future lunar materials processing infrastructure. This is possibly feasible for quite small mass-drivers owing to the much smaller gravity well.

The rocket-powered sled has been looked at as an option to boost winged and unwinged launch vehicles during take off. Again the infrastructure and running costs appear to outweigh the gain.

The basic principle of this concept is that the energy to combust the fuel in a combustion chamber comes from a laser or other high-energy beam, such as microwave, directed at the vehicle from a ground-based facility. The main challenge with this concept is to keep the beam focused on the vehicle at all times during the ascent.

The use of experimental aircraft flying parabolic trajectories provides a significant increase in processing time available for microgravity experimentation over and above terrestrial drop tubes and towers. Periods of 15 to 30 seconds of microgravity can be achieved during the central portion of the trajectory. During this free-fall period, gravitational effects in the range of 10^{-2} g can be obtained. In most cases, parabolic trajectories are repeated so that several periods of weightlessness are possible. In an aircraft, however, the variability of the reduced gravity makes precise experimentation difficult.

For example, the KC-135 can simulate up to 40 periods of low gravity for 25-second intervals during one flight. The aircraft accommodates a variety of experiments and is often used to develop future space flight experiment equipment and techniques. The plane climbs rapidly at a 45° angle (pull-up), slows as it traces a parabola (pushover), and then descends at a 45° angle (pull-out). The forces of acceleration and deceleration produce twice the normal gravity during the pull-up and pull-out legs of the flight, while the brief pushover at the top of the parabola produces less than 1% of the Earth's gravity. The KC-135 characteristics are given below. Other examples of parabolic flight aircraft are given in Table 15.8.

Bay dimensions:	3.04×16.4 m
Bay overhead clearance:	1.8 m
Maximum floor loading:	90 kg per 0.09 m^2
Acceleration:	2.5 g
Microgravity duration:	25 s
Number of maneuvers/flight:	40

Drop towers accommodate large experimental packages, generally using a drop shield to contain the package and isolate the experiment from aerodynamic forces during free fall in the open environment. In the drop towers, free-fall periods range from 2.2 to 5.1 seconds. An auxiliary thrust may be provided to overcome the initial resistance of air friction, but some facilities use an evacuated drop chamber. Accelerations acting on the experiments are less than 10-5 g. Two examples outside the United States are the German ZARM tower and the Japanese JAMIC facility.

15.9 REENTRY VEHICLES

Table 15.9 lists some general properties of reentry vehicles. Table 15.10 gives more specific information on ballistic vehicles, these designs are assumed to generate no lift; the only force acting on the skin as they descend is aerodynamic drag. Table 15.11 the lists semiballistic vehicles.

A prominent reentry philosophy is that of gliding entry, as in the NSTS Shuttle Orbiter, where the fuselage is inclined at a high angle of attack to the trajectory direction in order to generate lift. This is done to enable reentry to be performed in a high-altitude glide, where the heating rates are much lower. Ablative heatshields

TABLE 15.8 Further Examples of Parabolic Flight Aircraft

Aircraft	Base	μg time	μg quality	Number of parabolas/flight	Accommodation
Learjet model 25	NASA LeRC	15–20 s	10^{-2}g	5	$0.91 \times 0.73 \times 0.61$ m; 85 kg
KC-135	NASA MSFC	25 s	10^{-4}g – 10^{-2} g	40	$3.05 \times 16.2 \times 1.8$ m bay
F-104	NASA MSFC	30–60 s	10^{-2}g	1	$0.28 \times 0.35 \times 0.58$ m; 16 kg

TABLE 15.9 Reentry Vehicle Types

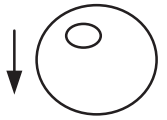
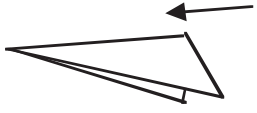
Reentry vehicle shape	Shapes and typical reentry angles	Reentry dynamic	Operational examples	Typical lift/drag ratio	Typical acceleration (m/s^2)	Typical cross-range	Control
Sphere		Ballistic	Foton	0	10 g	0°	None
Cone (symmetric)		Ballistic	Raduga	0	10 g	0°	None
		Semiballistic	Apollo, Gemini	0.2	4 g	1°	Navigation only
Cone (asymmetric)		Semiballistic		0.5	2 g	3°	Stable
Smooth curved wings and body		Lifting	BOR series HL-10 M2-F2 NSTS Shuttle Orbiter Buran	1+	1.5 g	10°+	Unstable
Complex; sharp-edged angular caret wing		Waverider	None: ground experiments only	2+	1.5 g	30°+	Unstable

TABLE 15.10 Further Examples of Ballistic Reentry Vehicles

Reentry vehicle	Country or agency	History/status	Characteristics
Descent capsule for Resurs-F series	Russia	First flight September 1979 (Cosmos 1127). Over 60 missions to date. Resurs-F series still operational.	Descent capsules can be refurbished and reused approximately three times. Film/camera return from Resurs-F photo-reconnaissance satellite. Diameter: 2.2 m Mass: 2,300 kg
Bion/Biocosmos	Russia	First flight October 31, 1973 (Cosmos 605) Launch site: Plesetsk Launch vehicle: Soyuz 11A511U	The spacecraft was based on the Zenit reconnaissance satellite. Launches in the program included Cosmos 110, 605, 670, 782, plus Nauka modules flown on Zenit-2M reconnaissance satellites. 90 kg of equipment can be contained in the external Nauka module. Bion was developed for biological studies of the effects of radiation. Design life: 30 days. Total mass: 5,400 kg. Total payload: 625 kg
Foton (or Photon)	Russia	First flight April 1985 (Cosmos 1645). Over 10 missions to date. Still operational.	Adaptation of recoverable Vostok spacecraft for zero-gravity materials processing tests. 400 W available to operate experiments. Materials processing tests. Space materials research. Design life: 16 days. Total mass: 6,190 kg. Total payload: 404 kg.
Nika series	Russia	Nika-B Nika-K Nika-T	Volume: 4.5 m^3 (Nika-T) Materials research.
MIRKA	Germany	October 9, 1997 Launch site: Plesetsk Launch vehicle: Soyuz 11A511U	German miniature reentry vehicle attached to exterior of Russian Resurs satellite. After release from Resurs landed in Kazakhstan October 23. Technology Research.
Generation series	Russia	Generation-1/2/3 Generation-4 (Yantar) Generation-5	Film/digital data cartridge return.
BremSat	Germany	February 3, 1994: BremSat 1 Launch site: Cape Canaveral Launch vehicle: Shuttle	Technology: Atomic oxygen, dust particle, microgravity, reentry experiments. Total mass: 63 kg. Perigee: 156 km. Apogee: 160 km. Inclination: 57.0°

TABLE 15.10 Further Examples of Ballistic Reentry Vehicles (*Continued*)

Reentry vehicle	Country or agency	History/status	Characteristics
Teknologia	Russia	In 1990, K.B. Salyut proposed an unmanned derivative of the TKS. The 20-tonne satellite would have a five-year life and was to be launched by 1993.	The mission would be a preliminary to a full-scale production TMP Skif-DM 90-tonne spacecraft. No backers for the concept were forthcoming. Proposed to conduct zero-gravity materials production experiments. Max dia: 4.2 m Total mass: 20,000 kg Elec system: Solar panels.
Biosatellite	U.S.A.	December 14, 1966; Biosatellite 1, Delta G September 7, 1967: Biosatellite 2, DeltaG June 29, 1969: Biosatellite 3, Delta N Launch site: cape Canaveral	The Biosatellite 1 was never recovered. Bio 3: to fly a 6 kg male pig-tailed monkey (*Macaca nemestrina*) named Bonnie in Earth orbit for 30 days. After only 8.8 days in orbit, the mission was terminated due to subject's deteriorating health. Biological capsule. Investigation of the influence of space flight on living organisms. Total mass: 542 kg.
Express 1	Germany/Japan	January 15, 1995: Express RV Launch site: Kagoshima Launch vehicle: Mu-3S Failure.	At first thought not to have reached orbit. Capsule was "lost" during re-entry but turned up in Ghana and was returned to the owners a year later. Materials research.
FSW (Jianbing) series	China	FSW-0 FSW-1 FSW-2	Oak heat shield. Earth observation and/or microgravity missions.
KH (Keyhole) series payload carrier	U.S.A.	KH-1 (Corona/Discoverer) KH-3 KH-4 (Corona) KH-5 (Argon) KH-6 (Lanyard) KH-7 (Gambit) KH-8 KH-9 (Big Bird)	Film return from KH series photo-reconnaissance satellite.
Orlets	Russia	Sixth-generation reconnaisance satellite Orlets-1: July 18, 1989 (Cosmos 2031)	After returning multiple film capsules, the spacecraft is deorbited. Military surveillance

TABLE 15.10 Further Examples of Ballistic Reentry Vehicles (*Continued*)

Reentry vehicle	Country or agency	History/status	Characteristics
FAST	U.S.A.	Placed in orbit on August 21, 1996. Launch site: Vandenberg Launch vehicle: Pegasus XL	Second Small Explorer mission. Prepared for mid-1994 launch date, FAST was placed into storage until a series of problems with the Pegasus launch vehicle could be corrected. Design life: 1 year Length: 1.8 m Max dia: 1.2 m To observe and measure rapidly varying electric and magnetic fields and the flow of electrons and ions above the aurora.
Taurus	U.S.A.	Launches: 5 Success rate: 100% First launch date: March 13, 1994 Last launch date: March 12, 2000	Pad-launched launch vehicle using Pegasus upper stages and Castor-120 first stage. First launch used slightly larger Peacekeeper ICBM first stage instead of Castor-120. Liftoff thrust: 131,180 kgf Mass: 73,030 kg Core dia: 2.4 m Length: 27.9 m

are not possible for reentry gliders as they would alter the vehicle shape as they melted away, leading to steering problems. The lower-surface wing area is also much greater than on a capsule, leading to the potential need for acres of ablative material and a prohibitive weight penalty.

The time of heating for gliding entry is much longer than that for a capsule, as there is less drag to slow the vehicle from high velocities at the lower air densities (higher altitudes). The best approach is to radiate heat out of the vehicle skin as fast as it comes in via convective heating, establishing a thermal equilibrium.

Radiation cooling rates depend on the radiator area and the radiator surface emmisivity, as well as the fourth power of the temperature difference between the radiator surface and the atmosphere. This is where the design philosophies of the shuttle and a waverider diverge. The NSTS Shuttle Orbiter deliberately has a reduced wing area because the original designers did not want the reentry glide to cover too much distance and overfly the Soviet Union. The NSTS Shuttle Orbiter design called for a vehicle capable of reentering and landing within the United States.

The NSTS Shuttle Orbiter's radiative cooling is not enough for thermal equilibrium, so over the time of reentry the tiles are subjected to a net heat build-up. Fortunately, due to the shorter reentry glide, the time of heating is reduced. However, the crew must remain inside for a long period of time after landing until the tiles cool down enough not to melt the disembarking stairs.

A series of lifting bodies flown at the NASA Flight Research Center from 1963 to l975 demonstrated the ability of pilots to maneuver and safely land wingless vehicles. These lifting bodies were basically designed so they could fly back to Earth from space and be landed like an aircraft at a predetermined site.

TABLE 15.11 Examples of Semiballistic Reentry Vehicles

Reentry vehicle	Country or agency	History/status	Characteristics
Mercury	U.S.A.	Spacecraft launched: Freedom 7 (Mercury 3) Liberty Bell 7 (Mercury 4) Friendship 7 (Mercury 5) Aurora 7 (Mercury 7) Sigma 7 (Mercury 8) Faith 7 (Mercury 9)	NASA used pigs to test human survivability in case of a land "splashdown"; they showed no apparent ill effects The first manned U.S. space mission.
Gemini	U.S.A.	Space craft launched: Gemini 3 (March 23, 1965) Gemini 4 (June 3, 1965) Gemini 5 (August 21, 1965) Gemini 7 (December 4, 1965) Gemini 6 (December 15, 1965) Gemini 8 (March 16, 1966) Gemini 9 (June 3, 1966) Gemini 10 (July 18, 1966) Gemini 11 (September 12, 1966) Gemini 12 (November 11, 1966)	Gemini aimed to teach astronauts such techniques as docking, rendezvous, long-term flight, and spacewalks that would be necessary for the flight to the Moon.
Apollo	U.S.A.	Space craft launched: Apollo 7 (October 11, 1968) Apollo 8 (December 21, 1968) Apollo 9 (March 3, 1969) Apollo 10 (May 18, 1969) Apollo 11 (July 16, 1969) Apollo 12 (November 14, 1969) Apollo 13 (April 11, 1970) Apollo 14 (January 31, 1971) Apollo 15 (July 26, 1971) Apollo 16 (April 16, 1972) Apollo 17 (July 7, 1972)	Apollo program aimed at landing the first man on the Moon.
ARD	France	Launch: October 21, 1998 Program: ISS Launch site: Kourou Launch vehicle: Ariane 5	The ARD was an 80% scale model of the Apollo Command Module. Technology test for a possible ISS crew rescue vehicle.

These vehicles, with their unconventional aerodynamic shapes, were the M2-F1, M2-F2, M2-F3, HL-10, X-24A, and X-24B. The information from the lifting-body program generated contributed to the database that led to development of today's Space Shuttle. The addition of fins and control surfaces allowed the pilots to stabilize and control the vehicles and regulate their flight paths.

PART 4

SATELLITES

Pascal Willis

Any object rotating about another is called a satellite. Here we are concerned with all of the technology satellites that have been launched into Earth orbit since the first in 1957. Hundreds are now in orbit. All were launched to gather or relay information; to do this effectively requires that the satellite stay in a prescribed known orbit at a given orientation. The stability of the satellites is therefore a central design concern.

There are four primary orbits that satellites are sent to:

1. Low earth: up to 250 km—mobile communications and reconnaissance
2. Polar: approximately 800 km—weather and navigation
3. 36,000 km—weather, navigation communication
4. Highly elliptical

A satellite in orbit is determined by its function.

The two principle modes of satellite stabilization are spin stabilization and three-axis stabilization. Spinning objects are naturally stable. Satellites are therefore often spun in orbit to maintain their position. The antenna, which must always point to a known location on Earth, is then spun relative to the satellite at the rotational speed to keep it in a fixed position. In three-axis stabilization, spinning wheels are used to control the orientation of the satellite by spinning one or two of the wheels faster to restore the satellite to its original position. The Earth's gravity and magnetic field are also used to give stability to some designs.

Solar cells are the main source of power for satellites. When in orbit, relatively large software area panels of cells are unfurled and orientated to face the Sun. These are considered in a later section.

15.10 COMMUNICATIONS SATELLITES

These deal with signals to and from telephones, television transmitters, and Internet providers. The benefit of using geostationary orbit is that only three satellites are needed to cover the majority of the inhabited globe. In contrast, 50–60 satellites are needed in low Earth orbit. The disadvantage is that geostationary transmission requires very significantly more power because of the increased transmission distance. The core electronics is called the transponder system and receives and transmits signals. It also has the facility to amplify, clean, and frequency shift signals prior to retransmission by the antenna. The antenna is now often too large to be launched unfurled, so, like the solar panels, it is unfurled in space. The signal to the satellite is the uplink, and the signal to earth is the downlink. This beams to a given area in earth called the satellite footprint.

15.11 SATELLITE NAVIGATION

Because so many acronyms are used in this subject, the most prominent are listed here first.

A-S	Antispoofing
BIPM	International Bureau of Units and Weights
DGPS	Differential GPS
DOP	Dilution of precision
EGNOS	European Geostationary Navigation Overlay System
ESTB	EGNOS test bed
GBAS	Ground-based augmentation system
GNSS	Global navigation satellite system
GLONASS	Global Navigation Satellite System
GPS	Global Positioning System
IAG	International Association of Geodesy
IERS	International Earth Rotation Service
IGS	International GPS Service
ITRF	International Terrestrial Reference Frame
ITRS	International Terrestrial Reference System
LAAS	Local area augmentation system
MSAS	Multi-Functional Satellite Augmentation System
NAVSTAR	Navigation System by Timing and Ranging
NIMA	National Imagery and Mapping Agency
PR	Pseudo-range
RAIM	Receiver autonomous independent monitoring
S/A	Selective availability
SBAS	Satellite-based augmentation system
TAI	International Atomic Time
TCAR	Three-carrier ambiguity resolution
UTM	Universal Transverse Mercator
WAAS	Wide-area augmentation system
WGS-84	World Geodetic System 1984

In the past, stars were use to determine location of ships at sea. Today, artificial satellites are used to determine precisely the location of any type of object on land, and in the air. All satellite navigation systems are based on the same principle. Basically the following components are needed:

- A specific receiver (user segment)
- A dedicated constellation of satellites (space segment)
- A tracking network of ground stations (control segment)

Each satellite possesses one or several precise atomic clocks and transmits one or several coded signals using some prespecified frequencies. The receiver on ground, by comparing the code that was generated locally using its own local oscillator and the actual received signal from the satellite, is able to determine simultaneously the epoch of transmission (as read by the satellite clock) and the epoch of reception (as read by its own local clock). The difference of those two epochs is currently called pseudo-range (PR) because it is a measure of the instantaneous distance from the satellite to the ground receiver biased by an unknown clock synchronization error (as well as additional propagation errors).

Without the user clock synchronization error, three pseudo-ranges would be sufficient to determine its location unambiguously (intersection of three spheres in space). By measuring simultaneously four pseudo-ranges at exactly the same epoch, you can determine your absolute location as well as your clock synchronization error instantaneously. So with four satellites in visibility, you can determine your position and your synchronization. This is now possible with present types of receivers that can track all in-view satellites simultaneously. Of course, more then the necessary four satellites would increase the accuracy of the result and also raise the confidence on these results.

In fact, in order to estimate the distance from the satellite to the receiver on the ground, you need not only the epoch of transmission of the signal by the satellite, but also the exact satellite location (orbit ephemeris). This information is usually broadcast to the user through the signal itself by the system providers. It is usually known as navigation data. For this purpose, an operational permanent tracking network of dedicated receivers determines the orbit of all satellites as well as their clock behaviors using the same principles and equations.

Satellite navigation systems usually use L-band signals, which are not affected by clouds. Positioning can then be obtained anywhere on the globe and any time, whatever the weather conditions. Satellite navigation system performance can be characterized using the following basic criteria:

- Coverage: the area on the globe where the system can potentially be used
- Accuracy: the difference between the estimated and the exact position of the user
- Availability: the possibility to have access of the system at a specific time and location
- Integrity: the ability of the system to provide timely warnings to users to avoid incorrect results

The aviation community can use satellite navigation systems for a broad range of applications. First of all, the on-board positioning capability, coupled with on-board digital charts, can be used for guidance of the airplane for all types of weather and visibility conditions. A possible precise positioning of the plane could also help in increasing the number of international air traffic lanes. It could also be use to allow different ways of approaching airports, leading to possible increase of planes landing at a specific airport, as described by McDonald (1991). Such systems could also be used for search and rescue situations where the last available exact location of the plane, as well as its past trajectory, could be essential for the rescue operation. Other applications can also be found, such as using multiple antennas on the plane to derive its orientation in space for certain applications like plane photogrammetry for mapping.

However, as we shall see, such systems have also some weaknesses and depending on their exact use can lead to very different types of performances.

Satellite Positioning Errors

The satellite positioning accuracy depends on several factors:

- Precision of the measurement: This depends on the signal characteristics of the system and also on the quality of the receiver as well as its adequacy to the plane trajectory in terms of acceleration.
- Satellite geometry: For certain configuration of satellites, the positioning accuracy can be strongly degraded, for example when satellites are aligned in the sky. This can be monitored in real time by the receiver and is know as DOP (dilution of precision).
- Orbital errors: These depend on the accuracy of the satellite orbits used in the position estimation. Broadcast orbits, available in real time, will never be as accurate as postprocessed orbits, especially when using a denser tracking network. To be used in real time, these data are issued from an extrapolation and are subject to errors.
- Propagation errors: As the signal from the satellite to the user on the ground goes through the atmosphere, corrections must be applied. Dual-frequency receivers can be used to estimate the ionospheric correction corresponding to the upper atmosphere.
- Multipath errors: When the signal does not come in a straight line from the satellite to the user and is, for example, just a reflection of the original signal by some metallic surface. Such effects can be seen easily on planes, especially during changes of direction.

Most of these errors can be considered systematic errors and will be more or less identical for a large area on the globe. With the use of differential techniques, most of these errors besides the multipath effects can be eliminated or at least reduced.

Augmentation Systems

To enhance performance of satellites systems, additional systems are being developed. There are basically of two different types:

- Satellite-based augmentation systems (SBAS), where corrections are broadcast by additional satellites
- Ground-based augmentation systems (GBAS), where corrections are broadcast by local stations

In both cases, the system broadcasts corrections for improved accuracy as well as information on the integrity of the satellite constellation (bad satellite clocks) and eventually a ranging signal that can be used by the receiver on the ground as an additional satellite to complement the availability of the previous system.

For SBAS, geostationary satellites are used as they provide service on a large area of the globe. For GBAS, the service will operate only in a very limited zone, such as the approach zone of an airport, where increased accuracy is required.

Other Applications of Satellite Navigation Systems

There are a large variety of applications of satellite navigation systems. We will mention here only the time and frequency applications for which most of the applications are now currently done with systems like GPS due to its relative low cost and facility of use. There are also several scientific uses of these systems in the field of positioning (geodesy, geophysics) as well as in the field of atmospheric sciences (Beutler et al. 1999).

Let us first review to some general concepts before describing these present satellite navigation systems. We first need to give some minimum background on geodetic concepts.

To define the position of an object on the Earth, we can define a reference system by an origin, the orientation of three orthogonal axes and a unit. Given such a terrestrial reference system, we can define the position of any object in the close surroundings of the Earth using three coordinates X, Y, and Z. However, it is more usual and useful to define geodetic coordinates latitude (ϕ), longitude (λ), and ellipsoidal height (h) as soon as you have defined a conventional ellipsoid of reference. In fact, maps even go one step further, using a map projection such as Universal Tranverse Mercator (UTM).

There is also another subtlety concerning the height as the ellipsoidal height, which we define here as a pure geometrical concept. It is then not affected by the topography of the Earth or the changes in density of the Earth. The intuitive concept of altitude is that the water flows in the direction for which the altitude decreases. There is then a fourth parameter that can be attached to an object—its altitude, which can be expressed as the height above a specific surface, called the geoid, which is an equipotential surface of the gravity field. As a first approximation, the geoid coincides with the mean surface of the oceans. While the ellipsoidal height is easily accessible to satellite navigation users, the altitude must be measured using other means, such as leveling (Schwarz and Sideris 1993).

Reference Systems and Reference Frames

All geodetic reference systems are equivalent, using a seven-parameter transformation (three to change the origin, three to change the orientation, and three to change the unit of length also called scale). However, reference systems can only be accessed through a theoretical definition. For directed accessibility to the users, geodesists use another concept called *reference frame,* or *datum* (Boucher 2001). A reference frame is a set of constants and conventions that makes it possible to practically access a specific reference system. Ground station coordinates been the general way for survey institutes to provide reference frame to users. The distribution of satellite positions in time is another way to provide a reference frame.

However, while reference systems are ''perfect'' by definition, reference frames possess some errors and inaccuracy because they are the results of a measurement. There is thus no exact mathematical transformation between two different reference

frames. The best one can do is to estimate the best seven-parameter transformation between them. The accuracy will be limited by the less accurate reference frame.

Multiplicity of Reference Frames

The multiplicity of references frames is a continual source of astonishment and confusion for the layman. But there are in fact several good reasons for the large number of reference frames in use. First of all, scientific improvements continually create better tools for positioning, allowing refinements in reference frames. There is then ongoing improvement of reference frames: from historical nationwide reference frames using ground optical geodetic techniques-to present worldwide reference frames making full use of satellite navigation systems as well as other fundamental space geodetic techniques. There are also some political and legal issues related to these problems. For example, national geodetic networks are the responsibility of each country. Changing terrestrial reference has a direct impact on land registry. Satellite navigation systems also usually belong to a particular country and need to be independent for operational reasons. There is then a natural tendency for each new satellite navigation system to provide orbits to the users in totally controlled reference frame in which the positioning results of the users will be expressed *de facto*. There is presently no International Bureau as such that would be responsible for the definition of terrestrial reference frame on a worldwide basis, as BIPM (International Bureau for Units and Weights) is for the definition of International Atomic Time (TAI).

On the other hand, when new refinements of reference frames are available, it is never possible to disregard all previous works and throw away large databases of geographical information for economical reasons. There is then a need to transform results obtained in a previous reference frame in a newer and improved one. However, the user must be aware that the precision of the transformed previous results depends not only on the precision of its initial results but also on the precision of the transformation, which is not only usually limited by the worse-defined reference frame.

WGS-84 and ITRF Reference Frames

Besides the multiple-country terrestrial references, there are at least two global reference systems that are important for plane navigation.

WGS-84 (World Geodetic System 1984) is a system defined and maintained by NIMA (National Imagery and Mapping Agency) (see Slater and Malys (1998). It is used to compute the GPS broadcast orbits and is then *de facto* the reference frame in which all real-time positioning GPS results are expressed.

ITRS (International Terrestrial Reference System) is a system defined and maintained by the International Earth Rotation Service as described by Boucher and Altamimi (1996), through a worldwide scientific cooperation using the best geodetic technique available. Since 1988, several realizations of this system have been distributed. The latest frame (ITRF-2000) contains about 200 high-quality geodetic stations all over the world, with present geodetic accuracy of 2–5 mm in position and about 1 mm/year for velocity, as presented by Altamimi et al. (2001).

However, the user must also be aware that the WGS-84 and ITRS systems are now aligned within less than a decimeter (there is no need to apply any significant

transformation between them). For meter-accuracy applications, there is then no need to distinguish them. However, for centimeter-accuracy and below, the global ITRF network offers easy access to the reference frame. Access to WGS-84 realization can only be done through GPS broadcast orbit and with an accuracy usually worse than 10 m.

Global Positioning System (GPS)

The Global Positioning System (GPS) is a U.S. satellite navigation system. Initially developed by U.S. Department of Defense (DoD) for military purposes under the acronym NAVSTAR (Navigation System by Timing and Ranging), the system is now available for civilian applications as a dual-use system. In March 1996, a Presidential Directive ''encourage[d] acceptance and integration of GPS into peaceful civil, commercial and scientific applications worldwide.''

The first GPS satellite was launched in 1978. The full GPS constellation of satellites was declared fully operational in 1994 and is presently composed of 24 satellites and 3 active spares orbiting at an altitude of 20,000 km (period of about 12 hours). The satellites broadcast signal on two frequencies (L1 = 1575.25 MHz and L2 = 1227.60 MHz) using civil and military codes. They also broadcast additional navigation data that are essential to the users, such as the satellites' orbits and clock information. The estimated satellite lifetime is about 10 years. Subsequently, the constellation needs to be regularly replenished. Additional technical description of the GPS system can be found in Hofmann-Wellenhof et al. (1992) and Parkinson and Spilker (1996).

In May 2000, the selective availability (S/A) degradation was removed, allowing enhanced performance for civilian applications (see Neilan et al. 2000). Navigation applications of GPS for terrestrial, maritime, or airborne users are presently countless, and their number is still growing.

For the future, a project for GPS modernization was officially announced in 1999. As a first step, the second GPS frequency (L2) will be fully accessible to civil users. Upgraded satellites transmitting civilian codes on L2 will be launched starting in 2003. In a second step, a third frequency signal, called L-5 (1176.45 MHz), will be added to enhance safety-of-life critical applications such as civil aviation, starting for the first launches of new satellites in 2005. Finally, a long-term commitment has been to made to improve present GPS capabilities at the horizon 2015 (GPS III).

Satellite-Based Augmentation System

In order to bypass the present limitations of the GPS systems, several satellite-based augmentation systems (SBAS) are being developed:

- Wide Area Augmentation System (WAAS) in North America: WAAS was developed by the FAA for GPS precision flight approaches, with the first tests being conducted in 1997, as described by Hansen (1998). Using a network of 25 stations, an integrity and correction message is generated and broadcast on the L1 frequency using geostationary satellites. The WAAS signal can be used in the United States, Canada, and Mexico but is still waiting for FAA certification.
- European Geostationary Navigation Overlay System (EGNOS) within Europe: EGNOS is a similar project in Europe, developed in common among the Euro-

pean Space Agency (ESA), the European Commission (EC), and Eurocontrol, the European organization for the Safety of Air Navigation. A network of 11 Ranging Integrity Monitoring Stations (RIMS) tracks the GPS and GLONASS satellites, verify their integrity in real time and generating corrections for more accurate positioning to be broadcast by Inmarsat geostationary satellites. Service has been available on a nonoperational mode since February 2000 and should be fully operational and certified in 2004, as presented by Gauthier et al. (2001).

- Multi-Functional Satellite Augmentation System (MSAS) by Japan Civil Aviation Bureau (JCAB) for civil aviation. MSAS is an adaptation of the same approach for Japan. The geostationary satellite was launched in 1999. Full service should be available in 2005.

All three systems have been developed in parallel and in close cooperation to allow seamless navigation worldwide. Present real-time positioning accuracies are in the 3–5 m range.

Ground-Based Augmentation Systems

The Local Area Augmentation System (LAAS) has been developed by the FAA to complement WAAS for more stringent navigation modes such as CAT-II and CAT-III, as presented by Swider, et al. (1997). Time corrections are broadcast to the user by very high frequency (VHF) using data from ground-based equipment. Expected accuracy is less than a meter.

International GPS Service and Precise Applications

Scientific users have developed free services based on worldwide scientific cooperation. Beginning in 1994, the International GPS Service (GPS) has been in continuous operation, providing users with precise GPS orbits and clocks for geodetic and geophysical applications requiring centimeter or millimeter accuracy position, as described by Beutler et al. (1999). With a network of several hundred permanent geodetic GPS receivers, eight data analysis groups, IGS provides GPS products through the Internet free of charge throughout the year, as presented by Neilan et al. (2000).

Initially limited to postprocessing scientific applications, this service is now moving to more and more precise real-time applications, providing users with rapid products or even predicted precise orbits and clocks for precise real-time applications.

GLONASS

GLONASS (Global Navigation Satellite System) is a military navigation satellite developed by Russia. In 1999, Russia's President Yeltsin offered ''national satellite navigation GLONASS as a basis for development of international satellite navigation systems.''

GLONASS is very similar to GPS, with a few noticeable differences. The altitude of the satellites is lower (19,000 km), the transmitted frequencies are satellite-dependent, allowing better resistance to jamming, and no degradation seems to exist to limit civilian performance. Since the first satellite launch in 1982, there have

never been enough satellites available for the system to be declared operational. Since the last launch in October 2000, there are presently only six to eight available satellites due to an insufficient number of replenishing launches.

However, several manufacturers have been able to commercialize dual GPS/GLONASS receivers. For those users, the GLONASS satellites are used to supplement the number of available GPS satellites. An increased number of satellites is important for integrity monitoring and also for accuracy improvement, as described by Langley (1997) and Willis (2001).

Galileo

More recently, in June 1999, Europe expressed willingness to develop a new independent satellite navigation service called Galileo, as described by Tytgat and Campagne (2000). As a joint initiative of the European Commission (EC) and the European Space Agency (ESA), this system will be established for the civil community and financed through public–private partnership and will be fully interoperable with present GPS and GLONASS systems.

In total a full constellation of 30 satellites (including 3 active spares) orbiting at 24,000 km altitude will be deployed. This constellation may also be completed with geostationary satellites using the experience gained in EGNOS. With first launches starting as early as 2004, a preoperational service should be available in 2006, allowing full operational capability in 2008 or earlier. Several services will be proposed to users:

- A free-of-charge dual-frequency positioning service for a large community of users (general public) with an accuracy of 4 m
- A subscription service for commercial and professional applications requiring better performance in terms of accuracy or integrity with a guarantee of service and certification for liability purposes
- Restricted access for safety-of-life critical applications or governmental uses

The Future of Satellite Navigation Systems

It is quite difficult to predict the exact future of satellite navigation systems while the number of applications and their diversity is still rapidly growing. However, there are certain general trends. First all all, it is almost certain that most applications will be multi-systems. We are heading towards a future Global Navigation Satellite System (GNSS), in which GPS will play a key role but other systems like Galileo or even GLONASS or augmentation systems, satellite-based or ground-based, will be used simultaneously. The interoperability issues will be solved by using signal structures sufficiently close to allow cheap receiver manufacturing and by adopting common basic references for time and reference frame, allowing totally seamless navigation service worldwide.

The increasing number of available satellites measurements, using as many different frequency signals, will help solve the present limitations of integrity and availability for civil aviation applications. On the other hand, new methods of processing, making full use of Internet connection, like Internet-based Global Differential GPS (IGDG), show real-time results of 10–20 cm accuracy, as described by Muellerschoen et al. (2000) when using all GPS frequency measurements (code and phase).

Finally, the natural complementarity of satellite navigation systems and telecommunication systems should lead to totally new services in the future that will be part of our lives.

15.12 METEOROLOGY

Accurate weather forecasting is an important feature of many industrial, leisure, and military activities. The state and development of earth weather systems is monitored by an array of satellites, the first of which was launched by the United States in 1960. The satellite data are merged with data gathered at ground and sea stations to form the data set from which forecasts are made. Meteorology satellites take images, using different parts of the electromagnetic spectrum to view different features, of weather systems, and also measure atmospheric pressure, temperature, and humidity.

15.13 EARTH RESOURCES SATELLITES

In 1972, Landsat 1 was launched by the United States to take the first combined visual and infrared images of the Earth's surface. By imagining the Earth in different wavelength bands, it is possible to assess different features of the soil, water, or plant life present. An infrared image, for instance, can tell if the moisture content in a crop has dropped below a critical level before this becomes visible to the eye on Earth. Earth resources can thereby be monitored. The most basic wavelength bands used are given in Table 15.12.

There is also now extensive monitoring of the Earth's oceans. ERSI satellites take microwave infrared and radar scans of the oceans, the choppiness of the seas, wave height, and the shape of the ocean surface. The Topex Poseidon satellite can monitor ocean height to within 4.3 cm.

15.14 MILITARY SATELLITES

Military satellites are used for communication, surveillance and monitoring description of enemy signals. The first GPS systems were developed by the military; they

TABLE 15.12 Wavelength Bands

Band	EM wavelength (10^{-4n})	Monitors
1	450–520	Coastal waters, soil vegetation
2	520–600	Shows healthy vegetation
3	630–690	Plant identification
4	760–900	Water resources
5	1550–1750	Moisture content of vegetation
6	10400–12500	Plant life heat stress
7	2080–2350	Sources of heated water

can now pinpoint individuals on the ground. From a wider perspective, early warning satellites are used to monitor missile launch sites and then track the trajectory of any missile launched. The Space Based Infrared System (SBIRS) is being developed to replace the Defense Support Program (DSP) satellites as the primary warning system for the United States. It not only monitors but also triggers defensive missiles in response to a hostile launch.

15.15 SATELLITE INSTRUMENT PACKAGES

A satellite consists of a number of subsystems and a payload—typically an instrument package consisting one or more sensors or passive collectors or microgravity experiments or commercial broadcasting transponders. The basic satellite subsystems are as follows:

- Structure
- Power
- Thermal control
- Data handling (on-board and downlink)
- Guidance, navigation, and control
- Attitude control—orientation and station-keeping

Some of these subsystems may be duplicated by the payload—it is common, for example, for payloads to have self-contained data handling for the preprocessing of instrument data. Other satellite subsystems may provide a service to the payload, such as power or the acquisition and downlinking of preprocessed payload data.

The Infrared Space Laboratory

In 1983, the U.S.-Dutch-British satellite started infrared space astronomy by mapping 250,000 cosmic infrared sources and large areas of extended emission. It has allowed astronomers to see both familiar objects in an unusual way and objects that are invisible at other wavelengths. The infrared space observatory, operating at wavelengths from 2.5 to 240 μm, observed astronomical objects that were hidden for optical telescopes, such as cool objects that are unable to emit visible light. Opaque objects, those surrounded by clouds of dust, are another specialty of ISO because the longer IR wavelengths can penetrate the dust. The single 0.6-m telescope collector in ISO fed infrared data via a pyramidal mirror to four instruments. The fields of view and the selection of wavelengths are varied to suit the nature of the object examined.

The four instruments are:

- The infrared camera, covering the 2.5–17 μm band with two different detectors. It can be compared to a normal photo camera, taking pictures of the infrared face of astronomical objects at a high resolution.
- The photo-polarimeter, designed to detect the amount of infrared radiation emitted by an astronomical object. The broad range of wavelengths at which this operated

(between 2.5 and 240 μm microns) allowed it to view objects as cool as the clouds of dust lying among stars and galaxies, whose temperature may be just a few degrees above absolute zero.

- The short-wave spectrometer covers the 2.4–45 μm band. It has provided valuable information about the chemistry of the universe, since many molecules emit strongly in the infrared. Moreover, it has been able to find out the physical conditions of those chemical constituents, such as temperature or density.
- The long-wave Spectrometer, able to operate at the 45–196.8 μm band. It has focused on cooler objects than the short-wave system. It is especially useful for studying the physical condition in very cold dust clouds in the space between stars.

An Ariane 44P launcher from Europe's spaceport in Kourou successfully launched the Infra red Space Laboratory on November 17, 1995. Initially it was supposed to be operational for 20 months, but thanks to good engineering and some good fortune its working life was stretched to more than 28 months.

Observing the cool universe requires instruments working at temperatures close to absolute zero. Keeping the temperature this low is the task of the large liquid-helium cryostat. Interestingly, this cryostat has made the observatory one of the coolest objects in the universe. The eventual shutdown of the observatory was due to the depletion of the liquid helium.

The most important recent developments for satellites include the dual trends towards constellations of mini or microsatellites on the one hand (such as the ill-fated Iridium project, which failed because of the technological limitations in delivering the service to the end user) and very large, costly, sophisticated instrument platforms (like Envisat) on the other, which create challenges in payload integration and testing. In order to minimize costs wherever possible, there is increasing interest in sourcing and space-qualifying payload components off the shelf—this trend pervades the entire space industry. Another future growth area in satellite payloads will be experiments to investigate applications for tethers and inflatable structures.

On the operational side, developments will include the flight of Web servers to enable housekeeping data to be accessed via browsers, and for high-cost payloads the possibility of in-orbit servicing. The best example of this to date includes the Hubble Space Telescope, which has used this option for the change of the solar arrays and for correcting a deficiency in the optical equipment.

REFERENCES

Altamimi, A., Sillard, P., and Boucher, C. 2001. "ITRF2000: A New Release of the International Terrestrial Reference Frame for Earth Science Applications," submitted to *Journal of Geophysical Research.*

Beutler, G., Rothacher, M., Schaer, S., Springer, T., Kouba, J., and Neilan, R. E. 1999. "The International GPS Service (IGS): An Interdisciplinary Service in Support of Earth Sciences," *Advances in Space Research,* vol. 23, no. 4, pp. 631–663.

Boucher, C. 2001. "Terrestrial Coordinate Systems and Frames," in *Encyclopedia of Astronomy and Astrophysics.*

Boucher, C. and Altamimi, Z. 1996. "International Terrestrial Reference Frame," *GPS World,* vol. 7, no. 9, pp. 71–74.

Gauthier, L., Michel, P. Ventura-Traveset, J., and Benedicto, J. 2001. ''EGNOS: The First Step in Europe's Contribution to the Global Navigation Satellite System,'' *ESA Bulletin,* vol. 105, pp. 35–42, February.

Hansen, A. 1998. ''The NTSB: A Stepping Stone to WAAS,'' *GPS World,* vol. 9, no. 6, pp. 73–77.

Hofmann-Wellenhof, B., Lichtenegger, H., and Collins, J. 1992. *GPS: Theory and Practice,* Springer-Verlag, New York.

Langley, R. 1997. ''GLONASS: Review and Update,'' *GPS World,* vol. 8, no. 7, pp. 46–51.

McDonald, K. D. 1991. ''GPS in Civil Aviation,'' *GPS World,* September, vol. 2, no. 8, pp. 52–59.

Muellerschoen, R. J., Bertiger, W. I., Lough, M. F., Stowers, D., and Dong, D. 2000. ''An Internet-Based Global Differential System, Initial Results,'' in *Proceedings of ION National Technical Meeting,* Anaheim, CA. January, pp 220–225.

Neilan, R. E., Moore, A. Springer, T., Kouba. J., Ray, J., and Reigber, C. 2000. ''International GPS Service 2000, Life without SA,'' in *Proceedings of ION GPS 2000,* Salt Lake City, September, pp. 438–446.

Parkinson, B. W. and Spilker, J. J., eds. 1996. *Global Positioning System: Theory and Applications,* American Institute of Aeronautics and Astronautics, Washington, DC.

Schwartz, K. P. and Sideris, M. 1993. ''Heights and GPS,'' *GPS World,* vol. 4, no. 2, pp. 50–56.

Slater, J. A. and Malys, S. ''WGS 84: Past, Present and Future,'' in *Advances in Positioning and Reference Frames,* ed. F. K. Brunner, International Association of Geodesy Symposia 118, Springer-Verlag, Berlin.

Swider, R., Braff, R., and Wullschleger, V. 1997. ''The FAA's Local Area Augmentation System (LAAS),'' *Journal of Navigation,* vol. 50, no. 2, pp. 183–192.

Tytgat, L. and Campagne, P. 2000. ''Galileo: A New GNSS Designed with and for the Benefit of All Kind [*sic*] of Civil Users,'' in *Proceedings of ION GPS 2000,* Salt Lake City, September, pp. 1362–1365.

Willis, P., ed. 2001. Special issue on GLONASS, *Journal of Geodesy,* vol. 75, no. 11.

PART 5

ORBITAL AND MISSION SPACECRAFT AND SPACE STATIONS

Nick Larter

David Deering

15.16 ORBITAL SPACECRAFT

An orbital spacecraft can be considered as a special kind of satellite which is not tied to one orbit during its missions. Although it may be in an orbit for some of its most important mission phases, it may move between orbits and orbital planes and at times may not be in orbit at all. An extreme example of the last is the deep planetary probes such as the Voyager spacecraft which may never orbit a larger body but may fly by several on a trajectory to take them eventually out of the solar system.

Spacecraft included in this Part are:

- Reusable launch vehicles during their mission phase in orbit, such as the Shuttle Orbiter
- Single-mission manned spacecraft consisting of a number of expendable modules and a recovery capsule
- Transfer vehicles
- Supply vehicles
- Upper stages
- Space laboratories—modules similar to a Space Station laboratory module but carried in the cargo bay of the Shuttle
- Manned planetary spacecraft: Apollo Command
- Manned planetary landers: Apollo Lunar Module

The NSTS Shuttle Orbiter will be used as the main example of an orbital spacecraft.

NASA's websites give excellent accounts of their space technology that may be referred to for further examples, The NSTS Shuttle Orbiter is designed as a space transport vehicle which can be used for 100 missions. The crew compartment accommodates up to 7 crew members and can handle 10 persons during emergency operations. The orbiter's 18.3 $\times$4.5 m cargo bay ferries payloads to and from low Earth orbit 185–600 km. It is similar in size and weight to modern transport aircraft. The three Space Shuttle main engines located in the aft fuselage comprise the main propulsion system. Fuel for the orbiter's main engines is carried in the external tank. Both the solid rocket boosters and the external tank are jettisoned prior to orbital insertion. In orbit, the orbiter is steered by the orbital naneuvering system, contained in two pods on the aft fuselage. The reaction control system, contained in these two pods and in a module in the nose section of the forward fuselage,

provides attitude control in space and during reentry and is also used during rendezvous and docking maneuvers. After completing operations in orbit, the orbiter reenters the Earth's atmosphere and glides to a runway landing. Nominal landing velocity is approximately 330 km/h. The orbiter is constructed primarily of aluminum and is protected from reentry heat by its thermal protection system.

The liquid hydrogen/liquid oxygen-fueled main engine is capable of operating at various thrust levels. Ignited on the ground prior to launch, it operates in parallel with the boosters during the initial ascent. After the boosters separate, the main engines become the sole propulsion element for the remainder of the ascent. The main engines develop thrust by using high-energy propellants in a staged combustion cycle. The propellants are partially combusted in dual preburners to produce the high-pressure hot gas that drives the turbopumps. Combustion is completed in the main combustion chamber.

TABLE 15.13 Main Characteristics of the NSTS Shuttle Orbiter

	Dimensions
Length	122.17 ft (37.24 m)
Wingspan	78.06 ft (23.79 m)
Height	56.58 ft (17.25 m)
	Mass
Inert mass	151K lb (69K kg) approx.
Gross mass	207K lb (94K kg)
	Structure
Type	Semimonocoque
Material	Aluminum
	Propulsion
Propellant	LOX-LH2
Average thrust	375K lb (1.67M N) SL
Engine designation	470K lb (2.10M N) Vac
Number of engines	3
Feed system	Staged combustion
Chamber pressure	2.970 psia (205 bar)
Mixture ratio	6.0
Throttling capability	65–104%
Expansion ratio	77.5:1
Restart capability	No
Control–pitch, yaw, roll	Hydraulic gimbaling (3 nozzles)
	Events
Nominal burn time	522 s
Stage shutdown	Command shutdown
Stage separation	Reaction control system
	Other
Crew	7
Flight duration	7 days (16 days for EDO)
Pressurized volume	2.525 ft^3 (71.5 in^3)
Crossrange	1,264 nm (2,034 km)
Power	14 kW continuous, 24 kW peak

The orbiter cabin is designated as a combined working and living area. The flight deck contains the displays and controls used to pilot, monitor, and control the orbiter and the mission payloads. Seating for as many as four crew members can be provided on the main deck. The middeck contains passenger seating for three crew members, the living area, an airlock, the galley, sleeping compartments, the toilet, and avionics equipment compartments. The lower deck contains the environmental control and life support system.

Table 15.13 shows the main characteristics of the NSTS Shuttle Orbiter. Table 15.14 compares the four currently operational vehicles. The fifth, *Challenger,* was destroyed in an accident in 1988. Figure 15.1 shows the NSTS Shuttle Orbiter.

15.17 TRANSFER AND SUPPLY VEHICLES AND UPPER STAGES

Single-use expendable upper stages have long been used to transfer satellites from low Earth orbit to higher orbits using Hohmann transfer maneuvers. Recently developed upper stages have become more sophisticated, with restartable engines and the transfer vehicles necessary for International Space Station logistics.

The Progress supply vehicle, workhorse of the old Salyut and Mir space station programs, is an unmanned version of the Soyuz system, though the descent module

TABLE 15.14 Comparison of the Four NSTS Shuttle Orbiters

Columbia	*Discovery*	*Atlantis*	*Endeavor*
The oldest orbiter in the fleet First on-line orbiter to undergo the scheduled inspection and retrofit program. Second orbiter outfitted with the multi-functional electronic display system or "glass cockpit."	At rollout, its weight was some 6,870 lb less than *Columbia*. Modified at Kennedy Space Center to enable it to carry the Centaur upper stage in the payload bay, although no Centaur flight was ever flown.	At rollout, its weight was some 6,974 lb less than *Columbia*. Use of thermal protection blankets on the upper orbiter body instead of tiles. The plumbing and electrical connections needed for extended duration orbiter modifications to allow up to 28-day missions.	A 40 ft diameter drag chute that reduces the orbiter's rollout distance by 1,000 to 2,000 ft. Extended duration orbitor enabled. Updated avionics systems that includes advanced general-purpose computers, improved inertial measurement units and tactical air navigation systems, enhanced master events controllers and multiplexer–demultiplexers, a solid state star tracker and, improved nose wheel steering mechanisms. An improved version of the auxiliary power units that provide power to operate the Shuttle's hydraulic systems.

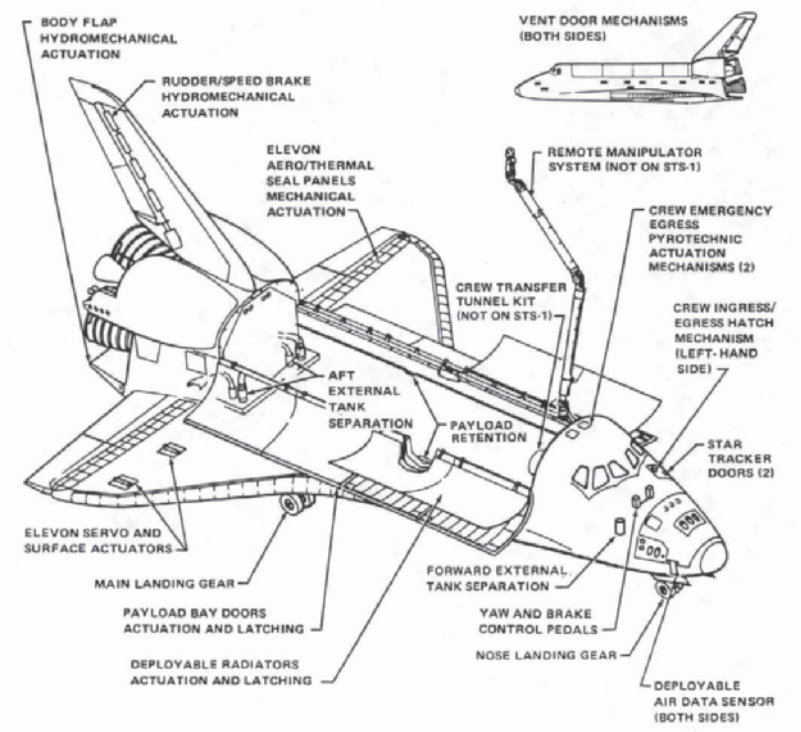


FIGURE 15.1 NSTS Shuttle Orbiter.

TABLE 15.15 Further Examples of Orbital Spacecraft

Orbital spacecraft	Country or Agency	Type	Characteristics
Soyuz	Russia	Operational system	Three-element expendable system with instrument and orbital modules and a ballistic return capsule (descent module)
Buran	Russia	Winged orbiter, no longer in operation	Shuttle-like orbiter clearly modeled on its NSTS counterpart, though with some design improvements, such as the absence of main engine hardware in the orbiter
Hermes	ESA	Defunct design	Spaceplane
HOPE	Japan	Design/ development system	Spaceplane
Shenzhou	China	Design/ development system	Soyuz-like configuration

(normally filled with space station waste) is destroyed in reentry rather than returned to the ground. Tables 15.15 and 15.16 give examples.

The biggest development in transfer vehicles in the future will arise as part of the development of more sophisticated in-orbit infrastructures. This will see the use of space tugs to shift materials and supplies around to different points in the infrastructure. These tugs will essentially be reusable versions of some of the transfer vehicles and upper stages now in use.

15.18 PLANETARY ORBITERS

For the purposes of this Handbook, mission spacecraft are considered to be those unmanned satellites that leave the orbit of the Earth. Manned mission spacecraft (of which the Apollo Service, Command, and Lunar Modules are the only examples to have seen operation) are covered below. Mission spacecraft have many design characteristics in common with satellites, the differences between the two focusing mainly on any or all of the following areas:

TABLE 15.16 Further Examples of Orbital Spacecraft

Vehicle	Country or Agency	Status	Characteristics
Automated Transfer Vehicle (ATV)	ESA	In development	Launched on Ariane 5; transfer of logistics modules to ISS like HTV
Progress	Russia	Operational	Soyuz-like
Inertial Upper Stage (IUS)	U.S.A.	Operational	Upper stage for LEO to geosynchronous transfer and boosting mission spacecraft out of Earth's orbit
Upper Stage Assembly (U.S.A.)	U.S.A.	Operational	Low-cost LEO to geosynchronous upper stage
Breeze—KM	Russia	In development	Sophisticated restartable upper stage, planned for Eorockot vehicle
Breeze—M		Operational	Sophisticated restartable upper stage for Proton launcher.
Fregget	Russia	In development	Sophisticated restartable upper stage.
Payload Assist Module (PAM)	U.S.A.	Operational	

- More sophisticated propulsion systems (as compared to Earth satellites)
- A nonsolar primary power source
- Novel braking and atmospheric entry systems for lander elements (*novel* meaning not in common use for terrestrial reentry)
- Extremely tight mass budgets
- Longer design life
- Augmented telecommunications—more accurate pointing, higher-specification antenna, etc.

Mission spacecraft can be classified into four main types:

- Orbiter
- Lander
- Wanderer
- Returner

The following paragraphs will discuss the first two of these types by considering prominent examples.

The Viking project consisted of launches of two separate spacecraft to Mars, Viking 1, launched on 20 August 1975, and Viking 2, launched on 9 September 1975. Each spacecraft consisted of an orbiter and a lander. After orbiting Mars and transmitting images used for landing site selection, the orbiter and lander divided and the lander entered the Martian atmosphere and soft-landed at the selected site. The orbiters continued their scientific operations from orbit while the landers deployed instruments on the surface. The fully fueled orbiter–lander pair had a mass of 3,527 kg of which 1,445 kg was propellant. After separation and landing, the lander had a mass of about 600 kg and the orbiter about 900 kg.

The primary objectives of the Viking orbiters were to transport the landers to Mars, perform reconnaissance to locate and certify landing sites, act as communications relays for the landers, and perform their own scientific investigations. The orbiter, based on the design of the earlier Mariner 9 spacecraft, was octagonal in shape and approximately 2.5 m across. The eight faces of the ring-like structure were 0.4572 m high and were alternately 1.397 and 0.508 m wide. The overall height was 3.29 m. There were 16 modular compartments. Four solar panel wings extended from the axis of the orbiter consisting of eight 1.57×1.23 m solar panels. They produced 620 W of power on Mars. Power was also stored in two nickel-cadmium 30 amp-hr batteries.

Propulsion was provided by a bipropellant (monomethyl hydrazine and nitrogen tetroxide) liquid-fueled rocket engine which could be gimballed up to 9°. The engine was capable of 1,323 N thrust. Attitude control was achieved by 12 small compressed-nitrogen jets. An acquisition Sun sensor, a cruise Sun sensor, a Canopus star tracker, and an inertial reference unit consisting of six gyroscopes allowed three-axis stabilization. Two accelerometers were also on board. Communications were accomplished through a 20 W S-band (2.3 GHz) transmitter. An X-band (8.4 GHz) downlink was also added specifically for radio science and to conduct communications experiments. Uplink was via S-band (2.1 GHz). A two-axis steerable high-gain parabolic dish antenna with a diameter of approximately 1.5 m was attached at one edge of the orbiter base, and a fixed low-gain antenna was extended from the top of the bus.

Following a 333-day cruise to Mars, the Viking 2 Orbiter began returning global images of Mars prior to orbit insertion. The orbiter was inserted into a 1,500 × 33,000 km, 24.6 hr Mars orbit on August 7, 1976, and trimmed to a 27.3 hr site certification orbit with a periapsis of 1,499 km and an inclination of 55.2° degrees on August 9. The lander separated from the orbiter on September 3, 1976, and landed at Utopia Planitia at 22:37:50 UT. Normal operations called for the bioshield connecting the orbiter and lander to be ejected after separation, but because of problems with this the bioshield was left attached to the orbiter. The primary mission of the orbiter ended at the beginning of solar conjunction on November 8 1976. The orbiter had developed a leak in its propulsion system that vented its attitude control gas. It was turned off on July 25, 1978, after returning almost 16,000 images in 706 orbits around Mars.

15.19 PLANETARY LANDERS

The Viking Lander is used to illustrate these spacecraft. The combined orbiter and lander of Viking 2 reached its Mars orbit in August 1976. Imaging of candidate sites was begun and the landing site was selected based on these pictures and the images returned by the Viking 1 Orbiter. The lander and its aeroshell separated from the orbiter on September 3. At the time of separation, the lander was orbiting at about 4 km/s. After separation, rockets were fired to begin lander deorbit. After a few hours, at about 300 km altitude, the lander was reoriented for entry. The aeroshell with its ablatable heat shield slowed the craft as it passed through the Martian atmosphere. During this time, entry science experiments were performed. At 6 km altitude at about 250 m/s the 16 m diameter lander parachutes were deployed. Seven seconds later the aeroshell was jettisoned, and 8 seconds after that the three-lander legs were extended. In 45 seconds the parachute had slowed the lander to 60 m/s. At 1.5 km altitude, retrorockets were ignited and fired until landing 40 seconds later at about 2.4 m/s.

The Viking 2 Lander touched down about 200 km west of the crater Mie in Utopia Planitia at 48.269° N latitude and 225.990° W longitude. Approximately 22 kg of propellants were left at landing. Due to radar misidentification of a rock or highly reflective surface, the thrusters fired an extra time 0.4 seconds before landing, cracking the surface and raising dust. The lander settled down with one leg on a rock, tilted at 8.2°. The cameras began taking images immediately after landing. The Viking 2 Lander operated on the surface for 1,281 Mars days and was turned off on April 11, 1980, when its batteries failed.

15.20 ORBITAL LABORATORIES

Most of the pressurized modules of the International Space Station and indeed of past space stations can be considered as orbital laboratories. The following is a description of the Spacehab. The Spacehab module is a pressurized laboratory designed to add-to Space Shuttle middeck experimental space. The Spacehab system consists of a module flown in the orbiter payload bay that is configured with middeck-type lockers, racks, and/or the logistics transportation system to accommodate a variety of experiments and equipment. The Spacehab module provides space for

crew members to work on experiments in a shirtsleeve environment. Three module configurations are offered.

Spacehab configuration 1 is connected to the orbiter using a modified Spacelab tunnel adapter. The Spacehab module provides crew members with a place to carry out their experiments and contains cooling, power, and command and data provisions, in addition to the following Spacehab housekeeping systems: power distribution and control, lighting, fire and smoke detection, fire suppression, atmosphere control, status monitoring and control, and thermal control.

Spacehab configuration 2 allows the Spacehab single module to be mounted in a new trunnion location to accommodate the orbiter docking system. All Spacehab module subsystems remain the same as configuration 1, except for a lower air exchange rate with the orbiter and the addition of two negative pressure relief valves.

Spacehab configuration 3 is a double module, consisting of one Spacehab module and one Spacehab module shell join by an intermediate adapter. Configuration 3 has the same tunnel configuration and attach points as configuration 2, except for two trunnions that are moved farther aft to accommodate the additional module. All Spacehab module subsystems remain the same as configuration 2, except for the addition of a dc fan and lights in the aft module segment.

15.21 SPACE STATIONS

If an artificial satellite can be defined as a body which once launched is inserted into one orbit which it then follows for the duration of its mission, then a space station can be considered to be a manned artificial satellite. The main differences between a space station and a normal unmanned satellite can therefore be seen to be as follows:

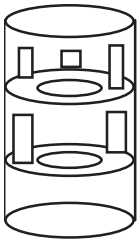
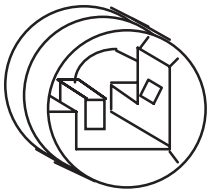
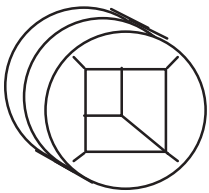
- Size: since it has to accommodate a human crew, a space station is larger than other existing satellites (though some classes of satellite which have been designed but not yet built, such as solar power satellites, would be at least as large as a space station). More fuel is therefore required for the space station to be able to keep in its nominal orbit.
- Complexity: supporting a crew requires the presence of additional special subsystems (for example; life support and nutrition and sanitation subsystems (see Part 7), as well as the presence of the corresponding logistical operations to supply food and fuel.
- Contamination: the outgassing of atmosphere from pressurized modules where the crew live and work, plus the intended venting of some gases and the continual operations in the proximity of the space station, which may be associated with logistics, construction, maintenance, or science, contaminate the space station's immediate environment, so many kinds of sensitive scientific sensors for astronomy or Earth observation cannot be located there.
- Residual accelerations due to crew activities, the docking and undocking of supply vehicles, and so forth which result in a noisier microgravity environment.

Three distinct types of design of space station have been realized operationally (Table 15.18), but only one of these, the ISS, is still in orbit and active. All three

TABLE 15.17 Summary of Single and Double Spacehab Module Characteristics

Description	Single module (SM)	Double module (DM)
Width	13.5 ft	13.5 ft
Height	11.2 ft	11.2 ft
Volume	308 ft^3	1100 ft^3
Launch weight	10,000 lb	20,000 lb
Cargo capacity	3000 lb	10,000 lb
Lockers	61 max	61 max
Rack/locker	1 rack/51 lockers max	61 lockers, 4 rooftop locations
Configuration	2 racks/41 lockers max	4 racks (2 powered) Floor stowage available for large, unique items
Windows	None	One
Crew required	2	2

TABLE 15.18 Past and Present Operational Space Station Module Types

Type	Sketch	Examples
Multistoried		Skylab
''Suburban garage''		Salyut series, Mir
Modular racked		Spacelab, Spacehab, NSTS shuttle orbiter middeck (orbital laboratory precursors), later Mir modules (e.g., Spektr), ISS

types are based on cylindrical modules, but the internal architecture of the cylinder differs radically between them.

International Space Station (ISS)

Assembly of the Space Station started in November 1998 with the launch of the first Station element, the Zarya module, into orbit, and will continue until at least 2003. Once the Station is complete—currently planned for 2006—it will offer an extensive range of facilities that cannot be found on Earth, and will enable mankind to continue to learn how to live and work in space for long periods. The Station facilities are contributed jointly by the international partners: United States, Russia, member states of the European Space Agency, Japan, and Canada; and will enable activities in the fields of physical and life sciences, space science, Earth observation, and technology innovation. Users do not have to wait until the Station is completely assembled before they can start using it. At the time of writing, utilization has already begun, with experiments being carried out on the Zvezda service module and the U.S. laboratory module Destiny, and this will continue with facilities and resources becoming increasingly available. The benefits the Station offers users include:

- The capability to perform an experiment or observation program over an extended period of time in weightless conditions
- The capability to perform iterative research on a short time scale through the provision of regular access to and return from the Station
- Provision of access to a significant level of resources
- The permanent presence of crew

A summary of the Station's key characteristics at the end of the build-up period in 2006 is provided in Table 15.19

The ISS is likely to represent the single operational space station for the foreseeable future. It remains to be seen how much this design will be replicated for

TABLE 15.19 International Space Station Key Characteristics

Parameter	Characteristic
Truss length	108 m
Module length	74 m
Weight	420 tonnes
Maximum power output	110 kW
Total volume	1,200 m^3
Atmospheric pressure	1,013 mbar
Orbital Altitude	350–460 km
Orbital Inclination	51.6°
Orbital velocity	8 km/s
Attitude	Local vertical/local horizontal
Crew (max)	7
Data rate (uplink and downlink)	72 kbit/s and 43 mbit/s
Ku and S-band coverage	68% and 50% of time
Anticipated lifetime	>10 years

other elements of a future in-orbit infrastructure—for example Earth–Moon or Earth–Mars staging posts—or whether some other design will supplant the current truss and module standard.

15.22 MANNED INTERPLANETARY SPACECRAFT AND LANDERS

The Apollo system is the only example of a manned interplanetary vehicle which has been built and used. The following is a brief description of the parts and operation of the Apollo Command and Service Modules.

The Apollo Command Module whose specification is given in Table 15.20 and shown in Figure 15.2 is a pressurized vessel encased in heat shields. It was the command, communications, and control center for the lunar missions. The crew compartment contained 210 ft^3 of habitable volume to accommodate the three astronauts, controls and displays, and many of the spacecraft systems. The forward compartment contained two reaction control engines and Earth landing equipment. The aft compartment housed 10 reaction control engines, propellants, and helium and water tanks.

TABLE 15.20 Apollo Command Module Specification

Crew size	3
Overall length	3.5 m
Maximum diameter	3.9 m
Habitable volume	6.17 m^3
Total mass	5,806 kg
structure	1,567 kg
heat shield	848 kg
reaction control system	400 kg
recovery equipment	245 kg
navigation equipment	505 kg
telemetry equipment	200 kg
electrical equipment	700 kg
communications systems	100 kg
crew seats and provisions	550 kg
crew mass	216 kg
miscellaneous contingency	200 kg
environmental control system	200 kg
propellant	75 kg
Main engine propellant	NTO/Aerozine-50
Reaction control system	
thrusters	12 × 410 N
propellant	NTO/MMH
specific impulse	290 s
total impulse	257 kNs
L/D hypersonic	0.3
Power: batteries	20.0 kW, 1,000.0 Ah
Environment	pure oxygen at 340 mbar

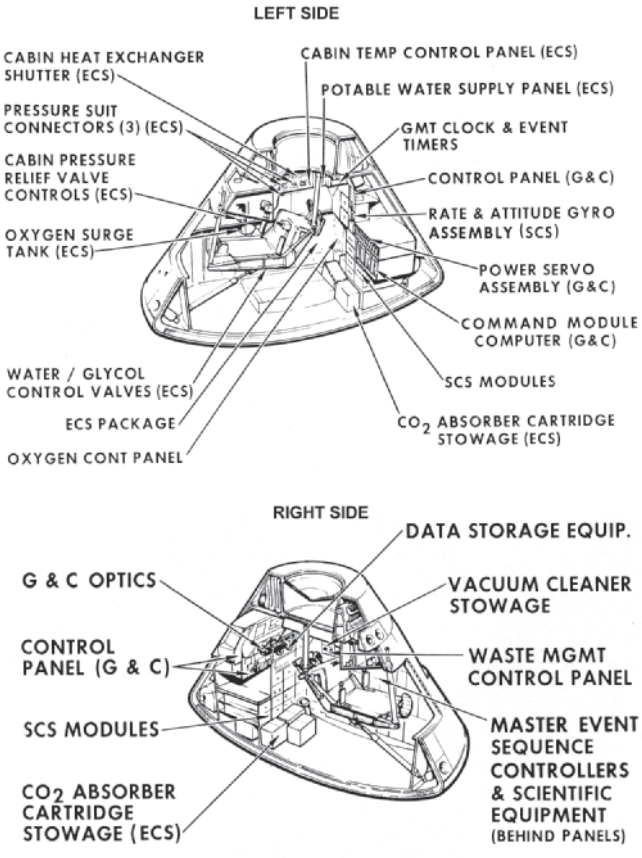


FIGURE 15.2 Interior of the Apollo Command Module

The blunt-end design for the Command Module was chosen to build upon experience gained with the similarly shaped Mercury and Gemini spacecraft. The spacecraft reentered the atmosphere with its protective heat shield (the widest end of the spacecraft) facing forward. Layers of special ablative material on the shield were purposely allowed to burn away during reentry to help dissipate the heat generated by atmospheric friction. Its height was 3.2 m, diameter 3.9 m, and weight 5,900 kg. It was manufactured by Rockwell and launched by a Saturn 5 rocket.

The Service Module was jettisoned before entering Earth's atmosphere. Pitch-control jets turned the command module so that the heat shield orientated to withstand the atmospheric heating. The commander maneuvered the module to within a 42 km wide reentry corridor. Drag-braking reduced speed to around sonic velocity at 9.14 km altitude. Drogue parachutes deployed at 7.62 km. From 4.57 km, three 24.4 m diameter ringsail parachutes lower the module gently into the sea. The atmosphere entry angle was 5.3° at 121.92 km. The maximum deceleration was of order 6 g.

The second Apollo module shown in Figure 15.3 to be considered is the Lunar Landing Module. Developed to implement the 1962 lunar orbit rendezvous decision, the lunar module was a two stage, self-sufficient spacecraft built by Grumman. It was carried to the Moon attached to the command module docking port, with its descent engine pointing forward. In this configuration, the descent engine could be fired as a back-up propulsion system to boost the lunar module out of lunar orbit in the event that the Apollo engine failed before the landing sequence.

It is difficult to predict what the next manned planetary spacecraft will look like. A very large number of paper studies have been carried out, featuring radically different designs. The key determinants on the design are likely to be the extent of the in-orbit infrastructure in place at the time of the next mission. This will also depend on whether it is launched from the Earth or partially constructed and assembled in Earth orbit and launched from there.

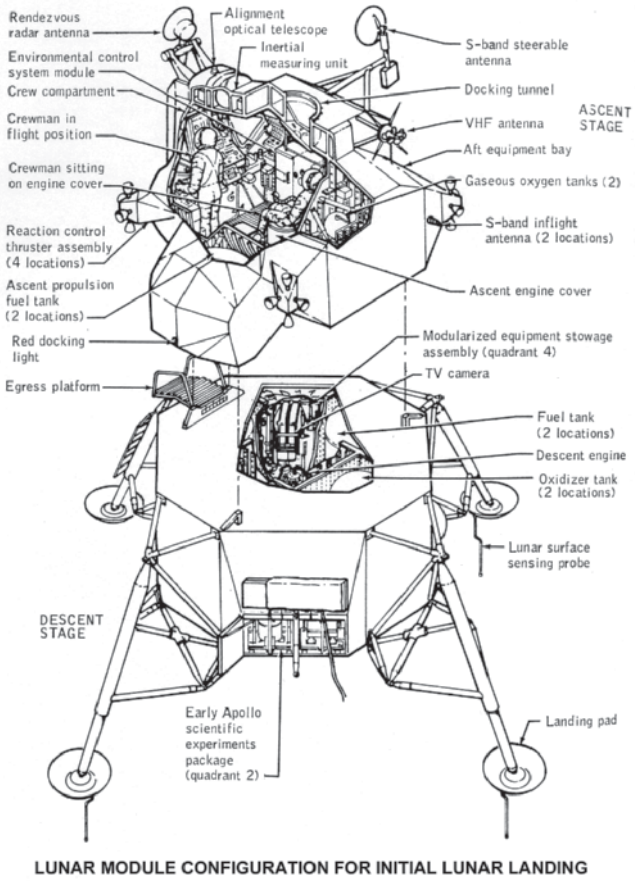


FIGURE 15.3 Three views of the Lunar Module

15.23 MANNED PLANETARY BASES

The closest that mankind has come so far to the establishment of a manned planetary base was in the plans for the later, aborted Apollo missions, which envisaged longer-duration stays on the lunar surface and the erection of temporary shelters. Since that time, progress toward the establishment of a manned planetary base has been set back by a succession of false starts in the United States, the *de facto* leader of any international effort, regarding the necessary political will and funding to see through such a program. Possibly more germane to the delays, though, is the ongoing technical debate as to the preferred destination, which has involved scientists and engineers throughout the spacefaring world, and the lack of resolution of which has tended to undermine whatever political and fiscal support was available. The ''Moon-versus-Mars'' debate hinges on the relative merits and drawbacks of the two planets. In the Moon's favor are the following:

- Proximity to Earth—requires less of an in-orbit infrastructure to service the program properly.
- Detailed proposals for mining and processing lunar resources to support low Earth orbit infrastructure development already exist which could help to justify the program.
- The Moon could be used as a testbed for techniques and hardware to be used on a later effort to reach Mars.

The case for Mars rests on the following:

- A return to the Moon would represent a step back into the past, whereas Mars is the future
- Mars is, relatively speaking, a less hostile environment than the Moon.
- Mars is scientifically richer.
- Mars has a greater diversity of natural resources, meaning that fewer resources would need to be transported there to support a base.
- The cost of reaching Mars in energy terms is little more than that of reaching the Moon, as most of the energy is expended in order to leave Earth's gravity well.
- The public's awareness of Mars is high thanks to the mystique that has built up over a century or more concerning life on Mars and Martians—any manned effort to reach Mars would therefore be likely to have good popular support.

Whichever destination is ultimately selected, it is clear that a good supporting infrastructure in low Earth orbit and beyond—at the Earth–Moon Lagrangian points, for example—will be an important precursor to any missions to either. The lack of any such infrastructural presence during the period of the Apollo missions was certainly a drawback and may have provided some technical reasons to discontinue the program. For the future, therefore, the presence of the ISS is a large bonus.

PART 6
PAYLOAD MANAGEMENT

Nick Larter

15.24 OVERVIEW

Table 15.21 categorizes the main types of space and space-related payloads. Payloads may be characterized first by the kind of spaceflight to which they are subjected:

- Nonorbital
- Suborbital
- Orbital
- Interplanetary
- Miscellaneous

The location of the payload is the next determinant—it may be on the outside of the carrier, exposed to the space environment, or it may be inside the vehicle, in which case it will generally be inside a pressurized compartment.

The third determinant is whether the payload is on a manned or unmanned carrier. This distinction is important because manned operations in the proximity of the payload generally lead to contamination of various kinds. The resolution of most kinds of sensors for space science or Earth observation is degraded through the deposition of foreign material on the sensor.

The fourth determinant is the kind of science being carried out. The fifth is the kind of carrier, which this influences both the kind of environment experienced (e.g., suborbital flights might delivery microgravity conditions but not exposure to space radiation) and the maximum duration of exposure that the payload will experience. Figure 15.4 shows this with reference specifically to microgravity.

The following subsections elaborate on each of the main payload types. For each payload type there is a brief overview, then one example of the payload is described in detail, and finally a short discussion is presented on likely future trends.

15.25 PLANETARY MISSION INSTRUMENT PACKAGES

One of the primary reasons to place instrumentation on another planet is to attempt to detect life. At present, most attention is focused upon Mars. The fundamental ethical problem is that of avoiding the inadvertent transport of viable biological material originating from Earth to Mars, where, conditions permitting, it could survive, propagate, and affect any indigenous life, life indicators, or prelife chemistry. When the purpose of such Mars missions is to search for such life indicators, it follows that there are subsidiary cases also of interest, such as:

TABLE 15.21 Main Types of Space and Space-Related Payloads

Flight type	Payload carrier type	Main disciplines
Nonorbital	reentry vehicle demonstrator (aerodynamic part)	engineering test
	balloon	Earth science
	balloon borne drop capsule	microgravity research
	parabolic flight	engineering test
	parabolic flight	microgravity research
Suborbital	reentry vehicle demonstrator	engineering test
	sounding rocket	Earth science
	sounding rocket	space science
	launch/reentry vehicle demonstrator	engineering test
	reentry vehicle demonstrator	engineering test
	sounding rocket	microgravity research
Orbital	satellite	engineering test
	satellite	Earth science
	satellite	communications
	satellite	space science
	free flyer	space environment research
	free flyer	space biology
	space station	engineering test
	space station	space environment research (including man-made contamination environment
	recovery capsule	engineering test
	recovery capsule	space environment research
	space station	space biology
	space station	microgravity research
	space station	space medicine/human factors
	space station	
	space station	
Interplanetary	interplanetary spacecraft	planetary science
	interplanetary spacecraft	space biology (exobiology)
	interplanetary spacecraft	space biology (exobiology)
Miscellaneous	closed systems	space biology
	closed systems	space biology
	planetary base analogues	space medicine/human factors
	hyperbaric chamber	space medicine/human factors
	bed-rest studies	space medicine/human factors
	neutral buoyancy tanks	engineering test

- The inadvertent transport of biological precursor materials, such as complex organic chemicals, and their potential effect on any indigenous life or life indicators
- The inadvertent transport of hazardous chemicals such as sterilants, which could, in the immediate surroundings of the spacecraft, affect indigenous life or life indicators

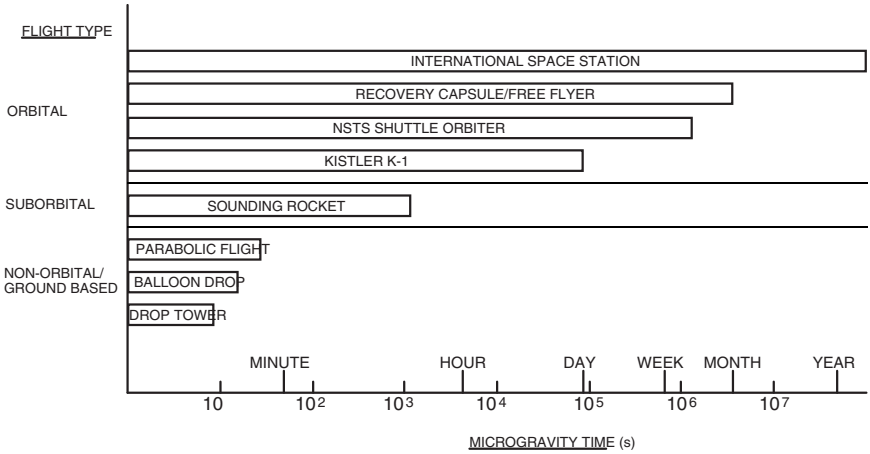


FIGURE 15.4 Microgravity duration offered by different carriers.

- The assurance that analytical equipment on the spacecraft dedicated to the search for life is contaminated neither by organic matter from Earth, which could lead to false positive results, nor by hazardous chemicals, which could lead to false negative results

The COSPAR (Committee on Space Research) planetary protection guidelines classify planetary bodies (including moons and asteroids) according to their susceptibility to terrestrial biological contamination and places sterility and verification requirements on visiting missions accordingly. There are four levels of classification and the most stringent requirements (class IV) relate to Mars. COSPAR deliberations suggest that the guidelines for orbiters and landers for the same planetary body should have different priorities—a new idea reflected in the resolution on Mars Planetary Protection Provisions originating from COSPAR Sub-Commission F, which were put to the Hamburg meeting in 1994. The COSPAR guidelines can be regarded as the global standard for planetary protection, which national agencies can then use as the basis for their own planetary protection specifications.

NASA's own specifications for planetary protection are compliant with the COSPAR guidelines. The Space Studies Board of the U.S. National Research Council (NRC) also has a role in formulating planetary protection guidelines, as for example followed the publication of its benchmark report *Biological Contamination of Mars*—Issues and *Recommendations* in 1992, which was subsequently adopted by COSPAR.

An interesting approach is to distinguish between the detector portion of a lander, which would be upset by traces of imported inorganic material and which therefore has to be absolutely clean, and the rest of the lander, which in the absence of viable organisms could only contaminate its own landing site, or rather a larger area if it crash landed, in which case the issue of the distribution of debris by winds would have to be addressed. What this implies is the possibility of two different standards for sterilization: one applicable to the detector portion, which after landing could be physically separated in some way from the rest of the lander and where the

standard is set by the properties of the detector; the other applicable to the lander as a whole, where the level is set by considerations of the contamination in the event of an uncontrolled impact or at a very much longer time scale as a result of its erosion and ultimate disintegration in the prevailing environment. This gives two cleanliness zones on the lander, and the key issue then is how to transfer the samples from the outer, less clean zone to the inner, cleaner zone.

The search for relict carbonaceous material also raises the idea of to what extent it is possible to construct the spacecraft without using carbon.

A key issue with any instrumentation to be used upon another planet is sterilization. The following questions must first be answered: How are space bioloads estimated? What organisms and residues are involved? What are the worst cases? What are the vulnerabilities of different classes of planets and bodies to contamination: What constitutes contamination and what challenge of biological material would be lead to it? What changes are necessary in spacecraft materials, processes, and procedures to accommodate sterilization? What are the budgetary and programmatic impacts of sterilization? What are the fundamental limitations on the application of techniques caused by necessary ground or flight operations? What are the limitations of current verification and certification processes?

The first principle of sterilization is that the objects to be sterilized must be clean, as the presence of contaminants such as organic matter greatly reduces the efficacy of any sterilization process. The stringent cleanliness requirements associated with spacecraft construction ensure that this prerequisite is invariably met. Sterilization techniques can be divided into chemical methods (sterilants applied in liquid or vapor phase) and physical methods (thermal, ionizing radiation, ultrasound, and ultraviolet light).

For spacecraft sterilization there are a number of problems in using chemical sterilants which stem from possible adverse effects on the spacecraft materials and electronic components and the deposit of residues. Chemical sterilants have been most widely used to reduce the problems with microbial build-up in long-term manned spacecraft, such as the Russian Mir station. The driver here is that the sterilant has no adverse effect on the crew.

Thermal treatment can be as dry heat (using ovens) or wet heat (using autoclaves). The steam sterilization process using autoclaves is the method of choice for most terrestrial applications, as the presence of the steam accelerates heat penetration. Typical wet sterilization temperatures are 115°C for 10 minutes, or 125°C for 15 minutes. For dry sterilization, baking at 125°C for hours or even days may be necessary.

Pyrogens (lipids of bacterial origin associated with proteins and polysaccharides) are one class of bacterial residue which is not destroyed or inactivated by normal sterilization processes. Heating to 115°C for 30 minutes inactivates pyrogens, but heating in an oxidizing atmosphere would be necessary to destroy any residual biomass and is very difficult to achieve.

The NASA Viking landers were baked for long periods as the primary means of sterilization. The Mars Express mission will use sterilization by baking at 125°C for 5 hours as one of its primary means of sterilization. Even so, the cost of rendering Viking added up to U.S. $200 million, at 1970s prices, to its mission costs, and residual areas of concern remained.

Ionizing radiation is widely used for sterilization in the food industry and for prepackaged medical supplies. It may be possible to use this technique, or sterilization by ultrasound, for some spacecraft instrument components that would be sensitive to thermal or chemical treatments.

Sterilization using ultraviolet light is most prevalent in applications where the aim is to maintain the sterility of previously sterilized equipment, in, for example, microbiological laboratories where ultraviolet is used in conjunction with a laminar flow cabinet where sterile media need to be handled. The main application to spacecraft sterilization is therefore in the assembly of small sterilized components or modules into larger modules.

Sealing subsystems is not always the answer. The Beagle 2 package for Mars Express incorporates new technologies such as miniature CCD cameras for the microscope. These CCD camera units are usually supplied as a sealed unit, but there is no guarantee that the interior is sterile. The provision of any sealed unit by a third party therefore provides special problems for the planetary spacecraft constructor, who is faced with unsealing and treating the units, which may not be feasible, or else procuring special units which are treated during manufacture, at greatly increased cost. The problem with sealed units is that in the event of crash landing the sealed unit could be broken, releasing whatever is inside. For this reason, in the past the potting of electronic boards to trap any organisms or other biological material was rejected. Current research is beginning to look at crash-resistant bioshields instead.

It is likely that the use of single-technique sterilization approaches, like the baking applied to Viking, will not be suitable for future spacecraft, owing to their complexity. An integrated approach is required where each element of the spacecraft is sterilized by the most appropriate means, verified, and then stored in a sterile state. Assembly of the spacecraft elements into larger modules is then carried out in aseptic conditions. For example, missions to Mars are subject to Class 4 planetary protection under the COSPAR guidelines, making it necessary to:

1. Physically remove as much biomass as possible from the vehicle by aseptic assembly.
2. Degrade any residual biomass, whether living or dead, spore, virus, pollen, etc., to a state in which it would be no longer be identified by the sensors fitted to the science payload.

It is difficult to achieve this second point by any means other than heating in an oxidizing atmosphere.

The following is an example of typical instrumentation carried on a planetary vehicle. On the Viking 2 Orbiter, scientific instruments for conducting imaging, atmospheric water vapor, and infrared thermal mapping were enclosed in a temperature-controlled, pointable platform extending from the base of the orbiter. The scientific instrumentation had a total mass of approximately 72 kg. Radio science investigations were also performed using the spacecraft transmitter. Command processing was carried out by two identical and independent data processors, each with a 4096-word memory for storing uplink command sequences and acquired data.

The Viking 2 Lander carried instruments to study the biology, chemical composition (organic and inorganic), meteorology, seismology, magnetic properties, appearance, and physical properties of the Martian surface and atmosphere. Two 360° cylindrical scan cameras were mounted near one long side of the base. From the center of this side extended the sampler arm, with a collector head, temperature sensor, and magnet on the end. A meteorology boom, holding temperature, wind direction, and wind velocity sensors, extended out and up from the top of one of the lander legs. A seismometer, magnet and camera test targets, and magnifying mirror were mounted opposite the cameras, near the high-gain antenna. An interior

environmentally controlled compartment held the biology experiment and the gas chromatograph mass spectrometer. The X-ray fluorescence spectrometer was also mounted within the structure. A pressure sensor was attached under the lander body. The scientific payload had a total mass of approximately 91 kg.

The purpose of the physical properties investigation was to determine the properties of the Martian surface and environment at the landing site. In particular, it attempted to determine such properties as bulk density, bearing strength, angle of repose, cohesion, angle of internal friction, particle characteristics, thermal parameters, eolian transportability, topography, and certain environmental properties such as wind, temperature, and solar flux levels.

The entry science atmospheric structure experiment (one of three that were part of the entry science investigation) studied the Martian atmosphere below an altitude of 132 km. A variety of instruments (accelerometers, radar altimeters, thermometers, pressure sensors) collected data to provide altitude profiles of pressure and temperature and acceleration of the lander capsule. From these data, atmospheric density and mean atomic mass could be calculated.

The biology experiment looked for the presence of Martian organisms by analyzing for metabolic products. Three distinct instruments (pyrolytic release, labeled release, and gas exchange) incubated samples of the Martian surface under a number of different environmental conditions. The pyrolytic release instrument sought to detect the photosynthetic or chemical fixation of CO_2 or CO containing C-14. The samples were incubated for several days in the presence of the radioactive gas mixture. Next, each sample was heated to 120°C to remove unreacted CO_2 and CO. The soil was pyrolized at 650°C and any organic products were collected in an organic vapor trap. Finally, the trap was heated to combust the organic material to CO_2 and any evolved radioactive gas was measured. The labeled release experiment sought to detect metabolic processes through radiorespirometry. Liquid nutrients labeled with radioactive carbon were added to the samples and the atmosphere above was continuously monitored to detect any radioactive. The gas exchange experiment measured the production and/or uptake of CO_2, N_2, CH_4, H_2, and O_2 during incubation of a soil sample. The sample was sealed and purged by He, then a mixture of He, Kr, and CO_2 was introduced as an initial incubation atmosphere. After the addition of a selected quantity of a nutrient solution (saturated with the diagnostic gas, neon), the sample was incubated. At certain intervals, samples of the atmosphere were removed and analyzed by a gas chromatograph with a thermal conductivity detector.

The molecular analysis experiment searched for chemical compounds in the upper surface layer of Mars and measured atmospheric composition near the surface. The soil analyses were performed using a gas chromatograph mass spectrometer that had high sensitivity, high structural specificity, and broad applicability to a wide range of compounds. Substances were vaporized from the surface material by a heating process with CO_2. The material was then carried into a Tenax gas-chromatographic column that was swept with hydrogen as a carrier gas. While passing through the column, substances were separated by different degrees of retention. The residual stream moved into the mass spectrometer.

The lander imaging experiment viewed the scene surrounding the lander, the surface sampler, and other parts of the lander, the Sun, Deimos, and Phobos to provide data for operational purposes and for geological and meteorological investigations. Two scanning cameras, capable of resolving at a high resolution of 0.04° or low resolution of 0.12°, were used on each lander. Each image acquired covered a vertical field of 20° (high resolution) or 60° (low resolution, color, and IR) and a

horizontal field that was commandable from 2.5° to 342.5° in 2.5° increments. Images were acquired from 40° above the nominal horizon to 60° below, and were commandable in 10° increments. The cameras were mounted 1.3 m above the nominal landing plane. The two cameras were separated by 0.8 m to obtain stereoscopic pictures, which were obtained over most of the scene.

The meteorology experiment analyzed the meteorological environment near the planetary surface. The atmospheric parameters determined were pressure, temperature, windspeed, and wind direction. Diurnal and seasonal variations were of particular importance. The sampling rates and durations for any one Martian day (sol) were selectable by ground command. The sensors were mounted on a boom. Three hot-film anemometers, through which an electric current was passed to heat two glass needles coated with platinum and overcoated with aluminum oxide, were used to measure windspeed. The electric power needed to maintain these sensors at a fixed temperature above the surrounding atmospheric gas was the measure of windspeed. Atmospheric temperature was measured by three fine-wire thermocouples in parallel. A thin metal diaphragm, mounted in a vacuum-sealed case, was used to measure atmospheric pressure.

The seismology experiment was designed to determine the level of seismic activity on Mars and its internal structure. The seismology instrument contained three mutually perpendicular seismometers. Each seismometer consisted of a moving coil and a fixed magnet. The data were compressed for transmission to Earth by averaging the amplitude of normal ground noise over a 15 s period. When an event occurred, a trigger activated a higher data rate mode that sampled the amplitude of the overall event envelope, which required only one amplitude sample per second to indicate its shape. The Viking 1 seismometer failed to uncage and could not be used in a seismic network with the Viking 2 instrument.

The magnetic properties experiment detected the presence of magnetic particles in Martian surface material. It used three pairs of samarium-cobalt magnets, two mounted on the backhoe of the surface-sampler collector head and one on top of the lander. Each pair consisted of an outer ring magnet about 2.5 cm in diameter with an inner core magnet of opposite polarity. The magnets were directly imaged by the camera system.

The lander radio science experiment used the S-band radio transmitter to acquire Doppler and range data for the lander, utilizing the same Deep Space Network facilities that were used by the orbiters. The resulting data were used to determine the location of the lander on the planet surface. They also provided more precise information about the orbital, rotational, and precessional motion of Mars than had previously been available.

The Viking entry science neutral atmospheric composition experiment was designed to provide the composition data for the various neutral species that were needed to define the physical and chemical state of the Martian atmosphere. Mounted in an opening in the aeroshell with its electron-impact open ion source recessed below the surface of the aeroshell, a double-focusing (electrostatic and magnetic) mass spectrometer was used to measure the concentrations of the atmospheric species that have mass-to-charge ratios from 1 to 49.

Surface samples were irradiated with X-rays in this experiment utilizing an energy-dispersive X-ray fluorescence spectrometer (XRFS) in which four sealed, gas-filled proportional counters detected X. The output of the proportional counters was subjected to pulse-height analysis by an on-board step-scanning, single-channel analyzer with adjustable counting periods. This instrument was located inside the lander body, and the lander surface sampler delivered samples to it. Reconstructed

spectra yielded surface composition data with accuracies ranging from a few tens of parts per million for trace elements to a few percent for major elements.

Lastly, the ionospheric properties experiment studied the composition, structure, and temperature of the ionosphere, which were probed during the descent of the lander capsule by means of retarding potential analyzer mounted flush with the front face of the aeroshell. To conserve battery power, the instrument was operated intermittently between 16,000 and 5,000 km altitude but continuously from 5,000 to 100 km. It was made up of a current-collecting plate with seven grids ahead of it. A fixed program of potentials was applied to the grids, and the collected currents were measured at 10 ms intervals. The instrument operated in three phases to measure energetic electrons, thermal electrons, and thermal ions.

15.26 SPACE LABORATORY RACK SYSTEMS

Any payload intended for a manned space station laboratory module is likely to coexist with the crew for a considerable time, and this, owing to the closed environment, imposes a number of design constraints on the payload. Further complicating factors for the design case are:

1. Many of such payloads concern space life sciences or microgravity research, integral to which is the handling and manipulation of samples and consumables of a biological or chemical nature
2. Current guidelines for crew working conditions in spacecraft stipulate a shirtsleeve environment. Any protective measures must therefore rest with the way the experiments or samples are packaged.

Payloads can affect the crew environment in a closed system in many ways. The first is the release of trace chemical contaminants into the atmosphere; depending on the contaminant released, the issue may be one of habitability (e.g.; odor, irritancy) or actual toxicity to the crew. These contaminants can arise from the offgassing of plasticizing chemicals from plastic materials, emissions from electronic equipment, or spillages of reagents or samples from the actual experiment. The second is hazard due to the breakage of equipment—fragile items such as glass electrodes for measuring pH and so forth are too hazardous to use without special precautions. The third is ambient noise. The cumulative effect of the running of many facilities within space laboratories is to make them very noisy environments. The fourth is support of microbial growth, which can result in additional contaminants entering the atmosphere, posing a risk to both crew health and equipment (several of Mir's subsystems were damaged by microbial growth by the end of its operational life).

In order to address the first two problems, the concept of *triple containment* has been developed. It essentially means that there have to be three layers of protection between the hazard and the crewmember. This is achieved in practice through the use of multiple levels of experiment containment and the use of a glove box with a lower internal pressure, if samples have to removed from the container and manipulated by the crew.

For hazardous chemicals, the concept of spacecraft maximum allowable concentration (SMAC) has been developed. This essentially takes the maximum safe

allowable concentration of the chemical in the atmosphere and then, given the known volume of a laboratory module, works out the largest amount of the substance that can be packaged in a single unit if the worst case, that all of it is released accidentally into the atmosphere, were to happen.

The problem with SMAC is that it cannot account for all hazards. The only way to control this is to design for such control from the beginning. This means designing laboratory modules which can be easily cleaned and sterilized where necessary. The inability to isolate a single module's atmosphere and the radial symmetry favored for the rack layout that makes access behind the racks difficult manifestly fails to provide the necessary circumstances to be able to clean and sterilize a module effectively. In Mir, later missions had to use peroxide-based chemical disinfectants and also test payload and experimental materials for the degree to which they could support microbial growth, in order to try and limit the problems which were arising.

The rack systems used in most space laboratories are based on the same IEEE 19 in. panel unit standard that is so familiar in aircraft avionic systems or any server room. While it makes sense to use such a standard, the trend toward arranging racks in a radial symmetry around the walls of the module creates a large number of issues which the designers have yet to address. Within the International Space Station laboratory modules the rack systems have to operate in and conform to various operational environmental parameter. The operational atmosphere is defined in Table 15.22

The internal lighting for the ISS consists, first of *general illumination* produced by a number of module lighting units distributed throughout the station, which may be manually controlled either remotely or locally. The illumination is adjustable between a minimum and a maximum level. A common level of illumination throughout all pressurized modules of the ISS is planned. Secondly, there is *portable lighting;* a number of these lighting units are available for temporary crew use. Finally there is *emergency lighting,* which is common throughout all pressurized modules. A common interior color scheme is used throughout all the pressurized modules (excluding those in the Russian segment) to ensure a consistent environment for the crew.

Owing to the length of the predicted operational lifetime of the ISS (at least 10 years), the control of contamination within the pressurized modules is crucial to

TABLE 15.22 International Space Station Operational Atmosphere

Parameter	Operational value
CO_2 partial pressure	Less than or 4 hPa
O_2 partial pressure	Approximately 200 hPa
Total pressure	Approximately 1,000 hPa
Temperature	18–27°C
Dewpoint	4.6–15.5°C
Relative humidity	25–70%
Ventilation velocity	0.076–0.203 m/s
Dilute gas	Nitrogen
Microorganism	Less than or 1,000 CFU/m^3
Particulate level	Class 100,000

maintain an efficient working environment for the crew, equipment, and user payloads. Contamination can effect the health of the crew, reduce the operational lifetime of equipment, and increase required maintenance activities.

To facilitate in-orbit interchangeability between different pressurized modules, internal payloads on the ISS are primarily accommodated within standard payload rack. The exception to this general statement is the two Russian research modules, which do not allow the accommodation of such racks.

Both NASA and the National Space Development Agency of Japan (NASDA) have developed International Standard Payload Racks, which may be utilized by users and payload developers, with interfaces and capabilities that are almost identical. The NASDA rack is the basic accommodation for European payloads. This is a nonsealed structure made of aluminum. Removable side and rear panels may be taken off during payload integration on the ground or to provide in-orbit access during the payload operations phase. The rack structure is 2013.4 mm high and 1046 mm wide and has a maximum depth of 858 mm. The six-post version supports up to 704 kg of payload mass and provides 1.2 m^3 of internal volume.

There are two power ratings of ISPR: a medium-power 6kW and a low-power 3 kW. The placement of a medium-power rack in a low-power location is not possible, but low-power racks may be placed in any location. International Standard Payload Racks are the largest individual entity that can be transported to and from orbit.

As an example of an ISS laboratory, consider the Biolab. The Biolab is the European Columbus Module laboratory, designed to support biological experiments on microorganisms, cells, tissue cultures, small plants, and small invertebrates. The major objective of performing life sciences experiments in space is to identify the role that microgravity plays at all levels of an organism, from the effects on a single cell up to complex organisms including humans.

Biolab is integrated into a single International Standard Payload Rack and is divided physically and functionally into the automated section (core unit) and the manual section. The main parts of the experimental activities are performed automatically in the automated unit, following a manual loading of the samples by a crew member. The core unit, features a large incubator equipped with two centrifuges, providing controlled levels of acceleration, including microgravity. The methodology of containing the biological samples in standard containers has been tried and proven successful in the numerous flights of the previous Biorack facility. The atmospheric environment in the containers can be adjusted during centrifuge operations, to best suit the need of the biological samples. On top of the incubator, the handling mechanism allows automatic operations on the samples, e.g., freezing and analysis of the samples. The main purpose of the handling mechanisms is to reduce significantly the time the crew needs to interact with the experiment and to allow full telescience capabilities with the scientists on the ground.

In the manual section, as the name suggests, the crew performs all the required experiment preparation and completion activities. As part of the manual section, the BioGlovebox provides a clean, controlled and enclosed environment for manual operations and, as it is under negative pressure, prevents any possible contamination of the Columbus Laboratory. The BioGlovebox is also equipped with an ozone generator, to perform sterilization of the working volume.

15.27 SPACE MEDICINE EXPERIMENTS

Ever since manned spaceflight began, scientists have been collecting data on the effects of the space environment on the crew. Some effects, such as fluid shift and

the effects of processing in the vestibular system, quickly became known and were characterized collectively as ''space sickness'' because of the visible effects. Other effects. such as the loss of bone calcium, became better understood over time, not least because of the weakened state experienced by crew members immediately after return to the ground, particularly after long-duration spaceflights. Some exotic effects of the space environment, such as the light flashes seen due to the impact of radiation particles on the retina, remained hidden at first because of crew reluctance to report them.

15.28 EXPOSED PAYLOADS

Exposed payloads support experiments to study aspects of the space environment. This encompasses both the study of the characteristics of the environment itself and the study of the effects of that environment (the near-vacuum, the thermal cycling regime, HZE particle impacts, micrometeroid impacts, the radiation environment during solar maxima, atomic oxygen corrosion) on space construction materials, biological matter, and so forth. Since virtually all exposed payloads to date have been located in Earth's orbit, a significant part of the experimentation carried out has also focussed on the characterization of the artificial environment due to space debris, for example, the debris load, the statistical distribution of particle sizes, and the effects of debris impact on materials.

Typically, exposed payloads have been located on free-flying satellites. Some have been specifically developed for this purpose (LDEF, EURECA), while on others the exposed element piggybacks a more general mission (e.g., the Biopan part of the Foton recovery capsule). Invariably, exposed payloads have to be retrieved. Although some useful information can be conveyed through remote imaging, the key is the return of the materials and collectors so that (in the case of space debris, for example) the damage to materials can be assessed and the debris particles, or at least the residue from their impact, can be collected. Hence, both LDEF and EURECA were launched and later retrieved by the NSTS Shuttle Orbiter. Biopan, in contrast is a clamshell arrangement which opens to the environment during flight and is closed before return on a recovery capsule.

15.29 SOUNDING ROCKET PAYLOADS

Sounding rockets are used for both upper atmosphere physics and microgravity research. Five to twenty minutes of microgravity can be delivered via a sounding rocket—effectively the free-fall period at the apex of the vehicle's flight parabola. This constraint dictates that effective fast telemetry is needed to acquire data from the short experiment period.

PART 7
HUMAN FACTORS

Nicholas Colford

15.30 SUBSTANTIVE OVERVIEW

Introduction

At the earliest point in the specification of a spacecraft, the presence of astronauts should be part of the design process. Once it is known what functions the mission is to perform, it should be decided which functions to allocate to astronauts and which to machines. The functions should be analyzed into activities and tasks and the necessary interactions between humans and machines be specified. The design of the relationship between the astronauts and the spacecraft is based on a detailed understanding of the characteristics of the human being. The following is just an outline of those characteristics and the reasons for their importance in designing the spacecraft. Following the description of the human being are some guidelines for designing a spacecraft in which the human being can perform safely and efficiently.

Anthropometry and Biomechanics

Anthropometry is the study of the variations in the size of body parts from one person to another and across the population. Biomechanics (at least the part that relates to human factors engineering) is the study of how the human body moves. Both disciplines are important to human factors engineering in space and on the ground.

Good human factors design accommodates the greatest possible percentage of the population. In order to do this, it is necessary to design for the extremes of the population range, not for the average. For example, switches must be located close enough to the operator so that the smallest person can reach them but enough headroom must be provided to allow the largest operator to fit in the workspace. For space applications the common approach is to design to fit 90% of the population, leaving the largest 5% and the smallest 5% of the population excluded. NASA's STD 3000 has extensive tables of anthropometry and biomechanical data.

In weightlessness, reaching upwards is easier than on the ground and reaching downwards is harder. Apart from that, the range of movements that the body is capable of does not change much. The amount of force that can be exerted decreases towards the limit of the range of movement. People tend to become taller by an average of 3% when they are in space and adopt a resting posture known as the neutral body posture. Increases in circumference of the upper parts of the body occur, as fluids do not drain downwards due to gravity, and this can cause discomfort and make clothes feel tight

Physiological Effects of Space Flight

Space flight has a gradual negative effect on many aspects of human physiology. Cardiac size decreases, muscles lose tone and reduce in size, and calcium is lost

from bones. Immediate changes take place, such as transient loss of red blood cells and the shift of fluids towards the head. These and other factors contribute to changes in human performance when in space.

Human Performance in Space

Human performance varies widely between individuals, just like body size and range of movements. In the same way as for anthropometry, it is essential to design to include the largest possible percentage of the range of performance rather than to take the average of the optimum level of performance as the baseline.

Vision

The intensity of sunlight in space, unfiltered by the atmosphere, is much greater than that which we experience every day on earth: around 25% more intense than a bright sunlit day on the ground. When orbiting the Earth, a spacecraft usually spends a little less than half of its time in the shadow of the planet. In the low Earth orbit that most manned spacecraft occupy, the orbit lasts around an hour and a half. (Only the Apollo Moon missions have so far taken astronauts beyond such an orbit.) This means that every 45 minutes the level of illumination will change from full sunlight to almost total darkness or vice versa.

The absence of atmospheric scattering of light means that shadows are more sharply defined and darker in space than on the ground. It also means that objects viewed at a great distance do not appear hazy, as they usually do on the Earth. These two effects can, either on their own or when combined, result in reduced capability to perceive the distance, size, shape, and motion of objects outside the spacecraft. The intensity of light and contrast are only issues outside the spacecraft or where there are windows. Displays in space vehicles with windows and on space suits have to compete with the harsh environment. Inside modules of a space station, where the outside view can be shut out, lighting is as in any closed space.

Although there is no definitive detailed information, many aspects of visual perception appear to be generally slightly disturbed to some degree by being in space. Reports have documented changes in perception of colors, contrast sensitivity, and reduced visual acuity at short distances. The reduced visual performance is often compounded by low lighting levels within modules fitted with few lighting units in order to conserve launch weight and reduce power consumption.

Auditory System

There is no ‘‘natural space environment’’ as far as hearing is concerned. Human hearing takes place within the artificially created environment of the spacecraft. Sound levels within spacecraft have to be managed to avoid pain, annoyance, degradation of performance and to provide an environment suitable for human occupation.

The immediate pain threshold for sound is around 135 dB, but the effect of sound on human well being and performance is gradual and occurs at levels well below that.

70 dBA: disruption of cognitive activity, especially short-term memory

50 dB: annoyance

75 dB: degraded astronaut performance

56 dBA: maximum limit for working environment

34–47 dBA: range acceptable in sleeping accommodation

In the context of acoustics and human factors engineering, noise is unwanted sound. Functionally, audible noise has to be kept low enough to maintain crew health, well being, and performance. Long-term (chronic) exposure to noise should be kept below 56 dBA for working and 47 dBA for sleeping, otherwise crew performance and well being will suffer. Short-term exposure to higher levels of noise must be managed and mitigated as much as possible.

Audible noise is much more easily tolerated when it lasts only a short time and, just as importantly, when the person hearing the noise knows that it will last only a short time. The noisiest times of space flight, launch and reentry, fall into this category. For limited periods such as these it is possible to accept higher-than-ideal noise levels in the cabin so long as measures are taken to ensure that crew health is not compromised and voice communication between crew members and with ground control is maintained intelligible.

The ability to detect a sound against background noise depends on many factors, such as how similar the sounds are and what the person is expecting to hear. However, there is a useful rule of thumb to obtain optimal performance. Measure the intensity at which the sound can just be heard above the noise and then use an intensity of sound halfway between that minimal level and 110 dB.

Olfaction and Taste

The capacity to smell and taste decreases in weightlessness. This is almost certainly due to the increased levels of congestion that occur in the nasal and sinus cavities. The mucus that forms in these cavities normally drains out due to gravity. In weightlessness it accumulates, causing congestion. Food tastes bland and smells seem weak.

Vestibular System

A human being perceives gravity primarily through acceleration sensors in the inner ear. The otolith apparatus senses linear accelerations, including gravity, while the semicircular canals sense rotation of the head. Weightlessness affects the otolith organs more than the semicircular canals, and the unusual combination of sensations can cause disorientation and nausea. Movement of the head increases the problem by adding to the semicircular canal stimulus. This can manifest itself as vertigo, a sensation of spinning or tumbling. It can also lead to extended periods of nausea and vomiting, known as space adaptation syndrome, or space sickness.

What the astronaut sees becomes very important in deciding where ‘‘up’’ and ‘‘down’’ are felt to be. Anything that upsets that idea (even seeing the Earth unexpectedly through a window) can create a sometimes violent sense of nausea. Seeing displays or other features inside the cabin in an unexpected orientation can also have the same effect.

These problems seem to affect about 50% of astronauts and generally last up to two or four days. The medication Scop-dex, made up of scopolamine and dexamphetamine, is often effective in reducing the symptoms.

Kinaesthesia / Proprioception

Secondary perception of weight and gravity is spread throughout the body. Proprioceptors are many different sensors that tell the person about the flexion of the limbs, the force being exerted by a muscle group, the pressure inside and around various organs, and so on. Forces and pressures around the body are all dependent on weight, and therefore in weightlessness the information from these sensors arrives in combinations of values never experienced on the ground.

In weightlessness, the ability to discriminate between masses is reduced by about 50% and absolute judgment of mass is less accurate.

Motor Skills

Human motor control relies a great deal on automatic motor skills that are tuned over years of practice in conditions of 1 g. They utilize much proprioceptive feedback which is disrupted in weightlessness, as explained above. As a result, some loss in motor skill capability occurs during the first few days of exposure to weightlessness. Fine motor control of limbs is affected more than gross movement, although manual dexterity is not significantly affected.

These motor control problems tend to diminish as the flight goes on and the astronaut adapts to the new environment. They reappear when the astronaut returns to Earth, but readaptation is rapid. Although readaptation to 1 g is rapid, presumably through remembering of previously learned rules, adaptation to other levels of gravity cannot be assumed to be the same. Mars or lunar gravity, for example, represents a new unknown gravity state and will require another period of adaptation.

Response Time

Reaction times increase in space. This is almost certainly due to the reduction in motor skill levels, requiring more conscious thought to perform any action.

Physical Strength

The amount of force that an astronaut can generate gradually decreases over time when in weightlessness, as muscle tone and muscle mass are lost. In addition, the ability to exert the force generated depends on the availability of suitably located handholds, footholds, or bracing points. The traction caused by body weight is lost, making it hard to move objects if there is no crew restraint point. However, when there is a crew restraint point the same lack of weight makes it easy to move very massive objects, although such objects must be moved slowly.

Physical workload should be designed very conservatively based on the fifth percentile capabilities and accounting for at least 10% reduction in strength or endurance after the fist few weeks of flight.

Skills Retention and Transfer

Whenever a person has been trained to do something, his or her proficiency begins to decrease from the moment he or she stops practicing it. This means that if a certain level of performance is required several weeks after the end of training, then the activity must be overtrained. That is to say, an astronaut must be more proficient than necessary to do a job at the end of training.

This issue is compounded by the fact that there is no simulator that completely recreates the conditions of being in space. When the skills are applied in an environment different than to the one in which they were learned, they are less applicable and less efficient. It is usual to allow anywhere between 25% and 100% more time to perform a task in space than on the ground, especially at the start of the mission.

15.31 DESIGN GUIDELINES

The above is an overview of the knowledge of human capabilities and requirements that has to be used in the design of a spacecraft. By following the sort of process outlined in the introduction, the design of the spacecraft should develop along with rules and guidelines as follows.

Design for Physical Factors

It is fair to say that design for physical feasibility of tasks in space is a mature science. It is not yet at the point where one can buy ''off-the-shelf'' solutions, but the correct application of the human factors engineering process to a job can almost be guaranteed to produce a workplace that enables the crew to get the job done. Whether that workplace will be comfortable or suit the preferences of all the people who end up using it is far less certain. This is because, in terms of user populations, very few users become involved in the design process and the empirical data that have been collected so far on the physical ergonomics or working in space are limited.

For physical activity, plan for between 50% and 200% more time to do a job in weightlessness. Delicate manipulative tasks require more concentration than on the ground because of altered mechanics of the hands and arms. The very basic ''in the bones'' knowledge of how to use tools has to be suppressed when working in weightlessness. Manual handling of large objects is much easier in weightlessness than on ground but still has to be performed slowly.

Design for Psychological Factors

This is especially important when designing a spacecraft for an eventual interplanetary mission or any other mission where the crew will spend months on board at a time. Perhaps the most important thing that the spacecraft will have to offer is a clear division between social space and the individual crew members' private space. Both are equally important in maintaining the individuals' well being and the cohesiveness of the group. At least one social space must allow the comfortable

accommodation of the whole crew, otherwise subgroups will start to form. Private space must belong to one person—hot bunking is not a good long-term solution. The communications system must allow for considerable amounts of private personal traffic. If this can be routed to the private spaces, so much the better.

Physical ergonomics for spacecraft has reached the point where it can confidently support the design of usable workplaces and tools. However, the process of optimizing the physical ergonomics of working in space is very much still in its infancy. Whether it ever matures will depend on how the utilization of space develops.

Cognitive ergonomics and psychological support to crews are two areas where terrestrial expertise and experience will be of great use in supporting eventual interplanetary missions.

Space tourism has been a subject of hypothetical discussion for many decades and, albeit to very limited extent, has become a reality in the ISS. The human factors of space tourism should be considered insofar as it parallels other forms of extreme tourism. There are two stages of space tourism, and these will be defined by the amount of effort that has to be made on the part of the designers to accommodate these tourists. The first phase, already begun, is that of the adventurer tourist who is willing to put up with all the rules, regulations, and training of a professional astronaut in order to experience space flight. The second stage will be in full swing only when the space tourist is able to step aboard a launch vehicle with no more preparation or training than is currently given in the preflight safety demonstration of airline cabin crew. The difference between these two situations is above all one of human factors. Certainly, the reliability and cost of travel into orbit have to be made acceptable to potential space tourists, but space tourism can only appeal to more than ''wannabe'' astronauts when the human factors of space flight have been improved.

If and when they ever take place, long-term lunar missions and interplanetary manned flights will challenge many fundamental aspects of human factors engineering. Perhaps the most challenging topic is the idea of ensuring crew well being and performance during a Mars mission. Quite apart from the length of the flight phases of such a mission, there is the question of how the crew could perform once on the Red Planet. Stepping down onto the moon (one-sixth Earth gravity) after a few days of weightlessness may well have been ''one small step for a man,'' but doing anything useful on the surface of Mars (one-third Earth gravity) after a couple of years in weightlessness is likely to be another matter altogether. Would a potential crew readjust to one-third Earth gravity more easily than astronauts returning to earth? Current plans for such a mission rely heavily on the crew's capability to refuel the ship for its return to earth. Despite sounding like a contradiction in terms, it is important not to forget human factors for satellites and probes. The ground control station is a man-machine system and requires all the same care and attention that should be lavished on the human factors of any process control facility. Fortunately, there exists a wide experience of literature on the ergonomics and human factors of control rooms.

Environment

The environment inside a spacecraft must respect the human capabilities described above. For example, the ambient sound levels must be kept low enough to prevent short-term damage and discomfort and long-term health problems. Lighting levels,

similarly, need to be high enough to allow astronauts to do the jobs required of them in each particular part of the spacecraft, but it must be possible to achieve lower levels of lighting for rest areas and darkness for sleep. Two other areas of environmental design are worthy of note.

Vibration. If the amplitude of any vibration is sufficiently large, vibration can affect human performance directly by moving the body or body parts. It also interfere with the actions that the person is trying to perform making the person physically miss the target, for example. In space flight the strongest vibrations occur during launch and reentry—the times when the crew member is most tightly held in contact with the structures of the spacecraft.

At lower amplitudes, vibrations are particularly felt where they coincide with the resonant frequency of an organ or body segment. Vibration can affect visual perception, especially at frequencies that make the head or the eyeballs resonate: 2–30 Hz and 40–60 Hz, respectively. Vibration of displays and controls also makes interaction difficult. In general, magnitudes of less than 0.02 g are considered comfortable.

Thermal. In order for the crew to be comfortable and work efficiently, the temperature inside a spacecraft should be maintained between 19 and 27°C. The lack of natural convection in weightlessness reduces the capability to expel heat when compared to a similar situation on the ground. This means that lower temperatures are generally tolerated more easily than higher temperatures. On the other hand, many life science experiment protocols require crew members to be partially clothed for extended periods of time and so a higher cabin temperature is called for. All in all, it is desirable to have the cabin temperature not only kept within a limited range but also to be able to raise and lower the temperature depending on the crew activity taking place.

Humidity and air flow speeds also affect the perception of temperature and the ability of humans to control their body temperature. Humidity affects the evaporation of sweat, which is an important heat-rejection method. Air flow, too, facilitates sweat evaporation. Relative air humidity should be kept within a range of 20–70%. Air should be kept moving through the cabin at speeds between 0.1 and 0.2 m/s.

Architecture

The visual environment of the spacecraft must be pleasing to the eye, striking a balance between boring and oversaturated. For long-duration missions, especially, the impact of colors can be critical. Generally light and nonsaturated colors should be used over large areas, with darker or more saturated colors being used only for details. Using more than five colors within a single area gives an impression of clutter.

A sense of up and down should be created consistently throughout the spacecraft to reduce disorientation and aid navigation within. This is best produced by using lighter colors for "ceilings" and darker colors for "floors." Similarly, lighting should come from "above" rather than "below." The sense or orientation should not be contradicted by the orientation of workstations, which should all correspond to the overall convention of up and down. This makes lighting easier to implement as well as reduces the likelihood of disorienting the crew.

Illumination levels required vary greatly according to the activity taking place. Reading requires at least 538 lux, while a corridor used only for transfer can be as dark as 50 lux. Lighting must be adjustable down to complete darkness for sleeping areas.

Lighting should approximate the full visible spectrum of terrestrial sunlight. General lighting should be diffuse and avoid creating glare and reflections, particularly around workstations. All lighting should be adjustable in intensity by the crew.

Workstation Location. Workstations need to be located such that astronauts can move between them as procedures require. Workstations that are used frequently, one after the other, should be located close to each other, if possible adjacently. The same is true of workstations that are used concurrently for closely related activities, to enable communication between the two operators. However, this closeness has to be traded off against the potential for one workstation to interfere with the work of the other. Light, vibrations, and (audible) noise from one may be problems for the other. Each workstation should be large enough to accommodate the whole of the activity required of the astronaut working there. Workstations also need to be located away from rest and sleeping areas.

Traffic Flow. The movement of astronauts from one place to another within the spacecraft must be designed for. First, the traffic flow must be optimized so that the crew does not spend large amounts of time moving from one workstation to another.

Astronauts need restraints at workstations, and mobility aids long traffic routes. Along a traffic route, handles and rails should be oriented along the direction of travel in order to maintain speed and fine-tune direction. At termination points they should be at right angles to the direction of arrival or departure to act as push-off and landing surfaces. Corners and junctions along routes should be supplied with handles and rails as if they were termination points.

Thought should be given to any equipment that is to be carried along traffic routes. There should be enough room to see around the package being carried. Packages require handles and have to be manageable with one hand only, as the other hand must always be free for movement or stability control.

Typical speeds for movement within a spacecraft range from 0.4 ms to 0.6 ms^{-1} for ordinary point-to-point translation to 0.15 to 0.4 ms^{-1} when carrying large of massive equipment. In play or gymnastics, astronauts can travel at 1.8 ms^{-1}.

Crew Restraints. For simple, short-duration tasks a simple handrail is often the most comfortable and convenient way of holding oneself steady in one place. If the job requires the use of both hands, then it becomes necessary hold oneself in place using the legs. A number of restraints nave been developed, ranging from simple fabric loops attached to the floor (foot loops) to harnesses and seat substitutes into which an astronaut can close or wedge himself or herself to obtain a stable restraint. Generally, the more complicated restraint systems are only considered worthwhile if the task being performed requires a great deal of stability or is of long duration, e.g., more than 5 or 10 minutes.

Restraints are necessary where the astronaut has to stay in one place and use both hands and where a stable head position is important. In these cases the restraint system must be able to accommodate the full range of size of crew members and enable them all to reach controls and see the necessary displays.

Workstation Design. Workstations must be designed around the neutral body posture. Displays should be located as close as possible to the zero-g line of sight, which is 15° lower than the normal terrestrial line of sight. The most important or frequently used displays should be closest to the sightline. Controls should be located below the line of sight and grouped as far as possible with the displays most closely associated with them. Controls and displays should both be compatible with the information transmitted through them. NASA STD 3000 lists extensively the sorts of controls and displays suitable for use in spacecraft, along with dimensions, colors, and suitable surface finishes.

Weightlessness poses some particular problems for control design. All controls have to be protected against accidental actuation by a passing astronaut or free-floating objects; In order to reinforce the perception of how far they are being moved, continuous controls, such as rotary knobs and wheels, should have detents in their motion. It is very important that the workstation have a consistent visual indication of up and down and that this orientation be compatible with the surrounding environment.

15.32 DISCUSSION OF SELECTED HUMAN FACTORS PRINCIPLES

Human factors engineering and ergonomics are essentially the same thing. The study of the relationship between the human being and his or her work and working environment. Advocates of the name *human factors engineering* (HFE) often say that ergonomics is the physical and mechanical part of HFE. Advocates of *ergonomics* sometimes say the same of HFE. Whichever of the two names it travels under, however, the discipline has the same principles. Perhaps the significant difference between the two terms is in their etymology. *Ergonomics* derives from the Greek *ergo,* meaning work, and *nomos,* meaning study. It is thus the study of work, implying a primary concern with what people are doing: their tasks and activities. Human factors engineering has undeniable connotations of taking into account human capabilities and limitations when designing and constructing things. Both schools, in fact, are concerned with both parts of the issue. On the whole, Europeans who do this work tend to think of themselves as ergonomists and their American colleagues tend to call themselves human factors engineers. In the space sector, *HFE* is much more common.

One of the most important contributions that human factors engineering can make to the design of spacecraft is to remind engineers and designers that they are building for a population of users, not for themselves. ‘‘This would be al right for me’’ is not enough to validate a design.

The human factors engineering of spacecraft is dominated by weightlessness. But it is a mistake to place too much emphasis on what is, after all, one of many factors affecting the work of astronauts during space flight.

One of the peculiarities of the human factors of spacecraft is that to a large extent and in many cases the equipment that astronauts use and the jobs they do are still made to measure. That is to say, a piece of equipment may be specified with a certain function in mind but by the time it is designed in detail, tested, and built it is already known who will be using it. This is less the case in terms of space infrastructure, such the Shuttle or modules of the Space Station. But it is

often the case for scientific experiment equipment, which is often produced and used mission by mission.

This subsection does not attempt to describe the human component of a spacecraft in sufficient detail to stand alone as a design guide. What it aims to do is provide an overview of how human beings tolerate the conditions of and can perform in the context of a space flight. More specifically:

- To provide a general understanding of the human factors that are important in spacecraft design and operation
- To indicate where to look for the more detailed and specific information
- To indicate the importance of combining the information about the human being and the objectives and operation

Unmanned spacecraft can do a lot and are very well suited for a large number (even the most forthright proponent of manned spaceflight might say the majority) of spacecraft missions. One might argue that the American and Russian emphasis on manned spaceflight is driven by a desire to convince the world that the most important thing to do is the thing that they can still do better than everyone else.

It may be of use to consider the human element in manned space flight in the same way as other subsystems with inputs versus outputs, and up to a certain point such a reductionist approach works. In order to apply this approach effectively, however, it is important to remember that the performance of the human being depends not only on what is input and output but on how it is input and output. It is also important to remember that the human being has motivations. One cannot separate what the human is doing from why he is doing it.

As component or subsystems to integrate into a spacecraft, human beings are very complex systems. Their physical input and output interfaces are many and varied. Their tolerance to environmental conditions is quite limited. Their information processing channels are on the one hand very limited but at the same time are capable of very subtle and complex communication. And to complicate matters further, human beings are autonomous information processors requiring a rich social environment in which to operate. Otherwise they simply do not work well at all.

Nutrition and sanitation are human factors that are covered specifically in Part 8. In addition to these, one must add sleep. Human beings should sleep 8 hours per 24; diurnal rhythm needs cues of light and dark or it runs slow.

15.33 HEALTH CARE

Human beings perceive their surroundings through the five external human senses of sight, hearing, smell, touch, and taste. These external senses work in tandem with the internal system of proprioception for feeling movement, posture, and orientation. Each sense has capabilities and limitations in terms of the range of stimuli that it can process and the amount and type of information that can acquired through it. In addition to the absolute limitations of the senses, which vary considerably from person to person, there are the interactions between the senses to consider. The human brain has the capability to process information from the different senses at the same time but can only handle a certain amount, especially when there is

conflict about what the information means. This subsection outlines the basic capabilities of the senses insofar as they relate to spacecraft design.

Visual Environment. The intensity of sunlight in space, unfiltered by the atmosphere, is much greater than that which we experience every day on Earth: around 25% more intense than a bright sunlit day on the ground. When orbiting the Earth, a spacecraft usually spends a little less than half of its time in the shadow of the planet. In the low earth-orbit that most manned spacecraft occupy, the orbit lasts around an hour and a half. (Only the Apollo moon missions have so far taken astronauts beyond such an orbit.) This means that every 45 minutes the level of illumination will change from full sunlight to almost total darkness or vice versa.

The absence of atmospheric scattering of light means that shadows are more sharply defined and darker in space than on the ground. It also means that objects viewed at a great distance do not appear hazy, as they usually do on the Earth. These two effects can, either on their own or when combined, result in reduced capability to perceive the distance, size, shape and motion of objects outside the spacecraft. The intensity of light and contrast are only issues outside the spacecraft or where there are windows. Displays in space vehicles with windows and on space suits have to compete with the harsh environment. Inside modules of a space station, where the outside view can be shut out, lighting is as in any closed space.

Although there is no definitive detailed information, many aspects of visual perception appear to be generally slightly disturbed to some degree by being in space. Reports have documented changes in perception of colors, contrast sensitivity, and reduced visual acuity at short distances.

Environmental Requirements (Physical)

Human beings can survive quite a wide range of natural terrestrial environmental conditions. Their optimal conditions for well being, comfort, and safe, efficient work, however, form a more limited range. Conditions inside a spacecraft are largely man-made, and control over most aspects is technically feasible. There are generally two ranges applicable for any environmental condition. The wider range is that in which a human being can survive; the narrower range is that in which a human being is comfortable and can work well.

Optimizing the human factors of a spacecraft consists of matching what the human being needs from an environment with what the spacecraft provides. The spacecraft must protect the human from the extremes of vacuum, heat, cold, and radiation that characterize the natural space environment. On top of that, the spacecraft must not expose the human to excessive levels of noise, vibration, and acceleration which result from the spacecraft's activities and mission. And with increasingly longer missions, spacecraft must ensure that the human being is in a comfortable social and psychological environment. Weightlessness, the factor that is most obvious in space flight, is an issue but should not be given excessive importance—certainly not at the expense of the other factors mentioned above.

Atmosphere

Human beings need an atmosphere in which to live. It prevents their bodily fluids from boiling, provides them with the oxygen they must consume, carries away the CO_2 they must expel, keeps them warm, and prevents them from drying up. The

preceding list reflects the order in which a total lack of atmosphere would kill someone were it not for the preceding items. The point is that all the atmospheric parameters need to be maintained: there is no point, for example, in maintaining a perfect overall air pressure if CO_2 is allowed to build up to dangerous levels. This subsection outlines the optimal and limit levels of the major factors conditioning atmospheric habitability.

The vast majority of human beings in the world are used to an atmospheric pressure of 14.7 psi, with the air composed of 78% nitrogen (N_2), 21% oxygen (O_2), 0.03% carbon dioxide (CO_2), and the remainder made up of many other gases in even smaller amounts. Although current practice in manned spacecraft is to reproduce this type of atmosphere, it is important to know about the acceptable levels of deviation from this normal situation.

Taking absolute overall pressure as the first consideration, atmospheric pressure must be maintained above 0.9 psi, otherwise bodily fluids will evaporate at bodily temperature. Since asphyxia will occur very quickly at pressures well above this value even with 100% oxygen, this value is important only for short periods of time. This is the pressure at which explosive decompression occurs.

In order to obtain the oxygen needed from the atmosphere, human beings need a pressure of 1.9 psi of oxygen. This amount, sometimes called the partial pressure of oxygen, or PPO_2, is amply supplied by the 20% of O_2 present in "normal" terrestrial air at 14.7 psi—3.1 psi PPO_2. An atmosphere of 100% oxygen at 3.1 psi would provide the same amount of O_2 to a human being. Performance suffers already with PPO_2 levels between 3.1 and 1.9 psi; night vision is lost at 2.7 psi, hallucinations, excitation, or apathy arises around 2.2 psi, and at 1.9 psi physical performance is impaired, loss of memory occurs, and paralysis sets in. Below 1.9 psi, the effects can become irreversible, with unconsciousness at 1.6 psi and death within 3 minutes at less than 1 psi. On the other hand, breathing an atmosphere with too much oxygen leads to chest pain and coughing and damages blood cells. The higher the PPO_2, the faster the onset of the symptoms: it can be 5 or 10 days for slightly raised levels (3.8 psi), while effects are observed within two days or less when breathing levels around 7.5 psi.

Human beings produce CO_2 through respiration, consuming O_2 at the same time. Since CO_2 is toxic, it must be expelled through the lungs and carried away by the atmosphere. If the amount of CO_2 in the atmosphere is higher than normal, then the ability to expel this gas is reduced, leading to the symptoms at the levels indicated in Table 15.23.

TABLE 15.23 Symptoms of Excess CO_2 in Atmosphere

Amount	Exposure	Symptoms
0.03%	Normal	None (0.0058 psi)
up to 3%	Chronic	Increased motor activity, euphoria, mental acuity for first day; headache, mental depression, loss of memory on second day; some recovery from third day onwards (0.58 psi)
up to 7%	Acute	Fatigue, impaired concentration, sweating face, headache, lowered body temperature (around 1°C), reduced aerobic capacity (around 12%)
10% plus		Acute vomiting, chills, hallucinations, loss of consciousness, convulsions, shock, death

Returning to breathing normal levels of CO_2 enables recovery from all but the final effect, and breathing pure oxygen will accelerate recovery. Nevertheless, withdrawal symptoms can be experienced during recovery, and these can be at least as severe as the symptoms of the poisoning itself.

Some other aspects of atmosphere design, although not strictly (or at least not directly) human factors, are worthy of note. Pure O_2 atmospheres, while allowing lower overall pressure to be used, greatly increase the risk of fire. The terrestrial N_2-O_2 mix means that life science and medical experiments are conducted in normal atmospheric conditions.

Although accelerations above normal gravity are experienced for only short periods of time in spacecraft, the importance of their effects is magnified because those periods tend to be the launch and reentry phases of spaceflight, where time-critical human performance and accuracy of action are most important.

In terms of human factors, the direction in which the acceleration is applied to the person is very important. This can be seen from the other point of view: it is important how the person is oriented with respect to the acceleration. Doctors and ergonomists like to use the anatomical reference system, made up of planes and orientation, and engineers like to use their x, y, and z axes. When x, y, and z axes are applied to human beings, the usual arrangement is the following: the frontal (or coronal) plane corresponds to the yz plane; the sagittal plane corresponds to the xy plane; and the transverse plane corresponds to the xy plane. Linear accelerations are described either in terms of the actual acceleration of the body a or the sensation of increased ''gravity'' that the person experiences as a result of the increased inertial reaction force from whatever the human is standing, sitting, or lying down on. For example, when a person experiences acceleration upwards, it is described as $-a_z$ acceleration and causes the sensation of $+G_z$ increased ''weight''; as an intuitive description, ''eyeballs down'' describes how the movement is sensed. Table 15.24 lists the different ways that linear and rotational accelerations are commonly described.

Acceleration affects human beings' performance directly by increasing the effective ''weight'' of the body and body parts in the direction of the G-force. This makes it hard to reach controls and manipulate them with accuracy and reliability. It also causes discomfort as the body is pressed against the seat, couch, or straps that restrain it against the acceleration. In space flight, high levels of acceleration are usually accompanied by vibrations, thus compounding the problem.

Indirectly it quickly affects the body by making blood run through the body in the G direction. The effect is most strongly felt in the case of G_z acceleration. In upward acceleration ($+G_z$) the blood tends to collect in the legs and lower body depriving the brain of oxygen. When seated, a person will suffer from tunnel vision after a few seconds of 3 to 4 G_z and become unconscious after about 5 seconds of levels around 5 to 6 G_z. During downward acceleration ($-G_z$), the problem is too much blood pressure in the brain. At levels of -2 to -3 G_z the feeling of facial congestion and headache is intense, and after about 5 seconds vision begins to fade to gray or red. Most people cannot tolerate more than 5 seconds at -5 G_z.

Forward acceleration (G_x), when seated with a headrest, is the most tolerable acceleration because there is less body tissue ''below'' the brain in the direction of the acceleration and so blood and oxygen supply are less compromised. Although discomfort and tightness of breathing are experienced at levels of 3 to 6 G_x, most people can tolerate up to 10 G_x while experiencing tunnel vision similar to lower levels of G_z accelerations. From 10 to 15 G_x, the symptoms increase but can be tolerated. Backward acceleration ($-G_x$) is less well-tolerated than forward accel-

TABLE 15.24 Linear and Rotational Accelerations

	Direction of acceleration		Inertial resultant of body acceleration	
Linear motion	Acting force	Acceleration description	Reaction force	Verticular acceleration
Forward	$+a_x$	Forward	$+G_x$	Eyeballs in
Backward	$-a_x$	Backward	$-G_x$	Eyeballs out
Upward	$-a_z$	Headward	$+G_z$	Eyeballs down
Downward	$+a_z$	Footward	$-G_z$	Eyeballs up
To right	$+a_y$	Right lateral	$+G_y$	Eyeballs left
To left	$-a_y$	Left lateral	$-G_y$	Eyeballs right
Angular motion				
Roll right	$+p$		$-R_x$	Cartwheel
Roll left	$-p$		$+R_x$	
Pitch up	$+q$		$-R_y$	Somersault
Pitch down	$-q$		$+R_y$	
Yaw right	$+r$		$+R_z$	Pirouette
Yaw left	$-r$		$-R_z$	

Note: Capital letter G expresses the inertial resultant in multiples of terrestrial gravitational acceleration, i.e., 9.807 m/s^2.

eration due to pain and instability arising from being pressed against restraint straps rather than the backrest of a seat or couch. If the head tilts forward, as is inevitable at levels around -8 G_x or above, it is subject to the sort of blood pressure as in G_z acceleration, leading to hemorrhaging.

Lateral accelerations ($\pm G_y$) are rarely experienced due to the flight dynamics of aircraft and spacecraft. The problems are largely due to discomfort from restraint straps, increased blood pressure in the head when the head tilts laterally, and engorgement of the dependent elbow causing pain. Around 15 seconds of $\pm .5$ G_y results in hemorrhaging and extreme headaches.

In addition to the magnitude and direction of the force, other factors, such as how gradual or abrupt is the onset of the acceleration, can affect how well tolerated it is. The particular condition of the individual also affects tolerance. The so-called G-suit (a pressurized pair of trousers) is useful in countering the pooling of blood in the lower body and legs and is therefore particularly useful in increasing tolerance of $+G_z$ acceleration. While age, physical condition, motivation, and diet have their roles to play, training can make a significant difference in one's ability to tolerate high G-forces. Of particular importance for spacecraft design is the fact that exposure to weightlessness results in rapid behavioral and gradual physiological adaptation. It must not be assumed that an astronaut, after even a few days of weightlessness, has the same tolerance to acceleration as at launch or at the end of the training course.

Just as linear accelerations are more or less tolerated depending on their direction, the effect of rotational accelerations depends on not only the direction of the axis of rotation but also the location of the axis with respect to the body. If the rotational axis is close to the center of the body, rotation rates in any axis of up to about 6 rpm are found to cause no problems for most people, and this tolerance can be raised to between 12 and 30 rpm with a gradual training program. Higher

rotation rates are found gradually to bring on similar symptoms to linear accelerations and numbness of the extremities with particular incidence of disorientation.

If rotation is to be used to create artificial gravity, the NASA human factors guidelines provide the following advice: Radial traffic should be kept to a minimum. Crew members should not traverse through the spin axis unless the center is nonrotating. Living and working areas should be located as far as possible from axis of rotation. Compartments should be oriented such that the primary traffic paths are parallel to the spin axis and workstation positions should be oriented such that during normal activity the lateral axis between the crew member's ears is parallel to the spin axis. At the same time the controls and displays should be organized such that right/left head rotations and up/down arm motions are minimized.

Impact acceleration limits are largely based on the impact that will fracture bones. After the magnitude, the direction of the acceleration vector and its duration have the greatest effect on the survivability of the impact. Impact survivability depends also on how well supported the body is during the impact, and this is part of the reason why impacts at right angles to the spinal column are better tolerated than impacts along the line of the spine. Table 15.25 shows NASA's recommendations for the limits of impact accelerations.

Acoustics is the study of that range of vibrations which is audible to the human ear. The range of frequency of human hearing is from around 20 to 20,000 Hz, although there are individual differences. Sound intensity, also called sound pressure level (SPL), can be measured in terms of the pressure of the vibrations, but since the ear perceives sound with logarithmic sensitivity, it is convenient to measure with a logarithmic unit, the decibel (dB).

$$SPL = 20\log(P_1/P_0)\ \text{dB}$$

where P_1 is the pressure of the sound being measured and P_0 is the pressure of a reference sound (20 $\mu N/m^2$). The value of the reference sound corresponds to the threshold of hearing for a healthy adult under ideal conditions. Sound pressure level is proportional to the square root of the power of the sound source. As a rule of thumb, a quadrupling of the sound source power corresponds to doubling in sound pressure level results in an increase of 6 dB in measured sound intensity.

As a human factor, sound is perceived in terms of loudness. It is an important channel for information and as such should be kept clear for this purpose in a manned spacecraft. In addition, it is an environmental factor that can have short- and long-term effects on the performance, capabilities, well being, and health of the crew.

The sensitivity of the human ear varies by over 40 dB according to frequency of the sound and is most sensitive in the range 3,000 to 5,000 Hz. In order that sound intensity measurements reflect this, the American National Standards Institute (ANSI) came up with weighting curves called A, B, and C. C is a flat weighting,

TABLE 15.25 Impacts (duration of force < one second)

Direction of impact acceleration	Impact limit	Rate of impact
$\pm G_x$	20 G	1,000 G/s
$\pm G_y$	20 G	1,000 G/s
$\pm G_z$	15 G	500 G/s
45° off (any) axis	20 G	1,000 G/s

B is a halfway house, and A is the curve closest to the human response. While a measurement of sound in dBC will give a mechanically ''pure'' value of energy, sound measured in dBA gives a more realistic impression of how loud the sound feels.

The immediate pain threshold for sound is around 135 dB, but the effect of sound on human well being and performance is gradual and occurs at levels well below that.

70 dBA: disruption of cognitive activity especially short term memory

50 dB: annoyance

75 dB: degraded astronaut performance

56 dBA: maximum limit for working environment

34–47 dBA: range acceptable in sleeping accommodation

In the context of acoustics and human factors engineering, noise is unwanted sound. Functionally, audible noise has to be kept low enough to maintain crew health, well being and performance. Long-term (chronic) exposure to noise should be kept below 56 dBA for working and 47 dBA for sleeping otherwise crew performance and wellbeing will suffer. Short-term exposure to higher levels of noise must be managed and mitigated as much as possible.

Audible noise is much more easily tolerated when it lasts only a short time and, just as importantly, when the person hearing the noise knows that it will last only a short time. The noisiest times of space flight, launch and reentry, fall into this category. For limited periods such as these it is possible to accept higher-than-ideal noise levels in the cabin so long as measures are taken to ensure that crew health is not compromised and voice communication between crewmembers and with ground control is maintained intelligible. It is important to note that noise perception and the amount of annoyance that it causes are quite subjective.

Sound and vibration are, of course, the same physical phenomenon. As factors impinging on the human being, they are considered differently because of how they are perceived. Vibrations with frequency below the threshold of hearing are ''felt'' rather than heard.

If the amplitude is sufficiently large, vibration can affect human performance directly by moving the body or body parts or interfering with the actions that the person is trying to perform. At lower amplitudes, vibration can affect visual perception, especially at frequencies which make the head or the eyeballs resonate: 2–30 Hz and 40–60 Hz, respectively. While sound is perceived almost exclusively through the air via the ears, vibrations are perceived through the whole body. They are transmitted to the body either through the air or through structures and crew restraints. In space flight, the strongest vibrations occur during launch and reentry—the times when the crewmember is most tightly held in contact with the structures of the spacecraft. They are particularly felt where the vibration coincides with the resonant frequency of an organ or body segment. Care must be taken to isolate the crew members from the vibrations of launch and reentry.

Radiation is one of the big medical issues for human space flight. Radiation can harm human beings, just as it can harm any living organism, causing almost immediate (acute) damage with high doses. Long-term exposure to lower doses can lead to gradual but no less serious chronic damage.

Radiation is categorized into two major types in its effect on humans. Ionizing radiation is made up of fast-moving atomic particles and very high-energy electromagnetic radiation, characterized by its ability to break atomic bonds within bio-

logical material. Nonionizing radiation, on the other hand, does not damage biological material in this way but in large amounts can harm people.

The particles making up ionizing radiation consist of electrons (also called beta rays), protons (or hydrogen nuclei), neutrons, helium atoms (alpha particles), and heavier atomic nuclei called HZE particles. X-ray and gamma ray electromagnetic radiation are grouped together with these particles because of their similar effect on organisms. As well as the different types of radiation, three categories of radiation are considered based on where it comes from. Galactic cosmic radiation (GCR) originates in stars, galaxies, and deep space. Solar particle event (SPE) radiation is emitted by the sun. Trapped radiation is held in the Van Allen belts by the Earth's magnetic field.

The amount of radiation that a spacecraft will encounter depends very much on where it goes and, within the solar system, what the Sun is doing while it is out there. The Sun's activity, visible as sunspots and solar flares, increases and decreases over an 11-year cycle. Increased solar activity generates a stronger interplanetary magnetic field. The activity level of the sun affects all three categories of radiation. Table 15.26 defines the radiation types.

Trapped Radiation

Radioactive particles, from whatever source, can become trapped in the Earth's magnetic field holding the highest densities of electrons, for example, at bay around 4,000 km from the Earth. The field is, however, tilted with respect to the spin axis and offset from the center of the planet, resulting in an area over the South Atlantic ocean where the trapped proton belts reach down as far as the atmosphere. As long as a low Earth orbit spacecraft is not in a polar orbit, it will pick up nearly all its radiation dose as a result of the occasions when it passes through the South Atlantic anomaly (SAA). A spacecraft in a low-inclination orbit (e.g., 28°) will avoid passing through the ''hottest'' part of the SAA, where the particle flux is higher. A spacecraft in a higher-inclination orbit (e.g., 58°) will pass though the ''hottest'' part. Paradoxically, in the higher-inclination orbit the craft spends less time in the higher flux with a result of a lower net dose. Increased solar activity also makes the Earth's atmosphere expand and absorb low-level protons, reducing further the component of radiation dose due to trapped radiation at LEO altitudes.

Galactic cosmic radiation (GCR) is a relatively low fluence of radiation arriving from all directions and is characterized by particularly high-energy particles. The increased strength of the interplanetary magnetic field at solar maximum means that some particles are kept out of the inner solar system, although the effect on the highest-energy particles is quite small. In free space, the flux of GCR during solar minimum is around 2.5 times greater than that at solar maximum. The earth's geomagnetic field also shields low-altitude orbits from GCR, but the protection is very dependent on inclination: for example GCR dose rates are two or three times higher at a geomagnetic latitude of 58° than at one of 28°.

While the solar maximum can bring about some reductions in the amounts of GCR and trapped radiation that orbital and interplanetary spacecraft might expect to encounter, it is not all good news. With the solar maximum comes a greater frequency of solar flares, provoking emissions of charged particles from the Sun particles (SPE radiation). Most flares are not large enough to produce worryingly large quantities of particles, and the Earth is not always in line with position of the flare on the Sun's surface for the particles to come our way. But once or twice in

TABLE 15.26 Sources and Characteristics of Electromagnetic and Particulate Ionizing Radiation in Space

Name	Nature of radiation	Charge	Mass	Sources
X-ray	Electromagnetic	0	0	Primary: solar corona, stars, galaxies, terrestrial atmosphere in auroral zone Secondary: Spacecraft structure in some parts of radiation belts and auroral zones after solar flares
Gamma ray	Electromagnetic	0	0	Stars, galaxies, unknown sporadic sources, and spacecraft atmosphere
Electron	Particle	−e	1 m_e	Radiation belts and auroral regions
Proton	Particle	+e	1840 m_e or 1 amu	Galactic cosmic rays, radiation belts, and solar flares
Neutron	Particle	0	1841 m_e	Primary: galactic cosmic rays atmospheric albedo neutrons Secondary: galactic cosmic ray interaction with spacecraft structure
Alpha particle (helium nucleus)	Particle	+2e	4 amu	Galactic and solar
HZE particle (heavy primary nucleus)	Particle	+3e	6 amu	Galactic and solar

each cycle the combination of size and location of a flare or several flares creates a very large energetic SPE. Solar flares occur without warning, and it is difficult to determine the magnitude and intensity of the event until the particles start to arrive.

Many parameters of ionizing radiation can be measured and described, but those most pertinent to human wellbeing are the dose, the quality factor, and the dose equivalent. In rather rough and ready terminology, these correspond to the very reasonable concerns of how much radiation there is, how nasty it is, and how much harm it will therefore do.

The dose, D, is the directly measurable amount of radiation in the environment. It is a measure of how much energy is absorbed by a certain amount of biological tissue or equivalent. The most commonly used unit in space applications is the rad, representing 100 erg/gram. (The equivalent SI unit is the Gray, Gy, which corresponds to 100 rad.)

The quality factor, Q, is a way of weighting each different form of radiation in order to take account of how much biological damage it does relative to the energy it transfers to the tissue. Q is itself derived from another parameter of the radiation, the linear energy transfer (LET) but the relationship between Q and LET is beyond the scope of this brief treatment of the subject.

The dose equivalent, DE, is simply the dose multiplied by the quality factor. It is expressed in units called rem (roentgen equivalent man). The corresponding SI unit is called the sievert (Sv): 1 Sv equals 1 Gy multiplied by Q, and therefore 1 Sv equals 100 rem.

Table 15.27 shows the values of Q for various types of cosmic radiation. It is reproduced for two reasons. The first reason is to show that there can be considerable difference between the physical, measurable property of the radiation and the effect it has on a person. The other is to take the opportunity to state that as more is learned about the relationships between radiation and living systems, these Q values are expected to be revised.

So much for the measurement and calculation of effective doses. What does this radiation do to people?

Astronauts are considered workers exposed to radiation as apart of their work. NASA has set limits for the amount of radiation to which its astronauts can be exposed as part of their work, and the time the astronauts spend in space is the major contributor to their dose.

Exposure interval	Depth (5 cm)	Eye (0.3 cm)	Skin (0.01 cm)
30 days	25 rem	100 rem	150 rem
Annual	30	200	300
Career	100–400	400	600

	Age			
Sex	25	35	45	55
Male	150 rem	250 rem	325 rem	400 rem
Female	100	175	250	300

TABLE 15.27 Quality Factor

Type of radiation	Quality factor, *Q*
X-rays	1
Gamma rays and bremsstrahlung	1
Beta particles, 1.0 MeV	1
Beta particles, 0.1 MeV	1
Neutrons, thermal energy	2.8
Neutrons, 0.0001 MeV	2.2
Neutrons, 0.005 MeV	2.4
Neutrons, 0.02 MeV	5
Neutrons, 0.5 MeV	10.2
Neutrons, 1 MeV	10.5
Neutrons, 10 MeV	6.4
Protons, greater than 100 MeV	1–2
Protons, 1 MeV	8.5
Protons, 0.1 MeV	10
Alpha particles, 5 MeV	15
Alpha particles, 1 MeV	20

The career depth equivalent dose limit is based upon a 3% lifetime excess risk of cancer mortality. The total equivalent dose yielding this risk depends on sex and age at the start of exposure. The career equivalent dose limit is approximately equivalent to:

200 + 7.5(age – 30) rem for males up to 400 rem maximum
200 + 7.5(age – 38) rem for females up to 400 rem maximum

All of the above serves to demonstrate that the radiation threat to astronauts should not be underestimated or taken for granted. When engineering a spacecraft, it is important to take into account the full range of its operational environments, the time when it will be in use, and the length of time that people will spend in it.

The amount of shielding that a spacecraft must provide to protect its crew depends upon the intensity and type of radiation environments in which it will operate and the amount of time that the crew will spend inside it in each environment. As the NASA guidelines show, the important parameter is how much radiation the person is exposed to, not how much shielding the spacecraft provides. This means that to decide whether a spacecraft's capability to protect the crew from radiation is sufficient, it is essential to analyze the situation from a point of view of the crew predicted dose, taking into account the trajectory of the spacecraft, the solar activity at the time of the mission, and the amount of time a person spends in the spacecraft. A space suit, for example, provides much less protection than does a module of the ISS. On the other hand, an astronaut will spend hours in the space suit as opposed to weeks or months in the module. The adequacy of the protection levels provided by these two different coverings has to be analyzed in the context of the mission, not just the hardware itself.

The vast range of temperatures that can be encountered in space goes for beyond the limits of human survival. Without protection, human beings can survive for short periods of time in temperatures a few degrees below 0°C and a few degrees above 40°C. Protective clothing can make temperatures as low as −30 or −40°C tolerable, but this is not a tenable situation for more than a few hours. A healthy human being maintains an internal temperature of 37°C and, as a result of metabolic activity, is constantly generating heat. As a factor in calculating the thermal budget of a manned space system, the heat output of a crew member can be estimated from around 2,600 kcal/man-day with a low physical workload up to a value of 3,200 for a day of heavy physical work.

In order for the crew to be comfortable and work efficiently, the temperature inside a spacecraft must be maintained between 19 and 27°C. The lack of natural convection in weightlessness means that heat rejection from the body is restricted when compared to a similar situation on the ground. This means that lower temperatures are generally tolerated more easily than higher temperatures. On the other hand, many life science experiment protocols require crew members to be partially clothed for extended periods of time and so a higher cabin temperature is called for. All in all, it is desirable not only to have the cabin temperature kept within a limited range but also to be able to raise and lower the temperature depending on the crew activity that is going on.

Humidity and air flow speeds also affect the perception of temperature and the ability of humans to control their body temperature. Humidity affects the evaporation of sweat, which is an important heat-rejection method. Air flow, too, facilitates sweat evaporation. Relative air humidity should be kept within a range of 20–70%. Air should be kept moving through the cabin at speeds between 0.1 and 0.2

m/s. Although not an thermal consideration, it should be noted, while talking about air flow, that it is important to maintain an air flow across the head of a sleeping crew member to prevent the accumulation of a ''cloud'' of air rich in CO_2 around the head.

Another thermal human factor is touch temperature. Just as the air temperature must be kept within limits for the astronaut's well being, so contact with surfaces at high or low temperatures must be avoided. Fortunately, weightlessness reduces the need to be constantly in contact with spacecraft structures. Current design guidelines require that any surfaces that could be at temperatures above 49°C or below 0°C be insulated or shielded from human contact. The only slightly more reduced range of 3°C to 45°C is the limit for surfaces where long-duration contact is permissible.

It can be argued that the use of the word ''weightlessness'' is misleading, in purely mechanical and gravitational terms, to describe conditions inside a spacecraft orbiting around the earth. It is the weight of the spacecraft (i.e., the attraction between it and the earth) that keeps it from flying away in a straight line. The spacecraft and everything in it, including astronauts, are falling freely together, giving, locally at least, the sensation of zero gravity. No one is being pulled down onto the floor, because the floor is being pulled down at the same rate as everyone. In practical terms, an astronaut in an orbiting spacecraft can consider himself or herself weightless. One talks of microgravity and free fall when describing the gravitational conditions in a spacecraft, but weightlessness is a particularly good term to use when talking about human factors engineering because it describes what the person feels.

For ergonomics and human factors engineering, weightlessness is important in two ways: the effect it has directly on the astronaut and a reduction in frictional forces around and about the astronaut.

A human being perceives gravity primarily through acceleration sensors in the inner ear. The otolith apparatus senses linear accelerations, including gravity, while the semicircular canals sense rotation of the head. Weightlessness affects the otolith organs more than the semicircular canals, and the unusual combination of sensations can cause disorientation and nausea. Movement of the head increases the problem by adding to the semicircular canal stimulus. What the astronaut sees becomes very important in deciding where ''up'' and ''down'' are felt to be. Anything that upsets that idea (even seeing the Earth unexpectedly through a window) can create a sometimes violent sense of nausea. Seeing displays arranged any old way can be such a stimulus.

Secondary perception of weight and gravity is spread throughout the body. Proprioceptors include different sensors that tell the person about the flexion of the limbs, the force being exerted by a muscle group, the pressure inside and around various organs, and so on. Forces and pressures around the body are all dependent on weight, and therefore in weightlessness the information from these sensors arrives in combinations of values never experienced on the ground. This contributes to the sensations of disorientation and nausea and also means that people are decidedly less capable of distinguishing between different masses in weightlessness than on the ground.

When relaxing all voluntary muscles in weightlessness, the human body tends to assume a characteristic position. This posture has been called the ''neutral body posture.'' Each person has his or her own neutral body posture, and although not many accurate studies have been published. It requires effort to maintain any posture different from this one. It requires a comparable force to maintain one's arms

lower than their neutral position as it does to keep them above it. And it is certainly more tiring since the muscles we normally use to stop our arms from falling down are much more frequently exercised than those we use to hold them down. The overall stature of astronauts has been found to increase by an average of around 3% in space.

The reduction in frictional forces is a simple result of the effective lack of gravity. With no weight pushing an astronaut onto the floor, for example, there is very little friction between the two. This makes it complicated to move around and hard to stay still. Type on a keyboard, for example, in weightlessness, and unless you are restrained in some way, even the tiny force you exert on the keys will gradually send you spinning away from the keyboard. This means that every human activity has to take into account the need for restraint. This can take the simple form of a handle, but it must then be remembered that only one hand is available to perform the activity. Suitable structures around which the legs can be wrapped or into which the legs can be jammed are also a popular solution.

The changes in perceptual and motor capabilities of the human being take place in the first week or so of exposure to weightlessness. As the astronaut gets used to the new situation, much of this capability is recovered. The changes in physical performance are a different matter and come on gradually over weeks and months of a mission. Long-term exposure to weightlessness results in a gradual loss of muscle tone and physical strength and a host of other medical conditions that, while not immediately affecting human performance, must be treated with medical countermeasures. The discipline of space medicine keeps an eye on these factors, implementing its solutions through the operational area of biomedical operations.

Human beings are able to survive in weightlessness at least for as long as the longest mission so far performed. It is by no means certain that they could do so indefinitely. Many health problems and physiological changes found to occur in astronauts during space flight have been attributed to the lack of gravity, and many more problems arise when an astronaut experiences gravity again upon return to earth.

Weightlessness is the most obvious feature that conditions the design and operation of manned spacecraft, but it is not the only major characteristic of a spacecraft. Other characteristics condition human well being, especially through psychological and social factors. Confinement in a closed volume, isolation from society, and the radiation environment are also characteristics of spacecraft. The fact that they are characteristics that spacecraft have in common with many terrestrial workplaces does not mean that they can or should be ignored. It is easy to think only about weightlessness when thinking about spacecraft, so it will be considered last.

There is another feature that, while it is not unique to spacecraft, has a profound effect on how human factors are considered in the engineering process: it is how few people ever actually use them. When developing a product for the mass market, an automobile, for example, you must consider that anyone could try to use it with any level of training, capability, and experience. When designing your spacecraft, you can rest assured, at least for the moment, that anyone who even gets near it will have been trained for years and be under orders to follow strict instructions and procedures.

The psychological factors affecting the people taking part in space flight are many and work at many different levels. This subsection looks at just the ones most directly linked to the people flying in the spacecraft and the engineering of the spacecraft and the systems that support it. The issues are greater than can be

covered in this part of a general handbook. It is hoped that the reader, as an engineer involved with a spacecraft, will find here enough information to get an overall idea of the issues and where they should be addressed in the process of developing a space mission.

The psychological factors can be directly engineered, so to speak, through selection and training of the crew. But there are limits to human performance, and people tend behave in certain ways, whatever engineers and mission managers may prefer. A manned spacecraft and its support systems have to be designed and built to accommodate humans' psychological characteristics just as much as their physical ones.

Psychological adaptation to space flight goes through two stages:

Stage 1 corresponds to the major physiological changes (the first two to six weeks). The new physical environment and lifestyle can combine to contribute to reduced performance and a reduced state of subjective well being.

Stage 2 covers the rest of the flight.

The monotony of work and limited variety that the spacecraft environment provides can lead to asthenia: symptoms include emotional hypersensitivity, irritability, decreased motivation, and reduced ability to sleep well despite feeling tired. Studies have shown that the most critical time for such problems is the third quarter of a mission, regardless of the actual duration—as the end of the mission approaches there is a feeling of change in anticipation of the return to Earth.

Visual tracking capability is degraded in space. Performance recovers a few weeks after launch but then gets worse again after the halfway point in a (short-term) mission. It is thought that in stage 1 this is due to not having yet learned how to control one's hands well, while in stage 2 the cause is thought to be reduced alertness contributed to by sleep disturbances, which are common. The ability to perform multiple tasks simultaneously appears to suffer in a similar way, with secondary tasks typically being sacrificed to ensure good performance of primary tasks.

This is a problem because many of the tasks that are attributed to humans are complex tracking tasks—piloting a vehicle, operating a robotic arm, operating tools, and manual maintenance tasks. It is a strong reason for designing equipment that can accommodate some imprecision and designing crew activity so that critical tasks are not superimposed on the same person at the same time. It also means that to ensure the acceptable performance of these tasks, the crew have to be overtrained prior to the flight—that is to say, they have to achieve a higher members level of ability than actually necessary in order to compensate for the fact that their performance once in flight may suffer.

The correspondence between decreased mental performance and reduced subjective well being has important implications for the design of crew operations in spacecraft. Whatever the crew members are capable of being trained for, it has to be remembered that their capability will be reduced in space. The mental workload to which they are subjected during the mission should not peak as high as their maximum capability on-ground. This is especially true of the times found to be critical—just after launch and in the third quarter of the mission. The risk of exacerbating the problem of boredom with too little stimulus has to be balanced against the risk of overloading the crew with too much work all at once.

Human beings require, on average, around 8 hours of sleep every 24 hours. The terrestrial day–night cycle conditions the human body to that frequency, with the

onset of bright illumination resetting the 24-hour cycle. The 24-hour cycle is optimal for long-term health, but it is not the only one which the body can be induced to follow. In the absence of a dark–light cycle, the human body will tend to settle on a 25- or even 26-hour cycle. Cycles of longer or shorter periods can be adhered to with increasingly negative effects as one departs from the 24-hour optimum.

What this means for spacecraft design is that a 24-hour day–night cycle is an essential part of any long-duration mission and is highly recommended even for short missions. It must be supported by patterns of work and rest and the ability to control the illumination levels inside the spacecraft. Noise should be kept below 47 dBA in sleeping quarters, as already mentioned in the subsection on acoustics.

The social structure of a spacecraft crew becomes more important as the number of crew members and the duration of the mission increase. The confinement of the spacecraft means that the individual crew members cannot escape from each other, and it is essential that they function together socially in a positive way. It is all too easy to say ''of course'' to the issues described here (after all, most of us have experience of overcoming difficult social situations) and dismiss them.

In order to work effectively, it has been found that spacecraft crews need the same three things that any team needs if it is to put into isolation and confinement to work: a complementary combination of individual characteristics of the personnel; training and creation of a coherent social group; and support from ground during the mission that at least does not aggravate the psychological stress of their isolated, confined situation and at best helps to alleviate it.

What individual characteristics make a good astronaut? The answer appears to differ to some extent for long-duration and short duration missions. Of course, crews should be made up of individuals who are highly motivated, intelligent, agreeable, and emotionally stable and have good social skills. But there is more to making a team than having capable individuals. Many individual characteristics are positive in themselves but can be incompatible with equally positive characteristics in other people. Good team building can be undermined by incompatibilities between crew members. A recent ESA report into psychological factors covers these issues in the light of increasing long-term missions.

It has been found that teams need to be made up of people who have complementary needs. The most critical of these needs is, perhaps unsurprisingly, dominance. Most conflict in teams stems from too many team members having too strong a need for dominance. That is not to say that a team has to have one leader and the others must be followers—there is a need for distributed leadership and redundancy of roles. It is an obvious problem, but the solution has to be carefully managed.

At the same time, there are other needs that crew members have to have in common; otherwise they will tend to work toward differing goals. These are the needs that can be satisfied by working together, such as need for autonomy, need for affiliation, and need for achievement.

It would be unwise to have the crew for a mission to Mars composed of four pilots and no doctor. This is an oversimplification of the complexities of a real situation, but achieving the right level of overlapping and complementary skills in the crew members is an important part of the engineering, so to speak, of the crew.

The knowledge that the crew members bring with them should be as complementary as possible. This means that they can learn things from each other and rely on each other to fill in gaps in their individual knowledge.

The crew must have some level of common values and expectations. It would not be desirable for a crew to be completely homogeneous, as this would result in

serious understimulus. But the crew must at the very least know what their fellow crew members find to be socially acceptable behavior. This issue will grow in importance: as missions become longer, there is more opportunity for value incompatibilities to become problems. And longer, more expensive missions are more likely to be international ventures, bringing together more heterogeneous crews. It is all too easy to fall into cultural stereotypes, such as Japanese being group-oriented and Americans being individualistic. But there are deep cultural differences which, precisely because they are deep, are easily overlooked. Superficial similarities in operational methods do not a homogeneous group make.

A group of astronauts who appear to make up a compatible and well-balanced crew having been selected, who the team still has to be built. This does not happen overnight. It has long been recognized in astronaut and cosmonaut training circles that the social group that makes up the crew must be stable before the flight. This means that the relationships between the individuals have to have had enough time to stabilize. One way of describing the process of stabilizing the relationships in a group is based on four observable phases: forming, storming, norming, performing. In the forming phase, the individuals have a lot of contact with each other and try to establish their social position in the group, but due to the newness of the group the interaction is quite formal and work-related. In the storming phase, the individuals lose their reserve and openly enter into conflict with each other; the social hierarchy of the group develops and leaders establish themselves, both formally and informally. The storming phase can only be considered finished when a social structure has emerged that all the crew accepts. With the structure set up, the group enters the norming phase and establishes its way of getting things done and working together. The performing phase occurs when the group has abandoned the ways of doing things that do not work very well. A spacecraft crew should not take off unless it is well into the performing phase of this progression.

A small but important caveat should be introduced amid the importance of forming a cohesive group. There is always the risk of the group becoming too cohesive and closed in on itself. This is especially true of groups in isolated situations. The group can become too certain of its own ability to do the right thing, and group members can become inhibited from going against the group consensus and they tend to see those outside the group as remote and irrelevant. All these tendencies can reduce the performance of a spacecraft crew and even put them at risk of ignoring important information either from inside or outside the group. Group cohesiveness is not, therefore, a factor to be maximized; rather, it has to be optimized.

Leadership of the crew is most important. A spacecraft crew is in a situation very similar to crews of small ships or mountaineering or polar exploration expeditions. The isolation and hostility of the environment are the major factors conditioning their social situation. Much is known about the social dynamics of teams in these situations, and the qualities required of a leader of such a group are fairly well known and agreed upon. The leader should take decisions together with the crew during normal circumstances but be able to decide quickly alone in a crisis; the leader should keep the crew focused on their objectives and encourage optimism; should be sensitive to the needs of the individual crew members and work towards ensuring that each member's short-term and long-term needs are rewarded by the activities allocated to them; and the leader should be able to spot the formation of subgroups and manage them to either eliminate them or at least avoid the formation of destructive rivalries. To some extent, training can enhance these leadership qualities just as much as any other social skills, but on the whole a crew leader has to be selected rather than trained. That said, the job of the leader and

the rest of the crew can be made a lot easier by good communications with the mission support on ground and the rest of the world.

Having good-quality communications with the ground can make a big difference to crew morale and reduce the feeling of isolation. It is not just a question of how much time per day communication is possible and the quality of the transmission that matters. Operational communications have to be managed not only in terms of the technical quality of the connection but also for the human quality of the communications. It is important for the team formation prior to the mission to integrate mission control personnel and flight personnel into one team, not two separate teams. The increasing frequency of routine missions lasting months rather than days or weeks means that personal communication with family and friends becomes very important. The length of the mission is a factor that works together with the perceived routines of the mission. Together they mean that crew members need a lot of freedom to communicate outside an operational context. Isolation may be easily tolerated when one is ''boldly going where no man has gone before,'' but even the most motivated crew will need the best support possible for routine missions lasting months.

Social Interaction

The social interaction between the crew and the mission support personnel is a particularly critical aspect for mission performance. It should go without saying that, as far as possible, the flight crew and mission control crew should form a team, and effort must be exerted before the mission to ensure that this takes place. Communications between flight crews and mission control can (and in the past have) become strained and lose efficiency, with serious implications for mission performance.

It is quite easy for empathy between ground personnel and flight crew to be lost during the mission. The most common cause is ground crew making requests or giving orders that disrupt the work of the flight crew. The way to minimize this is to keep open a dialogue allowing flight crew to tell ground what is and is not realistic. And ground has to take notice of this input. Ground crew have to remember that however well-formed a team was created with flight crew before the flight, their common experiences ended at liftoff and they are talking to people greatly altered since they last saw them face to face.

It is perfectly normal for small, isolated, and confined groups of people to become very self-centered and tend to apportion blame outside the group. They require more praise and affirmation for their actions as the duration of their isolation increases. It is important that ground personnel dealing with them understand this and accommodate this need. It is not that the flight crew are becoming selfish and grumpy people; they are simply responding to their environment. The less isolated they can be made to feel, the more chance that they will have to suffer less from this response.

It is all very well to try to avoid tension between flight crew and ground personnel. But when it does occur, the onus must be on the ground personnel to take the burden of the social tension. The astronaut in flight, after all, cannot slam the door and walk out of the office at the end of the shift.

Isolation is not only how far you are physically from the rest of society but also how far you are functionally. An astronaut on board the International Space Station may be only a few hundred kilometers from his home town when the orbit passes

overhead, but she is several days away from Earth by shuttle or hours away from an emergency landing who knows where.

Modern communications technology can go some way to mitigate the feeling and the effects of physical isolation. It is important to remember that isolation from different people is felt in different ways. Isolation from colleagues and friends is different than isolation from family and loved ones.

PART 8

NUTRITION AND SANITATION

Nick Larter

Any spacecraft that is intended to contain living matter, from the simplest experimental biological material to a human crew, requires a life support system. The complexity of the system will depend on what has to be supported and the duration and destination of the mission.

There are broadly two kinds of life support system. Environmental control and life support systems (ECLSS) are open or partly closed loop and based generally on regenerative physico-chemical processes. Biological life support systems (BLSS) are similar but, as the name implies, based on biological processes. A special subcategory of BLSS is the closed or controlled ecological life support system (CELSS), which, again as the name implies, is closed loop.

All currently operational life support systems are of the ECLSS kind, although some promising subsystems based on biological processes now exist which could ultimately be integrated into an otherwise physico-chemical ECLSS. One such example is a biological air filter. BLSS and CELSS are still experimental concepts, but the successful realization of both will be important for long-duration manned planetary missions and the establishment of planetary bases, not least because the subsystems to support crew nutrition and sanitation can be integrated into them to different degrees. In all current manned spacecraft these subsystems are largely separate from the ECLSS and from each other.

This Part will focus on the nutrition and sanitation subsystems of spacecraft, although atmosphere conditioning will also be considered in detail and other components of an ECLSS will be dealt with briefly to provide the necessary contextual information. Little historical detail will be covered, and the technical solutions presented will mainly be limited to those in use within currently operational manned spacecraft—the Russian Soyuz vehicle, the American NSTS Shuttle Orbiter, and the International Space Station (ISS). BLSS and CELSS issues will be briefly considered in the final paragraphs of this Part. The functional requirements for life support and habitability are shown in Figure 15.5

In the NSTS Shuttle Orbiter, the ECLSS maintains the orbiter's thermal stability and provides a habitable environment for the crew and avionics. The ECLSS also manages the storage and disposal of water and crew waste. ECLSS is functionally divided into four systems:

1. Pressure-control system, which maintains the crew compartment at 14.7 psia with a breathable mixture of oxygen and nitrogen. Nitrogen is also used to pressurize the supply and wastewater tanks.
2. Atmospheric revitalization system, which uses air circulation and water coolant loops to remove heat, control humidity, and clean and purify cabin air.
3. Active thermal control system, which consists of two freon loops that collect waste heat from orbiter systems and transfer the heat overboard.
4. Supply and wastewater system. The supply water system stores water produced by the fuel cells for drinking, personal hygiene, and orbiter cooling. The wastewater system stores crew liquid waste and wastewater from the humidity separator. The system also has the capability to dump supply and wastewater overboard.

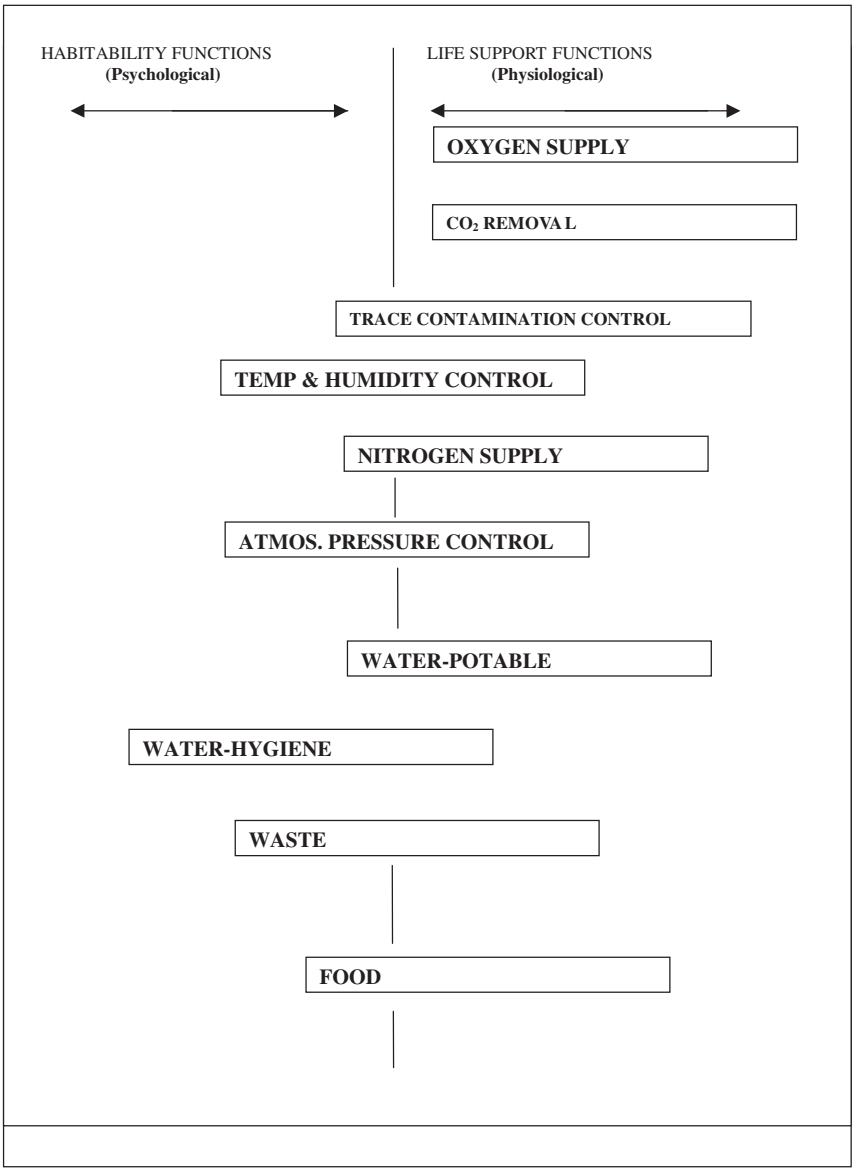


FIGURE 15.5 Functional requirements for life support and habitability.

The following subsections use the Shuttle Orbiter as an example of an atmosphere conditioning revitalization system. The NSTS Shuttle Orbiter air pressure is the same as Earth's at sea level: 1,033 g/c^2 (14.7 psf). Its air is made up of an 80% nitrogen and 20% oxygen mixture. The orbiter's environmental control system circulates air through filters to remove carbon dioxide and other impurities. Excess moisture is also removed, keeping humidity at comfortable levels. Temperature in the orbiter can be regulated between 16 and 32°C (61 and 90°F). The orbiter crew requires only ordinary clothing. People can move about, work, and relax unencumbered by bulky space suits.

The atmospheric revitalization system circulates air and water throughout the cabin to control ambient heat, relative humidity (between 30 and 65%), carbon dioxide, and carbon monoxide levels. It also provides cooling for cabin avionics.

Cabin air is circulated around the cabin. The heated air is then ducted, via cabin fans, to the cabin heat exchanger, where it is cooled by the water coolant loops. The water coolant loop system collects heat from the cabin heat exchanger, the inertial measurement unit heat exchanger, some of the cold-plated electronic units in the avionics bays, and the avionics bay heat exchangers. It transfers heat to the freon/water heat exchanger of the active thermal control system. The active thermal control system expels the heat overboard. Details are given below.

Except for ducting, all air loop components are located under the middeck floor. Based on the crew cabin volume of 2,300 ft^3 and a flow of 330 ft^3 of air per minute, one volume crew cabin air change occurs about every 7 minutes. The heated cabin air is drawn through the cabin loop and through a 300 μm filter by one of two cabin fans. Normally only one fan is used. Each fan is powered by a three-phase, 115 V ac motor. These 495 W motors produce a nominal flow rate of 1,400 lb/hr through the cabin air ducting. A check valve located at the outlet of each fan prevents air from backflowing through the nonoperating fan.

Carbon dioxide removal is accomplished by passing cabin air through one of two identical solid amine resin beds. The resin consists of a polyethylenimine sorbent coating on a porous polymeric substrate. Upon exposure to carbon dioxide-laden cabin air, the resin combines with water vapor in the air to form a hydrated amine which reacts with carbon dioxide to form a weak bicarbonate bond. Water is required for the process since dry amine cannot react with the carbon dioxide directly. While one bed adsorbs carbon dioxide, the other bed regenerates with thermal treatment and vacuum venting. The adsorption/regeneration process runs continuously with the beds automatically alternating processes every 13 minutes. An activated charcoal canister in the other CO_2 absorber slot removes odors. It is changed out midmission on flights over 10 days.

The scrubbed air is then forced through the rest of the system. Cabin air is then directed to the crew cabin heat exchanger, where heat is transferred to the water coolant loop. One of two humidity separators draws air and water from the slurper. In the humidity separator, centrifugal separation of the water from the air is used. The fan separator removes up to approximately 4 lb of water per hour. The water is routed to the wastewater tank, and the air is ducted through the exhaust for return to the cabin. The relative humidity in the crew cabin is maintained typically between 30 and 65% in this manner.

A small portion of the revitalized and conditioned air from the cabin heat exchanger is ducted to the carbon monoxide removal unit, which converts carbon monoxide to carbon dioxide. A bypass duct carries warm cabin air around the cabin heat exchanger and mixes it with the revitalized and conditioned air to control the crew cabin air temperature in a range between 65 and 80°F.

As above, the Shuttle is used as the primary example, in this case to describe potable ware conditioning and recycling. The supply water system provides water for flash evaporator system cooling, crew consumption, and hygiene. The supply water system stores water generated by the fuel cells, and the wastewater system stores waste from the crew cabin humidity separator and from the flight crew. Four supply water tanks and one wastewater tank are located beneath the crew compartment middeck floor.

The supply water system consists of four water tanks that are pressurized with nitrogen from the pressure control system. Each of the four potable water tanks has a usable capacity of 168 lb, is 35.5 in. in length and 15.5 in. in diameter, and weighs 39.5 lb dry. The three fuel cells generate a maximum of 25 lb of potable water per hour (about 0.81 lb of water per kwh). The product water from all three fuel cells flows to a single water relief control panel. The water can be directed to potable water tank or to the fuel cell water relief nozzle.

Temperature sensors are installed on each of the redundant paths; in addition, a pressure sensor is transmitted to telemetry and can be monitored. A pH sensor is located at the common product water outlet of the water relief panel. It provides a redundant measurement of fuel cell health and water purity. A single measurement of water purity in each fuel cell is also provided. If a single fuel cell pH sensor indicated high, the flight crew would be required to sample the potable water to verify the proper pH.

The hydrogen-enriched water from the fuel cells flows through the single water relief panel through two hydrogen separators to potable water tank. The separator removes 85% of the excess hydrogen. The hydrogen separators consist of a matrix of silver palladium tubes, which have an affinity for hydrogen. The hydrogen is dumped overboard through a vacuum vent.

The water entering the potable water tank, which is sterilized before launch, passes through a microbial filter that adds approximately one-half parts per million iodine to the water to prevent microbial growth.

As above, the Shuttle is used as the primary example, in this case to describe food preparation in space. The middeck of the orbiter is equipped with facilities for food stowage, preparation, and dining. The food supply is categorized as either menu food, pantry food, or fresh food. Meals are individually tailored, based on crew member preference. Menu food consists of three daily meals per crew member and provides an average energy intake of approximately 2,700 calories per crew member per day. The pantry food is a two-day contingency food supply that also contains food for snacks and beverages between meals and for individual menu changes. Pantry food provides an average energy intake of 2,100 calories per crew member per day. Fresh food items consist of perishable items such as fruits and tortillas. The types of food include fresh, thermostabilized, rehydratable, irradiated, intermediate-moisture, and natural-form food and beverages.

Three one-hour meal periods are scheduled for each day of the mission. This hour includes actual eating time and the time required to clean up. Breakfast, lunch, and dinner are scheduled as close to the usual hours as possible. Dinner is scheduled at least two to three hours before crew members begin preparations for their sleep period.

The galley provides a centralized location for one person to handle all food preparation activities for a meal. The galley has facilities for heating, rehydrating, and stowing food.

The length of Shuttle missions has steadily increased from the first mission in 1981 of 2 days, to 14 days for STS50 in June 1992. Missions beyond 10 days are

called extended duration orbiter (EDO) missions. In order to accommodate the weight and volume of trash generated by the food system on these longer missions, it was necessary to develop new food and beverage packages. A trash compactor was also developed to reduce the volume of the trash, and the new packages were designed to be compatible with the compactor.

A typical beverage package is made from a foil laminate. A septum adapter is sealed in the package after the beverage powder has been added. The septum adapter holds a septum, which interfaces with the galley water dispenser for the addition of water, and with a straw for drinking the beverage. Although the beverage package was designed for use on extended missions, it has replaced the square polyethylene beverage package on all Shuttle missions.

The rehydratable food package also is made from flexible material to aid in trash compression. The rehydratable package consists of a flexible bowl and lid with the septum adapter for adding water from the galley. Velcro on the bottom of the package holds it in the meal tray. After the required amount of water is added to the package, it is placed in the oven if the food is to be served hot, or directly onto the serving tray if it is to be served cold. The top of the package is cut off with a knife or scissors and the contents eaten with a fork or spoon. The EDO rehydratable food package was tested on STS-44 and used for all of the rehydratable foods on STS49 and 50. It has now permanently replaced the rigid square rehydratable package.

Food is packaged and stowed in the locker trays in Houston about a month before each launch. Stowed food lockers and shipping containers are kept under refrigeration. About three weeks before launch, the food lockers are shipped to Kennedy Space Center (KSC) in Florida. There they are refrigerated until they are installed in the Shuttle two to three days before launch. Besides the meal and pantry food lockers, a fresh food locker is packed at KSC and installed on the Shuttle before launch.

Again the shuttle is used as the primary example, this time to describe human waste disposal and cleaning. The waste management system is used primarily to collect and process crew biological wastes (Figure 15.6). It is located in the middeck of the orbiter crew compartment in a 29 in.-wide area immediately aft of the crew ingress and egress side hatch.

The system collects, stores, and dries fecal wastes and associated paper tissues. It processes urine and transfers it to the wastewater tank and processes EMU condensate water from the airlock and transfers it to the wastewater tank if an EVA is required on a mission. The system also provides an interface for venting trash container gases overboard and dumping atmospheric revitalization wastewater overboard in a contingency situation, and it transfers atmospheric revitalization system wastewater to the wastewater tank.

A waste management compartment door and two privacy curtains attached to the inside of the door provide privacy. The door also serves as an ingress platform during prelaunch (vertical) operations since the flight crew must enter the flight deck over the waste management compartment.

The system consists of a commode, urinal, fan separators, odor and bacteria filter, vacuum vent quick disconnect, and controls. The commode is $27 \times 27 \times 29$ in. and is used like a standard toilet. The commode contains a single multilayer hydrophobic porous bag liner for collecting and storing solid waste. When the commode is in use, it is pressurized and transport air flow is provided by the fan separator. When the commode is not in use, it is depressurized for solid waste drying and deactivation.

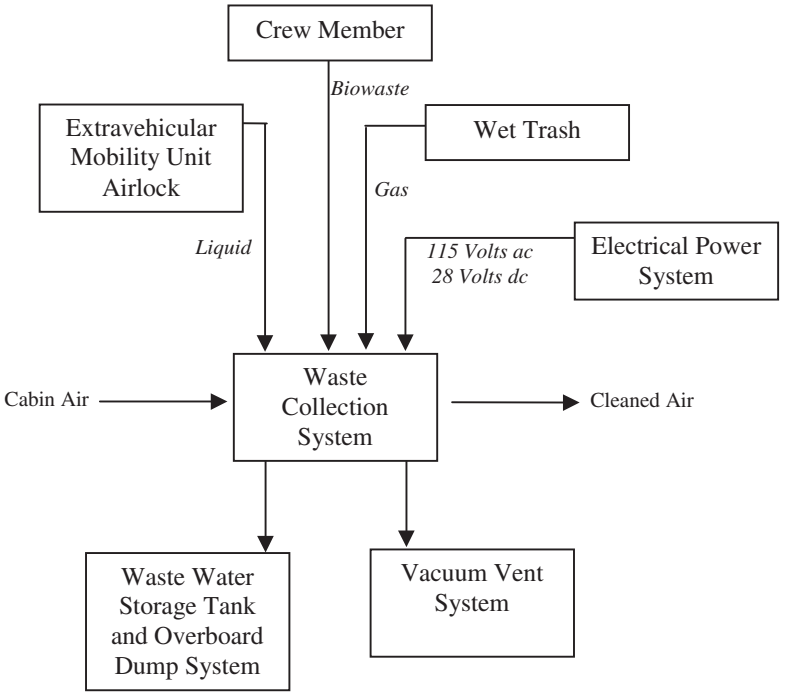


FIGURE 15.6 Waste management system schematic.

The urinal is essentially a funnel attached to a hose and provides the capability to collect and transport liquid waste to the wastewater tank. The fan separator provides transport air flow for the liquid. The fan separators separate the waste liquid from the air flow. The liquid is drawn off to the wastewater tank, and the air returns to the crew cabin through the odor and bacteria filter. The filter removes odors and bacteria from the air that returns to the cabin. The vacuum quick disconnect is used to vent liquid directly overboard from equipment connected to the quick disconnect through the vacuum line.

Sanitation is more important in the orbiter than on Earth. Space studies have shown that the population of some microbes can increase extraordinarily in a confined weightless area such as a spacecraft cabin. This could potentially spread illness to everyone on board. As a result, not only eating components but also the dining area, the toilet, and sleeping areas are regularly cleaned. Since there are no washing machines in space, trousers (changed weekly), socks, shirts, and underwear (changed every two days) are sealed in airtight plastic bags after being worn. Garbage and trash also are sealed in plastic bags.

Orbiter travelers have facilities and supplies available for sponge baths while in space. They can obtain water from the water-dispensing system. Water temperature can be set at any comfortable level from 18 to 35°C (65 to 95°F).

Because of weightlessness, water droplets would float about in the cabin. This could be not only a nuisance but also potentially hazardous to equipment and crew. To prevent this from happening, an airflow system directs wastewater into the or-

biter's waste collection system, where the wastewater is sealed in plastic watertight bags. Personal hygiene and grooming provisions are provided for both male and female flight crew members. Ambient warm water for washing comes from a personal hygiene hose (PHH) attached to the galley auxiliary port.

Whiskers cut off in shaving and floating about weightlessly in a cabin could be a nuisance and foul up equipment. Using conventional shaving cream and a safety razor and cleaning off the face with a disposable towel avoids this problem. Also available is a wind-up shaver that works like an electric razor and contains a vacuum device to prevent the escape of cut whiskers.

A personal hygiene kit is furnished for each crew member for brushing teeth, hair care, shaving, nail care, etc. Additional grooming and hygiene items are available for female crew members.

PART 9

SPACE SUITS

Nicholas Colford

The term *space suit* immediately brings to mind bulky, snowman-like figures with mirrored helmet visors floating against a backdrop of stars or bounding across the Moon. These days an astronaut spends most of his or her time in space wearing everyday clothes. Pressure suits are worn just for launch and entry, and the classic ''snowman'' suits worn outside the comfy environment of the spacecraft for a few hours at a time.

Nevertheless, space suits symbolize, perhaps more than spaceships, man's mastery of the hostile environment of space. The degree to which they succeed in allowing humans to work in space is a testimony to the technological progress made so far. Their limited operational autonomy and the degree to which they impede the full range of human activity is a sign of how alien the environment is in which they permit mankind to survive.

15.34 PRESSURE SUITS

Pressure suits in general and space suits in particular, protect the wearer from several aspects of reduced atmospheric pressure:

Dysbarism sickness: Blood embolism and the ''bends''—rate of decrease of atmospheric pressure

Oxygen starvation: PPO_2 in lungs

Explosive decompression: absolute atmospheric pressure

Hypothermia: temperature

Space suits also have other functions depending on what they are used for. Launch and entry suits serve as exposure suits in case of forced landings in water or cold climates and accommodate emergency equipment for survival and rescue. An EVA suit has to protect the astronaut from the extreme hot and cold of space, cosmic radiation, and micrometeoroids.

In the terrestrial atmosphere at sea level, the overall pressure is 14.7 psi. The 20% oxygen in the atmosphere contributes with a partial pressure of 3.1 psi. This pressure of oxygen ($P O_2$) is the critical parameter for supplying oxygen and maintaining human life. If the atmospheric pressure around the human is half that at sea level but the air contains twice the normal amount of oxygen, then the human can breathe just as easily. Similarly, if the atmosphere is pure oxygen then an overall pressure of 3.1 psi will provide a normal amount of oxygen.

Oxygen, though essential for life, gradually becomes toxic to humans at pressures above the normal 3.1 psi PO_2. At more than around 4.7 psi respiratory symptoms can be observed and cell damage, has been observed after long exposure to levels of 5 psi. Low levels of oxygen affect human performance in a most insidious manner, reducing immediately the capability to perceive the changes that are occurring. The lower limit NASA sets for acceptable performance is 2.7 psi. Hallu-

cinations can be expected at 2.2 psi and paralysis at 1.9 psi. The effects come on gradually as the oxygen already in the blood is consumed, but exposure to air with less than 0.1 psi O_2 results in almost immediate unconsciousness and death in around 2 minutes.

The importance of the above for the construction of space suits is that if the suit is filled with pure oxygen, it only needs to be inflated to 3.1 psi rather than the 14.7 psi that would be required to use terrestrial air.

Dysbarism occurs when the air pressure changes. Gases present in the atmosphere become dissolved in body tissues and establish an equilibrium concentration. If the ambient pressure drops, then some of the gas will come out of solution. The bloodstream is capable of carrying this gas away at a certain rate. If the pressure drop is fast enough, then more gas will come out of solution than the blood can transport away and bubbles of gas can form. The various symptoms of dysbarism sickness have been grouped into ''bends,'' joint and tissue pains; ''chokes,'' chest pain and coughing fits; ''skin manifestations,'' rashes and itching; ''circulatory collapse''; and neurological disorders such as delirium and coma. The effects tend occur more or less in the order listed above but can crop up with or without their normal precursors.

The above is a significant operational issue for space suits. Today's spacecraft operate with an atmosphere of oxygen and nitrogen mixed in the proportions found in the terrestrial atmosphere (80:20) and maintained at a pressure equivalent to sea level on the earth (14.7 psi). Today's space suits operate with an atmosphere of pure oxygen maintained at a pressure of 5.7 psi in the case of the Orlan, or 4.3 psi in the case of the EMU.

In order to avoid the risk of dysbarism sickness, both suits are used only after a period of breathing pure oxygen. The oxygen is breathed at the normal air pressure prior to one's closing oneself in the suit, sealing the airlock, and going down to the operating pressure. The prebreathe, as it is called, allows the nitrogen to escape from the body as the gas in the tissues tries to reach an equilibrium concentration with new atmosphere which contains no nitrogen. For the Orlan the prebreathe at 14.7 psi period is 30 minutes. For the EMU the equivalent period is 4 hours. This significant difference is in part due to the higher working pressure of the Orlan suit and in part to the Russian Space Program accepting higher levels of nitrogen in the blood. NASA Shuttle operations have an alternative prebreathe program in which the cabin pressure is lowered to 10.2 psi for the 36 hours prior to the EVA, from which pressure only a 40 minute pure oxygen prebreathe is required.

Pressure suits provide the astronaut with a nice, clean oxygen supply, but the astronaut inevitably poisons that supply with the CO_2 produced though respiration. If the suit operates an open loop with an external oxygen supply, then the rate of consumption from the external pure O_2 supply will ensure that CO_2 levels in the suit do not build up. If the suit has a closed-loop oxygen supply, then it will incorporate a scrubber to remove CO_2 from the air chemically before circulating it back to the suit interior for breathing again.

Space suits need to cater to other metabolic products of their wearers apart from scrubbing CO_2 out of the air. This is achieved with a diaper or absorbent brief for both solid and liquid waste. This hardly enhances the glamor of space flight, but given that space suits are worn mainly for takeoff and landing or EVA, the pressure of the job in hand is quite enough. The additional stress of bladder pressure or bowel pressure is one that the crew member can well do without.

One of the challenges in building a space suit is to have the suit follow the movements that its wearer makes without changing its internal volume. This is

important because if the volume changes, the pressure in the suit changes. One way of achieving this was discovered early in pressure suit design when it was noticed that the tomato worm can bend and curve without increasing the pressure inside its body. By mimicking the convoluted tubular surface of the worm's body, the ability was found to make segments of suit that could flex but did not change volume. These are known as tomato worm or convolute joints. Another solution is the stovepipe joint, in which two segments of a rigid tube are joined with a rotational joint whose plane is oblique, not at 90°, to the axis of the two segments. When the segments rotate around the joint they move out of axis, effectively providing a flexion in the tube they make up. Stovepipe joints are able to work at higher pressures than convolute joints without becoming harder to move. On the other hand, they can impose on the wearer sequences of movements to achieve a certain form, while the convolute joint bends immediately.

NASA's space suits underwent almost constant development and refinement throughout the Mercury, Gemini, Apollo, and Skylab programs. With the Shuttle program the suits have matured and become more operational items subject to occasional upgrading rather than being in constant flux as before. The period from Mercury to Apollo saw the evolution of the space suit into the form which is familiar today. This evolution included: The transition from a close-fitting crash helmet to a bubble helmet; The addition of the exterior thermal micrometeor garment for EVA; and The addition of the autonomous life support system (rather than having an umbilical cord from the spacecraft).

15.35 EVA SUITS

Both the Russian and the U.S.A. EVA suits carry the names of birds. The Russian Orlan is named after the large and majestic osprey or sea eagle, while the acronym for the U.S. extravehicular mobility unit (EMU) coincidentally (one hopes) is the name of a large but unfortunately flightless Australian bird.

The EMU is a semirigid EVA suit that NASA has used for Space Shuttle EVAs. It was originally designed for use with the Space Shuttle and therefore was intended to be serviced on Earth rather than stored for long periods on orbit. However it has been modified for use on board the ISS. The major parts of the suit are as follows.

The hard upper torso (HUT)/arm assembly is the part of the suit, apart from the helmet and gloves, worn above the waist. It contains the passive half of the mating ring with the lower torso assembly and the active parts of the mating rings with the helmet and gloves. The HUT is a rigid laminated structure which, as well as being a pressure vessel, provides routing for oxygen, cooling water, and electrical lines from the life support system to the helmet and LCVG. Each arm assembly is further divided into the upper arm assembly, the rotating scye bearing, the lower arm assembly, the rotating arm bearing, and the wrist locking mechanism. The pressure vessel of both upper and lower arm assemblies is a heat-sealed, urethane-coated nylon bladder. Axial loads are retained with redundant webbing attached to the bearings, and circumferential pressure loads are restrained by dacron cloth surrounding the bladders. Separate TMGs cover the HUT, upper arm assemblies, and lower arm assemblies.

The lower torso assembly (LTA) is the part of the suit, including the boots, worn below the waist. It consists of the waist assembly, the rotating waist bearing, the trouser assembly, and the two boot assemblies. The waist assembly contains the active part of the connection ring between the HUT and the LTA. Twelve latches keep the two parts together, and there are various alignment and operational aids to ensure that they can be easily and reliably mated during donning and doffing of the suit. The rotating waist bearing allows rotation of the hips with respect to the torso. The trouser assembly is made up of the thigh/brief assembly and then left and right versions of the thigh sizing insert, the knee assembly, the leg sizing insert, and the boot disconnect. As with the soft parts (arm asemblies) of the upper part of the suit, suit pressure is kept in heat-sealed, urethane-coated nylon bladders restrained by dacron cloth surrounding the bladders. The boots contain sizing inerts to optimize the fit with the wearer and on the outside have sole and heel structures that interface with the EVA foot restraints. There are TMGs for each part of the LTA, and on the right thigh TMG is a pocket to carry the EMU scissors.

The EV gloves provide a great deal of flexibility in very small structures, compared to the rest of the suit. The pressure bladder is a dipped urethane form and is prevented from deforming in the first place by a polyester cloth restraint system that deals with cirumferential and axial loads. In addition, a malleable stainless steel bar runs across the palm of the hand in a polyester sheath which, although it can be bent to fit and suit the form of the individual hand, prevents the palm of the glove from ballooning. There is a rotating wrist bearing in the wrist disconnect ring. Sizing and fit of the gloves is critical for performance of manual tasks in a pressurized suit. In addition to the nine sizes of glove, drawstrings allow the fingers to be shortened and the positioning and forming the palm bar. In really difficult cases, custom gloves are made. The astronaut has the option of wearing a nylon comfort glove under the EV glove for extra comfort. The EV glove is covered with a TMG, although the micrometeorite protection is removed and thermal protection is reduced to allow more flexibility of movement. A vulcanized rubber surface is applied to the palm of the TMG to improve gripping, and finger and thumb caps of similar rubber reinforced with kevlar provide provide a balance of thermal insulation and tactility. On the wrist of the TMG is a tether loop for tools, a roll-on sleeve to cover the disconnect, and, on the right-hand gloves only, a flap to cover the wristwatch.

The helmet and extravehicular visor assembly (EVVA) together form the only all-rigid part of the pressure suit. The helmet is a transparent bubble of ultraviolet-stabilized formed polycarbonate. Its disconnect ring latches to an active ring on the HUT, and when it is locked in position, oxygen can flow from a manifold on the ring into the helmet through the vent pad, which is a white polycarbonate flattened nozzle which deflects the air and makes it flow from behind the astronaut's head across the face of the helmet. A purge valve is located in the left-hand side of the helmet bubble. Inside the helmet two optional devices can be mounted. A Frenzel lens can help the astronaut to see the DCM mounted on the chest part of the HUT. The ''valsalva'' device is a small structure against which the nose can be pressed to facilitate clearing pressure from the ears during pressure changes. The EVVA is akin to a TMG for the helmet, providing visual shielding as well as thermal and micrometeor protection. It has a permanent clear protective visor at the front, hard-coated for scratch resistance, while the back shell of the visor is made of fiberglass and polycarbonate. From within the backshell the sun visor can be pulled down to completely cover the visor. The sun visor is made from polysulfone sheet, is hard-

coated and has a thermal/optical gold coating. Left- and right-hand side eyeshades can be pulled all the way down, and a center eyeshade can be pulled down to about halfway down.

The liquid cooling and ventilation garment (LCVG) is worn closest to the astronaut's skin. It is a tight-fitting, long-sleeved, long-legged combination-type undergarment. It contains a network of tubes that allows chilled water to circulate over the body, carrying away the heat produced by the person during the EVA. It also carries a second network of tubes that collect air at the extremities of the body (arms and legs) and deliver it back to the LSS. In addition to six sizes of main garment, there are extension cuffs and different-sized boots to better fir the full range of astronaut sizes. Elastic loops over the thumbs prevent the arms from riding up, and the boots prevent the legs from riding up. Antiabrasion pads for the shoulders and knees can be added if needed. The cooling and vent ducts come together across the back of the LCVG and are routed into a manifold and connector which mates with a corresponding connector inside the HUT in the chest area.

The operational bioinstrumentation system (OBS), or biomed, monitors the electrocardiogram of the wearer of the suit. The biomed is hosted in a small pocket on the LCVG and runs from three electrodes attached to the chest of the astronaut to a connector inside the HUT.

The communications carrier assembly (CCA) is a soft cap containing redundant earphones and microphones enabling voice communications with the Shuttle (or Space Station) and thence with the ground. Six sizes of cap cover the range of 5th to 95th percentile head sizes. The harness carrying the voice signals attaches to a connector inside the HUT.

Before we move out of the suit itself and onto the life support system, a few words on some of the mundane comfort items in the suit. In the HUT, just below the neck ring, an in-suit drink bag (IDB) can be attached. A bag holds 21 fluid oz. of water, which is available to the suit wearer through a mouthpiece that sticks up just above the neck ring. The bag also has a pocket where a food stick can be lodged for nibbling on during the EVA. At the other end of the same system, so to speak, urine collection and storage is provided. Female astronauts have the disposable absorption containment trunk (DACT), which works on the diaper principle, while male astronauts have the urine collection device (UCD), which is a conformal bag interfacing to the crewmember with a roll-on cuff and a one-way check valve. Both devices have a capacity of 32 fluid oz. of urine.

The Life Support System

The life support system provides the following functions for the astronaut during an EVA: provision of breathing oxygen; suit pressurization; cooling of the person; displays and controls of the suit; monitoring of suit performance and consumables.

The primary oxygen supply (POS) is the provider of the first two functions: breathing oxygen and suit pressurization. It consists of two oxygen tanks for a volume of 240 $in.^3$ The tanks are pressurized to 1050 psia and the system delivers oxygen to the suit at 4.9 psid. The POS also pressurizes provides regulated pressure in the feedwater tanks that supply water to cool the astronaut. The POS is controlled by the astronaut from the DCM worn on the front of the chest of the HUT.

There is a secondary oxygen system (or secondary oxygen pack, SOP) with oxygen stored in two smaller tanks at lower pressure than the primary system. As

an emergency backup, the SOP does not provide pressure for cooling water, but its purge operational mode does provide some cooling. When the suit's pressure drops below around 3.8 psid, the SOP starts providing oxygen to maintain suit pressure between 3.4 and 3.9 psid. The SOP will provide 30 minutes worth of oxygen even with a high metabolic work rate (1000btu/hr) in the absence of nonnominal suit leakage.

The gas in the suit is conditioned by running through the oxygen ventilation circuit, which regulates gas flow through the suit, cools the gas that circulates in the suit, and removes CO_2 and water vapour from the gas in the suit. The gas has CO_2 removed from it by passing through the contaminant control cartridge (CCC), which can remove 1.48 lb of CO_2, corresponding to 1,000 btu/hr for seven hours of EVA. The cartridge also has a filter to trap particles and charcoal to absorb trace contaminants. While the water that flows through the suit to cool the astronaut is constantly recirculated in closed loop, another separate loop of water, the feedwater loop, provides water to the sublimator, where it is lost into space cooling the water for the LCVG.

The space-suited astronaut can communicate with the Shuttle (or Space Station) and thence to the ground through the extravehicular communicator (EVC). Two AM radio transmitters and three AM receivers provide several modes of voice and data communication to carry conversations and telemetry data.

TABLE 15.28 Space-Suit Component Sizing

Component	Number of sizes
Helmet	1
HUT	4
Arm assembly	
Scye bearing	4
Upper arm assembly	5
Arm bearing	2
Lower arm assembly	6
Lower arm sizing insert	9
Wrist disconnect ring	1
Glove	
Standard	9
Custom	Unlimited
Lower torso assembly	
Waist disconnect ring	4
Waist assembly	6
Waist bearing	3
Thigh/brief assembly	3
Thigh sizing insert	6
knee assembly	2
Leg sizing insert	8
Boot disconnect	1
Boot assembly	2
Boot sizing insert (“slippers”)	6

TABLE 15.29 Important Pressure levels

	Overall air pressure	Oxygen pressure	
Sea level	101.4 kgPa 14.7 psi	21.2 kgPa 3.08 psi	
Minimum human need (air)	48.3 kgPa 7 psi	13 kgPa 1.9 psi	
Pure oxygen breathing range	24 kgPa 3.5 psi	34 kgPa 5 psi	
Mercury suit operating pressure	34 kgPa 5 psi	Gemini suit operating pressure	25.5 kgPa 3.7 psi
Apollo suit operating pressure	25.5 kgPa 3.7 psi	Shuttle suit (EMU) operating pressure	19.6 kgPa 4.3 psi
Zero prebreathe suit	55 kgPa 8 psi		
LES	Works up to 100,000 ft altitude; protects for 30 min over 100,000 ft		
ACES	Indefinite protection over 100,000 ft		

Sizing

LCVG: six sizes.

Helmet: one size.

Gloves: 4000 series, nine sizes with finger-length adjustment straps, made to measure for outliers.

Arm/leg assemblies: "different" size. An enhanced EMU will arrive with on-orbit size odification through interchanging of different sized segments and sizing rings. Enhanced EMU for Station and future Shuttle activities.

CCA: size sizes.

In-suit drink bag: two capacities.

UCD and MAG: no sizing information.

FURTHER READING

ASTM E595, *Standard Test Method for Total Mass Loss and Collected Volatile Condensable Materials from Outgassing in a Vacuum Environment.*

MSFC-HDBK-527/JSC-09604, *Materials Selection List for Space Hardware Items.*

MSFC-SPEC-522, *Design Criteria for Controlling Stress Corrosion Cracking.*

NHB 8060.1, *Flammability, Odor, Offgassing, and Compatibility Requirements and Test Procedures for Materials in Environments That Support Combustion.*

SECTION 16
EARTH'S ENVIRONMENT AND SPACE

Section Editor: Michael J. Rycroft

PART 1

THE EARTH AND ITS ATMOSPHERE

Michael J. Rycroft

16.1 THE EARTH IN SPACE

The Universe began with the Big Bang, about 14 billion years ago. About 4.6 billion years ago, the Sun formed during the gravitational collapse of an enormous cloud of gas and dust. Rotating around the Sun was a disk of material out of which the planets grew. These planets now orbit the Sun in the ecliptic plane. Our Earth is the third planet away from the Sun.

The different materials within the Earth gradually separated, with the densest (iron) forming an electrically conductive core. Dynamo action in the core generates the geomagnetic field. Above is the mantle, a region which convects very slowly. The Earth has a relatively thin crust, a few tens of kilometres thick, compared with the Earth's equatorial radius of 6,378 km. The crust and uppermost mantle are made up of 12 tectonic plates, hundreds of kilometers or so thick, which move relative to each other at up to one centimeter per year.

The Earth's atmosphere has evolved through outgassing from the interior via volcanoes. Oxygen was generated by dissociation of water vapor; the hydrogen that was also formed gravitationally escaped into space. Oxygen is also generated via photosynthesis by plants.

The Earth's hydrosphere (oceans), atmosphere, biosphere and solid surface interact with each other in many complex ways. This area of science is now termed Earth Systems Science (Ernst 2000).

16.2 PROPERTIES OF THE EARTH'S ATMOSPHERE

This subsection gives an overview of the physical properties of the atmosphere—its temperature, pressure, density, and composition—and how parameters expressing these vary with height (termed profiles). It also shows average wind distributions through the atmosphere.

The variation of temperature with height divides the atmosphere into different regions, as shown in Figure 16.1. The troposphere, the region of the Earth's weather, reaches up to about 14 km altitude in the middle latitudes, up to 18 km in the tropics, and only 9 km in the polar regions. The temperature decreases with increasing height; the lapse rate is typically 6.5 K/km. At the top of the troposphere, the tropopause, the temperature begins to rise with increasing height. This region is called the stratosphere. It is a stably stratified layer, with considerably less rapid vertical mixing than in the troposphere. Above the stratopause is the mesosphere, about which little is known. The mesopause is the coldest region of the atmosphere, typically at a temperature below 180 K. It is coldest in the local summer, paradoxically. Above lies the thermosphere.

EARTH'S ENVIRONMENT AND SPACE

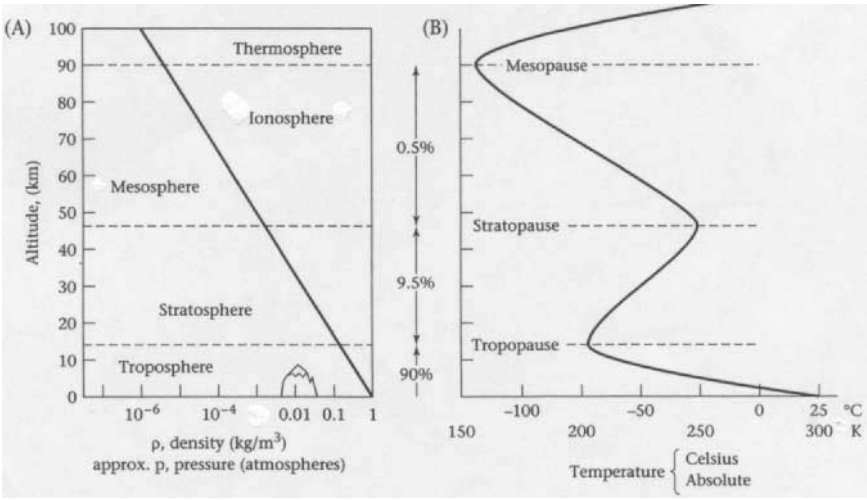


FIGURE 16.1 Variation of the Earth's atmospheric density and approximate pressure as a function of altitude in km. The density is 1.17 kg/m^3 at the surface, where the pressure is 1 atm or 1,010 hPa (or mb) (equivalent to a weight of 10.1 kg/m^2). The temperature of the atmosphere (shown in degrees Celsius (Centigrade) and also in degrees absolute, or Kelvin) varies considerably; it is higher where the Sun's energy is absorbed. At 0°C the absolute temperature is 273.15 K (from Ernst 2000).

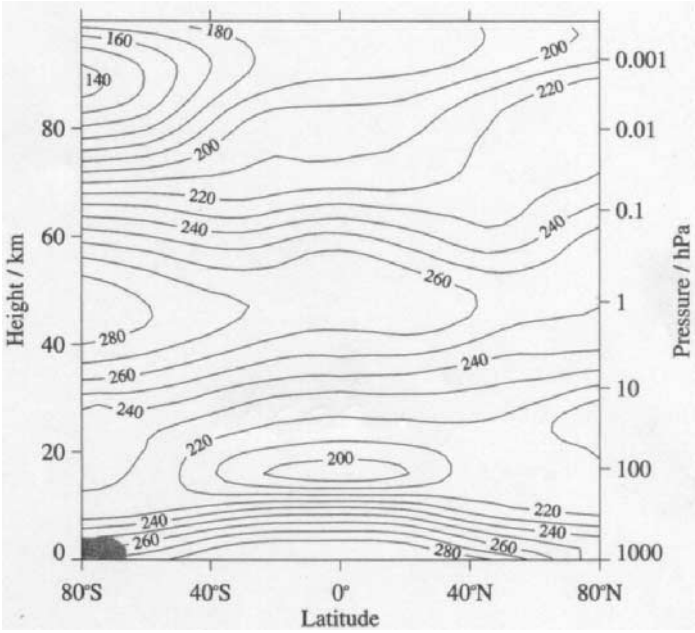


FIGURE 16.2 Zonal mean temperature (K) for January (from Fleming et al. 1990 and Andrews 2000).

Figure 16.1 shows, using a logarithmic scale, the density and also approximate pressure as a function of altitude. The pressure at 100 km altitude is about one millionth of its value at the surface (close to 1,010 hPa, on average). Some 90% of the Earth's atmosphere resides in the troposphere, and 99.5% in the troposphere and stratosphere. The upper atmosphere is tenuous indeed.

Conventional meteorological measurements of temperature over the Earth's land and sea surface are complemented by thermistor measurements aboard hydrogen- or helium-filled balloons, plus values derived from satellite instruments using infrared radiation to determine the temperature field in three dimensions. An average over longitude, the zonal mean temperature, in K is shown in Figure 16.2 (Andrews 2000). The cold summer mesopause is evident. Noctilucent clouds of tiny ice crystals often form here then.

The equator-to-pole temperature gradient drives the atmospheric heat engine, the weather machine. Winds blow over the Earth, attempting to reduce the temperature differences that exist between different places. Figure 16.3 shows average zonal winds for the month of January (Andrews 2000). At altitudes below 100 km the atmospheric composition is essentially the same everywhere. The predominant gas is nitrogen (N_2), 78% by volume. Next comes oxygen (O_2), 21% by volume. Argon, an inert (or noble) gas (A), is present at 0.93% by volume. Carbon dioxide (CO_2) now constitutes 0.036%, or 360 parts per million by volume (ppmv). In 1800 it was at 280 ppmv, and in 1900 295 ppmv, the increase being due to the burning of fossil fuels. Water vapor (H_2O) is present in amounts ranging from 0 up to 0.04%

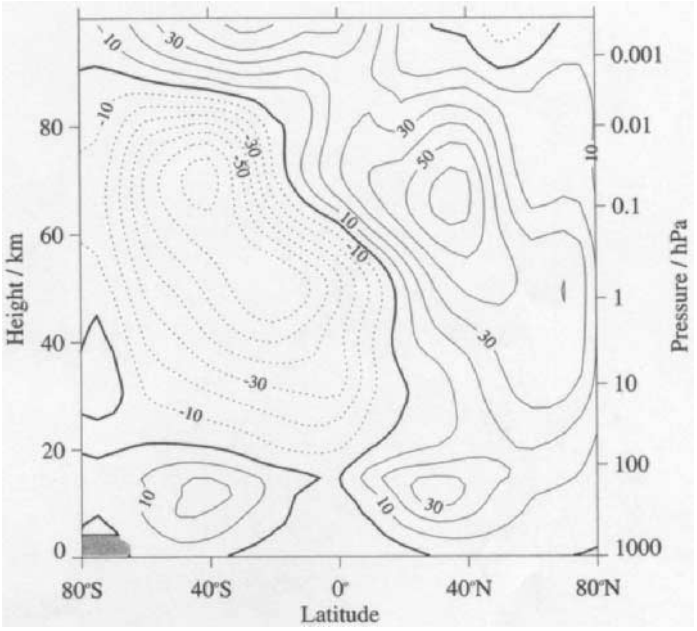


FIGURE 16.3 Zonal mean wind (ms^{-1}) for January (from Fleming et al. 1990 and Andrews 2000). Thin solid lines: eastward winds (westerlies); thick solid lines: zero winds; dashed lines: westward winds (easterlies).

EARTH'S ENVIRONMENT AND SPACE

(Wallace and Hobbs, 1977). Methane (CH_4), other inert gases, hydrogen (H_2), and ozone (O_3) are all at the 10^{-4}% level. Trace gases, at the 10^{-6}% level, include the oxides of nitrogen, carbon monoxide, sulphur dioxide (from volcanoes), and human-made chlorofluorocarbons (CFCs), which are also termed freons.

To put matters into perspective, the total mass of the Earth is 6×10^{24} kg, the mass of the oceans is 1.3×10^{21} kg, and the mass of the atmosphere is 5×10^{18} kg (Wallace and Hobbs 1977).

16.3 HOW THE EARTH'S ATMOSPHERE WORKS

This subsection discusses the underlying physical mechanisms, namely the absorption of solar radiation and gravity, which determine the profile of important atmospheric parameters.

The atmosphere becomes hotter where radiation from the Sun is absorbed. Visible light is absorbed at the Earth's surface, and the tropical regions become especially hot. Solar ultraviolet radiation is absorbed by molecular oxygen (O_2) and by ozone (O_3) in the stratosphere. X-rays from the Sun's corona (its outer atmosphere) are absorbed in the thermosphere. At 400 km altitude the temperature rises to ~800 K during solar minimum conditions and to ~2000 K at solar maximum conditions. The interval between one solar maximum and the next (occurring in 1957, 1968, 1979, 1990, and 2001) is about 11 years. This is termed the solar cycle.

The temperature of the Sun's surface, the photosphere, is close to 6000 K. It radiates a blackbody spectrum (shown in Figure 16.4(a)), which peaks in the visible part of the spectrum at a wavelength of 0.6 μm (600 nm), in the yellow. Figure 16.4(b) and (c) (Goody 1995) show that most of this radiation reaches ground level, with shorter wavelength radiation (wavelengths below 0.3 μm, the ultraviolet) being absorbed by O_2 and O_3 in the stratosphere.

Molecular gases such as water vapor, carbon dioxide, methane, and ozone absorb infrared radiation with wavelengths greater than 0.7 μm. In the far infrared, peaking at a wavelength between 10 and 20 μm, is radiation emitted into space by the Earth-atmosphere system at an average temperature near 250 K. Most of this terrestrial radiation is absorbed by molecular trace gases in the atmosphere. Reradiated both down and up, this mechanism accounts for the greenhouse effect. It maintains the Earth–atmosphere system at a temperature of 288 K, 33°C warmer than it would otherwise be. Thus, there is no doubt that the greenhouse effect is a good thing for humans living on the Earth. What could be a bad thing, at least for many, is an increasing greenhouse effect. Termed *global warming,* this is due to increasing amounts of gases such as CO_2, which are infrared active gases, produced by the burning of fossil fuels (Houghton 1997; Houghton et. al 2001). Other mechanisms also contribute to global warming.

Thermodynamics plays an important role in our understanding of atmospheric phenomena. The ideal gas law or equation of state relates the pressure p, density ρ, and temperature T, as

$$p = \rho RT$$

where R is the gas constant for 1 kg of gas.

For a mass of gas equal to its molecular weight in kg (28 for N_2, 32 for O_2), which contains 6×10^{26} molecules (N_A, Avogadro's number), the universal gas

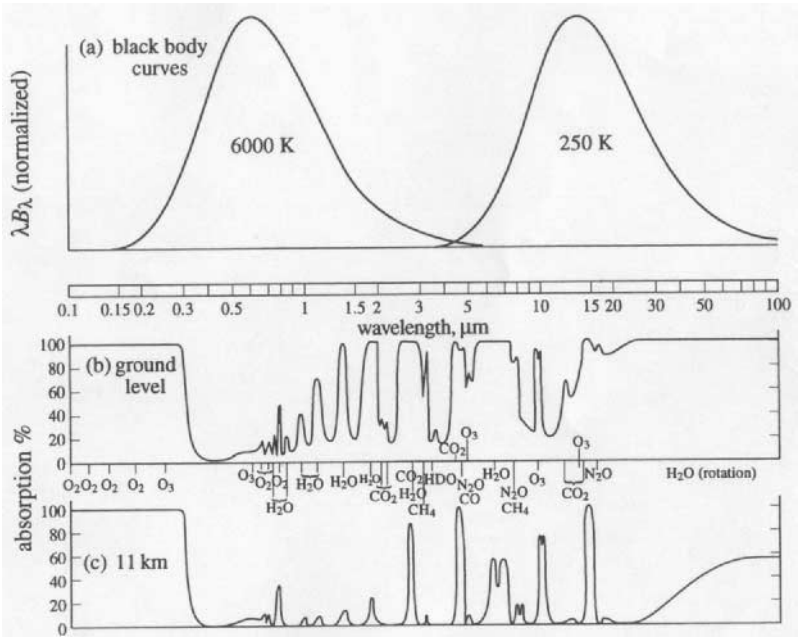


FIGURE 16.4 (a) Blackbody curves for 6000 K and 250 K, the Sun and the Earth, respectively; (b) atmospheric absorption spectrum for solar radiation reaching the ground; (c) the same for radiation reaching the temperate tropopause. The areas beneath the curves in (a), proportional to the energy fluxes, are the same over the globe for a year (from Goody 1995).

constant is $R^* = 8300$ J K^{-1} kilomole^{-1}. With the volume of this amount of gas being V, the ideal gas law becomes

$$pV = R^*T$$

For one molecule of gas, the universal gas constant is Boltzmann's constant, $k = 1.38 \times 10^{-23}$ J K^{-1}. Thus $R^* = N_A \times k$. Hence,

$$p = \frac{N_A}{V} kT = NkT$$

where N is the number density of molecules, each of mass m.

At sea level, $N = 2.7 \times 10^{25}$ m^{-3}. The reader can check that the atmospheric pressure at sea level, with $T = 288$ K, is $p_0 = 10^3$ hPa.

Going up into the atmosphere, the atmospheric pressure decreases. The atmospheric pressure is due to the mass of air above unit area (1 m^2). For the equilibrium of a slab of air of thickness dz,

$$\frac{dp}{p} = -g \frac{dz}{RT}$$

Integrating, for an isothermal atmosphere at constant temperature T, from p_0 up to the height z, where the pressure is p,

EARTH'S ENVIRONMENT AND SPACE

$$p = p_0 \exp\left(-\frac{gz}{RT}\right) = p_0 \exp\left(-\frac{z}{H}\right)$$

The pressure decreases exponentially with increasing height z above the Earth's surface. In a distance

$$H = \frac{RT}{g} = \frac{kT}{mg}$$

p becomes p_0/e. The pressure falls to $p_0/2.72$ over a height equal to the scale height. Inserting $T = 288$ K for near the Earth's surface, the scale height is 7 km.

Over a height range of 2.3-scale height, about 15 km, the pressure decreases by a factor of 10. Over twice that distance the pressure decreases a hundredfold. This theory thus explains the variation of pressure with height shown in Figure 16.1(A).

The mass of atmosphere per unit area (1 m^2) of the Earth's surface is in fact equal to the mass of atmosphere of uniform density at sea level N over a scale height H. This is about 10^4 kg, namely 10 tonnes, and explains the sea-level atmospheric pressure of 10^5 Pa (10^3 hPa).

The theory given here can be extended, with greater algebraic complexity, to consider a nonisothermal, and hence more realistic, atmosphere.

At heights above 110 km, termed the turbopause, the gas composition starts to vary with altitude. This is because lighter gases, e.g., atomic oxygen (O) and hydrogen (H_2), float above the heavier gases such as molecular oxygen (O_2). This process is called diffusive, or gravitational, separation. Thus, at a height of 300 km the neutral gas scale height is about 70 km, an order of magnitude greater than at sea level because the absolute temperature is five times greater and the molecular weight is halved.

At around this height the exosphere begins. This is where the atmosphere is so thin that the frequency of collisions between atoms/molecules is so small that the mean free path between collisions becomes comparable with the scale height. It is likely that an upward-moving hydrogen atom (H), produced from water vapor or methane in the stratosphere, will not collide with another particle. If its temperature is above 5200 K, the thermal velocity of a hydrogen atom exceeds the velocity required to escape altogether from the Earth's gravitational field. This is a mechanism by which the Earth loses mass.

The Earth gains mass, $\sim 10^5$ kg per day, by the influx of meteoroids to the top of the atmosphere. A meteor glows for a few seconds in the upper atmosphere ($\sim$100 km altitude) as it burns up. More massive meteoroids can reach ground level; these are called meteorites.

16.4 ATMOSPHERIC DYNAMICS AND ATMOSPHERIC MODELS

The atmosphere is warmed where solar radiation is absorbed. The atmosphere moves—winds below—to try to remove these temperature (or pressure) differences. Because the Earth is rotating (at an angular velocity Ω), a moving parcel of air is subjected to the Coriolis force. In the northern hemisphere, this makes the air parcel veer (to the right) of the direct line from high pressure to low pressure. This leads, for the northern hemisphere, to an anticlockwise circulation of air around a low-

pressure system, storm, or cyclone (Andrews 2000; James 1994; Wallace and Hobbs 1977).

Applying Newton's second law of motion, the acceleration of a parcel of air of unit mass is equal to

$$F\,\underline{v} \wedge \underline{k} - \frac{\Delta p}{p} + \underline{F}_{fr}$$

The first term is the Coriolis force, the second the pressure gradient force, and the third the frictional force at the surface. Here $F = 2\Omega \sin \phi$, where ϕ is the geographic latitude and k is a unit vector in the vertical direction $(+z)$.

For what is termed geostrophic flow, with steady state conditions and

$$\underline{F}_{fr} = 0, \qquad \frac{\Delta p}{p} = \underline{v} \wedge \underline{k}$$

Thus, the wind blows at right angles to the pressure gradient.

This is the explanation of Buys-Ballot's law: in the northern hemisphere, the low-pressure region is on the left if one's back is to the wind. The winds are stronger when the isobars (pressure contours on a weather map) are closer together.

The circulation of the atmosphere is shown in Figure 16.5 (Ernst 2000). It can be understood in terms of the theoretical notions presented. In the stratosphere/mesosphere, there is one convective cell (like a Hadley cell) from summer high latitudes to winter mid-latitudes.

In Figure 16.5, three large convecting cells of air (shown in cross-section on the left-hand side of the globe) define the circulation of the lower atmosphere in each hemisphere. The surface components of each atmospheric cell form the zonal wind belts that drive the surface circulation of the ocean. The limbs of the atmospheric cells include: (1) zones of rising moist air, low pressure (L), and high rainfall in the equatorial zone and along the polar fronts (at 50° to 60°N and S), and (2) zones of descending dry air, high pressure (H), and low rainfall over the polar regions and in the mid-latitudes (at approximately 30°N and S). Eastward rotation of the Earth and the Coriolis effect cause surface winds to veer to the right of their motion in the northern hemisphere and to the left of their motion in the southern hemisphere. The polar jet streams are not surface winds but rather flow eastward at high tropospheric altitude along the polar fronts in wave-like patterns. They influence the positions of the individual high and low-pressure systems north and south of these fronts.

Models of atmospheric motion are needed to produce weather forecasts. The laws of conservation of mass, momentum, and energy are applied to describe and predict (using observed initial conditions) the dynamics of the atmosphere. Solar energy is absorbed at the Earth's surface and through the atmosphere.

The exchange of energy and water vapor between the surface and the atmosphere must be described (Houghton 1997). Water vapor is important because of its latent heat, which it gives out when it condenses, resulting in cloud formation. Latent heat has to be supplied to evaporate water from the oceans.

The atmospheric processes involved in a model of atmospheric behavior are illustrated in diagrammatic form in Figure 16.6, from Houghton (1997). The interactions among atmospheric dynamics, radiation, and chemistry (which is especially important in the stratosphere), dealing with ozone (O_3), are illustrated in Figure 16.7, from Andrews (2000).

Ranging from short- and long-wave radiation to volcanoes, ice sheets, and both gas exchange and heat transfer with the ocean, Houghton (1997) summarizes the

EARTH'S ENVIRONMENT AND SPACE

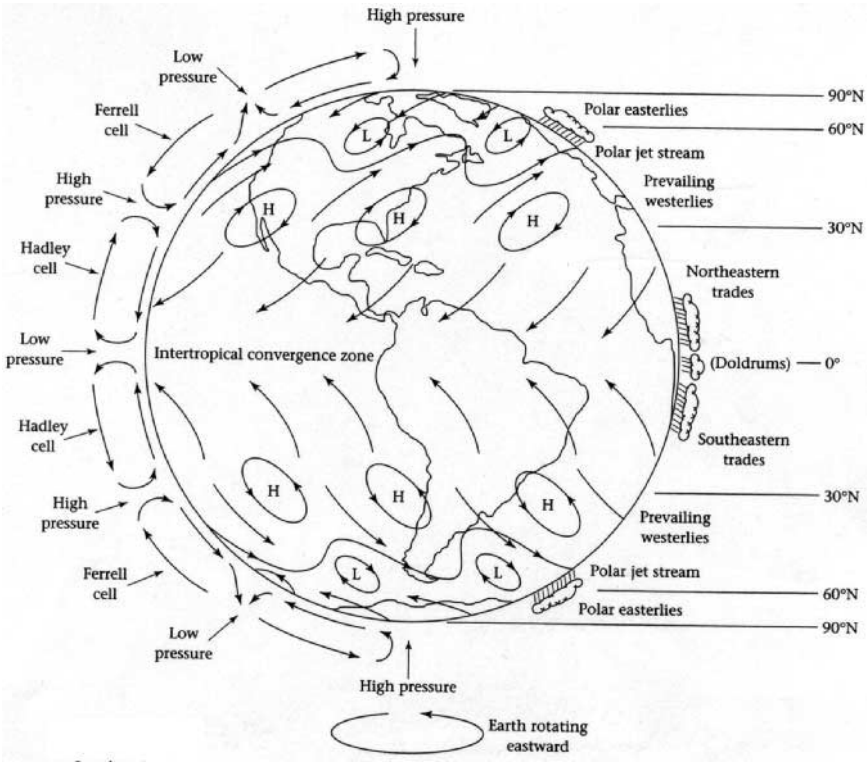


FIGURE 16.5 Schematic illustration of global atmospheric circulation and surface wind patterns (arrows on the Earth's surface) (from Ernst 2000).

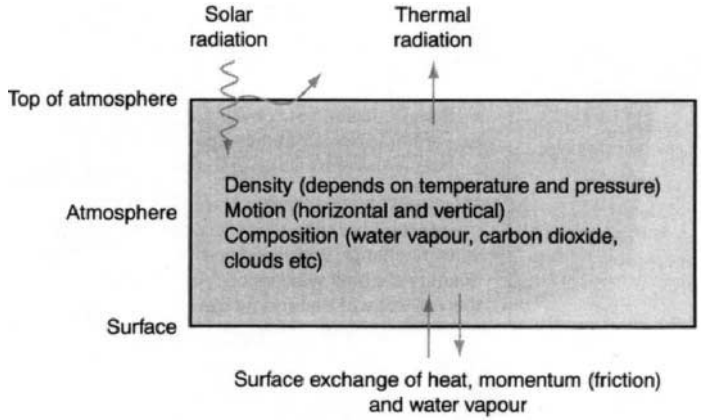


FIGURE 16.6 Schematic diagram illustrating the parameters and physical processes involved in atmospheric models (from Houghton 1997).

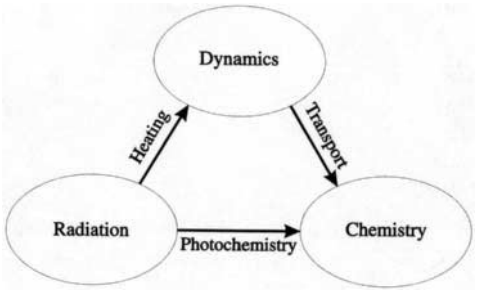


FIGURE 16.7 Diagram showing some of the interactions among dynamics, radiation, and chemistry in the atmosphere (from Andrews 2000).

physical processes involved in the Earth's weather and the climate system. These processes can be well modeled using powerful digital computers today. The anthropogenic impact, i.e., the effects of human activities on the climate (the long-term average of weather), is a great cause of concern today (Houghton et al. 2001). It is changes, both natural and human-caused, which are of especial importance. Figure 16.8, from Houghton et al. (2001), highlights these concerns diagrammatically.

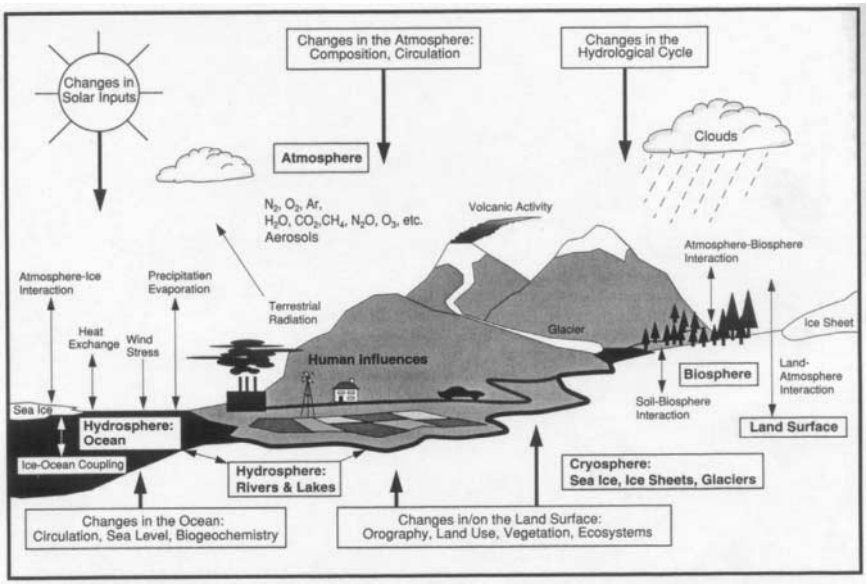


FIGURE 16.8 Schematic view of the components of the Earth's climate system (bold), their processes and interactions (thin arrows), and some aspects that may change (bold arrows) (from Houghton et al. 2001).

EARTH'S ENVIRONMENT AND SPACE

16.5 ELECTRICAL PHENOMENA IN THE ATMOSPHERE

Thunderstorms are at the high-energy end of the spectrum of atmospheric phenomena. Lightning discharges are dramatic events indeed, which can cause significant damage.

Within a thundercloud there are strong upward convective motions, with the lower part of the cloud containing water droplets, and the upper part ice crystals/particles. Heavy ice particles fall as hail, and falling particles collide with other droplets, causing electric charges to be produced.

An upward electric current in a thundercloud carries ~100 C of positive charge to the top, leaving −100 C near the bottom (MacGorman and Rust 1988). The potential difference between the top and bottom of the cloud may reach 100 MV or more. Electrical breakdown of the atmosphere can occur, leading to a lightning discharge to Earth, a cloud-to-ground discharge carrying either negative or positive charge to ground. Alternatively, a discharge may occur within a cloud, an intracloud discharge, or from one cloud to the next, an intercloud discharge.

The electric dipole moment within the cloud or that formed by its electrical image below the ground is destroyed in a short time (~μs). This acts as a strong impulsive source of radio signals at frequencies up to many MHz. These signals propagate in the Earth–ionosphere waveguide.

Some electric current continues up from the top of the thundercloud, charging the ionosphere to a potential of ~ +250 kV with respect to the Earth. The good-conducting ionosphere is almost an equipotential surface. Far away from thunderstorm regions, a fair weather current ~2 μA m^{-2} flows downwards. The global atmospheric electric circuit is completed by currents flowing through the land and sea, and finally by point discharge currents below the thundercloud.

Changes to the electrical conductivity of the atmosphere associated with changes in the flux of cosmic rays or energetic charged particles from the Sun or the magnetosphere may modify the properties of the atmospheric electric circuit (Rycroft et al. 2000).

Sprites are upward electrical discharges from the top of a particularly energetic thundercloud to the ionosphere; they are especially likely to occur after a positive cloud-to-ground strike. They glow strongly, at altitudes between about 70 and 90 km, for only a few milliseconds.

The audio-frequency components of an atmospheric discharge (sferic, for short) can propagate into the ionosphere and be guided by ducts of enhanced ionization along geomagnetic field lines to the opposite hemisphere. Traveling through a dispersive plasma (an electrically charged gas), they emerge as descending frequency tones and are termed *whistlers*.

The lowest radio frequency component of a lightning discharge at ~8 Hz excites the fundamental resonance of the dielectric shell of atmosphere between the good-conducting Earth and ionosphere. These Schumann resonances occur when the wavelength is comparable with the Earth's circumference.

REFERENCES

Andrews, D. G. 2000. *An Introduction to Atmospheric Physics.* Cambridge University Press, Cambridge.

Ernst, W. G., ed. 2000. *Earth Systems: Processes and Issues,* Cambridge University Press, Cambridge.

Fleming, E. L., Chandra, S., Barnett, J. J., and Corney, M. 1990. "Zonal Mean Temperature, Pressure, Zonal Wind and Geopotential Height as Functions of Latitude," *Advances in Space Research,* vol. 10, pp (12)11–(12)59.

Goody, R. 1995. *Principles of Atmospheric Physics and Chemistry,* Oxford University Press, New York.

Houghton, J. T. 1997. *Global Warming: The Complete Briefing,* 2nd ed., Cambridge University Press, New York.

Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van der Linden, P. J., Dai, X., Maskell, K., and Johnson, C. A., eds. 2001. *Climate Change 2001: The Scientific Basis,* Cambridge University Press, Cambridge.

James, I. N. 1994. *Introduction to Circulating Atmospheres,* Cambridge University Press, Cambridge.

MacGorman, D. R. and Rust, W. D. 1998. *The Electrical Nature of Storms,* Oxford University Press, New York.

Rycroft, M. J., Israelsson, S., and Price, C. 2000. "The Global Atmospheric Electrical Circuit, Solar Activity and Climate Changes," *Journal of Atmospheric and Solar-Terrestrial Physics,* vol. 62, pp. 1563–1576.

Wallace, J. M. and Hobbs, P. V. 1977. *Atmospheric Science: An Introductory Survey,* Academic Press, New York.

PART 2

THE NEAR-EARTH SPACE ENVIRONMENT

Michael J. Rycroft

16.6 BACKGROUND

The material presented in a section on the near-Earth space environment can be organized in several ways. First, descriptions of the several different aspects of the space environment can be given; this approach is followed here and by Mitchell (1994), Skrivanek (1994), and Tascione (1994). Detailed works on particular regions, such as the ionosphere (Kelley 1989; Rishbeth and Garriott 1969; Schunk and Nagy 2000) or the plasmasphere (Lemaire and Gringauz 1998), or on charged particles trapped by the geomagnetic field (Walt, 1994), have been published. Alternatively, the Sun–Earth connection, or solar–terrestrial physics, now popularly termed space weather, has been considered by Freeman (2001), Hargreaves (1992), Kivelson and Russell (1995), Gombosi (1998), Suess and Tsurutani (1998) and Song et al. (2001). This important subject is also discussed in Part 6.

Another approach, which might appeal to readers of this handbook, is to consider the impacts that the space environment has on satellites in orbit around the Earth or on spacecraft. This approach is adopted in Section 11, Part 2 and in DeWitt et al. (1993), Dyer et al. (2000), Fortescue and Stark (1991, ch. 2), Hastings and Garrett (1996), Holmes-Siedle and Adams (1994), Wertz and Larson (1999, ch. 8), and Tribble (1995). Further detailed information is available in all these references.

To understand the nature of the near-Earth space environment, some basic knowledge of plasma is required. Plasma is the fourth state of matter. The first state is a solid (e.g., ice); when it is heated, a liquid is formed (water). When a liquid is heated (i.e., given extra energy), a gas, or vapor (steam), is formed. When a gas is heated, an electron is detached from a significant fraction of the molecules to create an overall electrically neutral gas, a mixture of positively charged ions and negatively charged electrons. This partially ionized gas is termed a plasma if three conditions are met (see below).

Under the action of an applied electric field, the electrons move with respect to the much more massive ions. The electrons oscillate at the electron plasma frequency, whose value is f_{pe} (in Hz) $= 9\sqrt{N_e}$(in m^{-3}), where N_e is the electron density.

An electric charge moving at velocity $\underline{v}$ perpendicular to a magnetic field $\underline{B}$ experiences a Lorentz force perpendicular to both $\underline{v}$ and $\underline{B}$; $\underline{F} = q\ \underline{v} \wedge \underline{B}$. Thus, an electron whose negative charge has magnitude e gyrates in a circle of radius (the Larmor radius) equal to mv/eB about the magnetic field direction. The gyrofrequency, or cyclotron frequency, is $f_{Be} = eB/2\pi m$, where m is the electron mass; f_{Be} (in Hz) $= 28\ B$ (in nT). The electron gyroradius (in km) is about $100\sqrt{W}$(energy, in keV)$/B$ (in nT).

The Debye length is the distance in a plasma over which the electric field due to one particular positive charge is appreciable. It is given by λ_D (in m) $= \sqrt{T}$(in K)$/N_e$ (in m^{-3}). At distances greater than λ_D, electrons shield the remainder of the plasma from the effect of this particular ion.

The three conditions for plasma behavior are:

1. The number of electrons N_D in a sphere of radius equal to λ_D must be very much greater than 1.
2. The typical dimension of the problem of interest must be much greater than λ_D.
3. The electron plasma frequency must be greater than the electron-neutral collision frequency, so that plasma waves will not be damped out.

Figure 16.9 shows typical electron density and temperature values for different types of plasmas. Satellite-borne instruments can make satisfactory measurements in the magnetosphere and solar wind without affecting the plasma system being investigated, i.e., these three conditions are properly met.

16.7 THE PLASMA ENVIRONMENT

Gas in the thermosphere is ionized by ultraviolet and X-radiation from the Sun. The partially ionized gas, the plasma, so formed is termed the ionosphere. Representative profiles of the electron density as a function of height, by day and night, are shown in Figure 16.10. The specific radiations causing different ionospheric layers to be formed—the C, D, E, F1, and F2 layers—are also shown.

At 100 km altitude during the day, N_e is $\sim 10^{11}$ m^{-3} and the number density of neutrals is $\sim 10^{19}$ m^{-3}. Only 1 particle in 10^8 is charged, so the ionosphere is a very weakly ionized plasma there. At 300 km, about 1 particle in 10^3 is ionized, still a weakly ionized plasma.

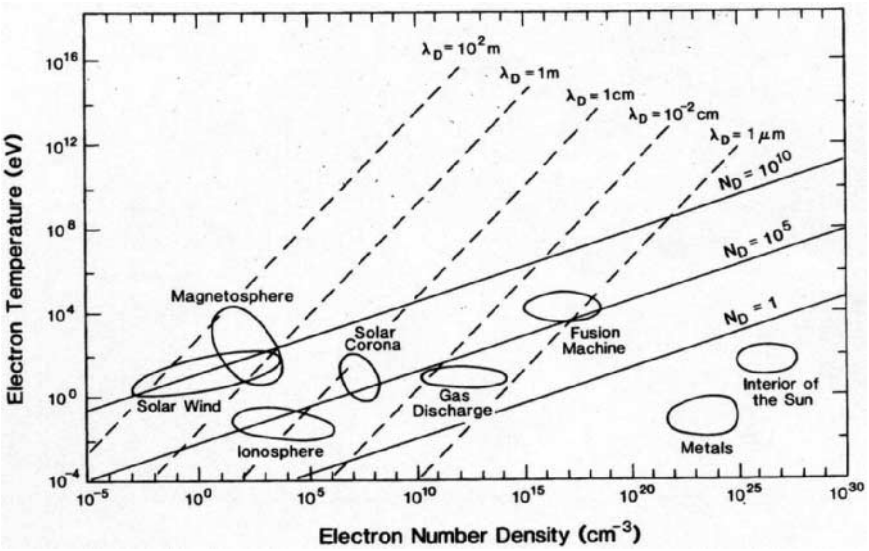


FIGURE 16.9 Values of the Debye length λ_D and the number of particles in the Debye sphere N_D for various plasmas having different electron temperatures and electron densities (from Kivelson and Russell 1995).

EARTH'S ENVIRONMENT AND SPACE

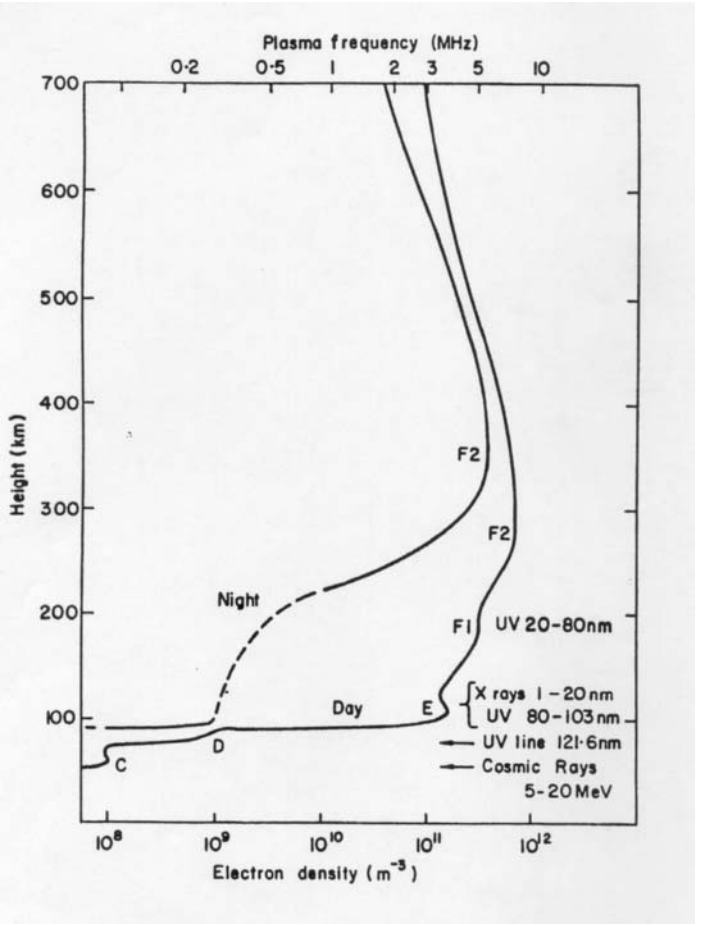


FIGURE 16.10 Typical ionospheric electron density particles, by day and night, and the different radiations responsible for the different layers.

At the top of the ionosphere, the strength of the ionizing radiation is great but there are only a few particles to ionize. Conversely, no ionizing radiation reaches down to the mesophere, but the gas concentration is large there. Thus, a layer of ionization, with a characteristic profile termed a Chapman layer, is produced between these two heights.

Figure 16.11 (Gombosi 1998) shows theoretical Chapman functions for an overhead Sun (S_0, for solar zenith angle $\chi = 0°$) and for greater solar zenith angles.

$$S = S_0 \exp\left[1 - \frac{z}{H} - \sec\chi \exp\left(-\frac{z}{H}\right)\right]$$

where the altitude z is given in units of scale height of the neutral species being

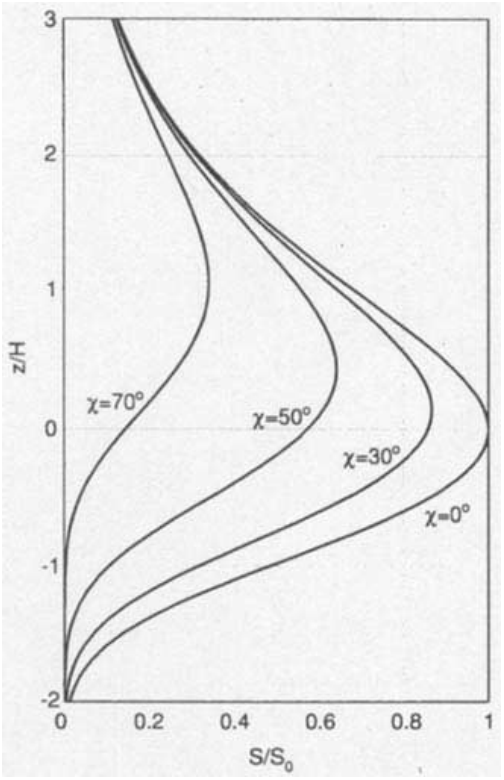


FIGURE 16.11 The normalized Chapman ionization function (from Gombosi 1998).

ionized, normalized to the altitude of greatest rate of production of ionization for an overhead Sun.

Of the various ionospheric layers, the E layer obeys the Chapman formalism most closely. Transport effects—motions due to neutral winds and/or to electric fields—are strong in the F layer. During magnetic storms, the neutral gas composition at each particular height varies from its usual composition, and so the ionization is also affected in a complicated way. Operating near the equator, a fountain effect produces the so-called Appleton anomaly of larger electron densities at ~20° latitude than at the equator. In the equatorial and auroral regions, plasma irregularities with scale widths less than 1 km but elongated up to ~100 km along the geomagnetic field are created by plasma instabilities.

The ionosphere and its irregularities affect the propagation of radio waves used for space communications and/or navigation to a certain extent. In order to propagate through the ionosphere from the ground to a satellite or vice versa, the radio frequency used must exceed the plasma frequency of the F_2 layer (~10 MHz).

High-frequency (HF, 3–30 MHz) radio waves are refracted (bent) by the ionosphere as they propagate through it. Actually two modes of propagation exist in a plasma, termed ordinary (O) and extraordinary (X) modes. At HF, the O and X modes travel at slightly different velocities, both of which differ a little from the

EARTH'S ENVIRONMENT AND SPACE

velocity of light in free space (c, which is nearly 3×10^5 km s^{-1}). Further, HF waves are scattered by ionospheric irregularities, giving rise to scintillations. These are rapid (less than 1 s) variations of signal amplitude (fading) and/or phase.

At higher frequencies (VHF, 30–300 MHz, and UHF, 300 MHz–3 GHz), both the refraction and scattering effects are less than at HF. This explains why radio signals in the UHF band, or even higher frequencies, are ideal for communication with satellites or for communications systems via satellites. Account has to be taken of both refraction and scattering effects for navigation methods using global positioning by satellites (GPS).

These effects also have to be accounted for when interpreting satellite altimetry results obtained using the radar principle. The necessary correction depends upon the electron density integrated along the ray path, which is termed the total electron content (TEC). Synthetic aperative radar (SAR) signals may also be affected by the ionosphere through which they have traveled.

Because the ionosphere is a birefringent medium, i.e., the O and X modes have slightly different refractive indices linked to their different velocities, the plane of polarization of a radio signal received from a satellite rotates. Due to the Faraday effect, such observations can be interpreted in terms of spatial variations of the ionospheric electron density. The principles of tomography are now being applied to study latitudinal variations of the ionosphere by observing radio signals from satellites in polar orbit.

Due to the Doppler effect, the frequency of the radio signal observed on the ground differs slightly from that transmitted from the orbiting satellite. Using two frequencies, the differential Doppler effect can be observed and used to derive some properties of the ionosphere.

The interaction between the ionospheric (or interplanetary) plasma and a satellite (or spacecraft) can lead to charging of the satellite/spacecraft. High-voltage systems aboard the satellite/spacecraft may discharge—or arc—to the surrounding plasma. The satellite/spacecraft is likely to sustain permanent damage in that event.

Such charging effects are much more prevalent at geostationary orbit (GEO, at a geocentric distance of 6.6 Earth radii) than in low Earth orbit (LEO, at heights typically between 300 and 1000 km). This is because at GEO the electron density is so low, usually between 10^6 and 10^7 m^{-3} (Lemaire and Gringauz 1998)

At altitudes above a few hundred kilometers, the ionospheric plasma is constrained to move along geomagnetic field lines from one hemisphere to the other. On field lines out to a geocentric distance of four Earth radii, the plasma density is relatively high, greater than 10^9 m^{-3}. Then the plasma density decreases dramatically, by a factor of 10 to 100, at the plasmapause. This field-aligned surface is usually on a field line crossing the equatorial plane at a geocentric distance between four and five Earth radii, termed $L = 4–5$.

16.8 THE NEUTRAL GAS ENVIRONMENT

Even though the atmosphere is very thin in LEO, it still exerts a drag force on a satellite and on other particles in orbit. The density of the Earth's atmosphere above 110 or 120 km altitude changes as the atmosphere expands in response to extra ultraviolet and X-radiation from the Sun with increasing solar activity.

Above this altitude, the atmosphere expands as it becomes hotter at times near solar maximum. This is evident from Figure 16.12. At 500 km altitude, the temperature is nearly 2000 K at solar maximum, but only ~800 K at solar minimum;

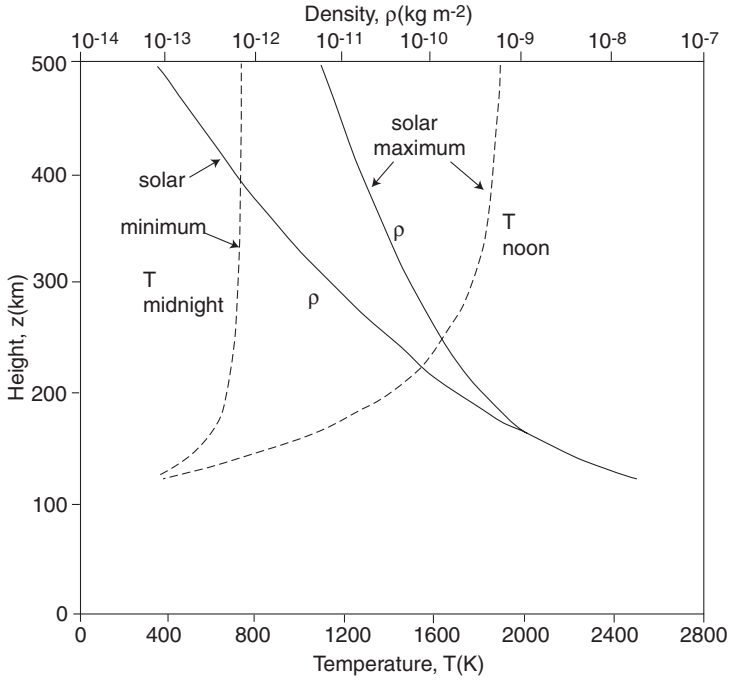


FIGURE 16.12 Variation with height (z) above 120 km of the atmospheric density (ρ), solid lines, upper scale, and the atmospheric temperature (T), dashed lines, lower scale, for both solar maximum and minimum conditions (based on CIRA 1965).

the neutral gas density is almost a hundred times larger at solar maximum than at solar minimum.

The drag force on a satellite

$$F_D = \frac{1}{2}\,\rho\, A\, C_D\, v^2$$

depends crucially on the neutral gas density ρ, the cross-sectional area of the object A presented as it moves at velocity v; C_D is the aerodynamic drag coefficient. For objects in LEO, v is almost 8 km s^{-1}.

As satellites or other objects in LEO reenter the Earth's atmosphere, their total energy decreases, their kinetic energy increases, and their orbital period τ decreases as time t progresses. Application of the law of the conservation of energy to satellites in a circular orbit (Hargreaves 1992) shows that

$$\frac{d\tau}{dt} = \frac{3\pi_\rho A C_D (z + R_E)}{M_s}$$

where z is the altitude, R_E is the radius of the Earth, and M_s is the mass of the satellite.

EARTH'S ENVIRONMENT AND SPACE

This equation was used in the early days of space research, the 1960s, to obtain the first information on $\rho(z)$, the neutral gas density as a function of height z, under different solar activity conditions.

At heights between 200 and 600 km the main atmospheric constituent is atomic oxygen (O). This is much more chemically reactive than molecular oxygen (O_2). It oxidizes the surface materials of satellites. For example, a front-silvered mirror becomes gray and its optical, thermal and mechanical properties are degraded (Tribble 1995).

In the ram direction the surface of a satellite glows. This is because the kinetic energy of the satellite is sufficient to excite the atmospheric atoms or molecules. These then radiate when they fall back to their ground state.

16.9 THE VACUUM ENVIRONMENT

At ~400 km altitude, the pressure is $\sim 10^{-6}$ Pa, 11 orders of magnitude less than at the Earth's surface. It is hard to make such a good vacuum on the Earth's surface, and so space is a good vacuum laboratory. The Space Shuttle wake shield facility utilizes this situation.

Under such conditions, lubricants between moving metal surfaces may not work well. Gas particles stuck on satellite surfaces outgas and may then stick on sensitive surfaces such as the lens of a telescope or thermal control surfaces, thereby reducing their performance.

The Sun's radiation is bright in space—there is no absorption of its ultraviolet radiation by ozone. This degrades the mechanical or thermal properties of some surface materials and may affect the performance of an optical instrument. Integrated across the spectrum, the solar constant of 1.368 kW m^{-2} and the terrestrial infrared radiation, heat a satellite. In the absence of gas to carry away heat by convection, excess heat can only be conducted through the satellite to a radiator that radiates energy away into the blackness of space (actually at 2.7 K, due to the cosmic microwave background radiation arising from the Big Bang).

16.10 THE RADIATION ENVIRONMENT

Charged particles with an energy much greater than the thermal energy of the plasma (less than 1 eV, equivalent to 11,600 K) constitute the radiation environment of space. Ions with an energy ~0.1 keV impinging upon the surface of a satellite eject atoms from the surface. Called sputtering, this process, over time, can remove a thin coating applied to part of a satellite's surface. Much higher energy charged particles can penetrate solar cells and microelectronic chips inside a satellite, causing them to change their state or even destroying them completely. Holmes-Siedle and Adams (1994) discuss many details. The charged-particle environment of near-Earth space, especially strong in the vicinity of the South Atlantic geomagnetic anomaly (see Figure 16.13), is hazardous to astronauts and cosmonauts carrying out extravehicular activities (EVAs).

Here some facts about the radiation environment are given. Deeper explanations are given in Part 6. The most energetic charged particles (electrons of greater than 1 GeV), galactic cosmic rays, come from outside the solar system. When the Sun

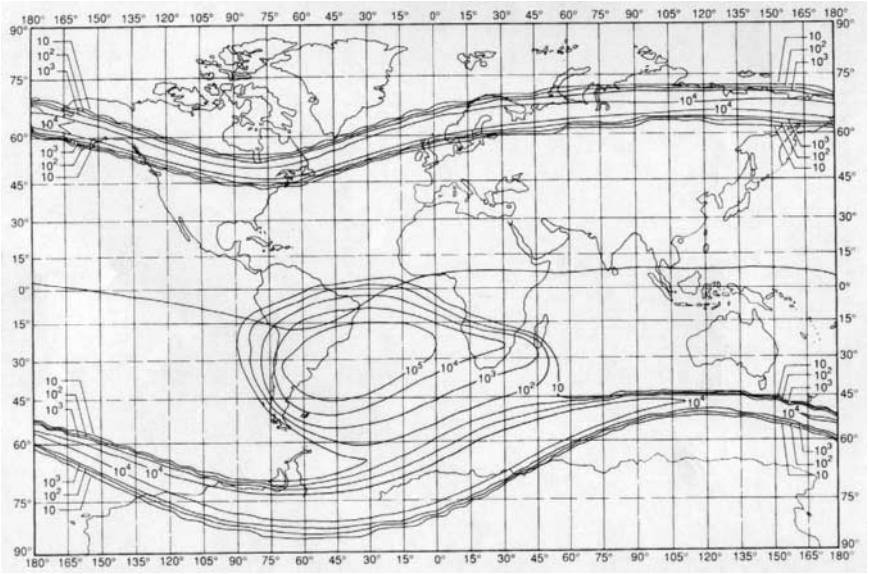


FIGURE 16.13 Contours of electron fluxes at >1 MeV, in units of $cm^{-2}\ s^{-1}$, at 500 km altitude, showing effects due to the outer Van Allen radiation belt and the South Atlantic geomagnetic anomaly (from Holmes-Siedle and Adams 1994).

is especially active, it can emit bursts of charged particles, protons, helium ions, and electrons (of greater than 1 MeV). These enter the magnetosphere at high latitudes. The usual interaction between the solar wind and the magnetosphere causes it to be populated by energetic ions and electrons (of greater than 1 keV); charged particles trapped on geomagnetic field lines are termed the Van Allen radiation belts.

An electron with an energy of some tens of keV, typical of the Van Allen belt population, undergoes three types of motion. These are indicated in Figure 16.14.

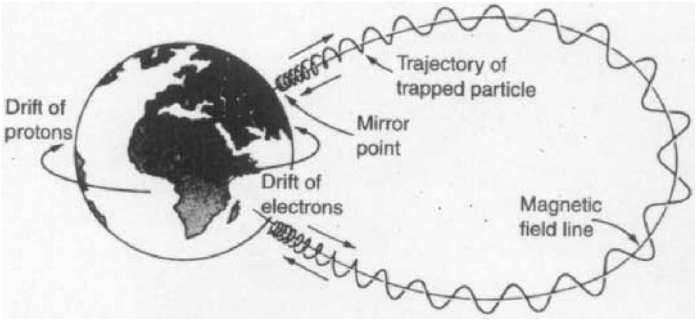


FIGURE 16.14 The motion of a charged particle trapped in the Earth's magnetic field (from Gombosi 1998).

EARTH'S ENVIRONMENT AND SPACE

The most rapid component of the electron's motion is gyration about the geomagnetic field line, on a time scale of less than 1 ms. The electron bounces from one hemisphere to the other on a time scale of seconds. The third component of motion, with a time scale of hours, is a longitudinal drift around the Earth, with electrons drifting eastward and protons westward. This constitutes a westward-directed ring current which reduces the geomagnetic field at the Earth's surface on the equator (nominally 31,000 nT) by some tens or hundreds of nT.

In meridional cross-section, contours of the fluxes of electrons (above 0.5 MeV), in $cm^{-2}\ s^{-1}$, are illustrated in the lower part of Figure 16.15. The inner and outer belts, separated by a slot region, are clearly seen. The contours come closest to the

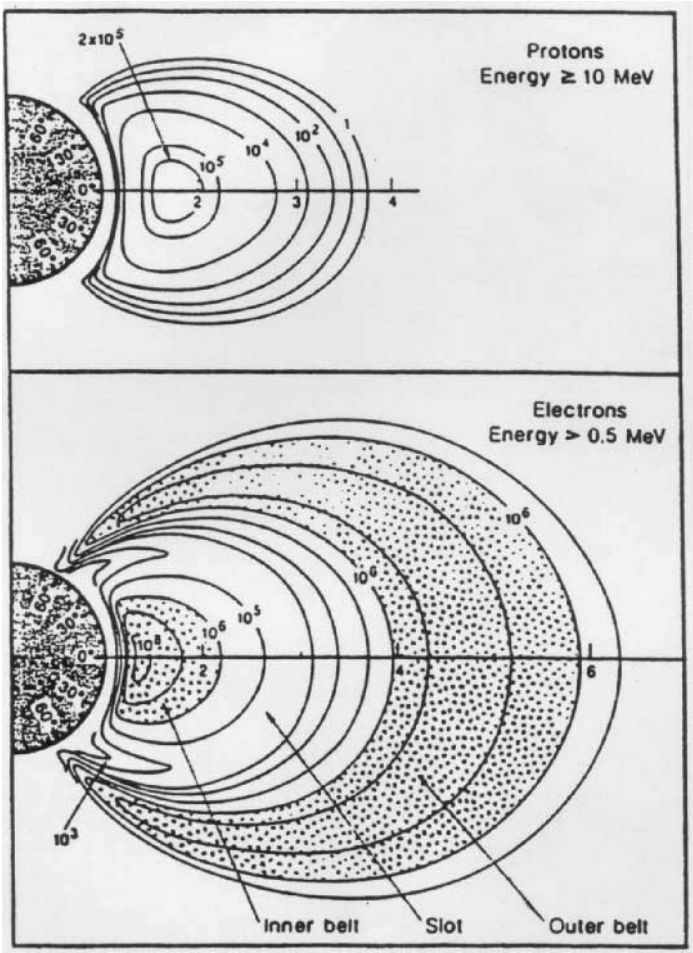


FIGURE 16.15 Fluxes of protons and electrons in the Van Allen radiation belts, in meridional cross-section.

Earth's surface near 60° geomagnetic latitude, to account for the high flux values at all longitudes evident in Figure 16.13.

The upper part of Figure 16.15 shows the proton fluxes, in units of cm^{-2} s^{-1}. While it is not apparent from these figures, the Van Allen belts are electrically neutral overall. The mechanism believed to be responsible for the proton belt is the decay of neutrons produced when cosmic rays hit the top of the atmosphere. This is termed the cosmic ray albedo neutron decay (CRAND) mechanism.

At any particular time, the charged particle fluxes may vary from these typical values by up to two orders of magnitude. In other words, the radiation belts, especially the outer one, are very dynamic in their response, on a time scale of minutes, to solar wind changes.

16.11 THE MICROMETEOROID AND SPACE DEBRIS ENVIRONMENT

The relative velocities, and hence the kinetic energy involved, in a collision between an orbiting satellite and a piece of debris are so large that a small particle (more than a millimeter across) can seriously damage the structure of a satellite. Particles with sizes larger than ~0.1 m can be observed from the ground, using either radars or optical telescopes.

The subject is of such importance that it warrants extensive discussion in Part 7.

REFERENCES

COSPAR International Reference Atmosphere (CIRA). 1965. North-Holland, Amsterdam.

DeWitt, R. N., Duston, D., and Hyder, A. K., eds. 1993. *The Behavior of Systems in the Space Environment,* Kluwer Academic Publishers, Dordrecht.

Dyer, C. S., Truscott, P. R., Sanderson, C., Watson, C., Peerless, C. L., Knight, P., and Mugford, R. 2000. "Radiation Environment Measurements from CREAM & CREDO during the Approach to Solar Maximum," *IEEE Transactions on Nuclear Science,* vol. 47, no. 1.

Fortescue, P. W. and Stark, J. P. W. eds. 1991. *Spacecraft Systems Engineering,* John Wiley & Sons, Chichester.

Freeman, J. W. 2001. *Storms in Space,* Cambridge University Press, Cambridge.

Gombosi, T. J. 1998. *Physics of the Space Environment,* Cambridge University Press, Cambridge.

Hargreaves, J. K. 1992. *The Solar-Terrestrial Environment,* Cambridge University Press, Cambridge.

Hastings, D. and Garrett, H. 1996. *Spacecraft-Environment Interactions,* Cambridge University Press, Cambridge.

Holmes-Siedle, A. and Adams, L. 1994. *Handbook of Radiation Effects,* Oxford University Press, Oxford.

Kelley, M. C. 1989. *The Earth's Ionosphere: Plasma Physics and Electrodynamics,* Academic Press, San Diego.

Kivelson, M. G and Russell, C. T. 1995. *An Introduction to Space Physics,* Cambridge University Press, Cambridge.

Lemaire, J. F. and Gringauz, K. I. 1998. *The Earth's Plasmasphere,* Cambridge University Press, Cambridge.

EARTH'S ENVIRONMENT AND SPACE

Mitchell, D. G. 1994. "Space Systems Engineering," in *Fundamentals of Space Systems,* ed. V. L. Pisacane and R. C. Moore, Oxford University Press, Oxford.

Rishbeth, H. and Garriott, O. K. 1969. *Introduction to Ionospheric Physics,* Academic Press, New York.

Schunk, R. W. and Nagy, A. F. 2000. *Ionospheres: Physics, Plasma Physics, and Chemistry,* Cambridge University Press, Cambridge.

Skrivanek, R. A., ed. 1994. *Contemporary Models of the Earth's Environment,* AIAA, Washington, DC.

Song, P., Singer, H. J., and Siscoe, G. L., eds. 2001. *Space Weather,* AGU, Washington, DC.

Suess, S. T. and Tsurutani, B. T., eds. 1998. *From the Sun: Auroras, Magnetic Storms, Solar Flares, Cosmic Rays,* AGU, Washington, DC.

Tascione, T. F. 1994. *Introduction to the Space Environment,* Krieger, Malabar, FL.

Tribble, A. C. 1995. *The Space Environment: Implications for Spacecraft Design,* Princeton University Press, Princeton, NJ.

Walt, M. 1994. *Introduction to Geomagnetically Trapped Radiation,* Cambridge University Press, Cambridge.

Wertz, J. R. and Larson, W. J. 1999. *Space Mission Analysis and Design,* Kluwer Academic Publishers, Dordrecht.

PART 3

THE SOLAR SYSTEM

Michael J. Rycroft

16.12 PHYSICAL PROPERTIES OF THE PLANETS

In this subsection the physical characteristics and properties of the bodies in the solar system are reviewed. Moving in almost circular orbits about the Sun, under the action of its gravitational force, are the planets and asteroids. Moving in elliptical orbits about the Sun are Pluto, comets from the Kuiper belt, and more distant comets from the Oort cloud. Asteroids and comets are remnants of the early solar system (Beatty et al. 1999; Lewis 1997; Lodders and Fegley 1998; Mendell 1999).

All the planets except Pluto orbit the Sun in the same plane, the ecliptic plane. Starting from closest to the Sun, the planets are Mercury, Venus, Earth, and Mars (rocky planets, termed the terrestrial planets), the asteroids (rocky material which did not form a planet) then Jupiter, Saturn, Uranus, and Neptune (large gaseous planets). The different planetary orbits are shown, together with their symbols, in Figure 16.16 (from Lewis 1997).

Table 16.1 summarizes, generally to two significant figures, numerical information about the Sun, the planets, and their motions. Information on one planetary satellite, the one orbiting the Earth— namely the Moon—is included. The distance

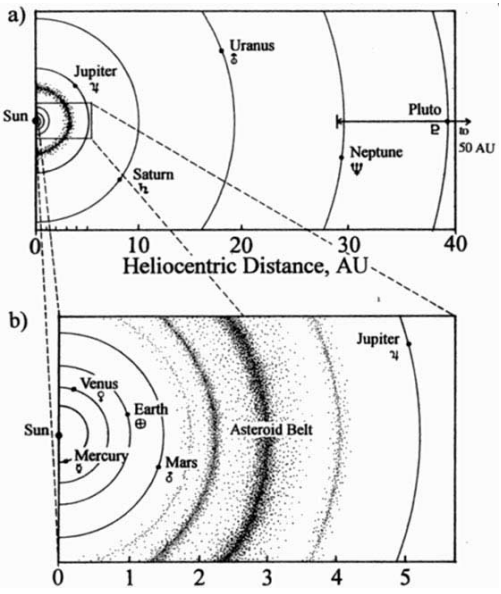


FIGURE 16.16 Diagram, to scale, of (a) the orbits of the outer planets and (b) the inner planets and asteroids (from Lewis 1997).

TABLE 16.1 Numerical Information on Bodies in the Solar System

Object	Distance from Sun (AU)	Mass (kg)	Radius (km)	Density (10^3 kg m^{-3})	Orbital eccentricity	Orbital period (year)	Rotation period (day)	Surface gravity (g)
Sun	—	2.0×10^{30}	7.0×10^5	1.4	—	—	25–32	28
Mercury	0.39	3.3×10^{23}	2.4×10^3	5.4	0.21	0.24	59	0.39
Venus	0.72	4.9×10^{24}	6.0×10^3	5.2	0.007	0.62	243	0.90
Earth	1.00	6.0×10^{24}	6.4×10^3	5.5	0.017	1.00	1.00	1.00
Moon	1.00	7.4×10^{22}	1.7×10^3	3.3	0.055	1.00	28	0.17
Mars	1.52	6.4×10^{23}	3.4×10^3	4.0	0.093	1.88	1.03	0.38
Asteroids	2.2–3.3	Total $\sim 10^{22}$	>1	few	~0.15	3.3–6	Fraction	<0.001
Jupiter	5.2	1.9×10^{27}	7.1×10^4	1.3	0.048	11.9	0.41	2.6
Saturn	9.5	5.7×10^{26}	6.0×10^4	0.69	0.056	29	0.43	1.1
Uranus	19	8.7×10^{25}	2.6×10^4	1.3	0.047	84	0.70	0.9
Neptune	30	1.0×10^{26}	2.5×10^4	1.6	0.009	165	0.74	1.1
Pluto	39	10^{22}	10^3	2	0.25	248	6.4	0.07
Comets	$1–5 \times 10^4$	Total $\sim 10^{27}$	<10	<1	0.5–1.0	3–10		Negligible

Source: Lewis 1997; Lodders and Fegley 1998; Mendell 1999.

at which a planet orbits the Sun is given in astronomical units (1 AU, the Sun-Earth distance, is almost 1.5×10^{11} m). The mass of the body is given in kg, its radius in km, and its density in kgm^{-3}. It is evident that the density of the inner (terrestrial planets) is several times greater than that of the outer (giant gaseous planets). The eccentricity of the orbit shows by how much the orbit departs from a circular orbit; the eccentricity is zero for an exactly circular orbit. The period for one orbit around the Sun is given in years. The period for one revolution of the planet around its axis of rotation is shown in days. It is remarkable that the larger outer planets rotate much more rapidly than the smaller inner (terrestrial) planets. Finally, the acceleration due to gravity on the surface of the planet is given in g; for the Earth, $1\ g = 9.8\ ms^{-2}$.

All these values have to be known to great accuracy in order to navigate a spacecraft to rendezvous with a planet or asteroid.

Table 16.2 highlights the major satellites of the planets in the solar system, and ring systems. For each stated satellite its radius is given in km.

New planetary satellites and ring systems discovered by instruments aboard spacecraft are considered in the next subsection.

16.13 SPACE AGE DISCOVERIES

One of the brightest planets in the sky, Mercury, was imaged directly by Mariner 10 in 1974 and 1975. Half its surface has been photographed. Like the Moon, it is heavily cratered. Because its spin axis is very close to perpendicular to its orbital plane, there are no seasons on Mercury. Its surface temperature is low (~100 K) at night, but high (~700 K) during the day (Lodders and Fegley 1998). The atmosphere of Mercury is very thin, the surface pressure being less than 10^{-9} hPa. The most abundant species is believed to be argon, with a number density less than $3 \times 10^{13}\ m^{-3}$ (Beatty et al. 1999). There is a weak magnetic field near its surface, whose properties are tabulated in Table 16.3. As discussed in Part 6, the magnetic dipole moment is given as a multiple of the Earth's magnetic moment, the surface field at the magnetic equator is given in nanoteslas (nT), the angle between the rotational and magnetic axes is shown in degrees (°), and a typical distance from the center of the planet to its magnetopause is given in units of planetary radius. The magnetopause is the boundary between the planet's magnetic field and the interplanetary magnetic field carried away from the Sun by the solar wind plasma.

TABLE 16.2 Planetary Satellites Known before Direct Space Observation

Planet	Satellite	Radius (km)
Mercury		
Venus		
Earth	Moon (see Part 4)	1.7×10^3
Mars	Phobos, Deimos	~10, ~6
Jupiter	16, including Io, Europa and Ganymede	1.8×10^3, 1.6×10^3, $2{,}6 \times 10^3$
Saturn	18, largest Titan; rings	2.6×10^3
Uranus	5, largest Titania; rings	790
Neptune	2, larger Triton	1.4×10^3

EARTH'S ENVIRONMENT AND SPACE

TABLE 16.3 Properties of Planetary Magnetic Fields

Planet	Magnetic dipole moment	Equatorial surface field (nT)	Angle between axes (°)	Magnetopause distance (planetary radii)
Mercury	7×10^{-4}	300	14	1.5
Venus	$<10^{-4}$	—	—	—
Earth	1.00	31,000	11	10
Mars	$<2 \times 10^{-4}$	<30	—	—
Jupiter	2×10^{4}	430,000	10	80
Saturn	600	22,000	<1	20
Uranus	50	23,000	60	20
Neptune	25	14,000	47	25

Source: Based on Beatty et al. 1999.

The main difference between Venus and Earth is the dryness of Venus. Its atmosphere of 96.5% carbon dioxide (CO_2) has a surface pressure some 90 times greater than Earth's. Since CO_2 is a very effective greenhouse gas, the surface temperature is about 750 K (Beatty et al. 1999; Marov and Grinspoon 1998). The thick clouds of aqueous sulfuric acid droplets at heights between 45 and 70 km prevent the Venusian surface from being seen in the visible part of the spectrum. However, the atmosphere is transparent to microwaves, and its surface has been impressively mapped by the radar instrument of the Magellan mission in the early 1990s. Both impact craters and volcanoes are very evident. Optical images of flat basaltic rocks on the surface of Venus were obtained by the Venera 14 lander in 1982. Venus has a significant ionosphere. No planetary magnetic field has been detected by magnetometers aboard spacecraft in the vicinity of Venus.

The origin of the Earth and the Moon is discussed by Canup and Righter (2000). The rather well-known properties of the Earth and the Moon are summarized, in the context of solar system studies, by Beatty et al. (1999) and Lodders and Fegley (1998). Detailed information on the Moon, as an object worthy of detailed study by both robotic and crewed space missions, is given in Part 4.

Mars, the red planet, may have harbored forms of life in past eons and may still be an environment where life exists. Because of the great interest in the search for life on Mars using space missions, the planet warrants a separate discussion in Part 5. Excellent images of Mars have been taken by the Hubble space telescope in orbit around the Earth. In the winter and springtime on Mars, carbon dioxide frost and a (water) ice cap are seen (Fischer and Duerbeck 1998). The atmospheric pressure on the Martian surface is only 6 hPa. The magnetic field of Mars is very small (see Table 16.3).

Images of some asteroids have been taken by the Hubble space telescope (Fischer and Duerbeck 1998) and by the Galileo spacecraft. The Near Earth Asteroid Rendezvous (NEAR Shoemaker) mission has obtained images of a few asteroids. At the end of its mission in February 2001 it landed on the asteroid Eros. The properties of asteroids are summarized by Kowal (1996).

The giant planets are gaseous, mainly hydrogen and helium, but with methane and ammonia at the 0.1% level by volume. If the mass of Jupiter were 13 times its actual mass, it would be classified as a brown dwarf (Beatty et al., 1999). It is therefore more of a "failed star" than a planet. Jupiter radiates 1.7 times the amount of solar radiation that it receives (Lodders and Fegley 1998).

Voyager 1 and 2 images of Jupiter's surface exhibit zonal bands of clouds, signifying strong jet streams. Weather systems are also clearly evident, the largest of which is the very long-lived Great Red Spot. Bursts of lightning, as well as whistlers, have been observed on Jupiter. Rings of light around the two magnetic poles of Jupiter, signifying its auroral zones, have been seen with the Hubble space telescope (Fischer and Duerbeck 1998) and by ground-based infrared telescopes. Jupiter's magnetosphere is enormous in scale.

In July 1994, pieces of the comet Shoemaker-Levy 9 hit Jupiter, causing material from below the clouds to rise up (Spencer and Mitton 1995). Comets are believed to be "dirty snowballs," carbon-rich ices left over from the early days of the solar system.

When the Voyager 1 spacecraft reached Jupiter in 1979, Jupiter's ring system was discovered in the vicinity of the four innermost satellites (Lodders and Fegley 1998). Of the next four satellites (the Galilean satellites), Io is known to be the most volcanically active body in the solar system. Both the Voyager and Galileo spacecraft have observed sulfur emitted by volcanoes. Europa and Ganymede have water ice on their surface and oxygen and ozone trapped in the ice. Spacecraft

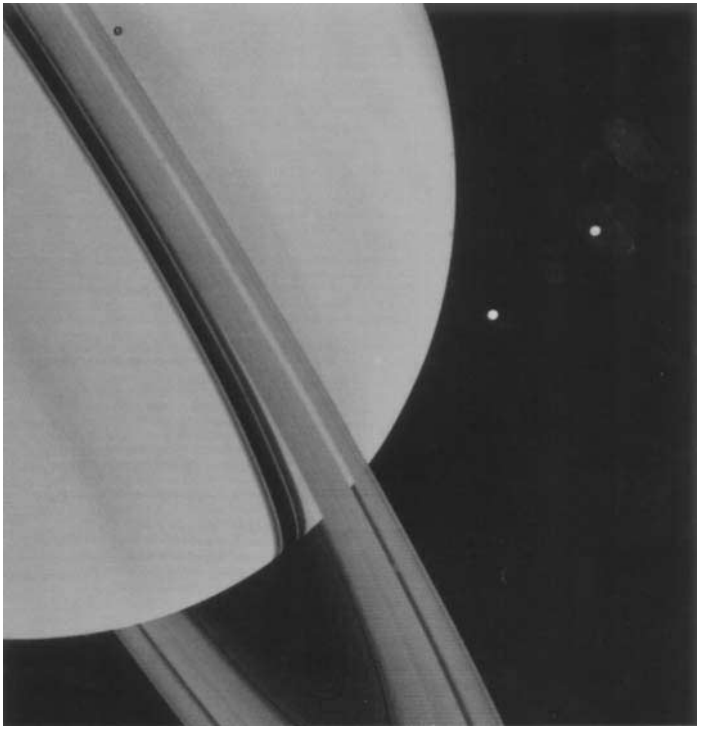


FIGURE 16.17 Voyager 1 image of Saturn taken in November 1980 of Saturn, its ring system, and its moons Tethys (above) and Dione (below) (NASA image). The shadows cast by the ring system and by Tethys onto the tops of the clouds account for the dark features on the planet's surface in this image.

EARTH'S ENVIRONMENT AND SPACE

magnetometer observations near Europa indicate the presence of conducting oceans below the ice. Such an environment may be conducive to life (Beatty et al. 1999).

Saturn is a somewhat smaller version of Jupiter; its surface exhibits a similar pattern of zonal winds and clouds. Ammonia and phosphine appear as absorption features in Saturn's infrared spectrum. Saturn's rings are the planet's most distinctive and dramatic feature (Figure 16.17). They are thought to be formed of ice and rocks, plus some carbonaceous material (Lodders and Fegley 1998). The particles forming the rings range in size from submicron ($<10^{-6}$ m) to some meters.

Both Voyager spacecraft visited Jupiter and Saturn, as is evident from their trajectories shown in Figure 16.18 (Rothery 1992). The Cassini mission was

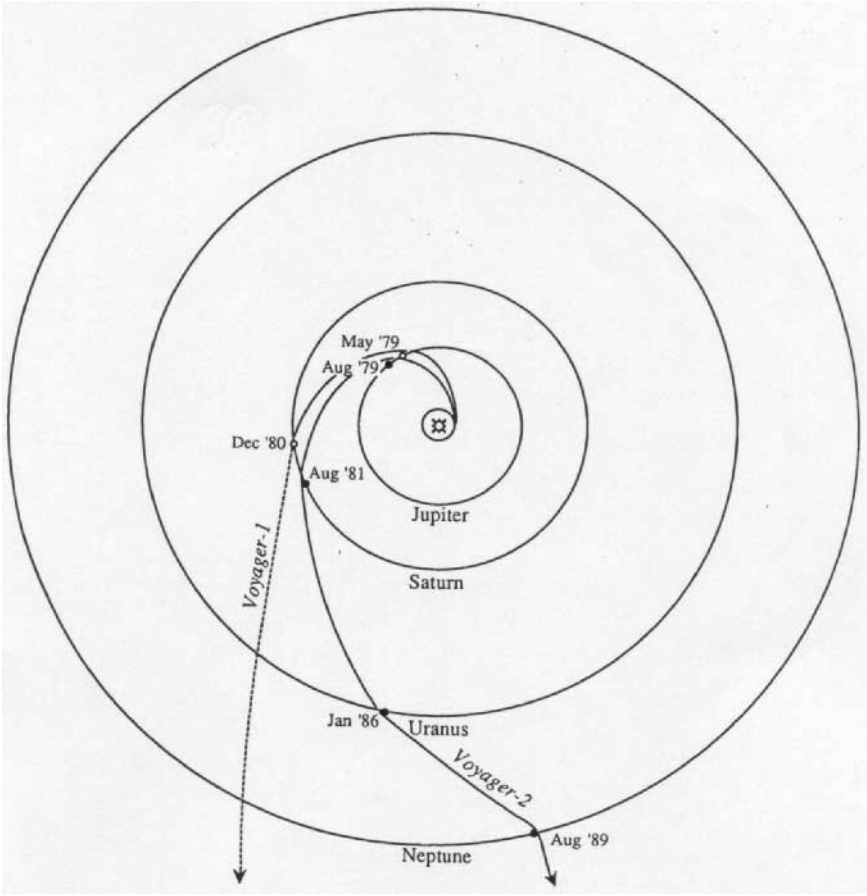


FIGURE 16.18 Trajectories of the two Voyager spacecraft. Voyager 1's encounter with Saturn took it out of the ecliptic plane. After Neptune, Voyager 2 continued on a course heading below the plane of the solar system (from Rothery 1992).

launched in 1997 to arrive at Saturn in 2004. It carries the Huyghens probe, which will parachute down to the surface of Titan (which is a little larger than the planet Mercury). The Sun's light reflected by Titan shows absorption features in the red, which are ascribed to methane present at the level of up to 10%. The dominant gases in Titan's atmosphere are inferred to be nitrogen (N_2) and argon (A).

The Voyager 2 spacecraft journeyed on to Uranus, reaching it in January 1986 (Miner 1998). Models of the interior that are consistent with Voyager 2 findings have been developed. Its characteristic blue-green color is due to methane (CH_4) in the cloudy atmosphere. The effective temperature of Uranus is 59 K.

As indicated in Figure 16.19 (from Miner 1998), the unique feature of the planet Uranus is that its axis of rotation is almost in the ecliptic plane. Uranus rotates in a retrograde (East to West) direction. The best-fit magnetic dipole axis makes a large angle (60°) with the rotational axis. The magnetic field is probably generated by dynamo activity in a partially fluid core extending out to 0.7 radius of Uranus.

The rings of Uranus were observed in detail for the first time in January 1986. Ten satellites of Uranus were also discovered in Voyager 2 images (Miner 1998); their radii vary from 10 to 80 km.

Observed by Voyager 2 and the Hubble space telescope, Neptune is not very dissimilar to Uranus. It has a cloudy atmosphere, but at the cloud tops water is not present because it is too cold (Lodders and Fegley 1998). Dynamo action in a conducting interior produces a magnetic field, which is offset from Neptune's center by 0.55 radius of Neptune and tilted at 47° to the rotational axis. Six small satellites were discovered by Voyager 2, which imaged its rings in August 1989 (Lodders and Fegley 1998).

Neptune's largest satellite, Triton, was also observed then. Its average surface temperature is only 38 K. Its nitrogen-rich atmosphere has a surface pressure of 1.6 Pa.

Pluto and its satellite Charon (whose radius is 0.52 radius of Pluto) are icy worlds. Pluto has predominantly nitrogen ice with a surface temperature of 40 K,

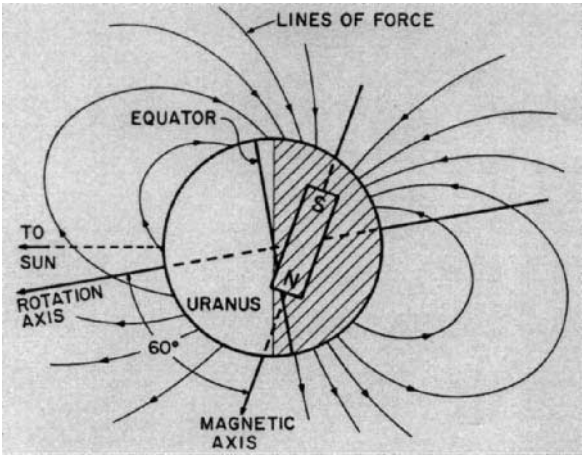


FIGURE 16.19 The simplest representation of the magnetic field of Uranus is a magnetic dipole tilted at 60° to the rotation axis and offset by 0.3 Uranian radii toward the dark North polar region (from Miner 1998).

whereas Charon's surface is of water ice, also at a temperature of 40 K (Lodders and Fegley 1998; Stern and Mitton 1998).

Comets have a nucleus, coma, dust tail, and plasma tail. The Giotto probe visited Halley's comet in 1986, taking the first close-up image of a comet's nucleus subliming. The nucleus is a 10 km sized conglomerate of ice and rocky dust particles (Lewis 1997; Lodders and Fegley 1998). Several other instrumented spacecraft intercepted the orbit of Halley's comet in 1986, and other comets have been investigated directly by other space missions.

REFERENCES

Beatty, J. K., Collins, C., and Chaikin, A., eds. 1999. *The New Solar System,* Sky, Cambridge, MA.

Canup, R. M. and Righter, K., eds. 2000. *Origin of the Earth and Moon,* University of Arizona Press, Tucson, AZ.

Fischer, D. and Duerbeck, H. 1998. *Hubble Revisited,* Copernicus, New York.

Kowal, C. T. 1996. *Asteroids: Their Nature and Utilization,* Wiley-Praxis, Chichester.

Lewis, J. S. 1997. *Physics and Chemistry of the Solar System,* Academic Press, San Diego.

Lodders, K. and Fegley, B., Jr. 1998. *The Planetary Scientist's Companion,* Oxford University Press, New York.

Marov, M. Y. and Grinspoon, D. H. 1998. *The Planet Venus,* Yale University Press, New Haven, CT.

Mendell, W. 1999. In *Keys to Space: An Interdisciplinary Approach to Space Studies,* ed. A. Houston and M. Rycroft, McGraw-Hill, New York.

Miner, E. D. 1998. *Uranus: the Planet, Rings and Satellites,* Wiley-Praxis, Chichester.

Rothery, D. A. 1992. *Satellites of the Outer Planets: Worlds in their Own Right,* Clarendon Press, Oxford.

Spencer, J. R. and Mitton, J. 1995. *The Great Comet Crash: The Collision of Comet Shoemaker-Levy 9 and Jupiter,* Cambridge University Press, Cambridge.

Stern, A. and Mitton, J. 1998. *Pluto and Charon: Ice Worlds on the Ragged Edge of the Solar System,* John Wiley & Sons, Chichester.

PART 4

THE MOON

Peter Eckart

16.14 ORIGIN OF THE MOON

The origin of the Moon has been pondered, modeled, studied, and endlessly debated, but still the mystery has not been ultimately solved. There are three classical theories: the coaccretion model, the fission model, and the intact capture model.

The coaccretion model assumes that Moon and Earth developed close to each other as separate bodies. This model is based on the dominant model for the origin of the solar system, suggesting that it condensed out of a cloud of hot gas, called the solar nebula. In such a model the planets are small pieces of the condensed material left over after star formation. Thus, our Moon and the moons of the other planets would be similar to the planets that form around a star, a moon being the left-over bits of a planet formation. One of the major problems of this model is the fact that the densities of Moon and Earth are so different. This appears unlikely if they were formed at the same time, in the same region, and from the same original matter. Also, there is a problem in accounting for the angular momentum of the Earth and Moon, explaining why material around Earth stayed in orbit rather than falling on Earth.

The fission model assumes that the Moon was initially a part of Earth. This theory is based upon the assumption that initially the original planet had a high rotational rate, and led to the suggestion that the Moon was flung off Earth very early in Earth's history. According to calculations, this offspring should have a mass of about 1–2% of the original body's mass, and the ratio of the diameters of the resulting spherical bodies should be about ¼—these are quite exactly the ratios of Earth and Moon. However, there are problems explaining how the Earth was spinning fast enough to fling off a blob of material from which the Moon formed. Because no explanation for this aspect could be found, the fission model had to be rejected.

The intact capture model assumes that the Moon's and Earth's flight paths crossed early in the history of the solar system, resulting in the Moon being captured by Earth. But, according to this model, the minimum distance of Earth and Moon would have to be less than 60,000 km. Also, the fact that Moon's orbit around Earth is almost circular and the symmetry of the Earth–Moon system make this theory appear very unlikely. Moreover, the probability of capturing a fully formed Moon is considered extremely tiny.

Thus, none of the three classical models of lunar origin are particularly successful at explaining all the properties of the Moon and all the orbital characteristics of the Earth–Moon system. Therefore, the giant impact (or Big Whack) model was created, which combines essential aspects of the three classical models. The idea that the Moon formed as a consequence of a giant impact of a planetoid on Earth, originally proposed in 1976, is currently accepted wisdom in planetary science. It assumes that during the last phase of Earth's accretion, after its core had formed and while it was still molten, an object the size of Mars smashed into this proto-Earth at an oblique angle (capture model). The resulting monumental explosion deposited large quantities of material in Earth orbit, and the primitive Moon formed

from this material (fission model). Although the giant impact model can explain most of the current phenomena of the Earth–Moon system, including the geochemical properties of the Moon and the orbital dynamics of the Earth–Moon system (coaccretion model), it is far from certain that the problem of lunar origin has been solved, and more data will be required either to prove or to reject it (Spudis, 1996).

16.15 ORBITAL PARAMETERS

The Earth–Moon-System

Viewed from above the North pole of the Earth, the Moon travels counterclockwise in a slightly elliptical path around the Earth. The eccentricity of this orbit oscillates between 0.044 and 0.067 (mean value 0.0549). Measured from center to center, the Earth–Moon distance ranges from 356,410 km at the Moon's closest approach (perigee) to a maximum value of 406,697 km at furthest separation (apogee). The mean value of the semimajor axis is 384,400 km. As the Moon orbits the Earth, it shows different phases, or lighting conditions of its surface, at various times of the month. The Moon's orbit around the Earth is synchronous, i.e., its rotational period is the same as its period of revolution around the Earth. Therefore, it always keeps the same hemisphere, the near side, pointing towards the Earth. Conversely, the other hemisphere, the far side, is always turned away from the Earth (Heiken et al. 1991).

When the Earth–Moon system is orbiting the Sun and the Moon is orbiting the Earth, the time interval between two successive new moon phases is 29.53059 days; this is a synodic month or a lunar day. For one complete orbit around the Earth, the Moon needs 27.32166 days, a sidereal month. During these 27.3 days, the center of gravity of the Earth–Moon system advances 27° on its way around the Sun. To be in alignment again with the Earth and the Sun, the Moon has to move another 27° on its orbit around Earth. A summary of lunar characteristics is given in Table 16.4 (Spudis 1996).

Lunar Librations

The plane of the lunar orbit is inclined to the plane of the ecliptic by 5°9′, and the Moon's polar axis is inclined 1°32′ to the plane of the ecliptic. This results in an inclination of 6°41′ of the Moon's rotation axis relative to a line perpendicular to the lunar orbital plane (see Figure 16.20). This geometric relationship causes what is called the latitudinal libration. Due to this libration it is possible to see an additional 6.7° of the lunar surface at the North and South poles, over a two-week interval. The Moon's longitudinal libration is caused by the slight noncircularity of its orbit around the Earth. The Moon revolves at a constant rate, so that after one quarter of the rotational period it has rotated through 90°. At the same time, however, it has moved through 97° of its orbit, so that seen from Earth the Moon's central meridian would appear to be displaced by about 7°. Because of this libration, it is possible to see an additional 7° along both eastern and western limbs of the lunar disc. Finally, there is a daily or diurnal libration, a parallax effect, due to the axial rotation of the Earth itself. The position of an observer on the Earth's equator will move laterally about 13,000 km during a 12-hour period. This gives a libration

TABLE 16.4 Lunar Characteristics

Property	Moon	Earth	Moon-to-Earth ratio
Mass	7.353×10^{22} kg	5.976×10^{24} kg	1/81 (0.0123)
Mean radius	1,738 km	6,371 km	1/3.7 (0.2727)
Surface area	3.80×10^{7} km^2	5.10×10^{8} km^2	1/13.5 (0.07429)
Flattening[a]	0.0005	0.0034	1/7.5 (0.147)
Mean density	3.34 g/cm^3	5.517 g/cm^3	1/1.65 (0.605)
Gravity at equator	1.62 m/s^2	9.81 m/s^2	1/6 (0.165)
Escape velocity at equator	2.38 km/s	11.2 km/s	1/4.7 (0.2125)

Parameter	Value
Mean distance from Earth	384,400 km
Mean eccentricity	0.0549
Mean inclination of the ecliptic	5°09′
Mean inclination to the lunar equator	6°41′
Length of sideric month	27.3217 days
Length of synodic month	29.5306 days

[a]Flattening = (equatorial diameter − polar diameter)/equatorial diameter.
Source: Heiken et al. 1991.

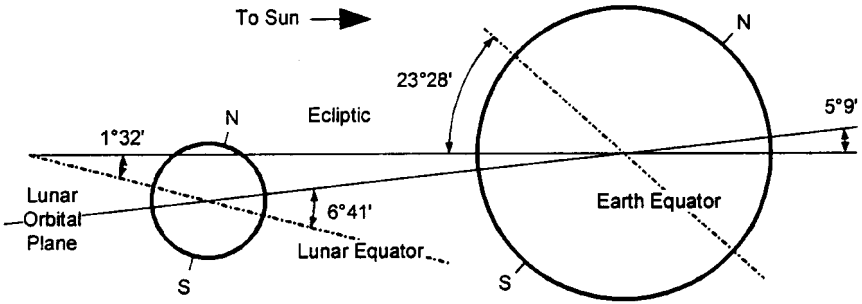


FIGURE 16.20 Schematic of lunar orbital characteristics (angular relationships have been exaggerated) (Heiken et al. 1991).

of 1° at the western edge of the Moon upon rising and 1° at the eastern edge when it sets. Because of these combined librations, a total of about 60% of the lunar surface can be seen from Earth. As a result of the lunar librations, all near side lunar sites can expect continuous communication with Earth; a site located near the terminator will see the Earth set below the lunar horizon, thus allowing direct communication only part of the time. Sites on the far side of the Moon cannot communicate with Earth without the aid of relay satellites or relay antennas on the lunar surface (Heiken et al. 1991).

Lunar Diurnal Cycle

A lunar day lasts about 29.5 days (see above). Therefore, at the equator and at all latitudes, except very close to the poles, a 14-day-night-14-day-light cycle results.

Due to the Moon's orbital parameters (see above), the Sun's elevation at the poles is only $\pm 1°32'$, hence a ½-year-light-½-year-night-cycle results. As a result, some craters in the South polar region are permanently shaded (estimated 2% of the lunar surface) (Heiken et al. 1991).

16.16 LUNAR GEOGRAPHY

The main features of the lunar surface are the highlands, maria, and craters. The light-colored highlands, or terrae, are the original primordial crust of the Moon. They have been continuously penetrated by meteoroids which created their very rough landscape. The dark maria or lowlands, which were formed by immense lava flows, are the most obvious and smoothest of all lunar features (see Figures 16.21 and 16.22). They are easily visible with the naked eye. Most lunar craters have a circular outer shape, sunken floors, a central peak, and rims rising to modest heights above the outer surface. It is generally accepted that meteorite impact is the predominant crater-forming mechanism. Craters appear all over the Moon, but there are fewer in number the younger they are (Heiken et al. 1991; Moore 1981).

In order to be able to locate particular points on the lunar surface, lunar maps follow Earth convention, with North at the top and the prime meridian referenced to crater Mösting A on the Moon's near side. Before 1974, lunar longitude was expressed in terms of 0°–180° East (+) and 0°–180° West (−) of the prime meridian. This convention is also used here because it is used in most literature. After 1974, a lunar longitude system of 360°, proceeding counterclockwise (increasing to the East) from the prime meridian, was adopted by international agreement for future cartography (Heiken et al. 1991).

16.17 LUNAR GEOLOGY

Like the Earth, the Moon is not a uniform, homogeneous planet. It consists of different rocks, formed in different ways at different times. To learn about the Moon

1 Oceanus Procellarum
2 Mare Imbrium
3 Mare Cognitum
4 Mare Humorum
5 Mare Nubium
6 Mare Frigoris
7 Mare Serenitatis
8 Mare Vaporum
9 Mare Tranquilitatis
10 Mare Nectaris
11 Mare Humboldtianum
12 Mare Crisium
13 Mare Fecunditatis

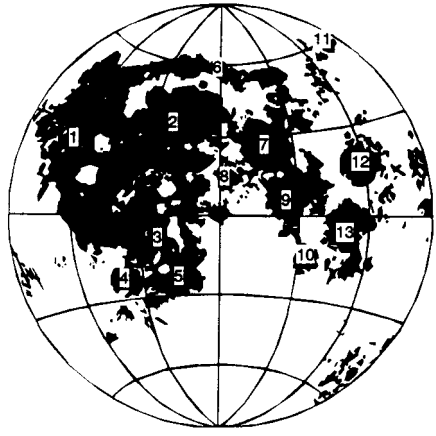


FIGURE 16.21 The maria of the lunar near side (Heiken et al. 1991).

14 Mare Marginis
15 Mare Smythii
16 Mare Australe
17 Mare Moscoviense
18 Mare Ingenii
19 Mare Orientale

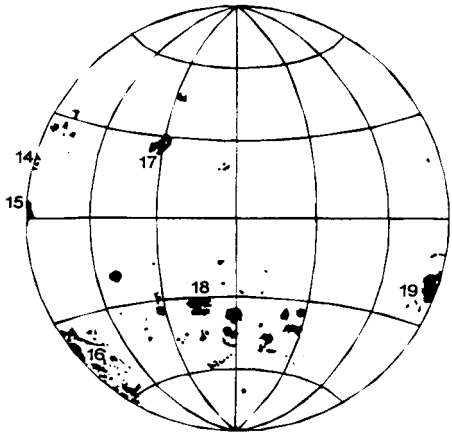


FIGURE 16.22 The maria of the lunar far side (Heiken et al. 1991).

and its history, one has to learn something about lunar rocks—what they are made of and how they are classified. Some of the Moon's rocks are familiar. For example, the basalt lava flows that cover the dark maria formed in the same way as many of the lava flows on Earth—by melting deep in the planet, followed by ascent of molten rock to the surface and eruption through fractures and vents. Within any group of rocks, e.g., basaltic lavas, the rocks are not identical. There are differences in mineral composition, in mineral shapes and sizes, and in each rock's chemical composition. The rocks of the Moon fall into four distinct groups: basaltic volcanic rocks, pristine rocks, breccias and impact melts, and lunar soil.

The basaltic volcanic rocks of the maria were produced inside the Moon by partial melting of the lunar mantle, followed by the buoyant rise of molten rock to the lunar surface and its eruption. Two types of volcanic rocks have been erupted to the lunar surface—lava flows and pyroclastic (volcanic ash) rocks. The volcanic plains that form the maria are rich in iron (Fe), and many are rich in titanium (Ti), most of which is in the mineral ilmenite ($FeTiO_3$). The common way of distinguishing mare basalts is to classify them according to their TiO_2 content into high-Ti basalts (>9 wt-% TiO_2), low-Ti basalts (1.5–9 wt-% TiO_2), and very low-Ti basalts (<1.5% TiO_2). Pristine rocks from the highlands are rocks that have original lunar compositions, uncontaminated by impact mixing. Pristine highland rocks are divided into three major groups: ferroan anorthosites (low-Na plagioclase, low-Mg pyroxene), Mg-rich rocks (plagioclase, high-Mg pyroxene), and KREEP rocks (K (potassium), rare Earth elements, phosphorus). However, most lunar rocks have been shattered, crushed, and recompacted by meteoroid impacts (breccias) and or formed by melting of older lunar rocks during micrometeoroid impacts (impact melts). Breccias containing a single rock type are called monomict breccias, while multiple rock types are called polymict breccias. Polymict breccias are the most common of lunar breccias. There are seven major groups: fragmental breccias, glassy melt breccias, crystalline melt breccias, clast-poor impact melt rocks, granulitic breccias, dimict breccias, and regolith breccias. Finally, there is the lunar soil, which is composed of fragmental (<1 cm), unconsolidated debris within the lunar regolith that covers the lunar surface.

EARTH'S ENVIRONMENT AND SPACE

Lunar Regolith

The lunar regolith is the source of virtually all our information about the Moon. *Regolith* is a general term for the layer or mantle of fragmental and unconsolidated rock material, whether residual or transported and of highly varied character, that nearly everywhere forms the surface of the land and overlies or covers bedrock. It includes rock debris of all kinds, including volcanic ash. On the airless and lifeless Moon, the lunar regolith basically results from two uniquely different processes—the continuous impact of large and small meteoroids on the lunar surface and the steady bombardment of the lunar surface by charged atomic particles from the stars and the Sun. All of the lunar landings and all of the photographic investigations show that the entire lunar surface consists of a regolith layer that completely covers the underlying bedrock, except perhaps on some very steep-sided crater walls and lava channels, where there may be exposed bedrock. Studies of returned samples have shown that the bulk of the lunar regolith consists of particles less than 1 cm in size, although larger cobbles and boulders, some as much as several meters across, are commonly found at the surface. Although lunar soil is lexicographically synonymous with lunar regolith, the term *lunar soil* usually refers to the sub-cm fraction of the lunar regolith. Lunar soil is a somewhat cohesive, dark gray to light gray, very fine-grained, loose, clastic material derived primarily from the mechanical disintegration of basaltic and anorthositic rocks. The mean grain size of analyzed soils ranges from about 40 mm to about 800 mm, averaging 60–80 mm. Individual lunar soil particles are mostly agglutinates, as well as various rock and mineral fragments. Agglutinates are individual particles that are aggregates of smaller lunar soil particles (mineral grains, glasses, and even older agglutinates) bonded together by vesicular, flow-banded glass. Agglutinate particles are small (usually less than 1 mm) and contain minute droplets of Fe (much of which is very fine-grained single-domain FeO), and troilite (FeS). Agglutinates make up a high proportion of many lunar soils, about 25–30% on average, although their abundance may range from a rare 5% to about 65%. Overall, lunar soils range in composition from basaltic to anorthositic, and they include a small (<2%), meteoritic component. Although the chemical compositions of lunar soils show considerable variation, physical properties such as grain size, density, packing, and compressibility are rather uniform. One manifestation of such uniform physical properties is that seismic velocities in regolith materials at both mare and highland sites are quite similar, varying between 92 m/s and 114 m/s. The current consensus is that the regolith is generally about 4–5 m thick in the mare areas but may average about 10–15 m in older highland regions, based on orbital radar data and modeling. Beneath this true regolith is a complex zone that probably consists of large-scale ejecta and impact-fractured, brecciated bedrock. Because of its surficial, unconsolidated, and fine-grained nature, it is likely that the regolith will be an important raw material to be used for lunar base construction and shielding, mining, road building, and resource extraction.

Soil Composition

It may appear incongruous for a planet that has no atmosphere, but oxygen is overwhelmingly the most abundant element in lunar surface materials. Table 16.5 presents the elements in their order of abundance, both in atom-% and weight-%. Thanks to the implantation of ions from the solar wind into the grains of lunar soil,

TABLE 16.5 Average Abundance of Elements in Lunar Soil

Element	Average abundance (atom-%)	Average abundance (mass-%)	Sites
O	60%	45%	Maria/highlands
Si	16.5%	21%	Maria/highlands
Al	10%	13%	Highlands
	5%	5%	Maria
Ca	5%	10%	Highlands
	4.5%	8%	Maria
Mg	5%	5.5%	Maria/highlands
Fe	2.5%	6%	Highlands
	6%	15%	Maria
Ti	<1%	<1%	Maria/highlands
Na	<1%	<1%	Maria/highlands

Source: Heiken et al. 1991.

even elements such as hydrogen (H), carbon (C), nitrogen (N), oxygen (O), phosphorus (P), and sulfur (S), as well as well as some noble gases, are available on the surface of the Moon. The solar wind is a plasma of chemical elements, expelled as ionized atoms from the atmosphere of the Sun. The principal element of the solar wind is H, the second most abundant element is helium (He), followed by C, N, and O. The concentration of H in typical lunar soil is about 50 g/kg. This corresponds to about 1.5 million liters of water per km^2 to a depth of 2 m, or 100 g/m^3 of H, assuming a soil density of 1.75 g/cm^3. So the problem is not that H and the other elements are scarce on the Moon, but whether they are economically accessible (Heiken et al. 1991).

Water Ice in the Lunar Polar Areas?

The very small inclination of the Moon's equator to the ecliptic led to the most interesting speculation that there may be water ice preserved in the lunar polar areas, for example in permanently shadowed craters of the South pole. This water ice may originate from comets that struck the Moon long ago. Debate over the presence or absence has developed over several years, but it still remains unclear whether or not the ice exists. In 1998, the Lunar Prospector mission provided further evidence for the existence of water ice, but still no ultimate proof (Heiken et al. 1991).

16.18 PHYSICAL SURFACE PROPERTIES

The physical properties of the lunar surface discussed in this subsection cover mainly the geotechnical properties, i.e., the surface properties needed to evaluate engineering problems, in particular the mechanical properties of lunar regolith. The experience of the first human being to set foot on the Moon, Neil Armstrong, gives an initial impression of these properties:

EARTH'S ENVIRONMENT AND SPACE

> I am at the foot of the ladder. The LM [Lunar Module] footpads are only depressed in the surface about 1 or 2 inches, although the surface appears to be very, very fine-grained, as you get close to it, it's almost like a powder; down there, it's very fine . . . I'm going to step off the LM now. That's one small step for [a] man. One giant leap for mankind. As the—The surface is fine and powdery. I can—I can pick it up loosely with my toe. It does adhere in fine layers like powdered charcoal to the sole and sides of my boots. I only go in a small fraction of an inch. Maybe an eighth of an inch, but I can see the footprints of my boots and the treads in the sandy particles.

The ranges of geotechnical properties of lunar materials are less than those that occur in surficial materials on Earth. This is because of the following factors:

- The familiar terrestrial geologic processes of chemical weathering, running water, wind, and glaciation are absent on the Moon. These processes tend to produce well-sorted sediments with uniform grain sizes.
- The three main constituents most likely to produce unusual or problem soils on Earth are absent on the Moon. There is no water, and therefore no clay minerals or organic materials.
- The variety of minerals in lunar soil is much less than that found on Earth. Many soil particles are simply fragments of rocks, minerals, and glass stuck together with glass (agglutinates).

The most significant variable of all geotechnical properties is the relative density. To a certain extent, it also controls the other physical properties discussed in this subsection (Heiken et al. 1991).

Particle Size Distribution

The particle size distribution in an unconsolidated material, such as lunar soil, is a variable that controls to various degrees the strength and compressibility of the material, as well as its optical, thermal, and seismic properties. The majority of lunar soil samples fall in a fairly narrow range of particle size distributions. In general, the soil is a well-graded (or poorly sorted) silty sand to sandy silt. The median particle size is 40–130 mm, with an average of 70 mm, i.e., approximately half of the soil, by weight, is finer that the human eye can resolve. Roughly 10–20% of the soil is finer than 20 mm, and thus a thin layer of dust adheres electrostatically to everything that comes in contact with the soil—space suits, tools, equipment, and lenses. Housekeeping is thus a major challenge for operations on the lunar surface (Heiken et al. 1991).

Specific Gravity

The specific gravity, G, of a soil particle is defined as the ratio of its mass to the mass of an equal volume of water at 4°C. Many terrestrial soils have a specific gravity of 2.7, i.e., the density of the individual particles is 2.7 g/cm^3, or 2.7 times that of water (1 g/cm^3). Values for lunar soils mostly range from 2.3 to >3.2 (agglutinate/glass particles: 1.0 to >3.32; basalt particles: >3.32; breccia particles: 2.9–3.1). A value of 3.1 is recommended for general scientific and engineering analyses of lunar soils (Heiken et al. 1991).

Bulk Density and Porosity

The bulk density, ρ, of soil is defined as the mass of the material contained within a given volume, usually expressed in g/cm^3. The porosity, n, is defined as the volume of void space between the particles divided by the total volume. Bulk density, porosity, and specific gravity, G, are interrelated as:

$$\rho = G\rho_w(1 - n)$$

with the density of water, $\rho_w = 1$ g/cm^3. The *in situ* bulk density of lunar soil is a fundamental property. It influences bearing capacity, slope stability, seismic velocity, thermal conductivity, electrical resistivity, and the depth of penetration of ionizing radiation. Consequently, considerable effort has been expended over the years in obtaining estimates of this important parameter. Taking into account all of the measurements, approximations, and analyses of the returned Apollo core samples and other measurements, the best estimate for the average bulk density of the upper 60 cm of lunar soil is 1.66 g/cm^3. For regions below 3 m no direct tactical data about the density of the lunar regolith exist. But it is known that the density approaches a maximum value at about 50 cm depth and increases very slowly beyond that (Heiken et al. 1991).

Shear Strength

The shear strength of a granular soil is typically defined in terms of the classic Mohr–Coulomb equation:

$$\tau = c + \sigma \tan \Phi$$

where τ = shear strength (kPa)
c = cohesion (kPa)
σ = normal stress (kPa)
Φ = friction angle

The shear strength therefore consists of two components: a cohesive component that is independent of applied stress, and a frictional component that is directly proportional to the normal stress. The shear strength governs such important engineering properties as ultimate bearing capacity, slope stability, and trafficability. As a result, estimates of lunar soil cohesion and friction angle have been the object of intensive research. Based on a variety of data sources, including the Apollo missions, best-estimate values for cohesion and friction angle have been developed as indicated in Table 16.6 (Heiken et al. 1991).

Bearing Capacity

The bearing capacity describes the ability of a soil to support an applied load, such as an astronaut, a vehicle, or a structure. Usually the topic of bearing capacity is divided into two categories: ultimate bearing capacity and allowable bearing capacity. Each of these is then subdivided further into static and dynamic quantities. The ultimate bearing capacity defines the maximum possible load that can be applied without causing gross failure, such as the overturning of a structure. The allowable bearing capacity defines a lesser load that can be applied without ex-

TABLE 16.6 Best-Estimate Values for Cohesion and Friction Angle

Location	Cohesion c (kPa)	Friction angle ϕ (°)
In general	0.1–1	30–50
Crater wall (inner)	0.17–1.0	45–25
Crater slope (outer)	0.52–2.7	45–25
Horizontal ground	0.34–1.8	45–25
Intercrater areas		
0–15 cm	0.44–0.62	41–43
0–30 cm	0.74–1.1	44–47
30–60 cm	2.4–3.8	52–55
0–60 cm	1.3–1.9	48–51

Source: Heiken et al. 1991.

ceeding a given amount of settlement. A settlement limit is usually imposed either for structural or operational requirements.

The static ultimate bearing capacity, q_{ult}, can be estimated on the basis of plasticity theory. It is therefore controlled by the soil density, its shear strength, and the size of the footing. Using the *in situ* bulk density estimates and the *in situ* shear strength estimates given earlier in this subsection, the static ultimate bearing capacity versus footing width can be calculated (see Figure 16.23). This means that the ultimate load (stress × area) for a circular or square footing is proportional to the cube of its width. Consequently, the ultimate bearing capacity of the lunar surface is more than sufficient to support virtually any conceivable structure.

The dynamic ultimate bearing capacity defines the maximum resistance to impact loading. This dynamic capacity is always greater than the static capacity be-

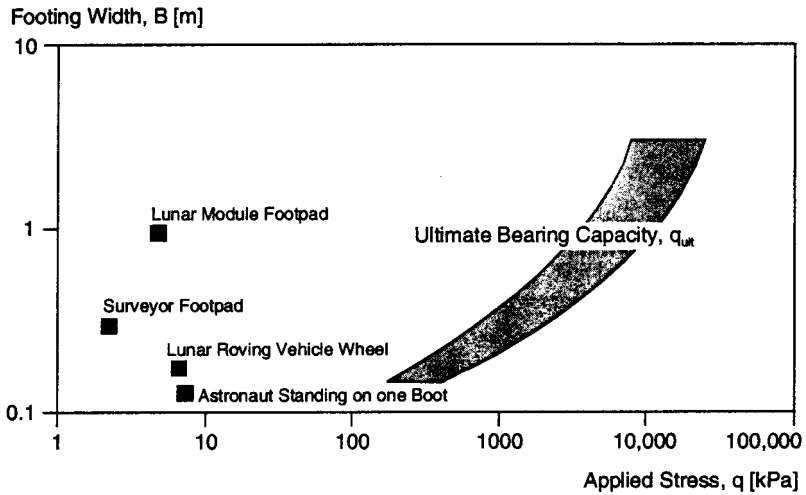


FIGURE 16.23 Static ultimate bearing capacity data for lunar surface material (Heiken 1991).

cause of the inertial resistance of the soil. Even after the experience from several manned lunar landings, many scientists and engineers still believed that an astronaut could hammer a rod or core tube into the lunar surface material to almost any depth. In fact, the practical limit of the Apollo core tubes was only about 70 cm, and it typically required about 50 hammer blows to reach this depth. An analysis showed that, if energy losses were neglected, the number of hammer blows required to reach a given depth would increase with the square of the depth. If energy losses were included, then there would be a depth beyond which no amount of hammering would drive a rod or core tube farther in (Heiken et al 1991).

Slope Stability

Many numerical methods have been developed to evaluate the stability of a soil slope, i.e., its ability to stand without support. On the Moon, the absence of water greatly simplifies the analysis of slope stability. The factor of safety, *F.S.*, against slope failure can be reduced to the expression:

$$F.S. = N\frac{\rho g_m h}{c}$$

where ρ = density of soil
c = cohesion of soil
g_m = acceleration of gravity on the Moon = 1.624 ms^{-2}
h = height of slope
N = stability number, which is a function of friction angle, Φ, and slope angle, θ

A constructed slope could be either an excavation, a compacted embankment, or a dumped pile (see Figure 16.24). The safe depth, i.e., the depth up to which no slope failure will occur, of an excavation in an intercrater area can be calculated by combining values for the *in situ* density with the *in situ* shear strength. Using a factor of safety of 1.5, which is more than adequate for design purposes, calculations show that a vertical cut could be made in lunar soil to a depth of about 3 m and a slope of 60° could be maintained to a depth of about 10 m. In order to construct an embankment, the soil must first be excavated, then transported, placed, spread, and compacted. As discussed above, *in situ* lunar soil is very dense, with a greater density than could be produced with mechanical compaction equipment. The processes of handling and manipulating the lunar soil would loosen it considerably, and it would not then be possible to compact the soil back to its original, undisturbed density. As a result, the density of the soil in a compacted embankment would be less than its original density, and the maximum possible slope angle of the embankment would be less than that of an excavation in undisturbed soil. Assuming a compacted relative density of 65–75%, a 10 m-high slope could be constructed at an angle of about 45°. If lunar soil were simply dumped in a pile, it would attain a relative density of about 30–40% and the factor of safety would be 1.0. The pile could be raised to a height of 10 m at an angle of nearly 40°. Not much is known about the stability of natural lunar slopes. The very limited cone penetrometer data obtained by both human and robotic missions have established that the soil on the slopes is actually somewhat weaker than the soil in the flatter intercrater areas, at least to a depth of 70 cm (Heiken et al. 1991).

EARTH'S ENVIRONMENT AND SPACE

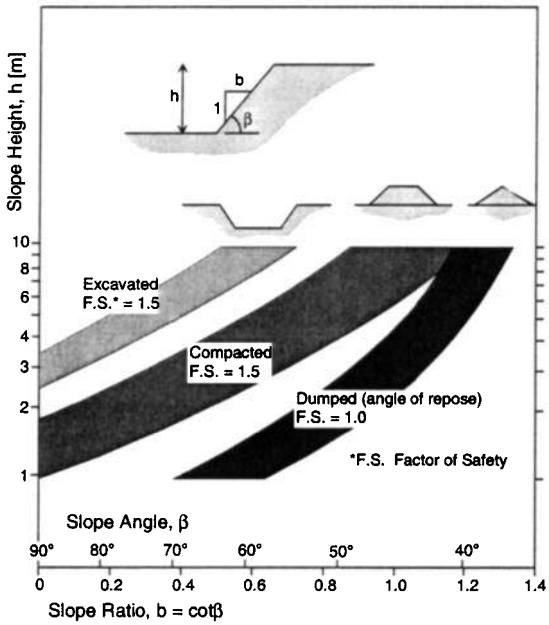


FIGURE 16.24 Calculated stability of artificial slopes constructed from lunar surface material (Heiken 1991).

Seismic Activities

The Apollo passive seismometers monitored the Moon's seismic activity for almost eight years. It was found that the release of seismic energy from the Moon is about seven orders of magnitude lower than that of Earth. The sources of seismicity on the Moon are:

- Monthly deep-focus moonquakes caused by Earth–Moon tidal stresses.
- Shallow moonquakes that are fewer but stronger and may be due to tectonic processes. They account for most of the seismic energy released in the Moon.
- Thermal moonquakes that may be due to thermal degradation of young lunar surface features. They last for about an hour.
- Moonquakes caused by meteoroid impacts, which vary widely in energy. Meteoroid impacts of all energies tend to be most common when meteoroid showers peak, particularly among the largest meteoroid impacts that tend to occur in the months of April to July. The largest recorded impacts, in July 1972 and May 1975, represented meteoroids of about 5 tonnes. Overall, seven meteoroid impacts of 1 tonne or more were observed within five years during lunar seismic monitoring.

Overall, the Moon can be considered seismically and tidally stable, with a magnitude of the largest events recorded being below 5 on the Richter scale (Heiken et al. 1991; Moore 1981).

Thermal Properties of Lunar Soil

Since the Moon is a small planetary body, there is good reason to believe that it has cooled considerably during its 4.6 billion-year history. Most of the present heat flux is probably generated by radioisotopes, mainly ^{40}K, ^{232}Th, ^{235}U, and ^{238}U, present in the interior to a depth of about 300 km. Experiments installed on the Moon during the Apollo missions provided extensive information on the temperature and thermal properties of the lunar surface layer to a depth of 3 m, including surface temperature variations, near-surface thermal properties, subsurface temperature variations, and thermal conductivity. At the Apollo sites, mean temperatures 35 cm below the surface are 40 to 45 K above those at the surface. This is primarily due to the fact that the upper 1 to 2 cm of the lunar surface have an extremely low thermal conductivity, with this conductivity being temperature-dependent. At a depth of about 2 cm, the conductivity increases greatly to values five to seven times greater than the surface value. This increase of conductivity appears to be mainly due to a large increase in the soil compaction and grain boundary contacts with depth. In lower layers of the lunar surface, the thermal conductivity of the lunar soil is on the order of 1.4–3.0 10^{-4} W/cm K. This is approximately a factor of 10 higher than the conductivity at the surface. Lunar surface thermal property values are summarized in Table 16.7 (Heiken et al. 1991).

Electrical and Electromagnetic Properties of Lunar Soil

The electrical properties of the lunar surface materials are those of silicates characterized by extremely low loss and low electrical conductivity. In the total absence of water, the DC electrical conductivity ranges from 10^{-14} mho/m for lunar soil to 10^{-9} mho/m for lunar rocks at 300 K in darkness. Upon irradiation with sunlight, there is a more than 10^6-fold increase in conductivity of both lunar soils and rocks. The relative dielectric permittivity k' for lunar materials as a function of bulk density, ρ, is approximately:

$$k' = 1.9\ \rho$$

The relative dielectric permittivity is controlled by bulk density and is independent

TABLE 16.7 Summary of Lunar Surface Thermal Property Values

Surface material	Parameter γ $\left[\frac{cm\sqrt{s}K}{J}\right]$	Density ρ $\left[\frac{kg}{m^3}\right]$	Specific heat c $\left[\frac{J}{kgK}\right]$	Conductivity k $\left[\frac{W}{mK}\right]$
Total range	5.97–334	500–3000	755–1007	2.14×10^{-3}–1.13
Range for particulate material heavily mixed with blocks	57.2–119	1200–2000	837	7.12×10^{-3}–1.8×10^{-2}
Blocks (rocks)	7.2	2500	837	9.22×10^{-1}
Range, excluding blocks	95.5–238	500–1100	837	4.18×10^{-3}–1.17×10^{-2}
Average maria	95.5–191	800–1500	837	4.18×10^{-3}–8.8×10^{-2}

Source: Heiken et al. 1991.

of chemical or mineralogical composition, variations of frequency above 1 MHz, and temperature variations within the range of lunar surface temperatures. The loss tangent for high-frequency electromagnetic loss in lunar soil is given by:

$$\tan \theta = 10^{[0.038(\%TiO_2+\%FeO)+0.312\rho-3.260]}$$

In both equations, ρ is the bulk density in g/cm^3. The extremely low electrical conductivities and low loss tangents indicate that lunar materials are very transparent to electromagnetic waves. For example, radio transmissions should readily penetrate through the lunar soils to a depth of about 10 m. As a result, radio communications on the lunar surface need not necessarily be by direct line of sight, but may penetrate low hills. The low conductivity and low loss are also responsible for the fact that lunar materials are readily chargeable and will remain electrically charged for long periods of time. The large photoelectric change in electrical conductivity at lunar sunrise and sunset can charge surface soil particles to the point that they will levitate and move. Such charged soils and mobile particles could readily coat surfaces and be hazardous to visibility and equipment operations during the lunar night.

16.19 LUNAR SURFACE ENVIRONMENT

Lunar Surface Temperature

The temperature of the lunar surface is determined by the amount of radiation absorbed by the Sun, as well as the heat load from the interior. Assuming constant thermal properties for the surface material, the surface temperature curve depends only on a single parameter, the thermal inertia, defined as:

$$\gamma = \frac{1}{\sqrt{k\rho c}}$$

where k = thermal conductivity
ρ = bulk density
c = specific heat

The thermal inertia of the lunar surface layer does vary with temperature and depth, but including these complications does not significantly increase the fidelity of the temperature model for engineering purposes. The lunar thermal inertia is so low that the daytime surface temperature is essentially in thermal equilibrium with absorbed incident solar radiation. At the equator this translates to a subsolar point temperature of about 390 K. The predawn temperature falls to approximately 110 K at the equator and to lower values at higher latitudes (see Table 16.8).

In general, the lunar surface temperature increases by about 280 K from just before lunar dawn to lunar noon. The temperature at lunar noon varies throughout the year because of varying distance from the Sun. The noon temperature increases about 6 K from aphelion (greatest distance from the Sun) to perihelion (smallest distance to the Sun). There is a large difference in mean temperature, i.e., the temperature averaged over a complete day-night cycle, just below the lunar surface. Estimated average surface temperatures and temperature extremes for different areas of the Moon are presented in Table 16.8. The temperature at the poles is basically

TABLE 16.8 Estimated Average Surface Temperatures and Temperature Extremes for Different Areas of the Moon

	Shadowed polar craters	Other polar areas	Front equatorial	Back equatorial	Limb equatorial	Typical midlatitudes
Average temperature	40 K	220 K	254 K	256 K	255 K	220 K < T < 255 K
Monthly range	none	10 K	140 K	140 K	140 K	110 K

Source: Heiken et al. 1991.

unknown, but it might be as low as 40 K in some permanently shaded areas, i.e., inside craters. This might be the case for about 2% of the lunar surface (Heiken et al. 1991).

Lunar Albedo

The albedo is defined as the fraction of light or electromagnetic radiation reflected by a body or particular surface. The average albedo of the Moon is very low (0.09), i.e., only 9% of the light received by the Moon is reflected back to space or to objects on the surface. For lunar surface materials, the brightness of any area, observed from the Earth at full Moon, yields a value that is virtually that of the normal (perpendicular) albedo for that area. The normal albedo is defined as the normal reflectance of a surface element which is illuminated normally. The value of the normal albedo for an area of lunar surface will depend on the local chemical and mineralogical composition, particle size, and packing density. In general, crater ray systems are the brightest features on the Moon, and highland areas are brighter than maria. Table 16.9 presents normal albedo values of the lunar near and far side.

Lighting Environment of the Lunar Surface

Natural lighting on the lunar surface is very dependent on location. Any spot on the lunar surface goes through the same light cycle during every lunar day. The incident light angle depends on the latitude of the lunar surface position. In addition

TABLE 16.9 Normal Albedo Values of the Lunar Near and Far Side

Regions	Minimum	Normal albedo maximum	Average
Front side			
Mare	0.06	0.09	0.07
Highland	0.11	0,20	0.15
Entire face	0.07	0.10	0.09
Far side	—	—	0.22

Source: Smith 1971.

EARTH'S ENVIRONMENT AND SPACE

to sunlight, some natural light is available in the form of earthshine, i.e., sunlight reflected from the Earth. Since the Earth is bigger than the Moon, and the Earth's global average albedo (0.39) is greater than the Moon's (0.09), the brightness of full earthshine as seen from the Moon is 58 times greater than the brightness of full Moon as seen from the Earth. Brightness is a function of albedo times radius squared. The earthshine light is site-dependent because the same side of the Moon always faces the Earth. Sites on the far side of the Moon never see any earthshine, while sites on the near site see the Earth go through all phases, but with varying efficiencies. Earthshine overlapped by sunshine has no benefit. Sites near the Earth terminator receive only minimal earthshine when not in sunlight. Overall, natural lighting will not be available for entire lunar cycles at most sites.

With no atmosphere to scatter light, the ambient light level of the Moon drops to only what is reflected by objects and the ground. There are three situations where the direction of the available light causes visibility problems that are attributed to the lack of ambient light:

1. Shadowing—low light angles (below 30° elevation) cause extremely long shadows that can hide obstacles and craters.
2. Washout—high light angles (within 30° of the normal) do not allow reflected light to reach the eye. An object simply does not appear until one is right next to it. There is also a lack of shadows on most of the lunar landscape, which causes objects to be less defined.
3. Back Lighting—in case of moving down-light, i.e., when the Sun is right behind the observer, an effect similar to washout results, but only in one direction. The surface appears featureless when looking about 10° to either side of the down-Sun or down-light direction.

During the Apollo missions astronauts found that their space suits could be used to reflect light into the shadows when a little extra light was needed.

The Lunar Atmosphere

The lunar atmosphere is very tenuous and not very well characterized. The undisturbed gas concentration is only about 2×10^5 molecules/cm^3 during the lunar night, falling to perhaps 10^4 molecules/cm^3 during the lunar day. This is about 14 orders of magnitude less than for the Earth's atmosphere, a difference so extreme that the Moon is often said to have no atmosphere at all. The major constituents of the ambient lunar atmosphere are neon, hydrogen, helium, and argon. Neon and hydrogen are derived from the solar wind. Helium is mostly derived from the solar wind, but about 10% may be radiogenic and lunar in origin. Argon is mostly ^{40}Ar that is derived from the radioactive decay of lunar ^{40}K. Table 16.10 lists the most probable abundances of these and other species in the undisturbed lunar atmosphere (Heiken et al. 1991).

Since the lunar atmosphere can be considered as a planetary exosphere, all neutral atoms move along ballistic or escaping trajectories in the gravitational field with no collisions between them. This implies that any localized injection of new particles from a landing spacecraft or from a lunar base will quickly spread all around the Moon, in a time on the order of the ballistic free flight time of the molecules, i.e., in less than three hours for oxygen atoms which have a mean thermal speed of 0.6 km/s. Light gases will all escape from the lunar gravitational

TABLE 16.10 Most Probable Abundances of Lunar Atmosphere Elements

Species	Abundances [molecules/cm^3]		Scale heights	
	Daytime	Nighttime	Daytime	Nighttime
^{20}Ne	4×10^3 to 10^4	10^5	100 km	25 km
He	8×10^2–4.7×10^3	4 to 7×10^4	511 km	128 km
H_2	2.5–9.9 $\times 10^3$	10^4–1.5×10^5	1022 km	256 km
Ar	2×10^3	$<10^2$	55 km	—
CH_4	1.2×10^3	—	—	—
CO_2	10^3	—	—	—
NH_3	4×10^2	—	—	—
$OH + H_2O$	0.5	—	—	—

Source: Heiken et al. 1991.

field. For example, atomic H disappears in less than 120 minutes from the sunlit lunar hemisphere where the temperature is 400 K and the thermal speed of H is 2.5 km/s, while the escape velocity of the Moon is only 2.38 km/s (ESA 1992).

The Lunar Magnetic Field

The magnetic field of the Moon is essentially negligible. During the Apollo missions, an extremely weak magnetic field of 3–300 nT (nanotesla) was measured. Additionally, an external field due to the solar wind of 5–10 nT is present when the Moon crosses the Earth's magnetic tail during four days per orbit. Although the Moon has no global magnetic field, its surface is dotted with small zones where the crust is strongly magnetized. These magnetic zones may be caused by unusual impact conditions or subsurface geology (Spudis, 1996).

Radiation Environment of the Lunar Surface

On the lunar surface, there are two kinds of incoming radiation—electromagnetic radiation and ionizing radiation. Basically all of the electromagnetic radiation in the solar system is emitted by the Sun. The solar electromagnetic radiation at a distance of 1 AU from the Sun has an average energy density of about 1,368 W/m^2. The most important aspect of electromagnetic radiation on the lunar surface is its potential use for solar power production.

Ionizing radiation consists mainly of protons, electrons, and some heavier nuclei. These particles interact with the Moon in different ways, depending on their energy and composition, resulting in penetration depths that vary from μm (micrometers) to m. Any kind of lunar surface habitat will have to be protected from the three different kinds of ionizing radiation in space:

1. Solar wind
2. Solar cosmic rays (SCRs)
3. Galactic cosmic radiation (GCR)

Meteoroid Environment

The term *meteoroid* is used for a naturally occurring solid body, traveling through space, that is too small to be called an asteroid or a comet. Meteoroids with diameters less than about 1 mm are commonly classified as micrometeoroids. The term *meteorite* is used for meteoroids that have fallen upon a planet and have been recovered. Almost all lunar rock surfaces that were exposed to space contain numerous microcraters. Studies of lunar rocks have revealed the average meteroid flux during the past several hundred million years. The average annual cumulative meteoroid model estimates for the lunar surface are:

For $10^{-6} < m < 10^{6}$:

$$\log Nt = -14.597 - 1.2131 \log m$$

For $10^{-12} < m < 10^{-6}$:

$$\log Nt = -14.566 - 1.5841 \log m - 0.063 (\log m)^2$$

where Nt = number of particles/m^2/s of mass m or greater
m = mass (in g)

The velocities of meteoroids that hit the Moon can be calculated to range from 13 to 18 km/s. The meteoroid flux at the lunar surface shows a significant enhancement from small meteoroids (<1 mm), traveling toward the Moon from the Sun, and a somewhat lesser enhancement of large particles (>1 mm), arriving from the direction in which the Earth is traveling. Whichever side of the Moon is facing into the direction of the Earth's motion in its orbit around the Sun will be more exposed to the larger and more hazardous meteoroids (Heiken et al. 1991).

Environmental Impact of Lunar Surface Activities

The Moon, in its pristine condition, serves as an important, well-preserved fossil of the solar system. Therefore, any lunar surface activity, in particular the installation of a lunar base and mining of lunar resources, will cause environmental concern. Most lunar scientific activities require that the unique lunar environment be preserved. Lunar base operations and mining might affect this environment in adverse ways. Specific potential environmental impacts include increased atmospheric pressure, which would compromise astronomical observations, and increased radio frequency background through lunar communication relay satellites, which could affect the use of the far side of the Moon for radio telescopes. Extensive mining efforts could scar the lunar surface irreversibly, which could destroy some of the potential for geological discoveries.

REFERENCES

Eckart, P. 1996. *Spaceflight Life Support and Biospherics,* Kluwer Academic Publishers, Dordrecht; Microcosm, Torrance, CA.

Eckart, P. 1999. *The Lunar Base Handbook,* McGraw-Hill, New York.

ESA. 1992. ESA Report No. SP-1150, ESA, Noordwijk.

Heiken, G., Vaniman, D., and French, B. 1991. *Lunar Sourcebook: A User's Guide to the Moon,* Cambridge University Press, Cambridge.

Landis, G. 1990. "Degradation of the Lunar Vacuum by a Moon Base," *Acta Astronautica,* vol. 21, no. 3, pp. 183–187.

Mendell, W., ed. 1984. *Lunar Bases and Space Activities of the 21st Century,* Lunar and Planetary Institute, Houston.

Moore, P. 1981. *The Moon,* Mitchell Beazley, London.

Spudis, P. D. 1996. *The Once and Future Moon,* Smithsonian Institution Press, Washington, DC.

PART 5

MARS

G. Komatsu

16.20 ORBITAL CHARACTERISTICS

Mars is the fourth planet from the Sun. It has a metallic core and a rocky mantle and crust. Its orbital characteristics and geometrical properties are summarized in Table 16.11. The main difference between the orbits of Earth and Mars is the eccentricity of Mars's orbit, 0.0934, giving the planet a larger range of solar illumination than Earth experiences.

A particularly interesting orbital property is obliquity. Because of the lack of large satellites stabilizing Mars, the obliquity is in chaotic motion and it can change up to 60°; some orbits can change dramatically in less than 45 Myr (Lasker and Robutel 1993). The large range of obliquity could have a profound impact on the geomorphic processes on the Martian surface.

16.21 SOLID GEOPHYSICAL PROPERTIES AND INTERIORS

The geophysical parameters of solid Mars are summarized in Table 16.12. Mars is a terrestrial-type planet smaller than Earth and Venus but larger than the Moon and Mercury. The surface area of Mars is approximately equivalent to Earth's land area. Like other planets in the Solar System, Mars was probably born from the gas surrounding the young Sun. Like the other planets formed near the Sun, Mars did not collect enough gas, and so it is rocky. The absence of extensive surface chemical measurement and seismological data limits detailed discussion of the bulk composition and interior structure of Mars. However, based on the mean density and the moment of inertia, it is estimated that Mars should have a core.

Mars Global Surveyor observations found that the magnetic field of Mars does not encompass the entire planet, implying that the Martian dynamo driven in the

TABLE 16.11 Orbital and Rotational Parameters of Mars

Semimajor axis	1.52366 AU
Eccentricity	0.0934
Inclination	1°.8504
Longitude of ascending node	49°.59
Longitude of perihelion	335°.94
Mean orbital period	686.98 Earth days, 669.60 Mars solar days
Obliquity	25°.19
Length of day	88642.663 ± 0.002 s (sidereal) 99775.2 s (solar)

Source: Based on Kieffer et al. 1992a.

TABLE 16.12 Geophysical Parameters of Solid Mars

Parameter	Value
Mass	6.4185×10^{23} kg
Mean radius	3,389,508 m*
Volume	1.6318×10^{22} m^3
Mean density	3.9335 ± 0.0004 g cm^{-3}
Surface gravity	3.758 ms^{-2} (at pole), 3.711 ms^{-2} (at equator)
Total surface area	1.4441×10^{14} m^2

Source: Based on Kieffer et al. 1992a, except * from Smith et al. 1999b.

metallic core is essentially extinct. Instead, magnetic field sources of multiple scales, strength, and geometry were discovered in the ancient southern highland crust (Acuña et al. 1999). They are observed as East–West bands of alternating polarity, the longest extending for over 2,000 km. These are perhaps remnant magnetization formed during early Martian history when the dynamo was active.

Because Mars has no sea level, the level where the average atmospheric pressure is about 6.1 mbar (hPa), the triple point of water, has been used as a datum. The MOLA (Mars Orbiter Laser Altimeter) on board the Mars Global Surveyor acquired topographic information about Mars with high accuracy (Smith et al. 1999b). Dominant features of Mars include southern highlands, the low northern plains, and the Tharsis province (Figure 16.25). The northern depression has been proposed to have been formed by an internal mechanism such as convective mantle overturn and crustal thinning which resulted from a thermal event, or by a giant impact.

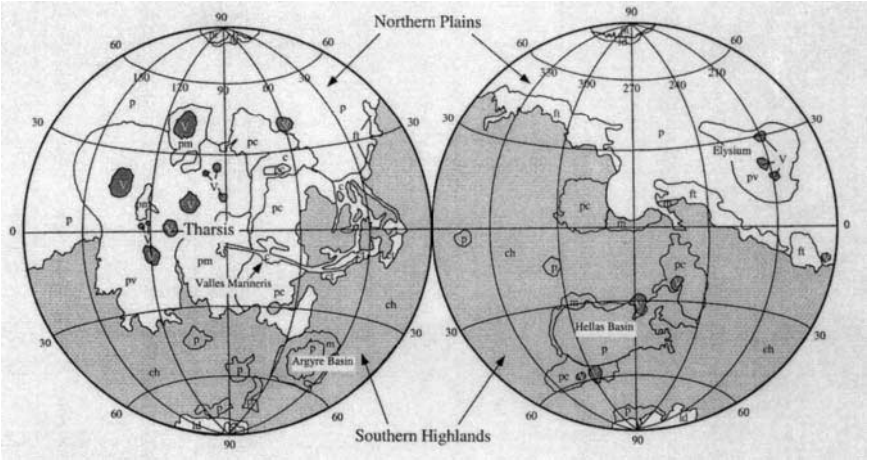


FIGURE 16.25 Simplified physiographic map of Mars, showing ancient (ch: cratered highlands, m: mountainous terrain), modified (p: plains, c: channel and canyon floors, ct: chaotic terrain, ft: fretted terrain), volcanic (V: volcanic constructs, pv: volcanic plains, pm: moderately cratered plains, pc: cratered plains), and polar (pi: permanent ice, ld: layered deposit), areas (modified from Mutch and Head 1975).

EARTH'S ENVIRONMENT AND SPACE

The gravity field of Mars has also been measured by the Mars Global Surveyor (Smith et al. 1999a). The rough, elevated southern hemisphere appears to be in a state of near-isostatic compensation, whereas the smooth, low northern plains display a wider range of gravitational anomalies that indicates a thinner, but stronger, surface layer than in the South.

16.22 SURFACE AND SUBSURFACE

Surface Materials

The global surface composition of Mars has been studied by Earth-based telescopes and spacecraft instruments. The early telescopic observations revealed that the albedo of Mars is about 0.15 and that the Martian surface has a red color and is divided into lighter and darker regions. This difference is explained by the degree of oxidation of basaltic rocks. The primary reason for the red color of the Martian surface has been generally attributed to the presence of iron oxide minerals, with the amount of the iron oxides varying depending on the minerals. In addition, clays have been detected in the bright regions and dust. It is also possible that materials in the bright regions are largely amorphous or poorly crystalline, and palagonite may be the main composition because of excellent spectral matching. Palogonite can be formed when volcanic eruptions occur in water-rich environments or during subsequent alteration of basaltic ash. The TES (Thermal Emission Spectrometer) on board the Mars Global Surveyor identified a few regions of high crystalline hematite concentration (Christiensen et al. 2000). Based on the detection of absorption features near 1 μm and 2 μm, the presence of pyroxene has been inferred in the dark regions of Mars, consistent with mafic composition of the surface estimated from the geomorphology of volcanoes. TES examined the dark regions and concluded that the highlands are basaltic whereas the northern plains appear to be more of andesitic composition (Bandfield et al. 2000). Carbonate deposits have been searched for extensively since they are closely linked with the presence of water on Earth. However, neither ground-based telescopic observation nor TES has identified an appreciable amount of carbonate on the Martian surface, within their detection limits.

Based on direct *in situ* analyses by Viking lander XRFS (X-ray fluorescence spectrometer) and supplementary information from the Martian meteorite Shergotty, a representative chemical composition of Mars soil has been derived (Banin et al. 1992; Table 16.13). Soil samples measured at the two Viking sites are remarkably similar to each other in terms of chemical composition despite the great distance between them. This may indicate that a thin semihomogeneous blanket soil material covers much of the Martian surface. There is no known terrestrial soil similar to the Martian soil, which is characterized by a very high concentration of sulfur.

The Mars Pathfinder APXS (alpha proton X-ray spectrometer) measurements of surface soils and rocks added another piece of information about the composition of the Martian surface. The soils at the Pathfinder landing site are similar to those at the Viking sites (Rieder et al. 1997). The measurement of rocks at the Pathfinder landing site yielded compositions close to andesite, a type of lava common on continental margins on Earth. However, it is not known whether the measured samples are igneous, sedimentary, or metamorphic rocks.

TABLE 16.13 Representative Chemical Composition of Martian Soil

Constituent	Selected average concentration (%)	Source
SiO_2	43.0	Direct soil analysis by Viking XRFS
Al_2O_3	7.2	Direct soil analysis by Viking XRFS
Fe_2O_3	18.0	Direct soil analysis by Viking XRFS
MgO	6.0	Direct soil analysis by Viking XRFS
CaO	5.8	Direct soil analysis by Viking XRFS
TiO_2	0.6	Direct soil analysis by Viking XRFS
K_2O	0.2	Analyses of Martian meteorite
P_2O_5	0.8	Analyses of Martian meteorite
MnO	0.5	Analyses of Martian meteorite
Na_2O	1.3	Analyses of Martian meteorite
Cr_2O_3	0.2	Analyses of Martian meteorite
SO_3	7.2	Direct soil analysis by Viking XRFS
Cl	0.6	Direct soil analysis by Viking XRFS
CO_3	<2	Estimated from Label Release simulation
H_2O	0.1	Variable: based on direct soil analysis on Viking GC-MS

Source: From Banin et al. 1992.

Eolian Processes

On present-day Mars, eolian (wind) processes are the most active geological force in shaping landforms. But there are many unanswered questions about these processes. Winds are capable of eroding and transporting a large quantity of materials over a geologically short time scale. Despite the tenuous atmosphere on Mars compared with that on Earth, there exist landforms such as yardangs and dunes, representative features formed by wind erosion and transport, respectively. Dunes are globally distributed, as revealed in the new Mars Global Surveyor images, but the main dune fields are found at high latitudes in both hemispheres and in craters of the southern highlands. These dunes are probably made of sand-sized material, but the source of the sand remains controversial (Greeley et al. 1992). How active are the eolian landforms on Mars? For example, observations made on a dune field over a 20 Earth-year period indicate that the upper limit of the dunes' motion is 200 times less than on Earth (Zimbeiman 2000). Even if they are still active, grain movement is slow. If this is the case, it is possible that many of the dunes are relict features formed during periods when the climate was somewhat different—for example, with a denser atmosphere. So-called wind streaks are bright, or dark, cone-shaped surface features emanating from obstacles such as impact craters. These features are considered to have formed through the interaction of wind, sediments and obstacles. Therefore, the directions of the wind streaks probably imply the direction of prevailing wind in the area. In fact, an excellent correlation has been found between the orientations of some wind streaks and wind directions derived from a GCM (Global Circulation Model).

Dust on Mars is globally distributed; it has been moved all over the planet by dust storms. There is a question of how the tenous Martian atmosphere raises dust particles. According to theoretical and experimental studies, the optimum size for a particle to be moved by wind on Mars is about 100 μm (Greeley et al. 1992).

EARTH'S ENVIRONMENT AND SPACE

The minimum wind velocities required to move dust-sized particles (<20 μm) are much larger than the wind velociities observed on Mars. Alternative dust-raising mechanisms include dust devils, the hitting of dust particles by saltating sands, and the suspension of dust aggregates and later pulverization into fine grains. In fact, dust devils have been observed on Mars, but it is not clear if they raise enough dust to initiate dust storms.

Impact Cratering

Impact cratering is a major geological process affecting the surface of Mars. The impact craters on Mars exhibit a very wide range of morphology due to the diversity of target materials, the presence of an atmosphere, and postimpact erosional processes. For example, many Martian impact craters are characterized by a lobate ejecta blanket, which is a sometimes overlapping set of lobes. This morphology has been attributed by many researchers to the presence of a volatile-rich layer in the target, but other researchers have suggested that ejecta interaction with the atmosphere is responsible for formation of the lobes.

The surface chronology of Mars has been defined mostly by the relative age based on the impact crater population density. For example, the southern highlands are the oldest, dating back to the end of the heavy asteroid/comet bombardment that formed Mars. The northern plains are much younger than the southern highlands; topographically they are generally flat. Image observations suggest that their surfaces are covered by young lava flows and sediments. The Martian history is divided into three periods—Noachian, Hesperian, and Amazonian, from old to young (Table 16.14). The absolute ages of the surface types are estimated by applying an inferred impact flux at the Mars orbit, the likely value ranging from six times the lunar rate to the same as the lunar rate. However, this approach is not yet established well enough for us to have great confidence in the ages derived. The big problem is that we do not know the variation of the influx of projectiles, such as asteroids and comets, through time. The resolution of this problem must await the results of the sample return mission, which should provide the absolute age of calibration sites. Impact craters on Mars exhibit a wide range of degradational morphology, which implies that geological processes have been actively operating on the planet. However, how active these processes have been, and when, are still a matter of debate.

Impact cratering may have played an important role in the atmosphere of early Mars. It has been suggested that very large basin-forming objects could have eroded the atmosphere to its present-day tenuous state (Melosh and Vickery 1989).

TABLE 16.14 Mars Time-Stratigraphic System

	Crater density	Absolute age
Amazonian-Hesperian boundary	400 > 2 km crater/10^6 km^2 67 > 5 km crater/10^6 km^2	1.8–3.55 Gyr
Hesparian-Noachian boundary	200 > 5 km crater/10^6 km^2 25 > 16 km crater/10^6 km^2	3.5–3.8 Gyr

Source: Based on Tanaka et al. 1992.

Vulcanology

Mars is a world of giant volcanoes. The largest is Olympus Mons, which is over 600 km in diameter and reaches more than 25 km above the surrounding plains. These giant volcanoes are called shield volcanoes because they have very small height-to-diameter ratios; this implies that the volcanoes were made primarily by eruptions of low-viscosity lava flows such as basalt. Due to the lack of Earth-type plate tectonics, the locations of the Martian shield volcanoes with respect to the magma chambers may have been fixed, and this may explain the extraordinary sizes of the volcanoes. Large shield volcanoes are concentrated in three regions. The largest one is Tharis that is centered on the equator and 105°W. Tharsis is a 6,000 km wide volcano-tectonic bulge overriding both the southern highlands and northern plains. The second is Elysium, which is centered on 25°N, 210°W. There are also eroded volcanic edifices distributed around the Hellas Basin in the southern hemisphere. Other major volcanic landforms include volcanic plains, very flat volcanic edifices called paterae, and tholi, which may be made of lava and ash layers. Recent Mars Global Surveyor observations of the Martian surface have revealed extensive coverage of lava flows in some parts of the northern plains. Their ages, based on crater counting, are on the order of 10 million years (Hartmann and Berman 2000), extremely young by Martian standards. This implies that Mars is practically still alive volcanically.

Tectonics

Global-scale tectonics on a planet are an indication of how the planet loses its heat generated internally. Earth loses heat through active plate tectonics and Venus seems to do it through plume activities. Smaller planetary bodies, such as the Moon and Mercury, are considered to have lost their heat primarily by conduction; they are practically dead internally. Mars does not have characteristic topographic features that indicate Earth-type plate tectonics, such as rifts, subduction zones, and collisional mountain chains. Because Mars is smaller than Earth, it is possible that Mars has already lost the majority of its internal heat.

For regional tectonics, the gigantic Tharsis bulge shows a wide range of structural features. In planetary geology, tectonic structures are used to infer the stress field acting in the region. For example, the Tharsis province is characterized by a wide range of structures such as grabens and wrinkle ridges, indicating tensional and compressional stress fields, respectively. The origin of Valles Marineris, the 400 km long canyon system, is often explained as being where the crust is spread by tensional stress. Many compressional landforms are observed associated with Olympus Mons, implying that the volcano is pressing down the crust because of its loading.

Polar Caps and Ground Ice

The polar regions of Mars are characterized by complex geology of ice caps, layered deposits, and peripheral deposits including dune fields. Polar ice is classified into two components, summer residual ice and CO_2 frost. The northern residual cap is made of H_2O, whereas the southern residual appears to be CO_2, but it is not known if the southern residual is underlain by H_2O ice. The layered deposits consist of

EARTH'S ENVIRONMENT AND SPACE

pairs of light and dark bands, but the composition of the bands is poorly constrained. They were probably formed as the result of cyclic climate change causing the deposition of dust and ice. The dunes around the northern polar cap seem to have some sources within the layered deposits, implying a genetic link between them.

According to theoretical models, very thick layers of ground ice can exist in the subsurface of Mars, at the depths of 1–3 km near the equator and 3–8 km in the polar regions (see, e.g., Fanale 1976). This type of modeling attempts to determine the upper boundary of ground ice layers, based on a diffusion process, by changing crustal temperature, which is related to surface temperature. If crustal temperatures are above the frost point of atmospheric water vapor, it is possible that ice may be lost by diffusion over time. Assuming that there was no replenishment of water since the early history of Mars, the upper several hundred meters of Martian ground ice at low latitudes should have been desiccated by now, unless the ice was protected from diffusion by some mechanism. In reality, we observe a number of landforms on the Martian surface at low latitudes, which can be attributed to processes due to shallow ground ice. These landforms include lobate debris aprons and thermokarst. This implies that there may have been hydrologically active events to replenish ice later in Mars history.

Hydrology

Water on Mars is one of the most exciting topics about the red planet. The issue can be addressed from two aspects—present conditions and past conditions. Present-day Mars is extremely dry; the atmosphere has 0.03% water content. The majority of the surface water inventory is concentrated in the polar caps, and a large quantity of water may lie underneath the surface as ground ice. The modern surface condition of Mars does not allow liquid water to flow freely, but new data from the MGS suggest that liquid water may have come out onto the surface very recently in geological terms to produce small-scale landforms such as gullies (Malin and Edgett 2000). The Mars explorations conducted mainly by the Mariner 9 and Viking orbiters have identified rich evidence of surface water activities and ground ice melting in the planet's past.

The discovery of numerous valleys on the Martian surface is the strongest evidence that liquid water played an important role in shaping the landscapes on the planet. Compared with their terrestrial counterparts, these valleys are less densely populated with water and have smaller drainage in general. Many of the valleys have short, stubby tributaries and a theater-headed terminus, characteristics similar to terrestrial sapping valleys. Sapping valleys are formed primarily by groundwater processes that undermine the rocks and cause the collapse of roof rocks. Although surface runoff cannot be excluded, sapping processes appear to be the dominant mechanism for the formation of valley networks on Mars. In any case, replenishing of the water to the source areas would be required, and hydrogenated circulation and precipitation have been proposed. For the latter, the past climate history of Mars would be different.

Channels on a grand scale exist on the surface of Mars. Called outflow channels, they were formed by ancient catastrophic floods, perhaps of water, with discharge rates a few orders of magnitude greater than that of the Amazon River. The majority of the outflow channels originate from the collapsed ground called chaotic terrain. It is considered that the ground ice layers became unstable due to magmatic heating,

and this caused a catastrophic release of water to the surface. There are also pieces of evidence indicating active ice-related processes that modified the outflow channels. There are many surface features indicative of past ponding of water, including shorelines such as wave-cut terraces and deltas, and sediments. The biggest may have been the northern plains ocean (see, e.g., Parker et al. 1989), but ponding perhaps also occurred in smaller topographic basins, inside crater rims, and in parts of the Valles Marineris. How often water filled the topographic lows is not well understood. A diverse group of landforms distributed widely in mid- to high latitudes has been interpreted to be glacial (Kargel and Strom 1992) and periglacial. Some of these interpretations have been opposed, but nonetheless ice seems to have played a major role in the geomorphology of Mars. The implications of ice processes on Mars are significant. In particular, glaciation requires a continuous supply of water to the accumulation zone in order to maintain the ice movement. And this means a hydrological cycle.

On Earth, hydrological elements such as oceans, lakes, rivers, and surface and subsurface ices are closely linked to each other. Water transport through the atmosphere is also very important. Whether this is also true for Mars is not known, and this point needs further investigation. Baker et al. (1991) proposed a hypothesis that connects many pieces of observational evidence of water-related landforms and envisions episodic global hydrological cycles relatively late in Mars history. According to the hypothesis, the episodic hydrological cycles were short-lived and therefore did not produce the range of erosional landforms seen on Earth. The hypothesis suggests that rapid formation of northern plain oceans by outbursts of water from below ground created the gigantic outflow channels. Perhaps the atmosphere was temporarily warmed due to greenhouse gases released at the same time as the water outflows, and precipitation may have occurred to supply water to valley networks and glaciers in the southern highlands.

16.23 ATMOSPHERE

Mars's tenous present-day atmosphere is mostly made of CO_2 and is characterized by its low water content (Table 16.15). The global annual average atmospheric

TABLE 16.15 Composition of the Martian Lower Atmosphere

Gas	Abundance
CO_2	95.32%
N_2	2.7
^{40}Ar	1.6
O_2	0.13
CO	0.07
H_2O	0.03
$^{36+38}Ar$	5.3 ppm
Ne	2.5 ppm
Kr	0.3 ppm
Xe	0.08 ppm
O_3	0.04–0.2 ppm

Source: Based on Owen 1992.

EARTH'S ENVIRONMENT AND SPACE

pressure at the surface is 5.6 mbar (hPa), but this value varies seasonally and spatially. The variation of atmospheric pressure is controlled by the sublimination and condensation of CO_2 on the polar caps. The ground temperatures at the equator range mostly from about 160 K to 180 K at night, to a peak temperature of about 260–280 K during the day (Carr 1996). The atmospheric dynamics are strongly affected by conditions unique to Mars, including the behavior of CO_2 described above, the radiative effect of dust loading, and variations of insulation due to the orbital eccentricity and the phasing of perihelion (Zurek et al. 1992). The average sunlight Mars receives is less than half the Earth receives. But the sunlight falling on the Martian surface varies greatly due to the planet's large orbital eccentricity and dust in the atmosphere.

The thermal structure of the Martian atmosphere is very different from the Earth's in some aspects. The major differences are caused by the lack of sufficient ozone and the variable dust loading in the Martian atmosphere (Zurek 1992). The ozone abundance in the Martian atmosphere is less than one three-hundredth of that in the terrestrial atmosphere, allowing strong UV light from the Sun to reach the surface. Because of the low ozone abundance, the Martian atmosphere lacks the stable, temperature-increasing layer equivalent to the stratosphere on Earth. Airborne dust in the Martian atmosphere plays the role of ozone by absorbing the Sun's radiation; hence the dusty Martian lower atmosphere somewhat resembles the terrestrial stratosphere. Obviously the dust loading in the Martian atmosphere is a strong function of season and region. Thus, the thermal structure of the Martian atmosphere can vary tremendously.

The most predominant large-scale atmospheric phenomenon is the dust storm. The largest of the observed regional dust storms and all of the planet-encircling dust storms have occurred almost exclusively in the spring and summer of the southern hemisphere (Kahn et al., 1992).

A number of pieces of observational evidence indicate that the past climate on Mars was not always the same as today's. For example, landforms such as valleys require the action of liquid water on or near the surface, implying that the atmosphere was once much thicker and warmer than now. Determining exactly when and how the climate was different is not an easy task without any of the proxy data that are still available in the case of Earth but not for Mars. There are two main lines of thought concerning past climatic variations. One is that the landforms indicative of liquid water date back from the early Mars, immediately after the formation of the planet. Therefore, the early history of Mars had a thick, warm and wet atmosphere, enough to allow liquid water (Pollack et al. 1987). The other asserts that Mars has experienced episodic climatic changes even in its late history; the hydrological processes occurring during such episodes produced the range of water-related landforms observed today (Baker et al. 1991).

Although the Martian magnetic field is not dipolar, but is crust remnant magnetization, sources with an equivalent magnetic moment as large as 1.3×10^{17} amp-m^2 exist (Acũna et al. 1999). The ionized atmosphere, the ionosphere, of Mars extends to high altitudes over the strong crustal magnetic fields. The crustal magnetic field is sufficiently strong to produce an asymmetric, time-varying obstacle to the flow of the solar wind around Mars.

16.24 SATELLITES

The two moons of Mars have general characteristics similar to those of other small bodies of the Solar System (Table 16.16), particularly the asteroids. Spectral studies

TABLE 16.16 Parameters of the Martian Satellites

	Phobos	Deimos
Radii of triaxial ellipsoid:	a: 13.3 ± 0.3 km	7.6 ± 0.5 km
	b: 11.1 ± 0.3 km	6.2 ± 0.5 km
	c: 9.3 ± 0.3 km	5.4 ± 0.5 km
Volume	5680 ± 250 km^3	1052 ± 250 km^3
Mass	1.08×10^{16} kg	1.8×10^{15} kg
Mean density	1.905 ± 0.053 g cm^{-3}	1.7 ± 0.5 g cm^{-3}
Semimajor axis	9378.5 km	23458.8 km
Eccentricity	0.0152	0.0002
Inclination	1°.03	1°.83
Orbital period	0.31891023 day	1.2624407 day
Length of day	Synchronous	Synchronous

Source: Based on Kieffer et al. 1992a.

indicate that Phobos and Deimos have a surface composition resembling that of carbonaceous asteroids, but with little bound or absorbed water. Phobos and Deimos have very different surface geologies. Phobos is heavily cratered and characterized by linear grooves that are related to the impact crater Stickney, whereas Deimos has a much smoother surface with subdued craters.

16.25 SEARCH FOR LIFE ON MARS

Mars has been the planet of choice when it comes to the image of extraterrestrial life in the minds of many people. Mars has very extreme environmental conditions in terms of its suitability for the survival of life, but many of these conditions are still the most similar in the Solar System to the Earth's. Studies of terrestrial life have increased our knowledge about the range of conditions in which life can survive (Rothschild and Mancinelli 2001). Life on Earth flourishes in hot springs, on sea floors, deep in subsurface rocks, and even in hypersaline lakes. This greatly increases the survivability of present or ancient life, if it ever existed, on Mars. The report of evidence of ancient life from a Martian meteroite ALH84001 (McKay et al. 1996), although not fully supported by the entire scientific community, stirred strong interest among scientists and the general public about life on Mars. However, the surface condition of Mars is extremely lethal to terrestrial-type organisms due to strong UV light from the Sun and the oxidizing conditions. Therefore, future investigations will focus on deep subsurface and hydrothermal systems as good candidates in the search for life on Mars. The earliest life on Earth is commonly considered to be hyperthermophile (Nisbet and Sleep 2001).

Whether life originated on Mars is another issue. For life, early Mars could have had a relatively safe environment, with oceans and a warm, wet atmosphere. However, the lack of knowledge on how life begins makes it difficult to assess Mars as a birthplace of life.

An interesting possibility exists for the transportation of life between planetary bodies. There is a group of meteorites called SNC meteorites that are considered to have come from Mars. These meteorites were ejected from Mars at a time in its history when giant impacts were common and so the inner planets were constantly

exchanging their surface materials. If microorganisms were shielded in the debris, it is possible that they survived the space trip between the planets. Therefore, life may have come from Mars to the Earth, or indeed vice versa.

16.26 EXPLORATION

Before the Space Age, the main means of observing Mars was to use ground-based telescopes. The history of Mars robotic exploration began with Mars 1 (Table 16.17). Important early results were obtained from missions such as Mariner 9 and the Viking orbiters and landers. Recently, Mars Pathfinder and Mars Global Surveyor have started a series of new explorations that is supposed to pave the way to eventual human flight to the red planet. The Mars Pathfinder lander was equipped with a panoramic camera, a meteorological device, and an alpha particle/proton/X-ray spectrometer for chemical analysis. The Mars Global Surveyor orbiter was equipped with a camera with high spatial resolution, 1.5 m. Other major instruments included a laser altimeter for a topographic survey and a thermal emission imaging spectrometer for mapping surface minerals. Future missions will focus more on detailed analyses of various environmental and geological properties and exobiological investigations of the planet. For example, the European Mars Express mission will conduct a survey of the upper several kilometers of the Martian crust using a ground-penetrating radar called MARSIS (Mars Advanced Radar for Subsurface and Ionospheric Sounding). Future rovers will have a long-range capability, up to some tens of kilometers, to enable flexible investigations of the geology in the landing areas. Sample return missions are also under consideration.

One of the major objectives in the upcoming Mars exploration program is the search for present life or evidence of ancient life on Mars. Many future missions will examine areas of recent water seepage (Malin and Edgett 2000), hydrothermal vents around volcanoes, and sedimentary basins and related water ponding.

Humans visiting Mars has been a dream for many people for generations. It has been envisioned in various science fiction books and movies and discussed extensively among space scientists and engineers. Without a political will such as that which drove the Apollo project during the Space Race, the possibility of human missions to Mars in the near future is slim. But other factors, such as the future

TABLE 16.17 List of Major Spacecraft Missions to Mars

Name	Country	Launch date	Accomplishments
Mars 1	U.S.S.R.	Nov. 1, 1962	First flyby in 1963
Mariner 4	U.S.A.	Nov. 28, 1994	First photographs
Mars 3	U.S.S.R.	May 28, 1971	First landing
Mariner 9	U.S.A.	May 30, 1971	Photographed extensively
Viking 1	U.S.A.	Aug. 20, 1975	Successful orbiter and lander mission
Viking 2	U.S.A.	Sep. 5, 1975	Successful orbiter and lander mission
Phobos 2	U.S.S.R.	July 12, 1988	Orbital investigation of surface
Mars Pathfinder	U.S.A.	Dec. 4, 1996	Lander and rover
Mars Global Surveyor	U.S.A.	Nov. 7, 1996	Extensive mapping of surface

discovery of life on Mars or major technological advancements, may change this situation.

The main technological obstacles to human flights to Mars are the long duration, the harsh space and surface environments, and fuel. But all of these are soluble, considering the ingenuity of the human race. The real question is when and why we will undertake the venture. The Apollo project was a huge success, and it created a view that the Earth is an oasis floating in the barren Universe. What kind of vision will human flights to Mars create? We now live in a more globally connected world than in 1969, the year of the Moon Landing. The journey to Mars may thus be conducted with the joint will of people worldwide rather than under conditions of competition between two nations. If this is so, the journey of humans to Mars will signify the arrival of a new way of human thinking.

REFERENCES

Acuña, M. H., Connerney, J. E. P., Ness, N. F., Lin, R. P., Mitchell, D., Carlson, C. W., McFadden, J., Anderson, K. A., Rème, H., Mazelle, C., Vignes, D., Wasilewski, P., and Cloutier, P. 1999. "Global Distribution of Crustal Magnetism Discovered by the Mars Global Surveyor MAG/ER Experiment." *Science,* vol. 284, pp. 790–793

Baker, V. R., Strom, R. G., Gulick, V. C., Kargel, J. S., Komatsu, G., and Kale, V. S. 1991. "Ancient Oceans and Ice Sheets and the Hydrological Cycle on Mars," *Nature,* vol. 352, pp. 589–594.

Bandfield, J. L., Hamilton, V. E., and Christensen, P. R. 2000. "A Global View of Martian Surface Compositions from MGS-TES," *Science,* vol. 287, pp. 1626–1630.

Banin, A., Clark, B. C., and Wänke, H. 1992. "Surface Chemistry and Mineralogy," in *Mars,* ed. Kieffer et al., University of Arizona Press, Tucson, pp. 594–625.

Carr, M. H. 1996. *Water on Mars,* Oxford University Press, New York.

Christensen, P. R., Bandfield, J. L., Clark, R. N., Edgett, K. S. Hamilton, V. E., Hoefen, T., Kieffer, H. H., Kuzmin, R. O., Lane, M. D., Malin, M. C., Morris, R. V., Pearl, J. C., Pearson, R., Roush, T. L., Ruff, S. W., and Smith, M. D. 2000. "Detection of Crystalline Hematite Mineralization on Mars by the Thermal Emission Spectrometer: Evidence of Near-Surface Water," *Journal of Geophysical Research,* vol. 105, pp. 9623–9642.

Fanale, F. P. 1976. Martian Volatiles: Their Degassing History and Geochemical Fate," *Icarus,* vol. 28, pp. 179–202.

Greeley, R., Lancaster, N., Lee, S., and Thomas, P. 1992. "Martian Eolian Processes, Sediments and Features," in *Mars,* ed. Kieffer et al., University of Arizona Press, Tucson, pp. 730–766.

Hartmann, W. K. and Berman, D. C. 2000. "Elysium Planitia Lava Flows: Crater Count Chronology and Geological Implications," *Journal of Geophysical Research,* vol. 105, pp. 15,011–15,026.

Kahn, R. A., Martin, T. Z., Zurek, R. W., and Lee, S. W. 1992. "The Martian Dust Cycle," in *Mars,* ed. Kieffer et al., University of Arizona Press, Tucson, pp. 1017–1053.

Kargel, J. S. and Strom, R. G. 1992. "Ancient Glaciation on Mars," *Geology,* vol. 20, pp. 3–7 (1992).

Kieffer, H. H., Jakosky, B. M., and Snyder, C. W. 1992a. "The Planet Mars: From Antiquity to the Present," in *Mars,* ed. Kieffer et al., University of Arizona Press, Tucson, pp. 1–33.

Kieffer, H. H., Jakosky, B. M., Snyder, C. W., and Matthews, M. S. 1992b. *Mars,* University of Arizona Press, Tucson.

Lasker, J. and Robutel, P. 1993. "The Chaotic Obliquity of the Planets," *Nature,* vol, 361, pp. 608–612.

EARTH'S ENVIRONMENT AND SPACE

Malin, M. C. and Edgett, K. S. 2000. Evidence for Recent Groundwater Seepage and Surface Runoff on Mars," *Science,* vol. 288, pp. 2330–2335.

McKay, D. S., Gibson, E. K., Thomas-Keprta, K. L., Vali, H., Romanek, C. S., Clemett, S. J., Chillier, X. D. D., Maechling, C. R., and Zare, R. N. 1996. "Search for Life on Mars: Possible Relic Biogenic Activity in Martian Meteorite ALH84001, *Science,* vol. 273, pp. 924–930.

Melosh, H. J. 1993. Blasting Rocks off Planets," *Nature,* vol. 363, pp. 498–499.

Melosh, H. J. and Vickery, A. M. 1989. "Impact Erosion of Primordial Atmosphere of Mars," *Nature,* vol. 338, pp. 487–489.

Mutch, T. A., and Head, J. W. 1975. "The Geology of Mars: A Brief Review of Some Recent Results," *Reviews of Geophysics and Space Physics,* vol. 13, pp. 411–416.

Nisbet, E. G. and Sleep, N. H. 2001. "The Habitat and Nature of Early Life," *Nature,* vol. 409, pp. 1083–1091.

Owen, T. 1992. "The Composition and Early History of the Atmosphere of Mars," in *Mars.* ed. Kieffer et al., University of Arizona Press, Tucson, pp. 818–834.

Parker, T. J., Saunders, R. S., and Schneeberger, D. M. 1989. "Transitional Morphology in West Deuteronilus Mensae, Mars: Implications for Modifications of the Lowland/Upland Boundary," *Icarus,* vol. 82, pp. 111–145.

Pollack, J. B., Kasting, J. F., Richardson, S. M., and Poliakoff, K. 1987. "The Case for a Wet, Warm Climate on Early Mars," *Icarus,* vol. 71, pp. 203–224.

Rieder, R., Economou, T., Wänke, H., Turkevich, A., Crisp, J., Brückner, J., Dreibus, G., and McSween, H. Y., Jr. 1997. "The Chemical Composition of Martian Soil and Rocks Returned by the Mobile Alpha Proton X-ray Mode," *Science,* vol. 278, pp. 1771–1774.

Rothschild, L. J. and Mancinelli, R. L. 2001. "Life in Extreme Environments," *Nature,* vol. 409, pp. 1092–1101.

Smith, D. E., Sjogren, W. L., Tyler, G. L., Balmino, G., Lemoine, F. G., and Konopliv, A. S. 1999. "The Gravity Field of Mars from Mars Global Surveyor," *Science,* vol. 286, pp. 94–97.

Smith, D. E., Zuber, M. T., Solomon, S. C., Phillips, R. J., Head, J. W., Garvin, J. B., Banerdt, W. B., Muhleman, D. O., Pettengill, G. H., Neumann, G. A., Lemoine, F. G., Abshire, J. B., Aharonson, O., Brown, C. D., Hauck, S. A., Ivanov, A. B., McGovern, P. J., Zwally, H. J., and Duxbury, T. C. 1999b. "The Global Topography of Mars and Implications for Surface Evolution," *Science,* vol. 284 pp. 1495–1503.

Tanaka, K. L., Scott, D. H., and Greeley, R. 1992. "Global Stratigraphy," in *Mars,* ed. Kieffer et al., University of Arizona Press, Tucson, pp. 345–382.

Zimbelman, J. R. 2000. "Non-active Dunes in the Acheron Fossae Region of Mars between the Viking and Mars Global Surveyor Areas," *Geophysical Research Letters,* vol. 27, pp. 1069–1072.

Zurek, R. W. 1992. Comparative Aspects of the Climate of Mars: An Introduction to the Current Atmosphere," in *Mars,* ed. Kieffer et al., University of Arizona Press, Tucson, pp. 799–817.

Zurek, R. W., Barnes, J. R., Haberle, R. M., Pollack, J. B., Tillman, J. E., and Leovy, C. B. 1992. "Dynamics of the Atmosphere of Mars," in *Mars,* ed. Kieffer et al., pp. 835–933.

PART 6

THE SUN–EARTH CONNECTION

D. N. Baker

16.27 INTRODUCTION

Humans have been fascinated since time immemorial by the Sun and its relationship to the Earth. Every civilization has speculated about the place of Earth in the cosmos. Solar and lunar eclipses have instilled a sense of awe from ancient times, and structures such as Stonehenge have shown the compelling inspiration that results as people try to understand Sun–Earth connections. Certainly in the past few hundred years, since Copernicus and Galileo, we have come to new scientific understandings of the Sun and its broader place in the realm of the stars. We have also come to a clearer picture of the Sun as a driver of magnetic disturbances at Earth. Especially in the past few decades, scientists have arrived at a much clearer picture of solar activity and its effects on Earth and on human technological systems.

The Sun reaches an activity peak every 11 years or so, during which it exhibits a wide variety of powerful disturbances. The shock waves and clouds of hot, ionized gas (plasma) emitted by the Sun can be extremely energetic and can accelerate interplanetary charged particles to millions or even billions of electron volts. If such disturbances happen to strike the Earth and its magnetosphere, the result can be violent geomagnetic storms that can have profound effects on the near-Earth environment that we call *geospace*.

As understanding of the global Sun–Earth system has evolved, human technological systems have also grown more sophisticated. A complex web of communication links connects virtually all points on the Earth's surface, power grids criss-cross the continents, and spacecraft pass through much of the volume in the near-Earth space above us. The ''cyberelectric cocoon'' that has been spun around the Earth has greatly enhanced global communications, navigational accuracy, national security, and weather prediction. However, reliance on these systems has also put human society into danger. Every technological system that utilizes radio waves or conducting wires or computer processors can be adversely affected by powerful disturbances in the solar–terrestrial system. Given the growing human dependence on advanced technologies, it is likely that the Sun–Earth connection will be increasingly important in the coming decades.

This portion of this section discusses the connected Sun–Earth system. The Sun is examined as the basic driver of the system, and the interplanetary medium is briefly discussed. The near-Earth space environment is described with an eye toward understanding the ultimate effects of solar variations on the atmosphere. Finally, near-Earth technologies are considered in the light of new geospace understandings.

16.28 THE SUN AND THE HELIOSPHERE

As illustrated in Figure 16.26, the Sun is a complex and highly structured object. At the center is the hydrogen-burning core. The nuclear fusion proceeding within

EARTH'S ENVIRONMENT AND SPACE

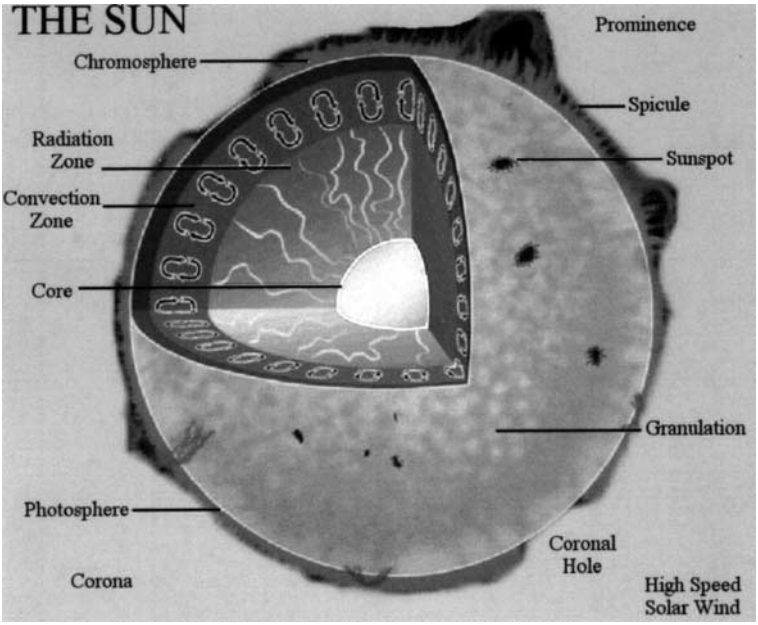


FIGURE 16.26 A schematic diagram showing the principal features of the Sun, including its interior and its outer atmosphere (adaptation of figure courtesy of P. Dusenbery, Space Science Institute).

this core gives rise to temperatures in excess of 10 million degrees Kelvin (K). Outside the core is the radiation zone, where energy is transported rapidly toward the outer layers of the Sun. In the convection zone, as shown in Figure 16.26, heated solar plasma percolates toward the surface and cools somewhat, thereby setting up large-scale circulatory eddies. Such convective motions produce powerful currents and related magnetic fields.

The visible surface of the Sun, called the photosphere, has a temperature of about 6,000 K. There are, however, regions in the photosphere of somewhat cooler plasma which appear slightly darker. Called sunspots, these are broad regions of strong, emergent magnetic field. The number of sunspots varies quite substantially over the course of a period of 10–12 years (average value 11 years) that is known as the solar cycle. At the time of sunspot minimum, the photosphere appears rather homogeneous and uniform (aside from small-scale granulations) and solar activity is very low. At the time of solar maximum, however, the number of sunspots is high, solar activity is very prevalent, with many bright, intense regions of X-ray and ultraviolet emissions, and there are frequent eruptive disturbances with powerful large-scale prominences.

The solar layer above the photosphere is called the chromosphere. It is normally visible from Earth only during solar eclipses. The hot, outermost layer of the Sun (directly above the chromosphere) is called the solar corona. There is a rapid temperature transition in the Sun's atmosphere, from 6,000 K at the surface to over 1 million K in the corona. The hot gases of the corona escape outward toward the cold void of interplanetary space and in so doing form the solar wind. The solar

wind is a highly ionized gas plasma formed mostly of protons and electrons with an important minor proportion of helium and other trace constituent elements up to iron. Quite importantly, the solar wind has embedded within it the interplanetary magnetic field, which is "frozen" into the hot plasma flow.

From the base of the corona to a few solar radii heliocentric distance, the solar wind is accelerated to supersonic speeds. How this occurs, in detail, is still an essentially unsolved mystery. Near the poles of the Sun, the solar wind emerges from regions called coronal holes. In these regions of weak X-ray emissions, the magnetic flux tubes extend far out into interplanetary space and hence are called "open" field lines. The solar wind from coronal holes is observed to have high flow speeds (600–800 km/s) and is rather cool. At lower solar latitudes, the magnetic flux tubes are generally "closed" in that they emerge from the solar surface and then return in looping arcs back to the surface. The solar wind plasma from the lower latitude portion of the Sun has lower characteristic speeds (300–500 km/s) and normally is hotter and denser than the plasma flowing out of coronal holes.

During certain portions of the solar activity cycle, especially in the approach to sunspot minimum conditions, coronal holes can extend to low solar latitudes in limited longitudinal sectors of the Sun. The high-speed flows from these regions can give rise to solar wind streams near the ecliptic plane. Such streams can have a significant effect upon the Earth as the streams propagate outward from the Sun (shown by the dark-shaded regions in Figure 16.27). On the other hand, at solar

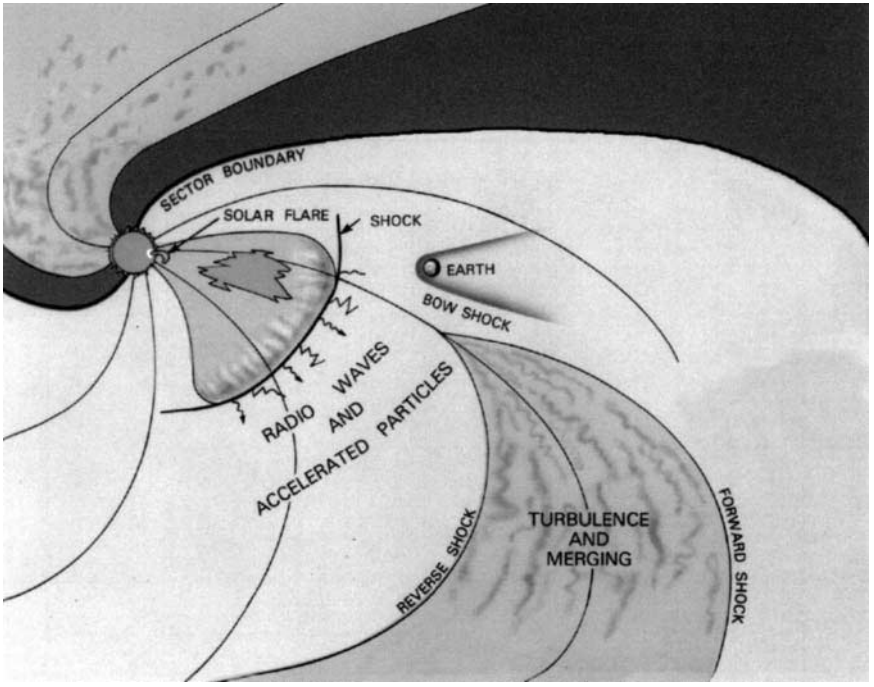


FIGURE 16.27 A schematic diagram of the heliosphere and some of its structural elements near the ecliptic plane (adaptation of diagram from NASA/GSFC).

EARTH'S ENVIRONMENT AND SPACE

maximum conditions, solar flares and large expulsions of coronal material (coronal mass ejections, CMEs) can occur on the Sun, as also illustrated schematically in Figure 16.27. The flares and CMEs tend to have a major influence throughout broad regions of the heliosphere, the huge volume of the Sun's extended plasma influence.

Figure 16.27 is an ecliptic-plane cut of the three-dimensional heliosphere, which has been mapped by such space missions as Ulysses (which has flown over the poles of the Sun at ~1 AU heliocentric distance) and by the Pioneer and Voyager missions, which are presently many tens of AU away from the Sun. Observations and modeling suggest that the heliosphere extends outward to perhaps 80–100 AU, where the solar wind flow ends in a termination shock. Beyond this is the boundary of the Sun's effective plasma influence, called the heliopause. Depending upon the speed of the Sun's motion through the local interstellar medium (LISM), i.e., whether it is supersonic or subsonic, there may also be a standing bow shock wave in front of the heliosphere.

As suggested by Figure 16.27, the magnetic and plasma structure of the heliosphere plays an important role in driving terrestrial magnetospheric processes. The time variations associated with solar flares, CMEs, and interplanetary shock waves mingle with the more stable features such as high-speed streams and sector boundaries. This entire pattern of structures and transient features tends to rotate with the 27-day period of the Sun. The Earth is bathed in many of these solar and heliospheric features in an ever-changing pattern of plasma variations and energy fluctuations.

16.29 STRUCTURE AND DYNAMICS OF THE MAGNETOSPHERIC SYSTEM

Each of the planets in our solar system with an intrinsic magnetic field (Mercury, Earth, Jupiter, Saturn, Uranus, and Neptune) has a region around it called the planet's magnetosphere. (Even the unmagnetized planets Venus and Mars have magnetosphere-like regions around them.) Figure 16.28 is a schematic diagram of the Earth's magnetosphere: this serves as a comparative basis for all the known planetary systems. On the sunward side—extending outward about 10 Earth radii (1 R_E = 6372 km, mean)—the Earth's magnetic field stands off the solar wind flow. The boundary between the solar plasma and the terrestrial influence is called the magnetopause. The solar wind flow distorts the Earth's field on the antisolar side into an elongated magnetotail. Inside the magnetotail, near the central plane, is a region of hot, dense plasma called the plasma sheet. At the center of the plasma sheet is a region of magnetic field reversal called the neutral sheet. On either side of the plasma sheet are huge cylindrical regions of strong magnetic fields and very low plasma densities called the tail lobes. The tail lobe magnetic fields are produced by solenoidal currents of large scale that flow across the width of the plasma sheet (i.e., the cross-tail currents) and then close over the top and bottom surfaces of the magnetotail to form the nightside magnetopause currents.

The magnetic flux tubes of the outer magnetosphere all tend to funnel back toward the Earth and connect to the uppermost part of the atmosphere (see Figure 16.28). The ionized upper layers of the atmosphere are called the ionosphere. There are powerful electrical currents flowing into, out of, and within the ionosphere. These field-aligned currents and electrojet currents provide a major avenue of energy coupling between the magnetosphere and the ionosphere and lead to strong Joule heating and energy dissipation.

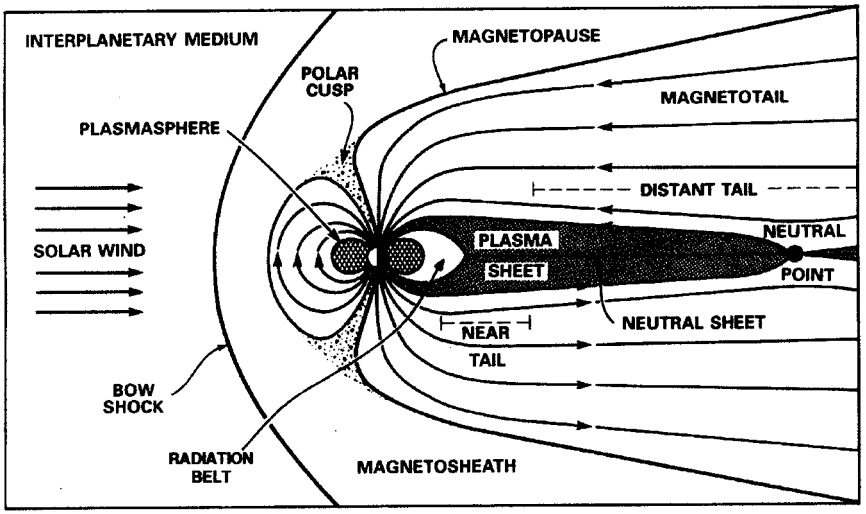


FIGURE 16.28 A noon–midnight cross-sectional view of the Earth's magnetosphere.

The inner part of the Earth's magnetosphere has a very strong, essentially dipolar magnetic field configuration. The strongly confined plasmas in this region drift under the magnetic field gradients and curvature forces to form the extraterrestrial ring current. This is also a region of high, steady plasma density, called the plasmasphere. As shown more explicitly in Figure 16.29, the part of the magnetosphere closest to the Earth also is the region of very energetic charged particle confinement, called the Van Allen radiation belts. The Van Allen belts contain highly time variable fluxes of relativistic electrons and high-energy protons. The inner belt (Figure

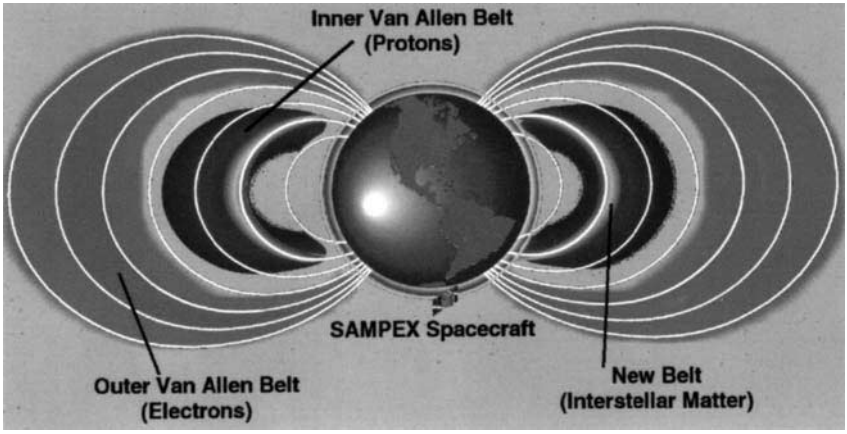


FIGURE 16.29 A cross-sectional diagram of the Earth's radiation belts (adaptation of figure courtesy of R. Mewaldt).

EARTH'S ENVIRONMENT AND SPACE

16.29) is mainly composed of very energetic protons, while the outer belt is composed principally of relativistic electrons.

16.30 THE SOLAR–TERRESTRIAL ENERGY CHAIN

The energy that drives the geospace environment has its origins at the Sun (as illustrated schematically in Figure 16.26). Nuclear fusion operating within the core subsequently drives convection in the outer layers of the Sun, and the associated dynamo action produces intense magnetic fields in the Sun's outer atmosphere. Eventually, emerging magnetic flux and hot, high-pressure gas are carried outward in the Sun's corona to form the expanding solar wind flow. This magnetized solar wind (after a transit time of several days) interacts with the outer parts of Earth's magnetic field and imparts—in a highly variable way—the energy to drive all forms of magnetospheric, auroral, and ionospheric disturbances (see Figure 16.30). These disturbances can be manifested by changes in auroral luminosity and by the range of phenomena covered by the general term *geomagnetic activity*.

As Figure 16.30 makes obvious, the ultimate deposition of energy from the solar wind occurs in the atmosphere. It is clear that solar–terrestrial coupling extends at least into the lower thermosphere and the mesosphere. However, mounting evidence suggests that solar–terrestrial coupling extends deeply into the stratosphere as well. Thus, in many ways the energy flowing from the Sun plays a key role in many aspects of life on (and near) the Earth's surface.

As noted above in the discussion of Figure 16.27, there can be another chain of connection between the Sun and the Earth, apart from the usual solar wind link: this alternate route is the one associated with transient events on the Sun. Large solar flares and CMEs can produce copious quantities of energetic charged particles, X-rays, gamma rays, and ultraviolet (UV) light bursts, as well as fast solar wind

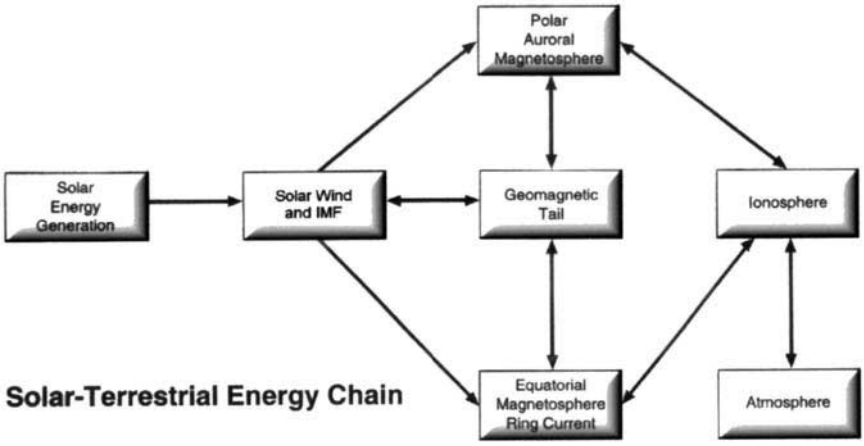


FIGURE 16.30 Schematic diagram showing the flow of energy from the Sun into the Earth's ionosphere–atmosphere system.

flows that can have the most extreme effects on the terrestrial environment. It is often these extreme solar events that are of the greatest concern in terms of severe consequences to human life and human technological systems.

16.31 DYNAMICS OF THE MAGNETOSPHERE-IONOSPHERE-ATMOSPHERE SYSTEM

The elementary episode of energy coupling from the solar wind to the magnetosphere and the subsequent dissipation of this energy is called a *magnetospheric substorm.* Initial phenomenological descriptions of substorms have been followed in subsequent decades by many debates and controversies about the nature and causes of these events. Figure 16.31 shows the flow of energy into and through the magnetosphere-ionosphere system during substorms. The energy to drive substorms comes from solar wind input, which is increased when the interplanetary magnetic field (IMF) turns southward. The left side of Figure 16.31 shows the substorm features related to the magnetotail. There is rather broad agreement that the magnetosphere exhibits a ''growth phase'' and thus stores energy in the magnetotail

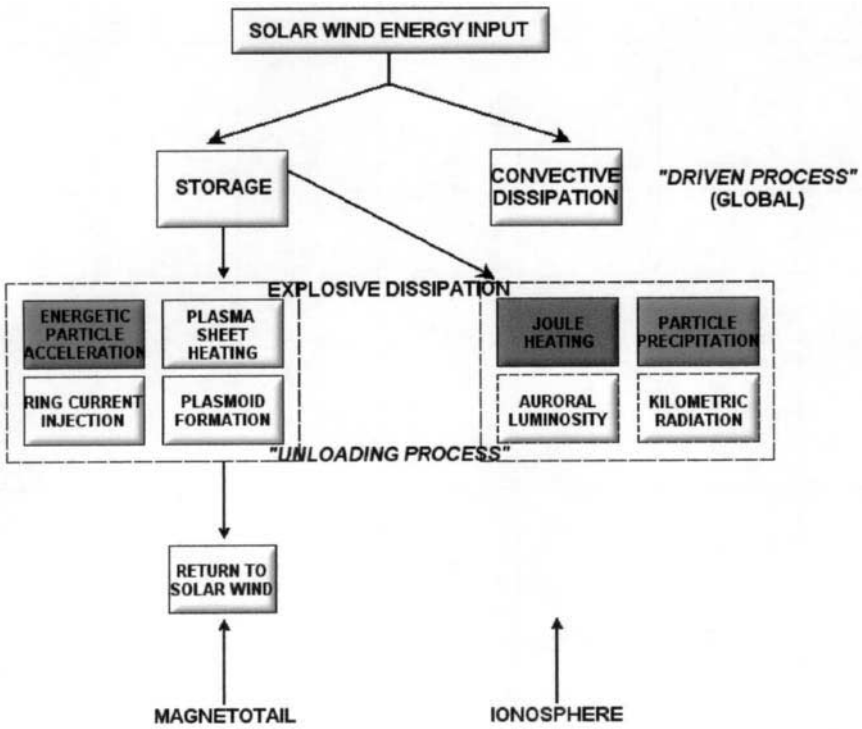


FIGURE 16.31 A detailed flow diagram of energy in the magnetosphere–ionosphere system during magnetospheric substorms.

EARTH'S ENVIRONMENT AND SPACE

prior to substorm onset. The right side of Figure 16.31 shows the substorm processes occurring in the ionosphere. The directly driven aspects of substorms are manifested in the ionosphere by dissipation associated with the global convection process. As this convective dissipation proceeds, the magnetotail is concurrently loaded with energy. This makes the storage process (the substorm growth phase) an integral part of the driven portion of substorms.

An essential aspect of substorms shown in Figure 16.31 is the explosive dissipation that constitutes the unloading process. This occurs in the substorm expansion phase and is localized initially on the nightside of the Earth. There are many specific dissipation processes that can be identified during the typical energy unloading event, including plasma sheet heating, ring current injection, energetic charged particle bursts, and plasmoid formation and release. In addition to plasmoids that carry magnetotail energy back downtail to the solar wind, some portion of the stored energy may also be so far tailward that it is not geoeffective. Explosive dissipation is strikingly obvious in the ionosphere, as well as in the magnetotail, during substorm expansion phases. There is substantial energy dissipation in a substorm for a typical period of one to two hours. The ionospheric energy dissipation is predominantly in the form of Joule heating and charged particle precipitation. However, there is also energy dissipation through auroral luminosity and kilometric radiation.

Figure 16.32 shows the sequence of magnetotail states that occur during the course of a magnetospheric substorm. The various panels illustrate the noon–midnight meridian plane. Magnetic field lines are sketched, and the gray-shaded region in each panel represents the plasma sheet. The diagram shows how during a substorm the magnetotail: (a) occupies an initial or ground state; (b) undergoes a stress-developing growth phase; (c) exhibits an explosive onset of tail magnetic reconnection; (d) diverts cross-tail current during the expansion phase and pinches off much of the plasma sheet to form a plasmoid; and (e) eventually relaxes and recovers to something approaching the initial state. The overall timescale for the sequence depicted in Figure 16.32 to occur is typically two to three hours.

As described above, the passage of solar wind containing southward interplanetary magnetic field imparts energy into the near-Earth space environment and produces dynamical responses of the magnetosphere, ionosphere, and atmosphere. A great deal is known about the responses of geospace to these inputs as a result of past satellite and ground-based observing programs. Episodes of southward IMF drive strongly enhanced convection of plasmas and magnetic fields within the magnetosphere; a strong extraterrestrial ring current is formed as plasma sheet particles drift deep into the inner magnetosphere on open drift paths. This strong ring current is a defining feature of geomagnetic storms. The drift inward takes some three to four hours. During this time the solar wind drivers have usually subsided, the convection has consequently decreased, and the charged particles have begun to move along closed drift paths. The symmetric ring current decays slowly through collisions with the hydrogen geocorona. Ring current ions also experience wave-particle interactions, which cause the recovery to quiet conditions. Figure 16.33 shows a rather typical geomagnetic storm interval in which the Dst index (a commonly used index of ring current strength and geomagnetic storm development) is plotted along with concurrent solar wind and IMF data. The September 1998 storm portrayed developed in a few hours and decayed over a several-day interval.

Enhanced convection stresses the magnetosphere to the point that instabilities are common both in the near-tail and in the mid-tail regions, producing dramatic reconfigurations of the tail's basic structure (as shown in Figure 16.32). In the near-tail, strong currents are diverted through the high-latitude ionosphere and produce

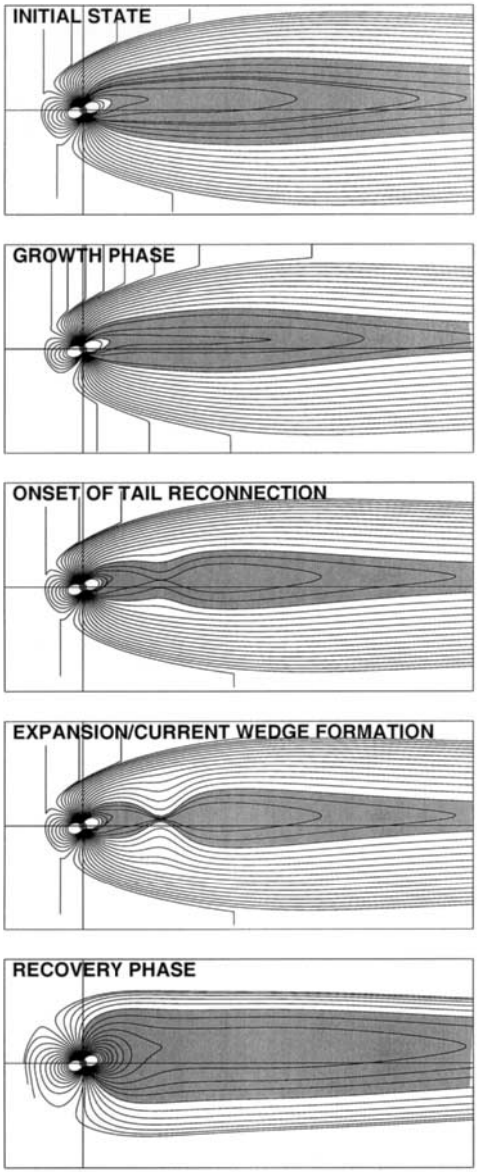


FIGURE 16.32 Diagram illustrating the sequence of steps occurring in the Earth's magnetotail during substorms.

EARTH'S ENVIRONMENT AND SPACE

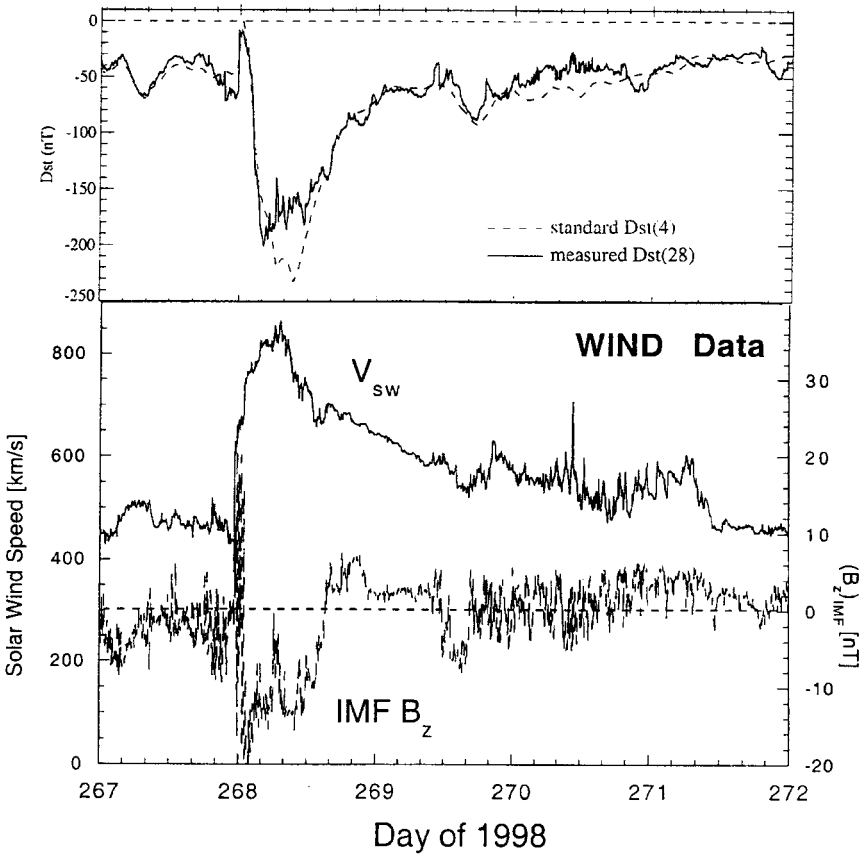


FIGURE 16.33 A typical geomagnetic storm development as seen in the Dst index (during late September 1998). Solar wind speed (V_{SW}) and southward component of the IMF data are also shown.

bright auroral displays. In the mid-tail, field lines merge, accelerating plasma earthward at high velocity and ejecting plasmoids downtail into the solar wind. The earthward flows from the mid-tail are thought to trigger the current disruption in the near-tail. Dipolarization of the magnetic field during substorms generates intense induction electric fields, which accelerate charged particles to high energies in the near-Earth magnetotail. Plasma sheet material moves into, around, and through the inner magnetosphere, modulating the intensity and composition of the ring current. Plasma composition measurements inside 30 R_E indicate that the plasma sheet contains substantial amounts of both solar wind and ionospheric plasma.

Radiation belt particles respond adiabatically to the magnetic field perturbations associated with the building stormtime ring current. If an enhanced solar wind dynamic pressure moves the dayside magnetopause to small radial distances as the storm develops, the radiation belt particles drifting rapidly on closed trajectories may encounter the magnetopause and be lost. The radiation belts rebuild to much

higher levels in association with high-speed streams (if they occur) following the storm maximum. The exact mechanism or collection of mechanisms by which magnetospheric charged particles are accelerated to relativistic energies are unknown at present, although several promising candidates exist. It appears that ULF waves in the magnetosphere which are driven by the high-speed solar wind streams may play a key role in the acceleration of substorm particle populations to relativistic energies. If a ring current develops while the radiation belts are enhanced, the associated magnetic field distortion scatters the radiation belt particles to small pitch angles where they encounter the high densities at the foot of the magnetic field lines and are lost. Dramatic losses from the inner belts may also result from the interaction of whistlers (generated by lightning) with the relativistic electrons there; this mechanism introduces an interesting direct coupling with weather systems in the lower atmosphere.

16.32 IMPORTANCE OF ATMOSPHERIC COUPLING

The neutral atmosphere is not a passive element in the coupled near-Earth system. It is heated at high latitudes by both auroral currents and charged particle precipitation associated with the substorm current systems, thereby expanding to higher altitudes and modifying its composition. The temperature and composition changes move in waves from the conjugate auroral regions and meet at the equator. The changing neutral atmosphere modifies the ionospheric plasma. These variations alter the conductivity of the ionospheric plasma and thus actively modulate the current flow between the ionosphere and magnetosphere. The heating of the high-latitude upper atmosphere in response to auroral inputs during active times can exceed solar inputs. The low-altitude ionosphere is enhanced in the auroral region through impacts of precipitating magnetotail particles on the neutral atmosphere. Large currents ($>10^6$A) flow along the auroral oval, subauroral electric fields are generated, and for short intervals these can extend all the way down to the equator. Thus, magnetosphere–atmosphere coupling is an important last leg of the Sun–Earth connection issue.

During the last three decades, important efforts have been underway to develop an understanding of variations in middle-atmospheric ozone (O_3) due to natural processes and to the effects of humankind. It is now reasonably well understood that the global balance of O_3 is governed in part by the balance between the production of O_3 and its destruction by reactions within the NO_y, Cl_y, HO_y, and O_y chemical families. One of the most important catalytic cycles in this global balance is that due to the following reactions:

$$NO + O_3 \rightarrow NO_2 + O_2$$

$$\underline{NO_2 + O \rightarrow NO + O_2}$$

$$O_3 + O \rightarrow 2O_2$$

The bulk of the odd nitrogen (i.e., NO_y, oxides of nitrogen) in the stratosphere is formed from the oxidation of N_2O by O (^{1}D), forming NO. It is known, however,

EARTH'S ENVIRONMENT AND SPACE

that middle-atmospheric NO_y is formed by ion chemistry initiated by the precipitation of energetic charged particles into the upper atmosphere (see Figure 16.34).

The ion and neutral odd-nitrogen chemistry initiated by energetic charged particles produces secondary electrons, e^*, which in turn ionize and dissociate the major atmospheric species. This ionization is followed by a series of recombination reactions involving nitrogen and its ions, which produce additional atomic nitrogen. The resulting atomic nitrogen may be in either the ground level or an excited level. Nitric oxide is formed by the reaction of atomic nitrogen with O_2. Reactions involving excited nitrogen are faster than those with the ground-state N atom. The destruction of odd nitrogen proceeds through the reaction of N atoms with NO to produce molecular nitrogen and atomic oxygen (Figure 16.34). In the sunlit atmosphere the photodissociation of NO is important; however, in the fall, winter, and spring in the polar region, the photolytic reaction is negligible and the resulting NO lifetime is sufficient for downward transport to bring the NO into the mesosphere and stratosphere.

Global atmospheric models driven by the measured flux of precipitating energetic electrons indicate that column densities of NO_y can increase by 20–40% near 25 km due to the electron inputs. This in turn can lead to O_3 reductions at those altitudes by up to ~20% compared to ozone levels expected without energetic electron inputs. These results imply that solar-wind-magnetosphere-atmosphere coupling is a key determinant of the overall chemistry of the middle atmosphere. This charged particle linkage therefore represents a solar-terrestrial coupling mechanism which, for the last solar cycle, was as important to stratospheric O_3 as were solar UV flux variations. Thus, such recent work shows that the solar-terrestrial energy

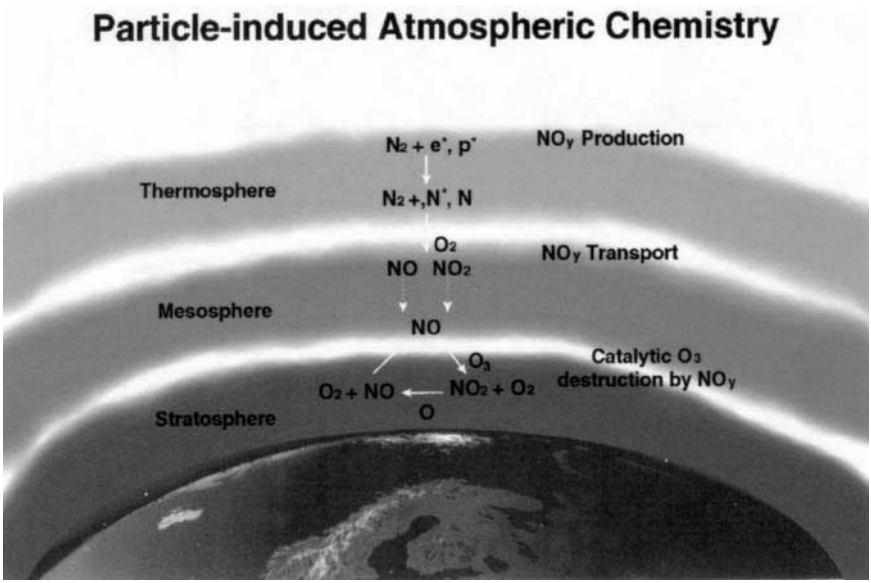


FIGURE 16.34 Charged particle-induced chemical changes in the various layers of the Earth's atmosphere and the catalytic destruction of ozone (O_3).

coupling chain extends down into the stratosphere, just above the troposphere, that part of the Earth's atmosphere which determines weather and climate.

16.33 SUN–EARTH CONNECTIONS AND HUMAN TECHNOLOGY

The effects of the Sun on the near-Earth space environment often are considered most significant in terms of their influences on human technology and humans in space. The National Space Weather Program strategic plan of the United States speaks about the wide-ranging impacts of space weather on satellites, power grids, communications systems, and many other technologies. Certainly, astronauts in space and airline passengers and crew on high-flying transpolar routes can be adversely affected by both solar disturbances and geospace responses to these.

Solar flares and solar energetic charged particles can damage solar cells and thereby greatly shorten satellite lifetimes. The heating of the upper atmosphere that occurs due to solar flares substantially increases atmospheric drag on low-altitude satellites. Ionospheric disturbances cause radio signal scintillations and "blackouts" in certain radio frequency ranges. Some of the largest solar-induced magnetic storms disrupt electrical power grids (as occurred in the Canadian province of Quebec in March 1989). All of these can have paralyzing large-scale consequences for human activities.

As another example, the Galaxy 4 spacecraft was a heavily used communications satellite at geostationary orbit; its sudden failure in May 1998 caused the loss of pager service to some 45 million customers as well as numerous other communications outages. Analysis by operators and builders continues as to the exact cause of the Galaxy 4 failure. Using a wide array of space data sets, researchers analyzed the magnetospheric and solar wind conditions during the April–May 1998 period. There was large solar and magnetic activity in early May. Strong evidence was found that the fluxes of highly relativistic electrons were substantially elevated above average conditions for a period of about two weeks prior to the May 19 failure of Galaxy 4. Thus, evidence was presented that the internal (deep-dielectric) charging mechanism acting on electronic chips may have played an important role in the Galaxy 4 failure.

There were several other failures and spacecraft anomalies in early to mid-May 1998. Whether the Galaxy spacecraft failure incident was or was not due to a space weather effect, it clearly showed the vulnerability of modern society to individual spacecraft failures. The large number of users affected by the loss of just the one Galaxy spacecraft shows how dependent society has become on space technology and how fragile modern communication systems can be. The Galaxy failure had a large impact because the spacecraft was optimally located over the central United States and could best handle digital pager signals. Therefore, 80% of all pager traffic was directed through it. Increasingly, phones, TV, radio, bank transactions, newspapers, credit card systems, etc. all depend upon satellites for some part of their communications links rather than being all ground-based. It seems unwise to have complex, societally significant systems susceptible to single-spacecraft failures.

EARTH'S ENVIRONMENT AND SPACE

16.34 SUMMARY

Study of the near-Earth space environment and its relationship to the Sun has come a tremendous distance in recent decades. Today, owing to a wealth of new observational tools and modeling methods, there is a much more profound understanding of the Sun, of major solar disturbances, and of how these disturbances affect the Earth. Over the course of the next solar cycle (see Figure 16.35), we can expect even further improvements in our understanding of Sun–Earth connections.

Study of the Sun–Earth system has improved because of the number and quality of observing platforms that are available compared to earlier times. The ability to observe the Sun nearly continuously, and especially to observe the most active regions on the Sun, is important for space weather predictions. When a major solar disturbance has occurred, it is also important to track the subsequent interplanetary motion of solar outbursts and to model the motion of solar charged particles and fields. Finally, as the solar outputs reach the vicinity of Earth, there is a need to use the large array of observation platforms available and the global models of the

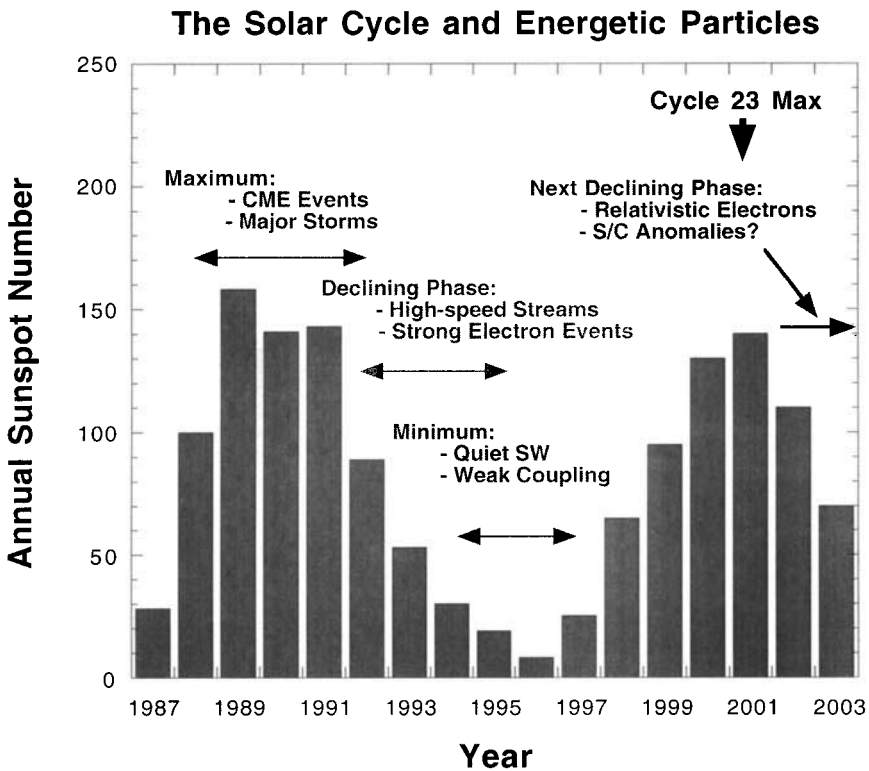


FIGURE 16.35 Portions of two solar activity cycles (1987–2003) as seen in sunspot number and showing their near-Earth consequences; after the maximum of the current solar cycle in 2001, will there be many examples of anomalous—dangerous—events to spacecraft?

magnetosphere-ionosphere-atmosphere system to predict and to assess their geospace consequences. All this is developing now and should, in the next solar cycle, improve quite dramatically (Figure 16.35). We can envision a program of Sun–Earth research with a thorough underpinning of basic physical understanding. Through a combination of data assimilation, numerical modeling and better definition of user needs, solar-terrestrial research may soon be at a stage of specification and forecasting of the near-Earth space environment that rivals modern tropospheric weather predictions.

FURTHER READING

Baker, D. N., "Critical Issues in Space Plasma Physics," *Physics of Plasmas,* vol. 6, pp. 1700–1708 (1999).

Baker, D. N. "Effects of the Sun on the Earth's Environment," *Journal of Atmospheric and Solar-Terrestrial Physics,* vol. 62, pp. 1669–1681 (2000).

Barth, C. A., "Nitric Oxide in the Lower Thermosphere," *Planetary and Space Science,* vol. 40, pp. 315–336 (1992).

Burlaga, L. F., Interplanetary Magnetohydrodynamics, Oxford University Press, New York (1995).

National Space Weather Program Strategic Plan, Office of the Federal Coordinator for Meteorological Services, NOAA, Silver Spring, MD (1995).

Parker, E. N., "Dynamics of the Interplanetary Gas and Magnetic Fields." *Astrophysical Journal,* vol. 128, pp. 664–676 (1958).

PART 7

SPACE DEBRIS

Rüdiger Jehn

16.35 INTRODUCTION

Space debris can be defined as any man-made Earth-orbiting or reentering object which is nonfunctional with no reasonable expectation of assuming or resuming its intended function, including fragments and parts thereof.

Since the launch of Sputnik on October 4, 1957, with more than 4,000 launches about 5,000 satellites have been placed in orbit. In total about 27,000 larger objects—satellites, rocket upper stages, mission-related objects like telescope covers or bolts, and fragments from in-orbit explosions have been observed by ground-based radar and telescopes. About 18,500 of these have burnt up in the atmosphere, leaving about 8,500 larger objects currently in Earth orbit. Figure 16.36 illustrates the steady increase of objects in Earth orbit.

Radar, optical, and infrared telescopes of the U.S. Space Surveillance Network track more than 1,000 objects every day, update their orbital elements, and maintain

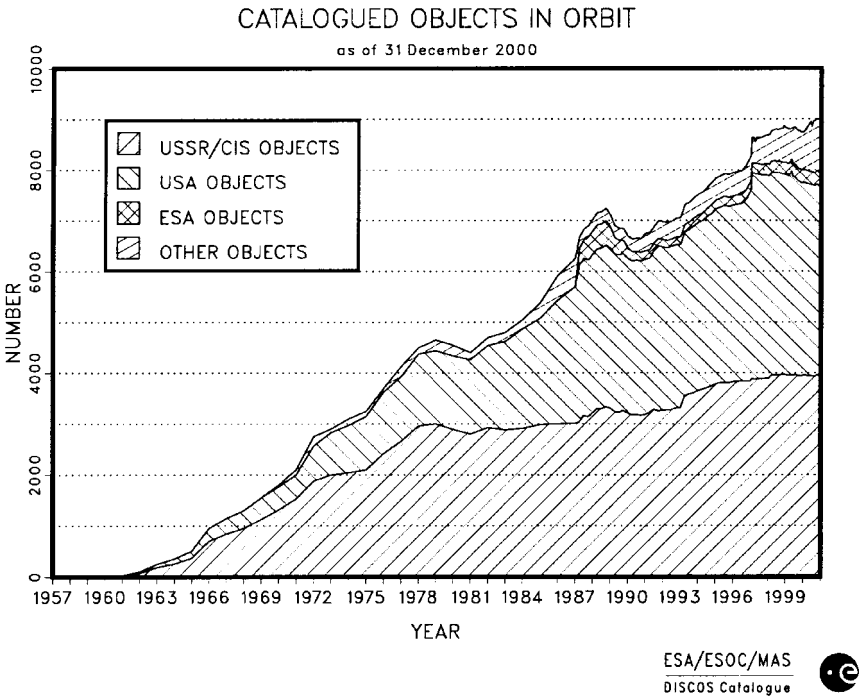


FIGURE 16.36 Catalogued objects in orbit. These are objects of at least 10 cm size in low Earth orbit and at least 1 m size in geostationary orbit.

the orbital data in a publicly available catalogue. Of these catalogued objects only 600–700 are operational satellites (7%). Nearly half of the catalogued objects are fragments from explosions in space (43%). As of today, more than 160 explosions in space are known. Intentional explosions have not been performed for more than 10 years, but fragmentations of rocket upper stages or propulsion units are still taking place four to five times a year. The reason is often residual propellant, which can ignite even after many years in space. There are about 1,350 discarded upper stages from launch vehicles (16% of the catalogued objects), 1,900 defunct satellites (22%), and more than 1,000 mission related objects like ejected covers or screwdrivers dropped by an astronaut (12%). Their number is steadily increasing, despite a reduced number of rocket launches per year (see Figure 16.37).

In low Earth orbit, objects down to a size of about 10 cm are maintained in the catalogue. In geostationary orbit, objects have to have a size of about 1 m to enter into the catalogue. The boundaries are not precisely determined. Whether an object can be detected depends on its material properties as well as on the observing geometry (elevation over the horizon, Sun illumination, number of observations, etc.). The sensitivity of radar and optical telescopes is shown in Figure 16.38.

As can be seen in Figure 16.36, the number of catalogued objects increases by about 210 per year on average. Temporary declines are correlated with solar activity. At high solar activity the high atmosphere is heated, increasing the local air density

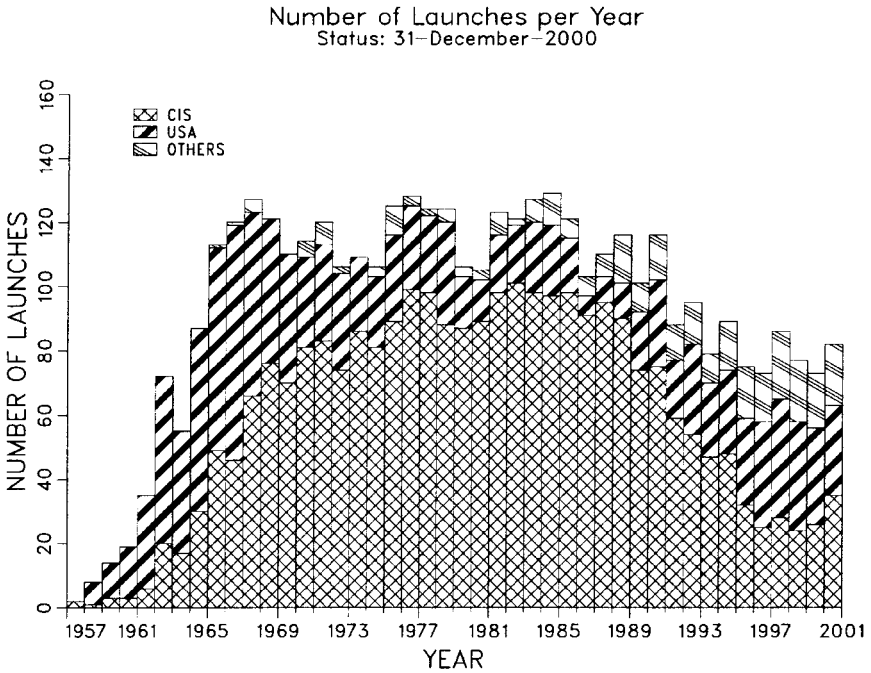


FIGURE 16.37 Launch record. The number of successful launches per year is listed (launch rate).

EARTH'S ENVIRONMENT AND SPACE

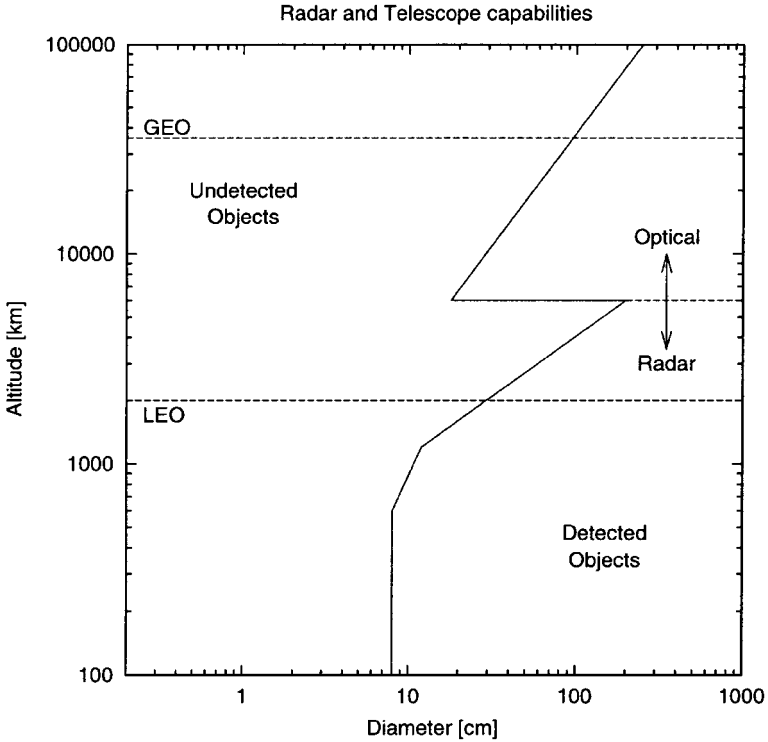


FIGURE 16.38 Sensitivity of the U.S. Space Surveillance Network. Objects to the left of the solid line are too small to be routinely tracked. Radar detect objects of about 10 cm in LEO, and telescopes are used to track objects of about 1 m in the geostationary orbit.

and thus increasing the air drag of satellites. This, in turn, causes the satellites to lose altitude more rapidly and finally to burn up in the atmosphere.

This is the only natural mechanism which cleans space. However, it is only effective in low Earth orbit. At altitudes above 1,000 km, the annual decrease in orbital altitude is very small. For a typical satellite in a 1,000 km circular orbit, it will take about 2,000 years until air drag pulls it down into the dense atmosphere, where it will finally burn up. This means that space activities at these altitudes will inevitably lead to a continuous increase in man-made objects.

Operational satellites can be damaged or even destroyed by a collision with space debris. Since the relative velocities are very high—9 to 13 km/s on average in low Earth orbit—centimeter-sized particles can produce considerable damage. In addition, new debris is generated during such collisions, in turn increasing the collision probability.

The biggest danger is due to untrackable objects. The number of these is much higher than the number of trackable, or catalogued, objects. The orbits are widely unknown and the spatial distribution can only be described statistically. The number of objects larger than 1 cm is estimated to be about 350,000.

16.36 SPATIAL DISTRIBUTION OF SPACE DEBRIS

The operational orbit of a satellite is chosen with consideration of various criteria such as mission objective, orbital altitude, orbital perturbations, requirements on orbit control, launch costs, radiation dose, etc.. Some orbits are of much higher interest than other orbits. Today, most spacecraft reside in low-eccentricity orbits below 1,500 km altitude. Near-polar and Sun-synchronous orbits are preferred for Earth observations. Eccentric 12-hr orbits with a critical inclination of 64.3° have been used many years. The geostationary orbit is also becoming populated with more and more satellites.

The spatial distribution of the trackable space debris is by no means homogeneous. It is closely linked to the orbits of operational spacecraft. For instance, in cases of highly eccentric orbits, the apogees are nearly all located in the northern hemisphere. Figure 16.39 shows the spatial distribution of the trackable space debris population as a function of altitude.

The spatial density has maxima between 900 and 1200 km altitude and around 1,500 km altitude. Higher up the density continuously decreases, with intermediate peaks at 20,000 km altitude (12-hr orbits, e.g., GPS and GLONASS) and at 35,800 km (geostationary satellites).

The approximate measured particle flux of space debris in low Earth orbit is shown in Figure 16.40 as a function of particle diameter. The dashed curve shows the meteoroid flux, which is of the same order of magnitude as the debris flux for

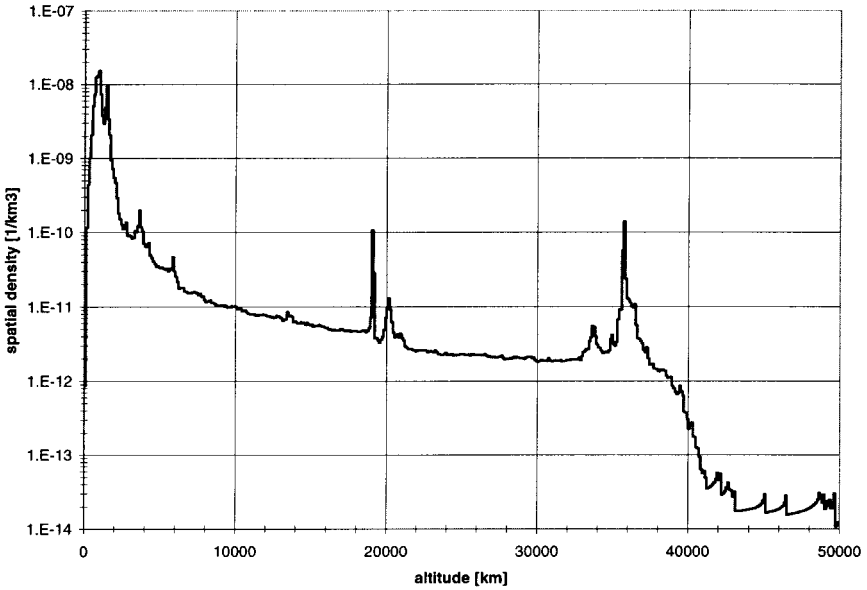


FIGURE 16.39 Spatial density of catalogued objects (as of January 2000; source: N. Johnson).

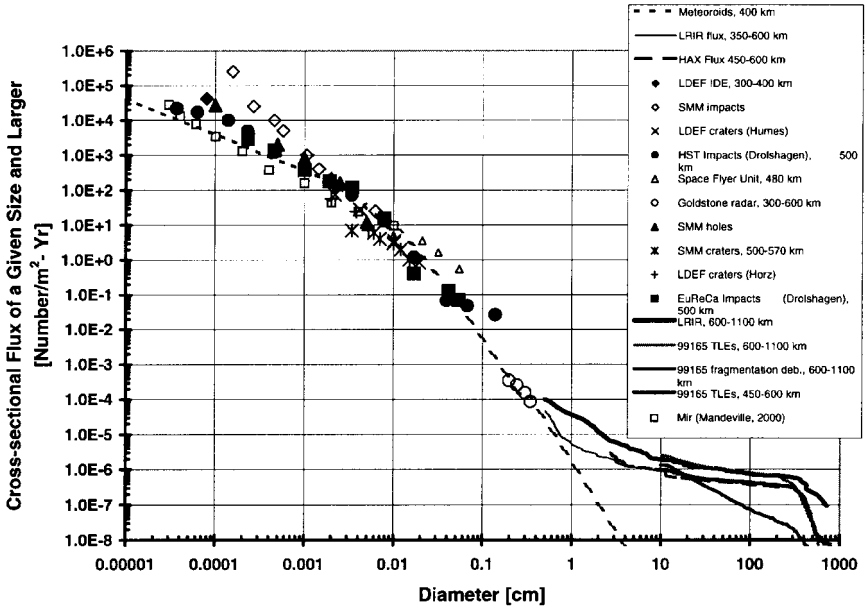


FIGURE 16.40 Approximate measured debris flux in low Earth orbit (source: N. Johnson).

particle sizes between 10 μm and 1 mm. At smaller sizes and also at larger sizes, the space debris flux predominates.

One source of 1 to 10 μm sized space debris is aluminium oxide, which is used as catalyst in solid rocket motors (SRM). During each of the more than 1,000 SRM burns which have been made up to today, trillions of such Al_2O_3 objects are released in space, and current models predict their number in orbit to be on the order of 10^{17}.

Space debris resides mainly in the neighborhood of orbits which are highly frequented by operational satellites. Therefore, these orbits are exposed to the highest collision risk. To estimate the collision risk of individual orbits the MASTER model was developed at ESA (Klinkrad et al. 2000). It describes the spatial distribution of space debris and meteoroids from low Earth orbit to the geostationary altitude for objects larger than 1 μm. The model population is based on:

- The catalogued population
- The simulation of more than 150 explosions in space
- The simulation of more than 1,000 SRM firings
- The release of 150 kg of sodium-potassium droplets
- The release of surface material (e.g., paint flakes)

In total, 350,000 objects larger than 1 cm are estimated to be in orbit, with 110,000 on average being below 2,000 km altitude.

16.37 THE COLLISION RISK

A simple method to estimate the collision probability is presented here. Assuming that a debris model provides the particle flux F and the satellite has a cross-sectional area of A, then the average number of impacts on the surface A during time T is:

$$N = \int_0^T FA dt$$

If we further assume that an impact is a rare event and follows a Poisson distribution, then the probability of at least one impact during time T is:

$$P = 1 - e^{-N} \approx N$$

For a constant flux F and cross-section A, the collision probability during time T is:

$$P = N = F\,A\,T$$

The collision risk in low Earth orbit is studied below for two examples: the European remote sensing satellite ERS and the International Space Station.

ERS is in a Sun-synchronous orbit at an altitude of 780 km and an inclination of 98.5°. Spacecraft in such an orbit face the highest collision risk. The reasons are twofold: the spatial density of debris reaches a maximum at this altitude, and most of the possible collision partners will impact at $\pm 20°$ around the flight direction of ERS. The reason for this special collision geometry is the high inclination, which leads to many near-collisions over the poles with debris in the same orbit, but with ascending nodes about 180° different from the node of ERS. The probability of colliding with an object of 1 cm is 5×10^{-5} m^{-2} yr^{-1} and the most probable collision velocity is 15 km/s. Figure 16.41 illustrates the distribution of collision velocities for three different orbital inclinations. The higher the inclination, the higher the average collision velocity.

The collision geometry as well as the total collision probability in the ISS orbit are quite different. Due to the ISS-inclination of 51°, most of the impacts occur at an azimuth between +30° and +60° and −30° and −60° (see Figure 16.42). The azimuth is measured in the local horizontal plane with 0° indicating the flight direction. This means the ISS needs the most effective debris shields around 45° on its "left" and "right" sides (left and right in the sense of a passenger in an aircraft). Little debris is coming from "above" and nearly no debris is coming from "below." The total 1 cm debris flux is 0.7×10^{-5} m^{-2} yr^{-1}, which is seven times less than the flux in the ERS orbit. Assuming a surface of 100 m^2 for an ISS module and a projected lifetime of 20 years, the probability that the module will be hit by 1 cm debris is 1.4%. The average impact velocity is 10 km/s (see Figure 16.41).

16.38 THE GEOSTATIONARY ORBIT

Since 1963, about 650 satellites have been launched into geostationary orbit. The majority of them were boosted from GTO into a nearly synchronous orbit by an apogee kick motor. For satellites equipped with solid rocket motors, the motors are usually ejected at the end. A considerable number of satellites are directly inserted

EARTH'S ENVIRONMENT AND SPACE

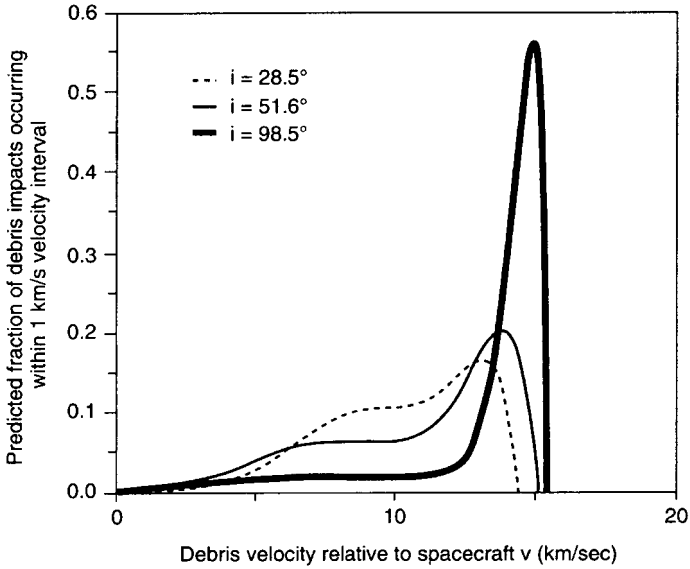


FIGURE 16.41 Calculated collision velocity distribution.

into the geostationary orbit by the launch vehicle, which itself remains in the vicinity of the geostationary orbit. In total about 900 objects are catalogued near the geostationary orbit, 300 of them operational and under orbit control (Hernández and Jehn 2001).

The Sun's, Moon's, and Earth's oblateness lead to long-period perturbations of the inclination of geostationary satellites. The period is 53 years and the maximum inclination is 15°. Thus, uncontrolled satellites build up an inclination of 15° after 27 years. They keep crossing the geostationary orbit twice a day with a velocity of 800 m/s with respect to the controlled geostationary satellites.

In order to eliminate this collision risk, GEO satellites are moved out of the geostationary orbit at the end of their missions. Three burns are recommended to raise the orbit by about 300 km, which is considered a safe distance to avoid future interference with active GEO spacecraft. The change in velocity that is required to raise the semimajor axis by 300 km is 11 m/s, and the propellant requirements correspond to propellant for three months of station-keeping. This means spacecraft operators have to stop operations three months before the spacecraft runs out of fuel and give up considerable revenue in order to reorbit their spacecraft. This is currently the only possibility to preserve the unique resource of the geostationary orbit.

However, only one-third of the aging satellites are properly reorbited into the recommended graveyard orbits 300 km above GEO. In recent years one-third were abandoned without any end-of-life measures, adding further debris to the already 125 objects librating from East to West and back with periods of up to 10 years. About one-third are moved out of GEO just high enough to free their orbital slots to make them available to new satellites.

There is a further collision risk when a couple of satellites occupy the same orbital slot (colocation). The orbital slots typically have extensions of 0.2° (140

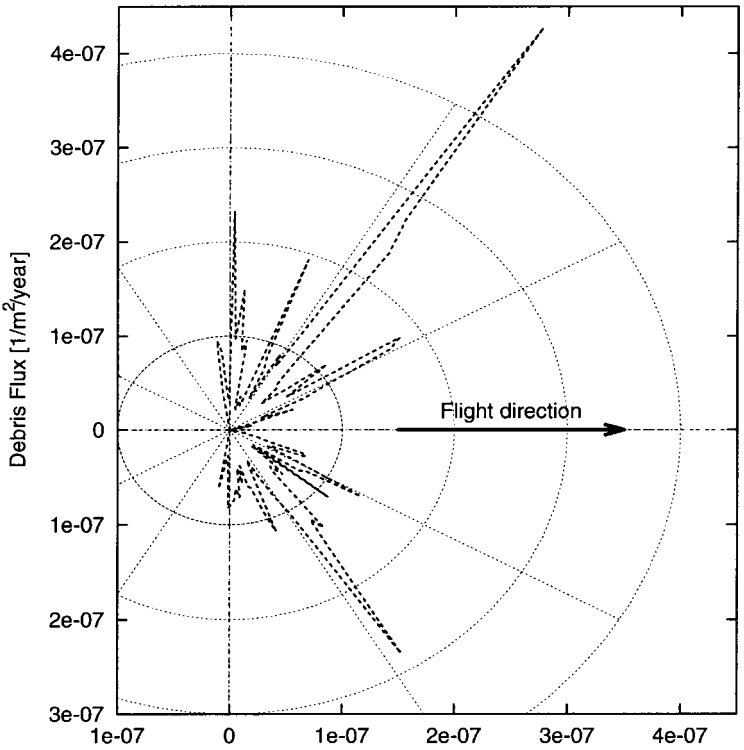


FIGURE 16.42 Space debris flux on the International Space Station.

km) in longitude and latitude. Since the GEO resource is limited, satellites have to share the same longitude slot (e.g., seven Astra satellites at 19.2° East). Special station-keeping strategies are necessary to eliminate the collision risk among the controlled satellites. The disadvantage is a slightly increased propellant consumption.

16.39 LONG-TERM EVOLUTION OF THE SPACE DEBRIS ENVIRONMENT AND MITIGATION MEASURES

All major space-faring nations have developed models to simulate the long-term evolution of the space debris environment. Currently, the steady increase in the number of objects in space is mainly due to satellite launches and explosions in space (five per year on average). There is an average increase of 200 catalogued objects per year. The solid line in Figure 16.43 shows the predicted number of 10 cm objects over the next 200 years if space activities continue as today (''business-as-usual''). The population will double about every 75 years.

EARTH'S ENVIRONMENT AND SPACE

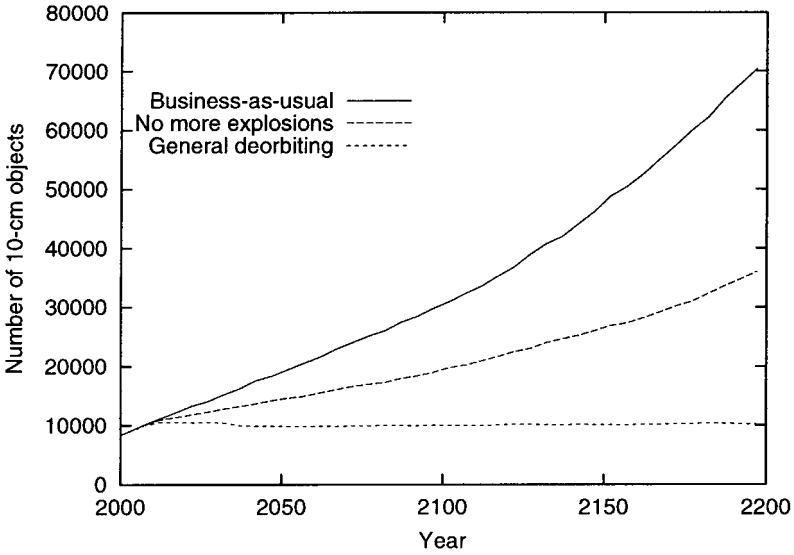


FIGURE 16.43 Simulated number of 10 cm objects in low Earth orbit depending on future space debris mitigation measures. Only if all explosions are prevented and if the deorbiting of old satellites and rocket bodies is introduced can the continuous increase of small objects in space be stopped (source: A. Rossi).

At some point in time the increase will no longer be dominated by launches and explosions. If a certain number of objects is reached (which means that a certain collision probability is also reached), the increase will be dominated by collisions. Simulations by Anselmo et al. (1999) predict that in the business-as-usual case the rate of catastrophic collisions will increase from 0.1/yr in the year 2000 to about 5/yr within the next 200 years. Kessler (1991) introduced the term *critical density*. The spatial density at a given altitude is called critical if the number of fragments created by collisions exceeds the number of objects removed from this altitude by air drag. If the critical density is reached, the number of debris particles will increase even if all satellite launches are suspended. Today the critical density is reached at altitudes of 800–1,000 km and at 1,500 km. Space debris mitigation measures are especially necessary at these altitudes.

As immediate steps, two measures must be taken: (1) prevention of in-orbit explosions, and (2) avoidance of mission-related objects. The first measure has the biggest impact on the future environment since about half of the catalogued objects in space are fragments from explosions. To prevent the breakup of rocket upper stages, they have to be made harmless after separation of the payload(s). The U.S. Delta rocket upper stage now performs a burn to depletion, and the Ariane upper stage and the Japanese H-1 second stage release their residual fuel to avoid a later explosion. Also, spacecraft batteries and critical items with a potential to explode have to be made safe. The second measure requires minimizing the number of objects released during spacecraft operations. Typical mission-related objects are adapters between two satellites of a dual launch, telescope covers, bolts, yo-yos, etc. Spacecraft designers and operators must design the mission such that objects stay attached to the spacecraft and do not form additional space debris. If no objects

are released after the year 2005 and if no more explosions happen after 2010, then the number of 10 cm objects will increase moderately and, within 100 years, reach only half of the value of the business-as-usual scenario.

However, to stop the ever-increasing amount of debris and stay below the critical space debris density, more ambitious mitigation measures need to be taken. In the long run, spacecraft and rocket stages have to be returned to Earth after completion of their mission. Currently, there are international efforts to come to a world-wide agreement to limit the orbital lifetime of retired spacecraft in low Earth orbit to a few decades (25 years is the limit in the NASA and European space debris mitigation standards). When such an agreement is applicable, all low Earth orbit spacecraft will be required to lower their orbits such that atmospheric drag will cause their decay within a specified time. And when solar electric propulsion has replaced chemical propulsion as the principal means of orbital transfer, the deorbiting of spacecraft in high Earth orbits up to the geostationary orbit will become possible without excessive propellant penalties. The dotted line in Figure 16.43 shows the number of 10 cm objects if all spacecraft and rocket bodies are deorbited after end-of-mission (which last 10 years in the simulation) starting in the year 2010.

REFERENCES

Anselmo, L., Rossi, A., and Pardini, C, 1999. "Updated Results on the Long-Term Evolution of the Space Debris Environment," *Advances in Space Research,* vol. 23, no. 1, pp. 201–211.

Hernández, C. and Jehn, R. 2000. *Classification of Geostationary Objects—Issue 3,* ESA/ESOC, Darmstadt, Germany.

Kessler, D. J. 1991. "Collisional Cascading: The Limits of Population Growth in Low Earth Orbit," *Advances in Space Ressearch,* vol. 11, no. 12, pp. 63–66.

Klinkrad, H., Bendisch, J., Bunte, K. D., Krag, H., Sdunnus, H., and Wegener, P. 2000. "The MASTER-99 Space Debris and Meteoroid Environment Model," 33rd COSPAR Scientific Assembly, Warsaw, Poland, July.

FURTHER READING

Flury, W., and Klinkrad, H., eds., *Space Debris,* 32nd COSPAR Scientific Assembly, Nagoya, Japan, *Advances in Space Research,* vol. 23, no. 1 (1999).

International Academy of Astronautics (IAA), *Position Paper on Orbital Debris,* IAA, Paris (1995).

Johnson, N. "The World State of Orbital Debris Measurements and Modelling," in *Space Safety and Rescue 1997,* ed. G. W. Heath, AAS Science and Technology Series 96, Univelt, San Diego, pp. 121–128 (1999).

National Research Council, *Orbital Debris: A Technical Assessment,* National Academy Press, Washington, DC (1995).

Office of Science and Technology Policy, *Interagency Report on Orbital Debris,* Washington, DC (1995).

Proceedings of the Third European Conference on Space Debris, ESA SP-473, Darmstadt, Germany (2001).

Technical Report on Space Debris, UN Doc. A/AC.105/720 (1999).

SECTION 17
AIRCRAFT SAFETY

Section Editor: Francis C. Gideon, Jr.

One of the primary considerations that must be on the mind of the aeronautical or astronautical engineer working at the computer or drafting table is the safety of the aircraft or spacecraft being designed. This section is meant to depart from the hard standards and design rules normally employed by the engineer and instead provide a general background on the subject of safety. The topics herein should be kept in mind, and as trade-offs are being made, the engineer should think ahead to how the design will affect flying safety.

PART 1

ORGANIZATIONAL SAFETY PROGRAM

Frank H. Snapp

17.1 INTRODUCTION

Organizations have safety programs for a number of reasons. Some form of accident-prevention program is usually a regulatory requirement. These requirements may take the form of laws, regulations, and directives that govern the design, construction, maintenance, and operation of aircraft. Many of these were instituted in response to accidents and implement the lessons learned from their investigation.

Accidents usually lead to extremely high costs. First, there is the cost associated with the loss of the aircraft itself. Considering the replacement cost of many of today's air vehicles, loss of one can be considerable to an organization of any size. There is also the cost of loss of life and property damage. Other costs include, but certainly are not limited to, hiring and training of replacement personnel, environmental cleanup, loss of use of equipment, increased wear on remaining equipment, legal fees and litigation, increased insurance premiums, and the cost of corrective action. If the financial position of the organization is in any way tenuous, loss of business and damage to reputation may destroy the organization.

There is a moral reason to protect workers, customers, and others associated with the operation from injury. Workers who sense that their job poses no excessive or unnecessary risk are more likely to be productive. Confident customers are likely to be satisfied and bring return business.

An effective acciden-prevention program is simply a smart, practical way for any organization to conduct business. Many of the same things that make an aircraft or other vehicle safe to operate also make it reliable and profitable. Accidents are showstoppers. They disrupt normal workflow and cause delays in all aspects of mission accomplishment, ranging from day-to-day operations to developmental programs. Preventing accidents minimizes costs and keeps the organization functioning smoothly.

17.2 SAFETY CONCEPTS (*Wood 1999*)

"Safety first" is a common slogan in many organizations. It has a few corollaries, such as "Safety is paramount" and "Safety first, last, and always." The problem with this slogan is, it just isn't true. Safety is not first and never was. If safety were the first priority, flying organizations would park their airplanes and not fly, because flying involves accepting some risk. The organizational mission is first, and its safety program supports the mission effectiveness of the unit.

Safety versus mission is a concept that implies that there is a conflict between mission accomplishment and the organizational safety program. One goal of an organizational safety program should be to have that safety program so thoroughly integrated into the processes of the unit that it is transparent to everyone involved.

Everything from vehicle design, through manufacture and maintenance, to operations should be designed to prevent accidents or minimize their severity.

Blood priority is the notion that it is very easy to institute changes and reduce hazardous conditions after a serious accident occurs, while in many organizations it is difficult to institute change before something serious happens.

Known precedent expresses the idea that there are no really new ways to have an accident. While technology changes over time, the causes of accidents remain largely unchanged. Consider the example of the *Titanic* accident. Although it is a maritime accident that occurred nearly a century ago, it contains a number of lessons for organizational leaders and safety managers in aviation-related organizations. Its fundamental causes—an organizational culture of arrogance and complacency, design flaws, poor decisions by leadership, and inadequate training of personnel—are present in virtually all aviation and space accidents.

Randomness of damage and injury expresses two notions. There is the notion, based largely on superstition, that accidents occur in threes. Analysis of statistics, however, shows they also occur singly, in twos, and in virtually every other combination of numbers, since they are random events. The second concept is that, due to the randomness of damage and injury, the same circumstances that cause a minor incident in one instance may cause a catastrophic failure, aircraft loss, or fatal injury another time. For this reason, most organizational safety programs include investigation and reporting of minor incidents, based on the philosophy that the same root causes of a minor incident could result in a catastrophe.

A *zero accident rate* is a laudable goal, but most organizations will probably never experience one, at least in the long run. There is always some risk associated with aviation and space operations, and the effort and expense required to reduce the risk to zero is beyond the point of diminishing returns.

Resistance to change is the last of the safety concepts. All organizations eventually become resistant to change as they develop. Management and individuals become comfortable with the procedures and processes. As organizations grow, instituting change becomes more difficult because seemingly small changes have to permeate the entire organization and overcoming inertia may present a formidable challenge. On the other hand, the whole point of the safety program is change. Without change, problems will persist and the causes driving minor incidents may eventually cause a serious or fatal accident.

17.3 ACCIDENT CAUSATION

Preventing accidents involves understanding how they occur. It is important for all levels of management in the organization to understand that accidents are not isolated events, but are the consequence of a sequence of events resulting from problems whose roots may be firmly entrenched in the organization. The sequence begins with underlying causes—those that existed long before the mishap and set the stage for its occurrence. Second in the sequence are those causes that are more immediate and constitute oversights omissions, hazards, and errors that continue the sequence of events. As the sequence continues, errors and omissions by leaders and workers make the accident almost unavoidable. The goal of the safety program is to identify the causes and eliminate them to prevent the accident or reduce its severity.

As an example, consider the sinking of the ocean liner *Titanic* at the beginning of the last century. Ask a layman why the *Titanic* sank, and he will usually reply that it hit an iceberg, sustained severe damage, filled with water, and sank. Actually, that account only explains only *what* happened, not *why* it happened. Furthermore, that explanation does not account for the extensive loss of life. Finally, it does not identify the problems that led to the accident so that they could be corrected.

To understand why *Titanic* sank, or why any accident occurred, for that matter, requires an in-depth look at the organization from the top down, as well as the specifics of the operation involved in the mishap. This examination must consider managerial attitudes and organizational traditions along with the hazards inherent in the mission itself. *Titanic* was built at a time when shipbuilders' confidence in their technology was so high they believed they could build an unsinkable ship. This confidence was misplaced, making designers and decision makers unable to see that the ship had a serious design flaw, in that the so-called watertight compartments were not watertight. They did not extend from the keel to the top deck and allowed water to spill from one compartment to the next, causing the ship to go under by the bow after the impact with the iceberg broke the integrity of the hull. Overconfidence by management also led to faulty decisions on such issues as the number of lifeboats the ship was equipped with and the extent of emergency procedures training for the crew and passengers. Misplaced confidence also led to a sense of complacency on the part of the ship's captain and officers during the voyage, causing them to disregard ice warnings and proceed at high speed in a known area of icebergs on a dark night. The company operating *Titanic* had a tradition of trying to set speed records on the maiden voyage of each new ship. This tradition probably clouded the captain's judgment when he decided to continue sailing at high speed after he learned the ship was operating in an area of iceberg activity. The level of technology and the state of the art are also root causes that cause or complicate accidents. At the time of the ship's sinking, no one had given much thought to the best use of wireless communications in maintaining continuous listening watch for ships in distress, or to standardizing emergency frequencies or distress signals. This caused a delay in the rescue efforts by ships in the area of the accident. Arrogance and overconfidence by managers and designers, organizational traditions, design flaws, and a lack of understanding regarding how to use new technology like radio were the additional underlying causes of the disaster. They either initiated the sequence of events that led to the accident, kept the sequence moving, or aggravated its severity.

While underlying causes set the stage for an accident, more immediate causes may initiate or continue the sequence leading to the accident. Proximate causes include such issues as inadequate training of personnel, incorrect, vague, or outdated procedures, and inadequate or nonexistent enforcement of established standards. To continue the *Titanic* example, the failure to ensure that every passenger and crewmember was familiar with emergency evacuation procedures meant that the lifeboats the ship had were not filled to capacity. This increased the number of fatalities resulting from the ship sinking.

The event that triggered the disaster was the captain's decision to disregard ice warnings and steam at high speed through a known area of icebergs. The last parts of the sequence are the mishap itself, followed by the destruction/damage and fatalities/injuries. In the *Titanic* example, colliding with the iceberg, although a significant event in the accident sequence, was the expected result of a captain's bad decision. It did not cause the accident. It was the result of a chain of events

that preceded it. There probably would have been fatalities when *Titanic* sank, but decisions by management prior to the voyage could have minimized them. More lifeboats, enough for each person on board, would have resulted in fewer losses of life. The losses in life could have been further minimized had the crew and passengers been trained sufficiently in emergency evacuation procedures. These conditions aggravated the mishap and resulted in increased fatalities.

Virtually all accidents have more than one cause. By definition, a cause is a deficiency that by itself or in combination with other problems resulted in the damage or injury. It may be an act, omission, condition, or circumstance such that, had it been corrected, eliminated, or avoided, the accident would not have happened. Missed opportunities to mitigate the damage and injury can also be considered causal. Normally, environmental conditions are not causes unless they are unavoidable. It is the function of the organizational safety program to identify and eliminate those causes before they cause accidents.

17.4 PRINCIPLES OF SAFETY MANAGEMENT (*Peterson 1998*)*

Unsafe acts, unsafe conditions, and accidents can all be traced to management problems. In the *Titanic* example virtually every cause of the accident stemmed from management's failure to identify and correct its own problems or to make the correct decision given hazardous conditions on the sea.

There are circumstances that are known to produce accidents and injuries. Many of these can be identified and avoided or controlled. These circumstances are inherent in aviation operations. The hazards may come from a long list of operations, including storage and handling of fuels and other hazardous materials, work performed at elevated heights, storage and handling of explosives, and aircraft servicing and cargo loading operations. Aircraft, servicing equipment, and cargo-loading equipment, along with their associated procedures, can be designed to minimize the risk of accident and injury. Ensuring this happens is a major goal of the organization's safety program. Every manager of an aviation operation knows weather conditions such as thunderstorms, low visibility, and wind shear present hazards to flight operations. While they cannot control the existence of these conditions, managers can, and usually do, provide air crews guidance on how to detect and avoid the hazards associated with hazardous weather. Some hazardous conditions and circumstances may not be so obvious. Some conditions, such as those associated with aircraft aging, may be detected and identified only by inspection or by analyzing statistical data from incidents and system failures that occur in the fleet for trends that indicate larger problems. This is a prime reason the organization's safety managers are involved in investigating and following through or corrective action for incidents involving system and equipment failures.

Safety should be managed like any other operational function. There should be goals and formal efforts to achieve those goals. The organization's safety program should begin with top management involvement and permeate the entire organization.

* *Author's note*: The principles are Peterson's; most of the discussion and illustrations are mine.

The organizational leadership must establish procedures that fix accountability for safety performance. Just as department managers must account for their production rates, quality control, and fiscal efficiency, they should be accountable for mishap rates. This does not mean the manager should necessarily be dismissed for an accident in his functional area, but it does mean he should be accountable for his decisions and their relationship to accidents.

The goal of the safety manager is to identify problems that can lead to accidents and work with the organization's leadership to correct these problems before the mishap occurs. This is the essence of the organizational safety program, and the safety manager has a number of tools at his disposal.

17.5 THE ACCIDENT-PREVENTION PROCESS

The first step in the accident-prevention process is setting standards. There are already a number of standards in existence. The regulatory environment provides many of them in the form of Public Law, Federal Aviation Agency (FAA) Regulations, Occupational Safety and Health Administration (OSHA) Standards, International Civil Aviation Organization (ICAO) Standards and Recommended Practices, and the International Organization for Standardization (ISO) Standards. Additional standards may be recommended by professional societies and trade associations such as the Flight Safety Foundation and the National Safety Council. Organizations that fly similar aircraft or missions may also be a source of information on setting standards. All imposed standards have one thing in common: they are minimums. Organizations usually expand on standards by setting their own, which take the form of company or local procedures.

Applying and enforcing standards is the second step in preventing accidents. This is the job of line managers. In order to be effective, standards must be known and understood by everyone involved in a particular operation. This step of the process includes such management functions as training programs, certification of pilots and maintenance personnel, and inspections and evaluations performed by safety managers along with those of other specialties in the organizations.

Identifying hazards is a critical step in an organizational safety program. While the safety manager may discover hazards in the course of his inspections, problems are usually brought to the safety staff's attention by employees in various parts of the organization. Many countries have laws that require employers to provide their employees with a means to notify management of on-the-job hazards. Requirement or not, it is important to provide employees with a hazard-reporting mechanism, since they are closest to the operation and usually well aware of any hazards in their workplace. An effective reporting program is easy to use and available to anyone who might detect a hazard. Of vital importance, there should be no retribution against the report submitter for identifying a hazard. For this reason, most documents associated with hazard-reporting programs usually allow the submitter to remain anonymous. Hazard report forms must be readily available and simple to fill out and should give the employee adequate space to explain the facility or equipment involved and the nature of the problem. When received by the safety manager, the report should be assigned some form of risk assessment, which allows managers to set priorities for resolution. Management should also have a mechanism for advising the report submitter of progress toward resolving the problem and allowing the submitter to participate in the process.

There are other tools that are useful in identifying hazards, including inspections, risk assessments, and accident reports from other organizations with similar equipment or missions. The organization's safety program should be structured to include these as elements of the safety program.*

Hazard elimination is ultimately the responsibility of line management, and a range of methodologies is available. The most desirable method is the engineering solution, as it eliminates the hazard. For example, if the lights on a frequently used taxiway no longer work because the underground wiring has become corroded, there is a hazard to aircraft during night operations. The obvious solution is to have the airfield manager hire a construction company and have the wiring replaced to restore the lights. Unfortunately, not all engineering solutions are this easy. They frequently involve redesign of the aircraft or one of its major components and require considerable time and money. Consider the complexity of retrofitting a large fleet of aircraft with new engines. From design through project funding to retrofit completion, the project will probably take years. It is often impractical to stop operations while the changes are made, so the hazard is usually controlled by one or more strategies that reduce the risk while continuing operations.

Hazard controls usually take the form of guards, procedural changes, frequent inspections, and operating limitations. Instead of redesigning a cockpit layout, for instance, management often decides to control a hazard by guarding switches or wiring them to a certain position to prevent inadvertent actuation. As airplanes age and develop structural problems, management may institute changes to procedures or institute interim operating limitations that reduce maximum g-load, airspeed/Mach, or gross weight pending modifications. These steps reduce the risk of operating to some degree but are less desirable than the engineering change because the hazard still exists and the control may not be completely successful.

Another approach to hazard control is to install alert and warning devices or to train individuals in the hazards present in their duties. Examples of alert and warning devices include stall warning systems and landing gear warning lights and horns. These systems are usually effective if designed and used correctly, but they certainly are not infallible. The operator may deactivate them, or their parameters may not match the operation being performed. Finally, a distracted or task-saturated operator may not hear or see the warning. The most obvious example in aviation is the gear-up landing.

The least desirable hazard control is protective equipment. This includes hearing protection, eye protection, respirators, gloves, aprons, helmets, and all forms of equipment intended to shield workers from hazards. This equipment is effective only if it is properly designed, available, and used by all that need it. The problem with protective equipment lies in the fact that the hazard often exists unchanged and the equipment only constitutes a barrier between the worker and the hazard. Moreover, the equipment has to be properly used by the individual, and often it is not.

Key components of any organizational safety program are the internal and external information systems of the organization to notify others of hazards and mishaps. Experience teaches that if one part of the organization has a hazardous situation or incident, other parts will have similar problems. For this reason, large organizations have internal systems to cross-feed information to all departments

Author's note: I have seen many similar lists of the hierarchy of hazard abatement steps in other publications. These are adapted from Wood 1999. The discussion and examples are a mix of Wood's and mine.

regarding accidents, incidents, and hazards. This aids the whole organization in identifying and correcting hazards.

17.6 THE ELEMENTS OF AN ORGANIZATIONAL SAFETY PROGRAM

The safety staff is the department of the organization that manages the organization's safety program. Ideally, the senior safety manager, chief, or director of safety should answer only to the organization's senior leadership. In large organizations, every level of line management should have a safety staff consisting of one or more individuals. These individuals are usually specialists in one or more of the aspects of the organization's mission. The typical flying organization safety staff has specialists in aviation and flight operations, in addition to personnel with expertise in industrial accident prevention and aircraft maintenance. Organizations that design and test aircraft and space vehicles usually have system safety and space safety managers and engineers as full-time members of the safety staff. Members of the safety staff should have specialized training in their area of expertise and, ideally, should be certified as safety professionals.

Safety councils and committees are forums for discussing safety issues. They keep the senior leadership in the organization informed on issues affecting the organization, hazards involved in operations and progress toward controlling or eliminating them, and progress in implementing lessons learned from mishaps. Some councils and committees are required by law, and their meeting frequency and membership are mandated by government regulation. Others are instituted by contracts with customers or established by management to solve a specific problem. In any event, the meeting should be chaired by a line manager in order to enable the proceedings of the council or committee to be binding on the organization.

Safety inspections help identify hazards, assist management in assessing the effectiveness of the safety program, and help evaluate compliance with safety directives. They enable management to determine the condition of work areas and the level of compliance with safety of work practices. They may be performed by the safety staff of the organization, or by an external government agency such as the Occupational Safety and Health Agency in the United States. The inspection may be a formal inspection, or an inspector may simply monitor an area or operation.

Formal inspections are usually scheduled well in advance. The unit being inspected should be provided a checklist that lists the areas to be covered by the inspector and the reference to the standards that apply to each item. The areas covered in an inspection usually include:

Management involvement in mishap prevention

Organizational culture and attitudes toward accident prevention

Qualifications and training provided to safety personnel

Problems noted during previous inspections

Quality and depth of mishap investigations

Programs to follow up on mishap recommendations

Hazard reporting and abatement programs

Mishap-prevention information flow to management and line employees

Safety councils and committees

Accident-response plan

Quality and effectiveness of job safety training for employees

Work practices

Conditions in the workplace

Formal inspections are usually thoroughly documented, either on a form provided by the organization or by memorandum. This documentation includes, at a minimum, the date of the inspection, the organization evaluated, hazards noted, their causes and an assessment of their risk, unsafe work practices found during the evaluation, and recommendations for corrective action. The report should also contain instructions to follow up on recommendations and instructions to document completion of corrective action.

Another form of safety oversight is the spot inspection. This involves simply observing activities in the workplace or monitoring an operation by the safety manager. Management should determine which areas of the organization pose operational risks, and the safety manager should set up a program whereby he can monitor these operations on a regular basis. Areas of greatest interest to management or those that pose the highest risk should be monitored more frequently. Normally, formal reports are not prepared for this kind of inspection if the safety manager finds no discrepancies. The safety manager should simply maintain a log of the facilities visited/operation monitored, date, point of contact and phone number, and any problems encountered.

Safety in aircraft maintenance involves oversight into all aspects of the maintenance operation. Maintenance supervisors are usually concerned with industrial safety, ensuring the equipment, facilities, and personnel are mishap free. They are also concerned with the airworthiness of the aircraft, as that determines whether or not the aircraft is safe to fly. Historically, problems in design, assembly or repair, training and supervision of mechanics, errors in maintenance data, and defective parts have caused maintenance-related aircraft accidents. Accident-prevention programs need to focus on these problems in addition to industrial safety issues.

Airport safety is a joint effort involving the safety staff, the agency that handles airfield management, and various agencies of the organization and outside contractors that perform maintenance on the aircraft. Information on such issues as facility design, runway and taxiway markings, lighting, navigation aids, and instrument departure and approach criteria can be found in ICAO or Federal Air Regulations. The safety program manager's day-to-day concerns are maintenance and construction, bird and wildlife control, and other issues relating to the care and use of the airport, such as pavement conditions, airfield markings, lighting, and navigation aids.

Any time construction is planned in or around the airfield, the organization should perform a review of planned construction activities. This review should include an analysis of the construction activities and their impact on day-to-day operations. There should be provisions to allow the construction crews always to be in contact with the control tower. Routes for contractors' vehicles should be specified to ensure there are no conflicts between aircraft on the movement area and construction equipment and vehicles. This usually involves specifying mini-

mum distances between these activities and aircraft. Local and transient aircraft need to be briefed on the operations during construction, and the airport manager should publish the necessary notices advising pilots of the construction operations.

Bird and wildlife control is a major issue in airfield safety, and bird strikes and other collisions with animals result in considerable monetary losses and occasionally catastrophic accidents. Airport wildlife control plans usually concentrate on denying food, water, and shelter to animals that inhabit the region. The plans usually emphasize controlling grass height, standing water and food sources such as location of nearby landfills, rodent populations, and bird-harassment measures that keep birds from congregating on the airfield. During the migration season or other times when bird activity is at its highest, airports establish systems for notifying air crews of the bird hazard and its severity.

Other airport safety issues include vehicle control on the aircraft movement area, incursions by vehicles and aircraft into active runways, foreign object control, snow and ice removal, and maintenance to navigation aids, airfield signs and markings, and emergency equipment such as barriers and arresting gear.

Accident-response preparation involves establishing a plan to rescue injured personnel, deal with fatalities, extinguish fires, and preserve evidence for an investigation. Airports are usually required to have accident-response plans and most organizations that design, manufacture, test, and operate aircraft have plans to respond to or support investigation of aircraft and space vehicle accidents. The organization's emergency planning should include response to the full spectrum of mishaps, from air aborts and in-flight emergencies to major accidents. Response by the safety manager to minor incidents allows him to collect data that will later be useful in risk analysis and prevention of more serious mishaps and to monitor the activities of other emergency response teams.

The planning for major accidents should be detailed in that it provides well-thought out guidance for all responding agencies. The plan should include hazardous material emergency response. It should include primary and backup notification methods for all agencies along with guidance for convening key personnel at any hour of the day or whenever flying or industrial/maintenance operations are taking place. This means that the plan must be periodically updated with regard to names of contacts, locations, addresses, and telephone and building numbers. Most plans include, but are not limited to, the following:

Immediate response—fire, rescue, and hazardous material, including contingency plans for response to remote locations

Procedures to notify key government agencies and personnel within the organization

Control and security at the accident site

News release and media affairs

Appointment of investigators/organization's participants in government accident investigation

Securing of air crew and vehicle maintenance records

Explosive ordnance disposal

Survey teams to map the site

Food, water, and shelter for personnel at the site

Communications

Administrative support to the investigation

Logistics—transportation, supplies, funding

Photography support

Legal advice

The plan should be exercised and the exercise response evaluated and critiqued on a regular basis. This allows the organization's management to validate all the elements of the plan and to discover and correct errors, omissions, and shortcomings. Exercises also allow management and all responding agencies to become familiar with their part of the plan and in order to minimize confusion and errors should there be an actual accident.

Safety awards can be a valuable part of the overall safety program. They recognize individuals and units that have contributed to the overall safety program and encourage participation by everyone in the organization. The award itself may consist of a certificate, plaque, aircraft model, coffee mug, or any other recognizable form. Whatever form it takes, it should have the individual's name and organizational unit on it. If an award is presented to a number of people, each should receive a separate award. It should be presented by a senior manager in a forum where a large number of the recipient's peers are present. The award should be something permanent that the individual can display in his work area. A properly run awards program need not be expensive. Safety managers often design awards using computer software written for certificates and save the design as a template. Making the award then requires entering the individual's name and a description of his contribution, printing the certificate on high-quality paper, and mounting it on a plaque or in a frame. The award is then ready for presentation.

In order to function effectively, the safety manager must have sufficient equipment and mobility to do his job. In an organization that involves flying or test range operation, this should include a vehicle equipped with a two-way radio. The safety staff should be provided with office space, supplies, computer equipment, and any needed special purpose clothing and equipment. Where the safety manager conducts training, he needs classroom space, visual information support, and any necessary training devices.

17.7 CONCLUSION

The purpose of the organizational safety program is preventing accidents. An effective program accomplishes this by identifying and eliminating potential accident causes before a mishap can occur. In order for the program to be effective, management at every level must actually be the leader in preventing accidents. The safety staff manages the elements of the program, the inspections, hazard, and incident investigations, the education and training function, and the other elements of the program. Correcting problems that lead to mishaps and creating a climate within the organization conducive to preventing accidents is a function of management and employees (including engineers) at every level.

REFERENCES

Wood, R. H. 1999. *Aviation Safety Programs,* Jeppesen Sanderson, Englewood, CO.

Peterson, D. 1998. *Safety Management: A Human Approach,* Aloray, Inc., Goshen, NY.

PART 2

AVIATION LAW

Phillip J. Kolczynski

17.8 INTRODUCTION

This Part will focus on aviation laws in the United States. Unlike engineering or science, laws vary substantially from country to country. Even in the United States, laws can vary significantly from state to state; accordingly, the most commonly applied principles will be discussed in this Part. The law abounds with exceptions, limitations, local rules, and other variations that may render the general principles discussed herein inapplicable to particular situations.

17.9 FEDERAL STATUTORY LAW

Many aviation laws in the United States are created by federal statutes. U.S. statutes are passed by Congress and signed into law by the President. Statutes establish major policies and principles to govern the aviation industry. One important statute is the Federal Aviation Act of 1958, 49 USC §40101 *et seq.* This act gives the Federal Aviation Administration (FAA) authority to regulate the industry. Another example is the Independent Safety Act of 1974, as amended, 49 USC §40102(a). This statute gives the National Transportation Safety Board (NTSB) the independent power to investigate all serious aircraft accidents in the United States or anywhere in the world where U.S. carriers are involved. There are a number of other statutes passed by Congress which provide authority in narrow and specific areas. For example, the Air Carrier Deregulation Act of 1978, 49 USC §41713(b), allows commercial carriers to function as private businesses without direct management by state government.

17.10 STATE CODES

In the United States, the aviation industry is primarily regulated by federal statutes and federal government agencies instead of state laws. Air travel is an interstate function and federal uniformity is needed to regulate air transportation.

Various state codes have been passed to govern matters peculiar to the states. State laws generally control airport ownership and usage, aircraft ownership and sales, liens on aircraft, etc. Similarly, contract matters are usually controlled by state law, as well as property matters, which are typically very local in nature.

17.11 REGULATIONS

Regulations are different from laws. Regulations are the rules or the standards by which the laws are implemented. In fact, most regulations can only be enacted

under the authority of some statute or code. The best-known regulations in the aviation industry are the Federal Aviation Regulations. They can be found in the Code of Federal Regulations, Parts 1–199. The FARs have been issued by the FAA by means of its authority under the Federal Aviation Act. Private pilots are most familiar with Part 91 of the FARs, which controls aircraft operation. Manufacturers are most familiar with Parts 21–43, pertaining to manufacturer certification standards and maintenance.

17.12 THE COMMON LAW

The common law consists of the decisions of courts regarding individual cases brought to court by litigants. Courts normally create common law by interpreting statutes and regulations. Some of the more activist judges actually create laws (court-made law) on a case-by-case basis. If they cannot find them in statute, judges use principles recognized in the historical jurisprudence of the country and commonly accepted in a Judeo-Christian society.

Many disputes which arise in the aviation industry are not simply decided by specific statutes, regulations, or the precedent of prior case decisions. Often there is no definitive rule of law, either statutory or in the common law, to resolve a particular dispute. The reason is that aviation disputes vary greatly on their facts. Indeed, in the aviation litigation, factual issues are often technical in nature. Thus, each case must be decided by careful evaluation of all evidence, the circumstances, and the equities, guided by appropriate principles of law.

In the United States, parties typically have a constitutional right to a jury trial. The jury typically decides factual questions and the judge decides legal issues. A jury's verdict is subject to review by the judge and must conform to the legal principles determined by the judge. Because aviation cases are fact intensive, it is often the jury, not the judge, who decides the liability for the cause of a crash.

17.13 INDUSTRY STANDARDS

For aviation law, the statutes and common law are usually not detailed enough to define whether an aviation engineer or an aviation safety manager has performed his job in compliance with the law. Thus, the courts must look to industry standards to determine whether the standard of care in any particular field has been complied with.

Industry standards can be found in the more specific Federal Aviation Regulations, as well as in FAA orders, directives and advisory publications. Manufacturer publications defining proper design, quality control, product usage, maintenance, etc. also set standards.

When issues arise that involve interpretation of the publications, technical witnesses and expert witnesses may be brought to court to explain them. When the regulations and publications do not address the problem sufficiently, expert witnesses explain the standard custom and practice in the industry. Custom and practice will then constitute the industry standard.

Industry standards are important because a trier of fact must decide whether or not an aviation professional has complied with the standard of care in the industry when liability is at issue.

17.14 THE SCOPE OF AVIATION LAW

Aviation law can be subdivided into the following fields.

Aviation Accident Investigation Law

The National Transportation Safety Board is the primary aviation accident investigative agency in the United States. The Independent Safety Act of 1974, and the various NTSB regulations promulgated thereunder, form the bulk of the aviation accident investigation rules. These rules control the proper procedures to be followed during accident investigations.

Various court cases have interpreted the rules to decide how accident reports may be used. In contrast to the situation in many countries, and even with military agencies in the United States, the entire NTSB accident investigative report of any civilian accident is available to anyone who is willing to pay copying charges. The probable cause determination, analysis, witness statements, photographs, etc. are fully accessible.

NTSB aviation accident reports are frequently used in evidence. There is a federal statute which has been interpreted by the courts as not allowing the introduction into evidence of the ''probable cause determination.'' The concern is that the conclusions of the official investigator will be given such deference by a jury that the prerogatives of the jury to decide liability will have been usurped. NTSB investigators do not investigate to determine liability.

Probable cause is not synonymous with liability. In air crash cases, more precision is required in allocating percentages of fault than in safety investigations, where the probable cause is the focus for safety analysis. Typically, assuming that hearsay objections have been satisfied, most of the NTSB report, including factual findings, analysis, witness statements, and investigator opinions on preliminary matters not involving probable cause, will be admitted into evidence. Often the entire aircraft accident report that might be prepared by other federal agencies will be admitted into evidence.

Air Crash Law

Laws which determine the rights and liabilities of victims and parties to air crashes are typically found in the common law. Air crash law is civil law, not criminal law. The standard for determining liability is the ''preponderance of evidence,'' not the ''proof beyond a reasonable doubt'' standard used in criminal cases.

Generally speaking, in most air crash cases, liability and damages are determined according to the laws of the court with jurisdiction over the case. Often, but not always, this is the state where the crash occurred. Laws on liability and damages may vary significantly from state to state. Accordingly, lawyers often fight over which state's laws should apply to the major issues in any serious crash case.

In an air crash case, the trier of fact (usually a jury) must decide which of the parties breached its standard of care (duties) and proximately caused injuries, damages, or death. The jury decides which defenses are valid. The jury must then allocate percentages of fault among all the liable parties. Finally, the trier of fact

must decide the nature and extent of any damages and how much compensation should be awarded.

17.15 FAA ENFORCEMENT ADMINISTRATIVE ACTIONS

Because the FAA has been granted authority to regulate the aviation industry, the FAA controls licensing, that is, the issuance of certificates to aviation businesses, manufacturers, and aviation professionals such as pilots. If the FAA believes there has been a violation of its rules, even if no one has been injured, the FAA may bring an enforcement action against the violator. Such enforcement actions are brought in special hearings under the administrative procedures approved by Congress.

For example, manufacturer's type or production certificate can be revoked or suspended. The engineer who holds a certificate or designation may face similar sanctions. In addition, aviation businesses may be subject to monetary fines or penalties by the FAA. The FAA also has the power to take emergency action canceling the certificate if safety is immediately jeopardized.

Instead of sanctions directed at certificate holders, the FAA may choose to issue Airworthiness Directives (ADs) requiring that aircraft cannot be used until changes have been made which bring the product in compliance with FAA safety standards.

17.16 INTERNATIONAL TREATIES/ICAO

International air commerce is almost completely controlled by various multilateral and bilateral treaties. One of the more famous treaties, the Chicago Convention, spawned the International Civil Aviation Organization (ICAO). The ICAO is an international organization based in Montreal, Canada, which is dedicated to aviation safety and uniformity in international aviation. The ICAO is a voluntary organization. Nations participate as member states by treaty.

International aviation safety and uniformity are self-regulating in the sense that member states agree to abide by international principles. To the extent that the member states of the ICAO have agreed upon uniform standards for navigation, engineer, design, communication and safety, etc., the rules and standards promulgated by the ICAO are binding on the member states and the aviation industry within those member states. The ICAO can act as a mediator to resolve disputes; however, the ICAO is not an international court with authority to punish offenders.

17.17 AVIATION BUSINESS DISPUTES

Aviation professionals and entities in the aviation industry often become involved in commercial disputes associated with their transactions. Aviation laws have been developed, sometimes by statute or code but more often in the courts, to resolve disputes involving sales, contracts, patents, trademarks, warranties, employment,

etc. These subjects are controlled by business laws which vary significantly from state to state in the United States and may vary greatly from country to country internationally.

17.18 U.S. FEDERAL AGENCIES THAT REGULATE THE AVIATION INDUSTRY

FAA

The Federal Aviation Administration (FAA) is tasked by the Federal Aviation Act to maintain aviation safety and to promote air commerce. Thus, the FAA wears two hats. Under one hat, the FAA promotes business growth in the aviation industry. Additionally, the FAA regulates and enforces as necessary to maintain aviation safety.

The FAA issues safety regulations and controls the National Airspace System. The FAA develops and operates a common system of air navigation, supervises the Air Traffic Control System, and administers a system of certification and inspection for airlines, airports, aircraft pilots, crewmembers, and maintenance. The FAA oversees the construction and development of the nation's airports and supervises the taxation system, whereby revenue generated by tickets goes into a trust fund used for the development of airports.

NTSB

The National Transportation Safety Board (NTSB) is an independent federal agency that is not under the auspices of the Department of Transportation. Instead, five members of a board, all appointed by the President, run the NTSB. The NTSB has one primary function—to investigate serious accidents to determine the probable cause and factors which led to the accident. Thus, the NTSB is exclusively in the safety business. The NTSB is *not* in the business of determining liability or fault for air crashes. The NTSB does not have the power to require that changes be made in products or operational procedures as the result of a crash. The NTSB can issue safety recommendations to encourage the FAA to make changes to improve safety.

The NTSB investigates all major air crashes and almost all general aviation crashes involving deaths or serious injuries. The NTSB delegates to the FAA investigations of relatively minor crashes to the extent Board lacks manpower and resources to perform all the investigations FAA personnel often assist the NTSB in major air crash investigations in matters for which the FAA has particular expertise.

The NTSB uses a party designation system to allow private industry participation in investigations. Operators, manufacturers, and other entities can have full participatory rights. The criteria for designation is that the entity be involved with the crash and have technical expertise to offer the NTSB investigation. Victims, owners, and surviving employee pilots are often not designated as party participants. The NTSB, by recent amendments to the Independent Safety Act, is deemed to have total and exclusive control of the wreckage and can conduct destructive testing if it desires.

DOT

The Department of Transportation (DOT) is primarily involved in the economic regulation of aviation commerce. The DOT supervises airline certification, international route authorization, small community air service, and various forms of air carrier marketing and reservations.

NASA

The National Aeronautics and Space Administration (NASA) is primarily involved in research and development in the field of aeronautics and with the exploration of space. NASA provides assistance to the NTSB and FAA on special matters involving aviation safety. NASA also administers the Aviation Safety Reporting System (ASRS) whereby pilots or other aviation certificate holders can submit reports of safety problems in the international air transportation system without fear of sanctions by the FAA.

DOD

The Department of Defense (DOD) is the agency of the United States Federal Government which supervises all military assets. Thus, to the extent that there are aviation research and development, design and manufacturing, maintenance, accident investigation, and other functions related to military aircraft, the DOD controls. The DOD has issued regulations and various directives which implement its power granted by Congress.

The U.S. military branches generally conduct two aircraft accident investigations; one focused exclusively on safety and the other done for administrative and disciplinary purposes. Two types of reports are prepared. One is fully releasable to the public and the other is available only to those in the safety chain of command and is not releasable even to the victims' families.

17.19 THE ROLES OF AN AVIATION SAFETY PROFESSIONAL IN AVIATION LAW

Because litigation often results from occurrences involving aviation professionals, it is important for safety managers and engineers to understand the various roles they can play in legal matters. Information provided below is intended to be of value to all aviation engineers and safety managers, regardless of their subspecialty, i.e., aeronautical engineering, mechanical engineering, system safety engineering, etc.

17.20 FAA REGULATORY COMPLIANCE

Engineers are the backbone of the aircraft manufacturing process. Often they are personally responsible for ensuring that product design, manufacture, and quality

control comply with FAA regulations and industry standards. Because manufacturing of aircraft in the United States is heavily regulated by the FAA, engineers necessarily perform a regulatory compliance function.

Aviation engineers also monitor products for safety improvements and are responsible for coordination with the FAA should it be necessary for product modifications or the issuance of airworthiness directives.

17.21 ACCIDENT INVESTIGATION

If the design or manufacture of an aircraft becomes a factor in an aircraft accident investigation, the aviation engineer may be called upon to act as a party representative of the NTSB on behalf of his or her company. If the accident is not serious enough to involve an NTSB investigation, an aviation engineer may still be called upon by the FAA or his company to investigate the cause of the product failure and recommend improvements.

17.22 LITIGATION SUPPORT

In accidents where there have been serious injuries, death, or substantial property damage, litigation is likely. Frequently, manufacturers are insured and insurance defense lawyers undertake to defend the manufacturer against allegations of product defect. Because such allegations are highly technical in nature, the lawyers must necessarily request assistance of the company's engineers in order to determine whether a product was defective or whether standards were complied with. Engineers are called upon to aid in litigation support or forensic functions such as advising lawyers, or to assist them with documents such as interrogatories, requests for production, or depositions. In this context, an engineer is being called upon for his or her expertise but is not necessarily designated as an expert witness.

17.23 LITIGATION AVOIDANCE ("PREVENTIVE MEDICINE")

Engineers are often involved in a form of preventive medicine, sometimes called litigation avoidance. Some of the functions they perform may include reviewing designs prone to misuse, analyzing operating manuals to correct ambiguities, and identifying problems requiring warnings. Others may be involved in failure analysis or various forms of computerized troubleshooting to anticipate failure scenarios. Engineers may function as technical representatives who are in a position to ensure that the product is being used properly and with necessary safeguards. A critical function involves postmarketing product review to whether reports of incidents or user complaints may reveal impending problems.

Litigation avoidance may also involve careful review, storage, destruction, and/or dissemination of documents to avoid unnecessary legal exposure.

17.24 THE ENGINEER AS A TECHNICAL WITNESS

When litigation arises, engineers are frequently questioned under oath by lawyers concerning their technical knowledge. Engineers are often required to provide oral depositions as fact witnesses because they are the persons responsible for product design and manufacture. An oral deposition is an out-of-court meeting in which the attorneys representing the various parties have the right to interrogate the witness under oath concerning matters relevant to the litigation. After an oral deposition, a transcript of the questions and answers is prepared. The technical witness is required to sign for his or her answers under oath.

An engineer testifying as a technical witness is not necessarily functioning as an expert witness. In the civil court system in the United States, there is a distinction between a highly competent technical person who is called to testify in deposition or a trial on factual matters within his personal knowledge and an expert witness. The same person may be qualified as an expert witness to testify not only on factual or technical matters within his personal knowledge, but also to render opinions based on his expertise.

17.25 THE ENGINEER AS AN EXPERT WITNESS

Frequently, parties involved in litigation regarding serious air crashes have need to call upon engineers as expert witnesses. Often the engineers who function as expert witnesses are no longer employed by the company with which they gained their industry experience. They are independent consultants who are retained by attorneys to analyze the accident and render opinions regarding the breach of industry standards and/or causation. They are paid on an hourly basis for their research and analysis. Experienced experts have learned to make clear in testimony that they are being paid for their time, not their opinions.

At times lawyers for a manufacturing business will call upon engineers employed by the company to testify as expert witnesses in cases. This is more the exception than the rule because there is a concern that jurors or judges will presume that an engineer employee of a company will be biased in favor of the company when rendering expert opinions. Nevertheless, on some occasions the most qualified expert on a particular technical issue may be an engineer who is employed by the manufacturing company.

Engineers gravitate towards forensic expert witness work upon retirement because they find the field fascinating. They are called upon to use their wealth of training and experience to analyze crashes carefully and render opinions as to the causes of product failures and the accident. For an in-depth article on aviation expert witness work, see http://www.aviationlawcorp.com/content/successfulexpert.html.

17.26 THE ENGINEER AS ARCHITECT OF DEMONSTRATIVE EVIDENCE

Aviation lawyers must appreciate the significance of the expression "A picture is worth a thousand words." In aviation law perhaps more than any other specialty of

law, lawyers have become adept at utilizing sophisticated demonstrative evidence to explain aviation and technology to jurors. Accident reconstruction tapes, computer animations, flight test videos, etc. are often used as evidence. The lawyer cannot create any form of evidence; he must retain expert witnesses to design demonstrative evidence to persuade the jury.

The courts are sensitive to the persuasive value of sophisticated demonstrative evidence. A strict standard of evidence is employed to determine their admissibility. In most jurisdictions a technical professional functioning as an expert witness must convince the court that the demonstrative evidence is substantially similar to the actual accident. Thus, the recreation or other demonstrative evidence must be technically accurate for it to be admissible.

17.27 AVIATION PRODUCT LIABILITY

Strict Liability

Aviation product liability is a major issue in many air crash cases. After pilot error, mechanical malfunction is considered one of the more likely causes of air crashes.

When product liability cases were first heard by the courts, judges believed that victims would have an extremely difficult time proving a technical case against a manufacturer. In addition, because a manufacturer had released a defective product into the stream of commerce, many judges believed that the manufacturer, not the victim, should not bear the greater risk with regard to proof of liability. Accordingly, the concept of strict product liability was developed as a way of easing the plaintiff/victim's burden of proof against the manufacturer.

With strict product liability, the victim need only prove that the product was defective and that the defect in the product caused the crash and resulting injuries. It is not usually necessary for a victim in a product liability lawsuit to prove negligence by individual engineers, although some lawyers include negligence claims for tactical reasons. The focus in strict product liability is on the product and whether or not it is defective, instead of on the engineer and whether or not he fulfilled the standard of care. Because of the eased burden of proof and the perception that manufacturers are ''deep pockets,'' manufacturers are sued after many air crashes.

Design Defect and Manufacturing Defect

A design defect is a flaw which pervades the whole product line. A manufacturing defect is a problem in a single product or a few products which results from improper fabrication, inadequate quality control, or some other anomaly that does not infect the entire product line.

When is a product defective? There are two primary standards or tests used in the United States in different jurisdictions. One is called the risk-utility balancing test, in which the design of a product is evaluated by balancing the risk of design against the usefulness of design. If the risk outweighs the usefulness, the product may be considered defective in design. A key way of balancing the risk in a design is to look at whether an alternative safer design was mechanically feasible at the time the suspect product was designed. If the victim can prove to the jury that an

alternative safer design existed but the manufacturer did not use it, then the jury can find that the product was defective. Factors the jury must consider include not only the mechanical feasibility of the alternative safer design, but the relative cost and functionality of the alternative safer design.

The other test used in some states is not as specific. The jury is simply asked to evaluate whether the design of the product was "unreasonably dangerous." This test necessarily requires the jury to evaluate the product subjectively and determine whether it was unreasonably dangerous under the specific factual circumstances involved.

Pre-Sale Duty to Warn

Courts have recognized that some products cannot be made completely safe—aircraft are a prime example. There is always some life-threatening risk associated with complex products used to transport people in the air. Accordingly, a manufacturer's obligation is to design a product that is technically sound but also warn of those dangers that accompany the design. A manufacturer may be strictly liable not only because of technically deficient design but also because the manufacturer failed to warn of some of the risks associated with the use of that design. Manufacturers typically fulfill their duty to warn by promulgating complete instructions for use, affixing placards and other warnings to the product at issue, and issuing cautions and warning notices with the product when it is sold.

Post-Sale Duty to Warn

In various jurisdictions a manufacturer may be held liable if a product is determined to be defective even though the manufacturer did not know it was defective when it was sold. Thus, if a manufacturer receives notice of a safety hazard, the manufacturer will have a post-sale duty to warn of the problem. The typical post-sale duty to warn involves a service bulletin or other written warning issued by the manufacturer to the purchaser of the product.

In some jurisdictions, the courts have held that manufacturing has a duty to issue post-sale warnings not only to the immediate purchaser of the product but also to subsequent users. It may seem unfair to manufacturers that they would be held liable to a second- or third-generation purchaser with whom they have no business relationship. However, the original philosophy of product liability law is to ensure that the manufacturer, who put the product into the stream of commerce and made a profit out of it, should bear the risk of any consequences. The fact that many manufacturers carry insurance is also a factor in the courts' balancing of equities in this regard.

Proximate Causation

An important principle is the concept that a manufacturer should be held legally liable for injuries resulting from a product failure only if the product failure was a proximate cause of the injuries. There are situations in which there is a product failure but the operator has primary responsibility for the resulting injuries. For example, emergency procedures may establish a standard of care whereby the pilot

is expected to troubleshoot the product failure and land the aircraft safely notwithstanding the failure. In such situations a manufacturer's defense lawyers may be able to prove that the proximate cause of the crash was pilot error and not a defect, even though there was a product failure. A similar situation arises when a manufacturer designs, manufactures, and sells a safe product but the product is inadequately maintained, improperly repaired, or modified by a subsequent owner. Thus pilot error, maintenance malpractice, government negligence, or any one of a number of other causes may be asserted by the manufacturer's lawyers to show that the product defect was not the proximate cause of the crash.

Crashworthiness

I used to defend a general aviation manufacturer with engineers who proudly told me they designed airplanes "to fly, not to crash." I told the old-timers that under modern product liability law, they were wrong. In addition to designing for reasonably foreseeable use, they also had to anticipate reasonably foreseeable misuse.

Manufacturers must design aircraft with a certain degree of crashworthiness. Thus, cockpit integrity, seat design and fuel system vulnerability, etc. are taken into account along with operability. The aircraft must be designed for foreseeable minor crashes where g-forces are not excessive so that the occupants can avoid major injury. There is no precise line of demarcation as to when a manufacturer has designed a product sufficiently to be crashworthy. Courts use a balancing test, weighing the foreseeability of serious injury associated with the design against the feasibility of employing safeguards to prevent unnecessary injuries in survivable accidents.

Post-Accident Remedial Measures

After a product failure, the manufacturer often undertakes a product modification to make it safer. A realistic concern of the engineers involved in the redesign is the fear that the modification will constitute an admission that the product was not designed safely in the first place. Court rules have been promulgated to deal with such post-accident remedial measures. Litigants are not permitted to introduce evidence of post-accident remedial measures to prove defect or any form of liability.

Unfortunately, as for many good principles, there are exceptions. Courts allow victims' lawyers to introduce evidence of post-accident remedial measures for very limited purposes. Jurors are instructed not to consider the post-accident remedial measure as evidence of liability but simply as evidence of the feasibility to make such changes. Unfortunately many jurors are not able to distinguish between these concepts, and the remedial measures must be explained by defense lawyers.

State of the Art

Generally speaking, the courts in most jurisdictions hold manufacturers to a standard of design consistent with the state of the art at the time the product was sold. Normally, a manufacturer will not be held liable for a crash in the year 2002 if an aircraft designed in 1972 was not as safe as design capabilities of the year 2002. There are some limited exceptions to this concept. If product safety retrofits and

modifications become the norm in the industry, the manufacturer who fails to follow the industry trend may be exposed to product liability.

Product Modification

One of the strongest defenses a manufacturer has against product liability is the proof that after the sale of the product somebody modified the product so that it no longer conforms to the manufacturer's design. Any product modifications must be made in conformity with an FAA supplemental type certificate and approval of the manufacturer. Operators who, even with good intentions, modify aircraft or safety checklists without manufacturer approval may incur liability for subsequent accidents.

Adequacy of Warnings

When defects or other problems cause an accident after a product has been sold, it is necessary for the manufacturer to issue warnings to prevent recurrence of similar accidents. Can post-sale warnings concerning newly discovered defects protect a manufacturer against product liability?

Under the laws of many jurisdictions, post-sale warnings may protect the manufacturer against product liability if the warnings are "adequate." Manufacturers may be quick to issue notices regarding safety improvements or design improvements. They are sometimes reluctant to acknowledge that there is a defect in the product or to frighten users with safety warnings.

To be "adequate," a warning must sufficiently focus on the precise nature of the problem and conspicuously alert the user to the risk of continued usage without improvement. Courts also examine whether manufacturers, when dealing with parties of unequal bargaining power, offer to pay for the fix. The reasoning of the courts is that the manufacturer is responsible for deficiencies in the product and should not impose upon the user the financial burden of remedying a problem originating with the manufacturer.

Airworthiness Directives

The aviation industry in the United States does not use the product recall system that is used by the automotive industry or other industries. Instead, the manufacturer provides the FAA with the technical data to show that it is necessary for the FAA to issue an airworthiness directive. An airworthiness directive is an order from the FAA tantamount to a law. The aircraft may no longer be used until the deficient designed is brought into conformity with the requirements of the airworthiness directive. Airworthiness directives may require immediate compliance, or if safety permits, delayed compliance under specified terms.

Government Approval / Minimum Standards

Generally, in most jurisdictions in the United States, the fact that the product conforms to FAA minimum standards does not alone immunize the manufacturer

against a product liability lawsuit. Certainly the defense lawyers may introduce evidence that the product conformed to federal standards in defense of the product. However, courts have held that compliance with FAA minimum standards is not the same as showing that the product is safe and nondefective under product liability standards.

The Government Contractor Defense

Military products and certain civilian government products are often designed for combat or special operations. The government may prefer to spend money on weaponry or other operational capabilities instead of on enhanced safety. Manufacturers have been sued by victims of military air crashes who have claimed that the lack of safety features in the aircraft caused the injuries. Manufacturers argued that they should not have to pay for such injuries when they could have designed the aircraft more safely but the federal government contract limited expenditures to certain specified features.

In 1988, the U.S. Supreme Court ruled that government contractor products may be immune from product liability. For immunity, the government contractor must prove three things:

1. That the product conformed to reasonably precise specifications approved by the U.S. government
2. That the product as manufactured actually conformed to the reasonably precise specifications
3. That the government contractor warned the government of any dangers associated with the product that the government was not already aware of

Companies that manufacture products outside the government contract procurement process do not enjoy this special defense.

The General Aviation Revitalization Act

General aviation manufacturers in the United States lobbied Congress for many years for special legislation to protect them against product liability lawsuits. Manufacturers such as Cessna, Beech, and Piper claimed that product liability lawsuits and insurance costs were driving them out of business. In 1994, Congress responded with unique legislation, the General Aviation Revitalization Act (GARA), 49 USC §40101. GARA provides special protection for general aviation manufacturers who sell aircraft with a seating capacity of less than 20 seats, not intended to be used in air commerce. Under GARA, the general aviation manufacturer will become immune to suit 18 years after the date of sale. If a product failure causes an accident in the first 18 years, the manufacturer can be sued.

GARA shortened the liability tail of general aviation manufacturers for their older models. This law does not necessarily apply to the entire aircraft. If a new engine, new landing gear, or other new component is installed on the aircraft after original sale, the 18-year time frame starts from the date of the sale of the subcomponent part.

17.28 CONCLUSION

I have spoken with engineers who consider the law, or sometimes lawyers, to be counterproductive to aviation safety. This problem arises where witnesses are afraid to discuss problems or manufacturers are afraid to make changes due to liability. Accident investigators sometimes express frustration that they are not allowed to pursue their safety goals because of concerns that statements for safety investigatory purposes will be used in subsequent litigation and will result in liability.

It is true that the policy goals of the law and those of aviation safety interests sometimes collide. Laws focus on rights, liabilities, and compensation for injuries. Safety professionals care about preventing the recurrence of the accident. Despite this conflict, there are safety enhancements from aviation litigation.

Aircraft manufacturers and their engineers exemplify the highest standards of professionals and ethics. Yet there have been product defects which have gone uncorrected until accidents occurred. Sometimes the same product failures occurred repeatedly, resulting in multiple accidents and deaths. Federal accident investigators are often too limited in budget and manpower to investigate all accidents as thoroughly as they desire. Sometimes lawyers and insurance companies have more resources to determine the cause of crashes than federal investigators. Litigation, or the fear of litigation, has caused parties to make safety changes that they might have been reluctant to make without this threat. Air crash litigation, although criticized as costly, has in some cases spurred aviation safety.

In my experience, the true professionals in both the engineering profession and the legal profession are able to work in harmony to satisfy the respective obligations of their professions. This is particularly true when they have made a point of studying and understanding the responsibilities and goals of each other's professions.

PART 3

ACCIDENT ANALYSIS DATA

Donald W. Pitts

Since the mid-1960s, the commercial aviation industry has achieved an extraordinarily high level of safety. For the past three decades the annual hull loss rate for U.S. operators of large transport aircraft has averaged 1 aircraft per million departures. In comparison, the worldwide rate of aircraft destroyed has averaged 1 to 3 accidents per million flights (see Figure 17.1). Clearly these statistics highlight a marked improvement over the rates recorded in 1959, when U.S. commercial operators suffered 26 major accidents per million flights. Formulation of the Federal Aviation Administration (FAA) in 1958 and its safety charter clearly illustrate that a high level of safety oversight is assumed and expected by the American public.

The uniqueness of aviation accidents, coupled with the large numbers of deaths associated with a single aircraft loss, spotlight these events with a high level of public awareness—even though quite infrequent, they gather tremendous attention as events leading up to the catastrophic loss of life are recounted by the media. Figure 17.2 illustrates how widely the number of fatalities can vary with a fairly constant accident rate. The traveling public and the concerns they raise to lawmakers ultimately demand demonstration of due diligence by responsible agencies in the aftermath of an aviation accident. In the United States, this was most apparent

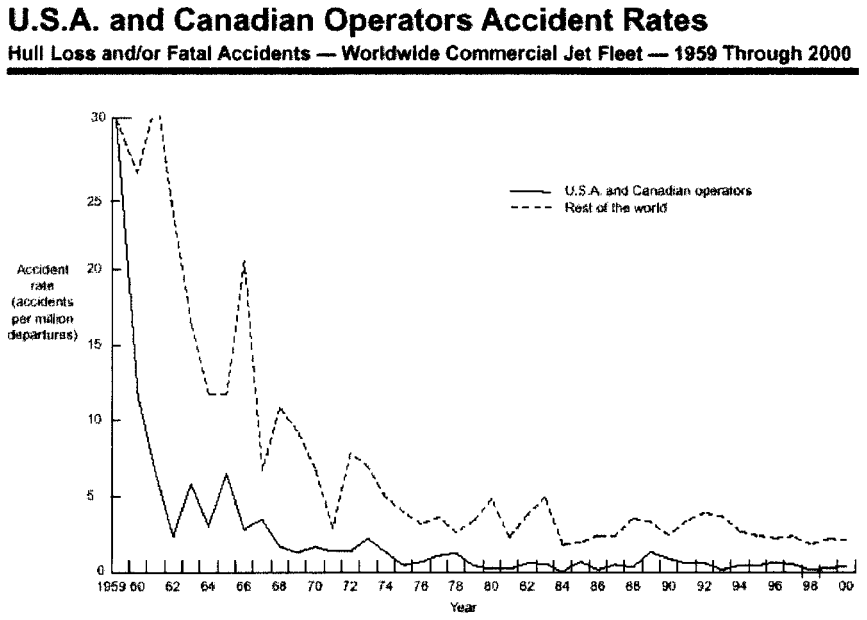


FIGURE 17.1 (From Boeing 2001, p. 15.)

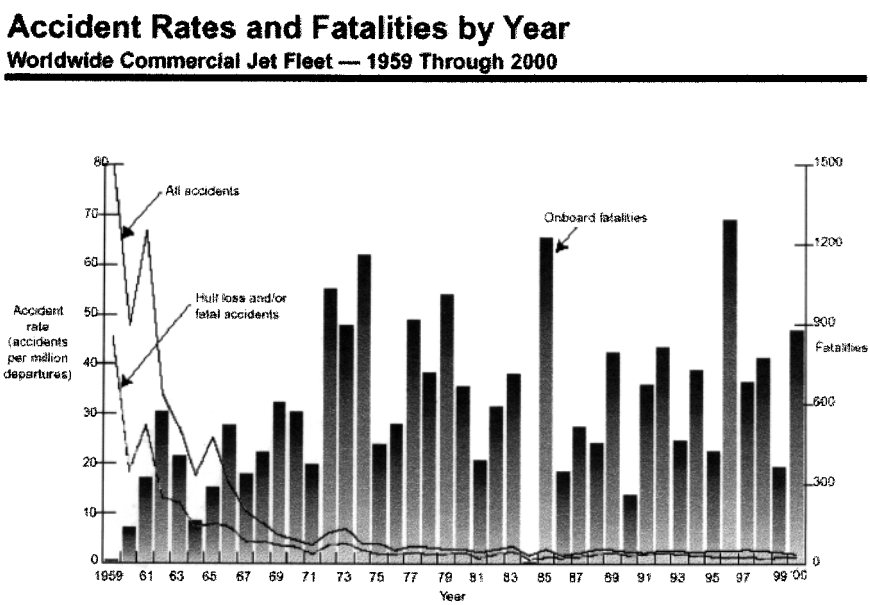


FIGURE 17.2 (From Boeing 2001, p. 13.)

in 1996 with the formation of the White House Commission on Aviation Safety and Security. This commission, chartered by President Clinton, responded to the public outcry following the crash of a ValuJet DC-9 into the Florida Everglades, followed closely by TWA 800, the complete destruction of a Boeing 747 departing JFK International Airport. In these case studies, a general sense of inadequate safety standards and/or increased threat from terrorist activity lowered the public's confidence in aviation's system wide safety. Even though the risk of exposure to a life-threatening event is significantly higher in other endeavors, public acceptance of that same level of risk is not tolerated in aviation—a higher standard of prevention is expected.

Many attribute the significant improvements in aviation safety over the past 40 years to technological advances resulting from reliability of power plants, i.e., the departure from piston-powered to turbine-powered aircraft. Additionally, the better understanding of accident causal factors made available by flight parameters captured on data recorders has led to significant improvement in aviation safety. With the assistance of recorded data, post-accident analysis provided safety recommendations which, when acted upon, add layers of protection to the system as a whole and generally raise the safety bar, lowering the risk. The compelling force for change comes in the form of safety recommendations from various agencies and regulations written into law by the FAA as Federal Aviation Regulations (FARs). This reactive process, of incident/accident, investigation, recommendation, and regulation, has clearly improved the safety record of aviation compared to that of the record of the 1950's. As successful as this record is, it still falls short of the moral obligation safety professionals face daily as they pursue the idealistic goal of zero losses while quietly acknowledging there are no risk-free endeavors. This is the

challenge aviation experts (and those who manage them) face as they strive to meet the public's expectation of no human tragedy, demonstrated by a continuous effort to improve safety records.

As a result of a 1958 FAA mandate, the first accident-protected analog flight data recorders were placed into service, with four flight parameters captured on a foil recording medium. Today, the latest generation of digital flight data recorders (DFDRs) provides hundreds of flight parameters and records not only dynamic flight values but also the status of various systems and subcomponents of the aircraft. These recorders capture factual data which can be used to better understand causal factors in aviation accidents. As a result of these data, better information has been obtained over the years, identifying the largest contributing factor to the accident chain of events to be the human element—found causal in 70–80% of accidents. Investigations today strive to go beyond simply placing the blame at the feet of the individual(s) holding the last link in the chain of events leading up to an accident. In the past, those reviewing incidents with a compliance mindset along the lines of strict regulatory enforcement of FARs have focused criticism of wrongdoing on these lesser human performances, failing to allow credence to the honest mistakes that professionals will make. This has proven ineffective in prevention of future occurrences of similar events and had discouraged the reporting of anecdotal information necessary to reveal hidden systemic deficiencies. In this scenario, punishment is rendered and the books are closed on the event, while the larger community of aviators does not gain benefit and a hazard typically lies dormant, awaiting another prey. Over the years, identifying what has been irreverently referred to in investigation circles as the ''Guilty Bastard'' and administering an appropriate corrective action (punishment) has done little or nothing to improve human performance and human susceptibility to errors caused by lack of education, poor aircraft design, faulty maintenance, natural phenomena, inadequate procedures, poorly developed missions, or inadequate standards and controls. Proper corrective action and follow-on control measures cannot be developed if the precursors that lead to an accident chain of events are never brought to the attention of those with the ability to understand the need for or the authority to implement change. Post-accident review from investigative agencies such as the National Safety Transportation Board (NTSB) with their subsequent safety recommendations have clearly been a positive influence on the overall system safety; however, only in recent times has significant progress been made in addressing the man–machine interface and the systemic hazards that may have triggered an accident.

17.29 ZERO ACCIDENTS: THE QUEST FOR BETTER PREVENTION

The primary goal of any safety program must be the prevention of incidents and accidents. Identifying situations that require corrective action before problems occur not only supports the zero accident goal but can also reduce operational costs and significantly enhance training effectiveness. This translates into operating efficiencies and cost savings. To achieve the highest levels of accident prevention, the hazards faced by those at risk must first be identified accurately enough to allow development of control measures to mitigate that risk. This is no easy task.

Figure 17.3, an equilateral triangle divided into four sections, depicts Heinrich's Pyramid, a safety model. The top two regions of the triangle represent accidents

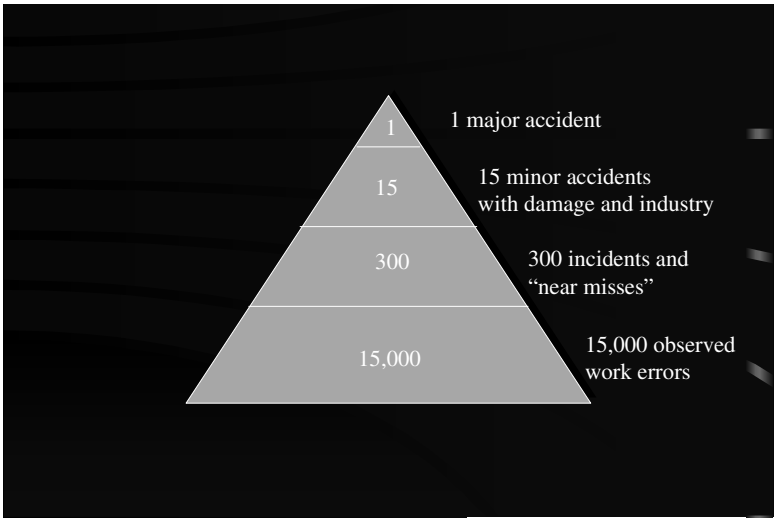


FIGURE 17.3 Accident pyramid.

(major and minor). The third section represents incidents, while the base contains observed but unreported work errors. In the case of one major U.S. airline, these everyday unreported occurrences could stem from as many as 2,500 daily flight operations. Clearly, accidents will draw scrutiny, while incidents, depending upon the severity, will also be reviewed by investigators. However, the largest body of information lies in the unreported occurrences which, for the most part, go unnoticed except by those who personally experience the event. This model represents root cause(s) of a catastrophic event—the accident resulting in loss of life. The point of this characterization is that most of the information highlighting root causes is not noticed during the everyday normal conduct of business. The primary reason for overlooking this information is failure to identify/recognize the hazard or simply not enough resources to spend reviewing otherwise innocuous events. Absent this precursor knowledge, safety professionals are condemned to providing after-the-fact accident analysis in a post mortem effort to understand what happened and why it occurred. Understanding the ''what–why'' relationship is critical to every safety program, be it reactive or proactive. Realizing that rarely is a deliberate or intentional violation found to be involved in an incident, some means is needed to induce those involved to come forward freely and provide details of the incident. The concept of self-disclosure has been considered to be indicative of a constructive attitude worthy of protection in the pursuit of safer operations.

Airline executives realize that technology and operational know-how reduced the accident rate of aviation to today's extremely low levels. However, it is also well understood that as consumer demand for commercial aviation travel increases, the total number of accidents will also increase given the mathematical relationship of increasing the numerator applied to the current (tolerated) accident rate. Without a reduction in the rate, this rise in demand is forecast to generate more total accidents than considered acceptable by the public. Furthermore, three decades of statistically significant data suggest that dramatic improvements in aviation safety are

unlikely if the traditional regulatory compliance/enforcement reactive approaches are taken toward future safety programs. The common belief among aviation safety experts is there will not be a significant reduction in the actual accident rate until front-line operators (managers, supervisors, and employees) are in the mindset to provide experts the precursor information they need to prevent the next accident. To encourage this information flow, an improved safety culture that shares values, procedures, and skills must be present. This allows stakeholders jointly to identify and respond to hazards without fear of reprisal, prosecution, or disciplinary action by any authority reviewing the information. To move beyond the current plateau and truly become proactive, safety processes will need a more powerful tool—one which will allow analysis of previously identified and potential failures with the focus on how best to apply limited resources to mitigate future risks.

One hurdle proactive safety departments face is that public laws hold the FAA responsible for the proper investigation and disposition of all suspected cases of noncompliance with the FARs. The FAA Administrator has a statutory responsibility to enforce safety regulations, and in previous times the FAA drew its approach for corrective measures from civil and criminal law. In these cases the underlying assumption was one of commission of error through negligence or recklessness. Fortunately, today it is recognized that highly skilled professionals can and will make mistakes, regardless of regulatory prohibitions. What must be better understood are which factors or influences contributed to the human error.

17.30 HAZARD IDENTIFICATION: THE PURSUIT OF INFORMATION

U.S. military services led the way in aviation accident investigation following their first aircraft accident on September 17, 1908. An aircraft piloted by Orville Wright with observer Lieutenant Thomas Selfridge crashed while demonstrating the machine to U.S. Army officials at Fort Myer, VA. Lieutenant Selfridge was killed. Wright survived with broken bones and was able to provide key information to the investigators. Post-accident review indicated that elongated propeller blades of a never-tested design led to excessive vibration, which caused the propeller to strike a guy wire on the aircraft, tearing the wire from its fastening in the rudder and breaking the propeller two feet from its tip.

Obviously, critical information to the investigation resided with the first-hand knowledge of the surviving pilot, who provided insight necessary to take corrective action and thus prevent recurrence. Without this valuable input a similar catastrophe could have afflicted another unwary operator. By the 1940s, both the military and industry recognized the value of voluntary incident reporting in an effort to prevent accidents. The lessons-learned approach in military aviation influenced their civilian counterparts. A need for a U.S. Incident Data System was raised during the FAA enactment hearings in 1958. Following a United Airlines incident in October 1974, which foreshadowed a TWA accident in December 1974, a study of the National Air Transportation System was conducted by a task force focusing on the FAA's safety mission. In May 1975, Advisory Circular 00-46 was issued. With that circular, the FAA first implemented the Aviation Safety Reporting Program. In 1976, it was later modified and implemented by both NASA and the FAA as the Aviation Safety Reporting System (ASRS). This program receives, processes, and analyzes voluntarily submitted reports from pilots, air traffic controllers, and other aviation

industry stakeholders. Since the mid-1970s, NASA has collected hundreds of thousands of reports outlining human performance errors in the National Aviation System (NAS). The purpose of this program is to identify deficiencies and discrepancies in the NAS, with particular concern paid to the quality of human performance. Once a deficiency or discrepancy is identified, the objective is to improve the current aviation system and provide data for future planning and improvements system-wide. This program identifies potential safety problems while providing limited protection to airmen reporting under provisions of FAR Part 91. A paradigm shift from regulatory compliance and punitive-based incentives to assurances of nondisciplinary action against self-disclosed operational errors was evolving.

The value of ASRS as an aviation safety research database is recognized worldwide; however, even though the program processes large quantities of data, its ability to correct identifiable aviation hazards is severely limited. Data from ASRS implied that numerous significant events were occurring in the NAS that were unrecognized by the airlines or the regulators. Yet typically ASRS could not report details of specific events back to the airlines or the FAA due to requirements of confidentiality and jurisdiction. Furthermore, the airmen involved were unlikely to report this precursor information to anyone other than NASA without greater regulatory assurances that the report would not be used to precipitate a follow-up review and potential action taken against them. Therefore, detailed analysis of incidents that could potentially lead to accidents was not always available to those who *could* implement corrective measures. Be that as it may, progress was being made, moving from regulatory compliance and punitive-based incentives to assurances of nondisciplinary action against self-disclosed operational errors—thus providing a better analysis tool of the root causes during hazard identification.

Taking the philosophy of nonpunitive safety processes one step further was the Allied Pilots Association, which represents the professional interests of the pilots serving American Airlines. In 1994, then-chairman of the National Safety Committee Captain K. Scott Griffith envisioned a program to provide a program whereby pilots would actively identify and report safety issues to a tripartite review composed of the pilot's union, management, and the FAA. A key element of this critical information source was confidence that neither the airline nor the FAA would use those same reports to take enforcement action against an errant pilot. A trial program was allowed for 18 months based on pilot self-reporting and a cooperative effort to improve safety through a partnership among the three signatories to the memorandum of understanding. The program was called Aviation Safety Action Program (ASAP). ASAP extended the regulatory assurances, so critical to ASRS success, and further provided for the collection, analysis, and retention of safety data that would otherwise be unavailable to either the FAA or the airline's management. ASAP reports did not preclude continued data collection by the human performance experts at NASA but rather multiplied the effort by providing an additional stream of information to the ASRS database as all ASAP reports automatically generated an ASRS report. The Event Review Team, made up from the three parties of the tripartite, would then further analyze the ASAP report.

Since the only way of achieving a goal of zero accidents was to apply corrective action before an accident occurred, the objective of ASAP was to have an unfiltered view of operational hazards with aggressive and timely corrective action. In order to ensure the benefit of self-reporting, the FAA offered certain nonpunitive enforcement-related incentives to encourage individual employees and the certificate-holding airline to report incidents of inadvertent noncompliance with the FARs *even* when the infraction went unnoticed. ASAP is based on the principles of identifi-

cation and corrective action rather than immunity. The only immunity associated with ASAP is the relationship existing from participation in the NASA ASRS program, which is independent of ASAP. The ASAP program was granted several extensions over the next five years and proved to be so successful that other airlines have now adopted this model. ASAP was ultimately incorporated into FAA Advisory Circular 120–66A and forms the foundation of self-reporting hazard identification programs which allow safety experts to formulate preventative recommendations based on quantifiable data gathered in a nonpunitive environment.

17.31 THE NEED FOR DATA-DRIVEN PROCESSES

Insight from operational flight data recorders and information derived from crash-survivable accident data has proven to be invaluable in the identification of accident causes and contributing factors. Developed over the past decade, a strong consensus exists among aviation safety professionals that making safety improvements based on reactive accident data, and to a more limited extent some incident data, can sustain us at our current level of safety, but it may be insufficient to adequately anticipate future problems. A need exists for more robust data, recorded automatically without dependence on a dedicated act or effort of the observer. These data would then be compiled into aggregate data and plotted in a manner that highlighted events falling outside some predetermined normal distribution, thus pointing to precursors of incidents or accidents that could then be reviewed for systemic concerns. This identification of root causes provides high leverage potential where future accident-prevention plans could be evaluated based on the numbers of safety events that would be addressed, the severity of those events, and expected effectiveness of the plan. Using this approach, accident prevention would move beyond tactical to more strategic operation. This far-reaching ''macro'' view builds the foundation of a comprehensive safety risk management program.

17.32 FLIGHT OPERATIONS QUALITY ASSURANCE: THE PURSUIT OF KNOWLEDGE

In January 1995, the Department of Transportation's Aviation Safety Conference recommended that the FAA encourage and facilitate the voluntary implementation of a program using airborne digital recording equipment to record flight data for routine monitoring of operations. This Flight Operations Quality Assurance (FOQA) data, when combined with other data and operational experience, would then be used to develop objective information that enhances flight, safety, and maintenance decisions.

FOQA-like programs developed by European air carriers over the past three decades indicate that FOQA data are a valuable source of information that significantly enhances aviation safety. Those airlines currently using FOQA agree that the insights derived from established programs have prevented serious incidents and accidents. Based on this past operational experience and the recommendations of the Flight Safety Foundation, the FAA began to consider FOQA programs as a tool for continuously monitoring and evaluating operational practices and procedures.

Consistent with the lessons learned in ASAP, which offered an alternative to traditional FAA legal enforcement in proactive safety programs, the FOQA approach was chosen in 1998 with the announcement of the FAA's Safer Skies initiative. This program, based on a comprehensive review of the causes of aviation accidents, resulted in adoption of a focused priority safety agenda designed to bring about a fivefold reduction in fatal accidents—a truly aggressive goal. With Safer Skies, the FAA agreed to concentrate its resources on the most prevalent causes of aircraft accidents and vowed to use the latest technology to help analyze U.S. and global data to find the root causes of accidents.

Safer Skies vowed to use partnerships between the FAA and the aviation industry. These programs would include ongoing analytical programs to determine causes of accidents. Once understood, intervention strategies would be developed and evaluated to determine which mitigating factors provided the largest ''bang for the buck'' impact on safety. Once implemented, effectiveness was to be reviewed during internal evaluations in a feedback loop. Thus, the initiative will use data in new ways that allow operators, manufacturers, and the FAA to focus on breaking causal sequences and taking action before an identified chain of events leads to an accident.

To support the data requirements of such an effort, a program which used empirical data such as that recorded on flight data recorders was needed. Unlike post-accident investigations, the program sought to provide decision-makers better tools to manage operational risks before an incident or accident. Through the use of computer software programs processing massive quantities of data, safety experts would review distributions of operational activities, thus moving from reacting to mishap investigation information to acting on precursor data. This results in a proactive approach focusing on systemic concerns, allowing managers to anticipate problems and avoid costly surprises. As one Air Line Pilots Association representative put it, ''If you can't measure it . . . you don't know about it. If you don't know about it . . . you can't fix it.''

A FOQA program will identify latent conditions in flight operations and help quantify active trends. This trending and the associated change program activities provide feedback to operational practices and training. Previously unknown relationships are discovered using atypical event data review empowering ''drill down'' analysis. Additionally, FOQA programs provide feedback for development of standard deviations in future risk management of like events—what we don't know can hurt us. FOQA is the Rosetta Stone of a proactive safety program, providing the key to understanding how to optimize system performance.

17.33 CONCLUSION

The aviation industry as a whole must be aware and constantly vigilant of the pressures exerted on those corporate executives fiscally responsible for conducting business in the airline industry. Revised capitalization plans, innovative management practices, and maintenance and training savings are tactics used in an effort to achieve a competitive business advantage in a marketplace offering a slim 2–5% profit margin. These executive decisions most certainly are preceded with empirical data from ''bean counters'' who fully appreciate the value of a dollar but rarely foresee the unintended consequence of reducing or failing to implement a safety layer of protection.

The benefits of proactive data-driven processes are clear. The National Civil Aviation Safety Review Commission stated that ''whenever possible, FOQA should become part of safety risk management programs.'' Their recommendations went on to state that FOQA systems have the potential to become the basis for aviation safety decisions at three levels: the company, the air crew, and the air transportation system as a whole. They further went on to state that FOQA information on a national level could identify faults in systems, procedures, airport operations, air space structures, aircraft certification, and human–automation interface. Accomplishing the goal of a dramatic reduction in the accident rate will require this kind of strategic plan, one that places emphasis on a cooperative interaction of information sharing and collaborative development of solutions to safety issues. As an aircraft designer, how will you ensure your design fits into the FOQA program?

REFERENCES

Boeing Commercial Airplane Company, 2001. *Statstical Summary of Commercial Jet Airplane Accidents, Worldwide Operation, 1959–2000.* Seattle: Boeing Commercial Airplane Company, June.

PART 4
RISK MANAGEMENT

John D. Phillips

17.34 INTRODUCTION

Risk management principles and methods have long been associated with liability avoidance, but they have recently have been adapted to managing safety and other risk issues. This new application of risk management concepts and techniques has had a dramatic effect on the way safety issues are managed, but unfortunately, the approach is so simple that its value is often underestimated. On the surface, risk management is essentially a common-sense approach based on variations of standard decision-making models. The application of a well-organized and consistent process, combined with the use of helpful techniques, makes risk management a powerful business asset. The basic idea is to manage safety-related risks so that an optimal level of risk is achieved. This acknowledgment that some level of risk must be accepted, given virtually any situation, is one of the key differences between risk management and the traditional, compliance-based, safety approach. It is important to recognize that risk is inherent in every operation, system, and process. Few, if any, things are perfect, and likewise not many are absolutely safe. So how safe is safe enough? The trick is to make informed decisions about what level of risk is acceptable given a particular situation, while still accomplishing whatever is intended. Yes, there are laws, standards, and other regulatory requirements that must be complied with, but they do not cover every situation and generally only provide for a minimum level of safety. Risk management is the overarching decision-making methodology that supports the attainment of a proper balance between risk and opportunity.

17.35 RISK MANAGEMENT GUIDELINES

Several guidelines must be consistently adhered to when applying risk management. The first is that no unnecessary risks are acceptable. This seems obvious, but if a risk is not identified or is not fully understood, it is accepted by default. The corollary to this rule is that necessary risks must be taken—that is, when benefits genuinely outweigh risks, the residual risks should be accepted. Again, this assumes that there are no unnecessary residual risks. Another key principle is that risk decisions must be made at an appropriate level. Risk decision-making will be discussed in more depth later in this Part, but it is important to note that people need to understand which risk decisions they are empowered to make and which will require higher-level approval. Whether this understanding is based on general guidance or formal policies, people should be aware of their authority with regard to risk decision-making.

17.36 THE RISK MANAGEMENT PROCESS

The actual application of risk management can be thought of as a circular or repetitive cycle. It begins as early in the planning or design stages as possible to be

most efficient and effective. The process then continues throughout an activity or life cycle of an item and can be as formal or informal as is dictated by the situation. Also, the breadth of issues dealt with can vary depending on the desires of the individual. The scope of risk management application is based on the definition one uses for the term hazard. A hazard can be narrowly defined as a safety-related concern or, ideally, more broadly, to include anything that might have a negative impact on an activity, system, or process. Hazards may include security, financial, environmental, political, and other types of issues that could potentially produce an undesirable outcome. As one might imagine, given the variety of issues under consideration, the range of undesired events is varied as well and could include, for example, injury, loss of an asset, failure of a system, negative public opinion, tort liability, etc. Of course, all decisions involve weighing pros and cons in order to arrive at the best course of action; risk management provides a systematic approach that results in more consistent and better-informed decision-making. While a variety of risk management models are used in both the private and public sectors, the six-step model (see Figure 17.4) is a particularly good one.

Step 1: Identify Hazards

The first and arguably the most important step in the risk management process is hazard identification. Hazard identification is critical to the overall process since unless hazards are identified, nothing is likely to be done to counteract or eliminate them. The longer a hazard goes undetected, generally the more disruptive and costly the corrective action will be. Luckily, there are a wide variety of techniques that can be employed to identify hazards, and most are relatively easy to use regardless of whether an operation or system is being evaluated. The selection of method(s) to use in hazard identification is dependent on the type, complexity, and/or risk of the situation and the time, resources, and experience of the person doing the assessment. The following techniques are only a sampling of those that are available.

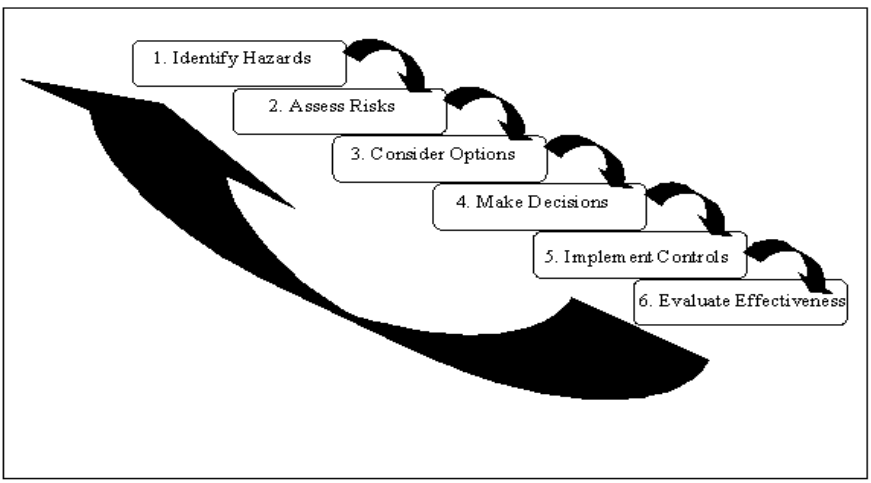


FIGURE 17.4 Risk management process model.

While one person can use any of these methods, it is desirable in many cases to involve several people who are knowledgeable about the operation or item being assessed. The output of this step, regardless which techniques are employed, is a list of hazards that can be further assessed in subsequent steps.

Hazard Identification Techniques

Preliminary Hazard Analysis. The preliminary hazard analysis (PHA) is a simple technique that can be used alone or as a quick look before using more detailed methods. It is used in nearly all risk management assessments. The PHA is performed by breaking the activity or system into logical segments, whether pieces or stages. Each segment is then considered and hazards are detected using the experience and intuition of those involved as well as any other available resources such as mishap database information, standards, etc. The benefit of the PHA results from the forced consideration of all aspects of an operation or system so that hazards are not overlooked by unorganized review.

Logic Diagram. One of the most popular and effective hazard identification methods is the logic diagram. This method can be much more analytical than others described herein and is intended to illustrate that a more deliberate approach may be taken. While there are several types of logic diagrams, two types, the negative logic diagram and the positive logic diagram, are primarily used in hazard identification. Either type of diagram can be used to assess operations (i.e., fueling procedures, hazardous material handling) or systems (i.e., aircraft and aircraft components). The diagram can be performed at a fairly macro level or can be very detailed, depending on the need. Procedures for each are described below.

The negative logic diagram begins with a negative or undesired outcome (i.e., wheels-up landing). The evaluator then lists the ways this outcome might occur (i.e., pilot fails to lower gear, mechanical failure of gear to lower, etc.). Once all the reasons have been identified, all possible causes for each way the outcome might occur are listed (i.e., pilot failed to lower gear). The causes for these causes can also be listed, and so on (e.g., pilot forgot because he failed to use checklist, etc.). How deep a person should go varies, but generally three or four levels is adequate. Going to this level of detail allows more effective targeting of corrective actions later in the process since the actual source of a potential problem can normally be identified. Once complete, hazards are taken from the lowest-level causes in the diagram.

A positive logic diagram, as might be expected, is just the opposite of the negative logic diagram. Instead of an undesired event at the top, the desired outcome is used as the starting point. Each thing that must be in place for that outcome to occur should then be listed. Again, as in the other diagram, subfactors may be listed to obtain increased granularity. When interpreting the diagram, instead of the lowest-level factors being hazards, they are actually the key things that must be in place for the positive outcome to occur. The user, in order to identify the hazards, must therefore consider what might prevent these things from occurring. A very simple example is that in order to have a fire, fuel, oxygen, and a heat source are required (assuming a fire is a desired event, such as internal to an engine). The hazards are those things that would prevent one or more of these items from existing in the proper quantities.

What If. The ''what-if'' technique is probably the simplest, yet one of the most effective, particularly if a group of people is assisting with the risk assessment. The technique involves beginning with the statement ''what if'' and continuing with a

description of what ''might'' occur. An example, involving security on commercial aircraft, might include statements such as ''What if a passenger threatens a crew member with a knife?'' or ''What if an employee smuggles weapons onto the aircraft before the flight?'' These statements can be used to identify hazards that could allow or contribute to these events.

Change Analysis. Another widely used technique is change analysis. Change analysis is a real time saver and is used to identify risks resulting from changes to operations, processes, or systems. It is usually used when reevaluating a situation that has already been assessed, but it can be used anytime in looking at the risk implications of a change or modification. Change analysis begins by listing the differences between the current situation and what is planned. Once the differences are identified, their impact can be assessed. An example of a change where this technique could be used is when changing from one type of equipment to another, such as switching from electric- to gas-powered equipment. Change analysis is also a good technique to use in order to evaluate a situation periodically even when no intentional change is being made to identify those inadvertent changes that have occurred that may be a new source of risk.

Step 2: Assess Risks

Once all possible hazards have been identified, they must be prioritized based on their significance. This assessment, although subjective, if done consistently, provides a good stratification of issues to deal with. This is especially important when time or other resources do not afford the luxury of working every issue. The two variables used to rank each hazard are severity and probability. Both of these variables can be divided into levels, from high to low, and each level may be defined, if desired, so that more accurate and interpretable risk assessments are made. When definitions are used, specific rates, percentages, and other figures can be worked into the definition in some cases, whereas in other instances more general descriptions are more desirable and appropriate. The first variable, severity, is an estimation of how serious the negative event possibly produced by the hazard would be (i.e., a death, loss of an aircraft, loss of communications capability, etc.). Probability is the likelihood of the undesired event occurring. In order to determine a risk level based on the combination of severity and probability, a risk matrix is commonly used. A standard risk matrix is shown in Figure 17.5. To use the matrix, simply identify the severity on the left, find the probability of the outcome across the top, and then locate the intersection on the matrix for the risk level. Hazards, once their risk is determined, can be ranked from high to low, allowing priority risks to be addressed first. The overall risk of the operation, process, or system can be computed as a sum of all identified risks or simply assigned the value of the highest risk.

Step 3: Consider Options

The next step is to consider which control options might be chosen for the identified risks. The various control options, with a description of each, are provided below. This list offers a variety of alternatives that, while not applicable to every situation, are all worthy of consideration. These options can be applied to individual risks and/or the entire operation, process, or system being assessed. Begin this step with

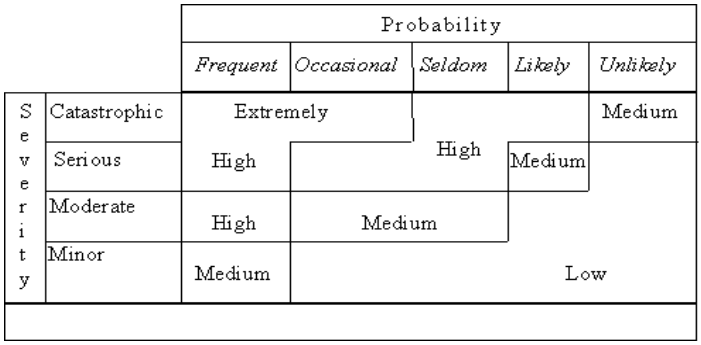


FIGURE 17.5 Standard risk matrix.

the highest risk and review the options for control. When considering options, look for opportunities to control effectively more than one risk issue with a control. This grouping of risks is a great way to save both time and resources when applying controls. If taking the risk decision to someone else, be sure to list all practical options. The person conducting the assessment should not limit the options too narrowly or exclude any unless truly not applicable, otherwise you are providing the decision-maker with an incomplete picture upon which to make a decision. Yes, it is appropriate to make recommendations to a decision-maker, but do not make the decision yourself by not giving the whole story. This will also help to avoid being embarrassed by a decision-maker who asks why an option is not being presented. As control options are considered, the risk reduction value of the control must be computed. Reassessing the risk, based on implementation of the control option reveals the residual risk. Before briefing a decision-maker, be sure also to list the risk reduction value of recommended options.

Risk Reduction Options

Delay. It is frequently possible to delay a planned activity and in doing so, avoid a risk altogether or at least reduce it. An example is to reschedule an operation when more favorable weather conditions exist.

Transfer. The transference of risk may sound at first like a sneaky way to avoid problems, but it is actually a solid option in many cases. If, for example, your company lacks experience, training, or equipment required to perform a task safely, it may be best to hire someone who has the experience to perform the work.

Reduce. The reduction of risk is the primary goal of the control options step. The reduce option offers a hierarchy of options that can be recommended in an effort to lower the probability or the severity of an undesirable outcome. Reduce options include:

- Engineering solutions: The most effective risk reduction option, where possible and cost effective, is to engineer a system so that risks are either eliminated or significantly reduced. This option generally allows avoidance of human error potential and is best employed during the earliest phases of system design. Of course, engineering modifications may be made to systems later, but this can mean much higher implementation costs that may delay or even make the change

impractical. Engineering changes can be planned to reduce both probability and severity of risk.

- Administrative controls: Administrative controls include training and procedural steps taken to reduce the probability an adverse event will occur. Examples include tool control procedures, wing walkers when moving aircraft, and access restrictions.
- Safety devices: The incorporation of safety devices, while not as effective as engineering controls, can also reduce the probability and/or severity of a risk. For example, alarms or other warnings of impending danger may allow an operator to take action in order to avoid or lessen system damage.
- Personal protective equipment: Although generally the least desirable alternative, requiring personnel to wear personal protective equipment (PPE) to reduce the severity of a risk may be the only practical or possible choice. One of the primary flaws with PPE is the reliance on individuals to use it properly. Another often overlooked consideration with regard to PPE is the long-term cost of purchasing and maintaining it.

Reject. Rejecting a risk is often the only viable course of action for a given situation. Occasionally, even with reasonable control measures in place, the overall risk may outweigh the benefit or opportunity potential of the operation, process, or system, and the proper decision is not to move forward. Also, it is common for necessary control measures, from a purely financial standpoint, simply not to be cost effective. That is, the return on investment, given the expense or other constraints required to properly control unacceptable risks, may make plans impractical and the decision to reject the risk altogether a smart one.

Step 4: Make Decisions

After control options have been identified for the risks under consideration, it is time to make the risk decision. There are two key considerations to risk decision-making. The first is that if the residual risks, after implementation of control measures, are greater than the anticipated benefits, a decision not to proceed must be made. Remember, however, as previously mentioned, that if the benefits outweigh the risks, then the risks should be accepted. Secondly, the risk decision should be made at the right level. That is, people must understand what levels of risk they are empowered to accept and when decisions are to be elevated to a higher level. Do not take it for granted that people naturally know when to elevate decisions. Operational personnel tend to be have a ''can do'' attitude and may take risks, in an effort to get the job done, that put a company at significant risk. In some cases it is desirable and practical to establish a formal risk decision-making hierarchy. For example, low risk decisions may be delegated to first-level supervisors while higher-level issues are taken to specific management levels for a decision.

Step 5: Implement Controls

Once a risk decision is made, control options are implemented. It is important that implementation be properly managed and controlled to ensure things are done as planned. Necessary resources must be provided to implement the control options,

and personnel should understand the importance of the controls. Lastly, do not assume anything. Oversight and verification of proper control implementation, particularly for key tasks, is critical.

Step 6: Evaluate Effectiveness

The last step in the risk management process is evaluating the output of the overall effort and assessing the need to reengage the process to deal with either changes or undesirable results. This assessment can be a cursory review to identify any changes that warrant action or a formal performance measurement effort. In either case, the overall purpose is to identify when to reevaluate the operation, process, or system by starting the process over at step 1. When appropriate, risk performance and control measure effectiveness can be fairly easily measured at this stage, particularly since, if a proper risk assessment was completed, all risks and their controls have already been identified. Simply put, one should have a good idea of what to measure based on the key risks that were identified during the assessment. Measuring these areas can often allow proactive preventive action to be taken prior to an accident or other negative event occurring. An example of a proactive or leading metric in a case where a particular training level is desired to maintain an acceptable risk level might involve monitoring personnel training levels to allow managers to identify when training may be impacting risk. Of course, other metrics can provide risk performance output information, such as accident rates, numbers of security incidents, or other indicators of undesirable events. These output indicators, however, are after the fact and should not be used exclusively. While they are important, it is better to measure things that allow action before something bad happens. Use the information from the risk assessment to focus on high-risk issues—those things that can really result in a problem. When performance is not at a proper level, it is likely due to one of two reasons. One reason may be that the risk assessment overlooked one or more risks. A previously undetected risk will probably be obvious once the causes of incidents are reviewed. The second and most likely cause of trouble is change. Something about the operation, process, or system has changed, and this change has introduced a risk that is not being adequately controlled. In either case the answer is to reengage the risk management process at step 1 and work completely through the process again. It is for this reason that the risk management process is referred to as continuous, since we assume an organization will always strive to improve and ensure management of risk.

17.37 SUMMARY

Risk management provides a logical and systematic means of identifying and controlling risk. It supports informed decision-making and, with minimal training, can be employed by virtually anyone to assess any activity or system. Properly implemented, risk management becomes an accepted and integrated part of the way day-to-day business is done. While some aspects are subjective, given the objective nature of the overall risk management process, it can be a powerful and valued tool—and the design engineer is right in the middle of an effective risk management program.

PART 5

AIRCRAFT ACCIDENT AND INCIDENT INVESTIGATION

Robert L. Sumwalt, III

17.38 INTRODUCTION

Aircraft accidents and incidents can be tragic. They are often associated with loss of human life or injuries, costly damage to equipment, and an increased financial burden to the operator. In spite of the negative aspects, however, a properly investigated mishap can provide an opportunity to uncover and correct deficiencies in the system. And though the mishap may be tragic in itself, an even greater tragedy would be to endure an accident or incident and then fail to investigate it properly and implement remedial actions.

The International Civil Aviation Organization (ICAO) is the civil aviation arm of the United Nations. Currently there are 187 contracting states (member countries) belonging to ICAO. Since its formation in 1944, ICAO has outlined international *Standards and Recommended Practices* (SARPs) for these countries to follow. SARPs are contained in the Annexes to the Convention on International Civil Aviation. Annex 13 pertains to aircraft accident and incident investigation.

17.39 WHY INVESTIGATE ACCIDENTS AND INCIDENTS?

When it comes to the question of why to conduct a mishap investigation, Annex 13, paragraph 3.1 is quite clear: ''The sole objective of investigation of an accident or incident shall be the *prevention* of accidents and incidents. It shall not be the purpose of this activity to apportion blame or liability.'' This thought must remain crystal clear throughout an investigation. Although popular with news media, finger-pointing, laying blame, or assigning fault is not the purpose of contemporary mishap investigations. Instead, the purpose is clear-cut: find out what happened so that the problems can be corrected to prevent similar mishaps from occurring again. When the outcome of an investigation is an exercise in finger-pointing, no real prevention measures are likely.

17.40 SYSTEMS APPROACH

''Historically accident investigation reports usually depict clearly what happened and when, but in too many instances they stop short of fully explaining how and why the accidents occurred'' (ICAO 1993). Another shortcoming of investigations has been concluding that the mishap's cause was simply some performance failure or error by a front-line operator such as a pilot, air traffic controller, or mechanic. While the cited individuals may have been the last person(s) to make an error prior

to the accident, it must be clearly understood that people do not make errors in a vacuum; they are merely one element in a very complex system. To understand fully what caused a particular mishap, the investigation must dig beneath the surface—look beyond the obvious—to determine what systemic factors may have caused or led an individual(s) to act inappropriately. An investigation that looks beneath the surface "helps identify the underlying deficiencies that might cause other incidents or another accident to happen" (ICAO 1993). Almost all the "primary" causes of accidents shown in Figure 17.6 have underlying causes associated with them—and many of these are engineering related.

People are greatly influenced by their surroundings. For example, the actions of flight crews are affected by factors present in the cockpit, such as interactions with other crew members, flightdeck automation, and the physical human–machine interface. But also affecting crew performance are factors outside the cockpit, such as the operator's corporate culture, regulatory influence, environmental conditions, and aircraft design. When combined into the aviation context, these individual pieces form the aviation system.

The aviation system is not unlike aircraft systems that engineers, mechanics, or pilots are accustomed to studying. For example, consider an aircraft electrical system. The individual pieces of that system include generators, constant speed drives (CSDs), wires, busses, transformer-rectifiers, circuit breakers, etc. When the individual components are combined, they form an electrical system. If any components do not function properly, the operation of the system is affected.

The same is true with the aviation system. Several components combine to create the aviation system, and any element can affect other elements. For example, if an engineer designs a procedure that is unclear or very difficult for a mechanic to

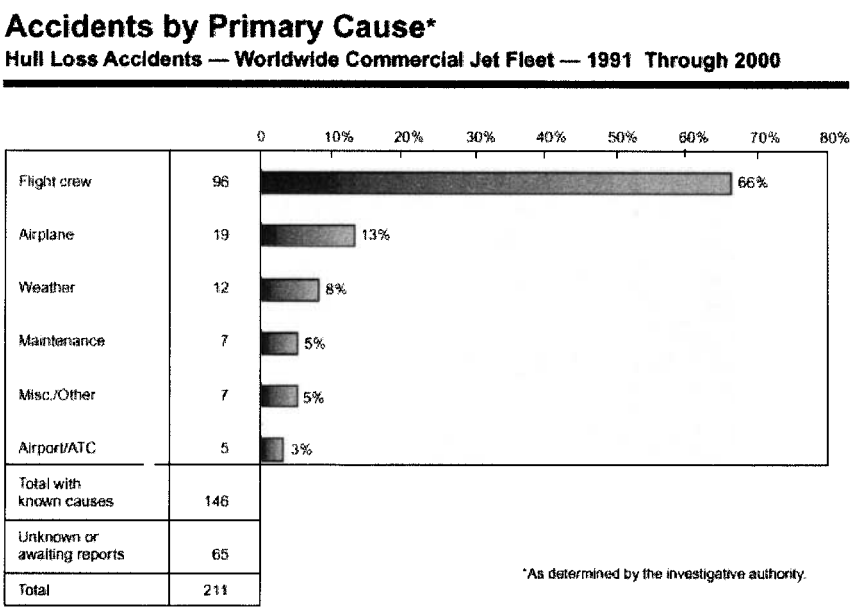


FIGURE 17.6 (From Boeing 2001, p. 21.)

accomplish, then well-intentioned workers on the shop floor may eventually develop methods to work around the engineer's prescribed task. This may lead to errors in the mechanic's performance, which may then lead to a flight crew encountering difficulties in flight. Suppose in this example that the flight crew mishandles the problem, resulting in an accident or incident. A superficial investigation might only conclude that the mishap was caused by inappropriate pilot actions. While this finding would be partially correct, without digging deeper, the root problem (in this case, the poorly implemented engineering procedure) would remain uncorrected. Thus, at some point in the future, others may be snared by this same problem.

A systems approach to accident and incident investigation acknowledges that *all* elements of the system may have a role in the occurrence. By using a systems approach it is possible to identify the underlying causes through a better understanding of how various system components interacted and integrated to result in a mishap. Unless the underlying systemic deficiencies are identified, it is not likely that corresponding corrective actions and recommendations will be aimed properly to prevent future mishaps.

The easiest factors to identify are the proximate errors, or those that are closest to the mishap. However, if an investigation stops with only finding the proximate errors, then the system flaws remain in place. Uncorrected, these flaws may lead to another mishap when encountered by a different person facing similar circumstances.

Case Study: Why a Systems Approach Is Important

In 1979, an Allegheny Airlines Nord 262 crashed immediately after takeoff on a snowy afternoon in Clarksburg, West Virginia. Thirteen years later, USAir (as Allegheny was now called) lost a Fokker F28 while the crew attempted takeoff during a snowstorm at New York's LaGuardia Airport. Although these accidents occurred in different cities, both involved the same airline and both involved takeoff attempts with contaminated wings.

What distinguishes these accidents is the way that U.S. National Transportation Safety Board (NTSB) approached each investigation. Following the 1979 accident, the NTSB stated "The probable cause of the accident was the captain's decision to take off with snow on the aircraft's wing and empennage surfaces" (NTSB 1979). While the captain's faulty decision certainly enabled the accident, this probable cause statement did little to explain the factors that could have set the stage for poor decision-making. Essentially, the investigation concluded by identifying the last person who made a mistake. Because the investigation's concentration was on *who* caused the accident, not *why* it happened, the flaws in the system remained uncorrected. This is supported by the fact that in the 13 years between the Clarksburg and New York accidents, 10 air carrier accidents occurred worldwide due to takeoff attempts with contaminated aircraft wings.

In contrast, the NTSB used a systems approach in its investigation of the 1992 New York icing accident. By examining the entire system that surrounded that crew on that snowy evening, the Safety Board uncovered systemic deficiencies that allowed the crew's errors to result in a fatal accident. The NTSB's analysis of this accident cited, in part, the "failure of the airline industry and the Federal Aviation Administration to provide flight crews with procedures, requirements, and criteria compatible with departure delays in known icing conditions" (NTSB 1993).

Since systemic deficiencies were highlighted in this accident's aftermath, proper corrective solutions could be applied, all aimed at preventing future accidents. In response to this progressive investigative approach, the *system* was changed to improve safety during ground icing conditions: U.S. Federal Aviation Regulations (FARs) were overhauled; extensive training for flight and ground crews was introduced; widespread procedural changes were enacted, such as use of hold-over charts to help crews determine how long de/antiicing fluid would remain effective under existing atmospheric conditions; improved de/antiicing fluids were developed, which are now widely used; and ATC procedures for minimizing ground delays after deicing were introduced. The list goes on. The point is quite simple: when investigators focus on answering *why* an accident occurred and remove focus from *who* caused it, true safety improvements can result.

17.41 THE IMPORTANCE OF INVESTIGATING INCIDENTS

There are basically two types of incidents. One is where equipment is damaged or injuries occur but the mishap is not severe enough to be classified as an accident. The second type is where neither injuries nor actual damage to property occurs. Often in these cases if only one thing had been different a disastrous full-blown accident would have occurred. Consider each of these situations:

- A single-engine aircraft is cleared to taxi into position on a runway at night. The air traffic controller then forgets that he has cleared that aircraft onto the runway and clears an airliner to land on that same runway. As luck would have it, the runway has a displaced threshold for landing and the airliner flies over the single-engine airplane and lands without ever seeing the traffic on the runway. No damage to equipment or property, and no injuries.
- A military aircraft mechanic uses the wrong fasteners to secure an engine cowling, causing the cowling to separate in-flight. As it rips away, the cowling causes damage to the airplane's left wing. Minor equipment damage and no injuries.
- An overnight express company's Cessna Caravan slides off an icy runway during an early morning arrival at a small airport. Minor equipment and property damage, no injuries.
- Three passengers receive minor injuries while evacuating a 737 after smoke billows from air conditioning vents. No equipment or property damage, and minor injuries.

Because little damage and few injuries occurred in these examples, there is sometimes a tendency to forgo a comprehensive investigation. But by investigating the near miss or the incident, it is possible to get a free lesson. That is, an organization can document and correct system failures without having to suffer the anguish or consequences of an accident.

"Many incidents occur every day which may or may not require reporting by the investigation authority; some come very close to being accidents. Because there is no injury or little damage, these incidents might not be investigated. (ICAO 1993). The need for an investigation by either the investigation authority or operator

must be emphasized, because an incident investigation can often produce better accident prevention results than can an accident investigation.''

Incidents Are Precursors

''It is felt that incidents are precursors to accidents, and most accidents have many precursors that may have led one to predict the accident. The challenge is to identify these precursors, minimize their individual risk, implement strategies that protect against these broad classes of risk, and assure that specific chains of events containing these precursors cannot link up in unexpected ways that lead to an accident'' (FAA 1995).

Analysis of industrial workplace mishaps shows that for every fatal accident there are as many as 300 incidents and near misses. Although these figures are not specific to aviation, the concept is one that aviation safety experts feel can apply to their field. The issue is not whether or not the exact numbers apply to aviation; the point is that incident data can point us to areas that can lead to accidents. Incidents can provide tremendous opportunities to identify systemic deficiencies and because they can be viewed as precursors, or predictors, of accidents, they must be properly investigated.

Reason's Model. Finding systemic deficiencies can be equated to peeling an onion: layer after layer is peeled to get closer and closer to the core. Dr. James Reason has developed a useful methodology (see Figure 17.7) that, in effect, helps air safety investigators peel back layers of the system to determine the root cause of a mishap (Reason 1990).

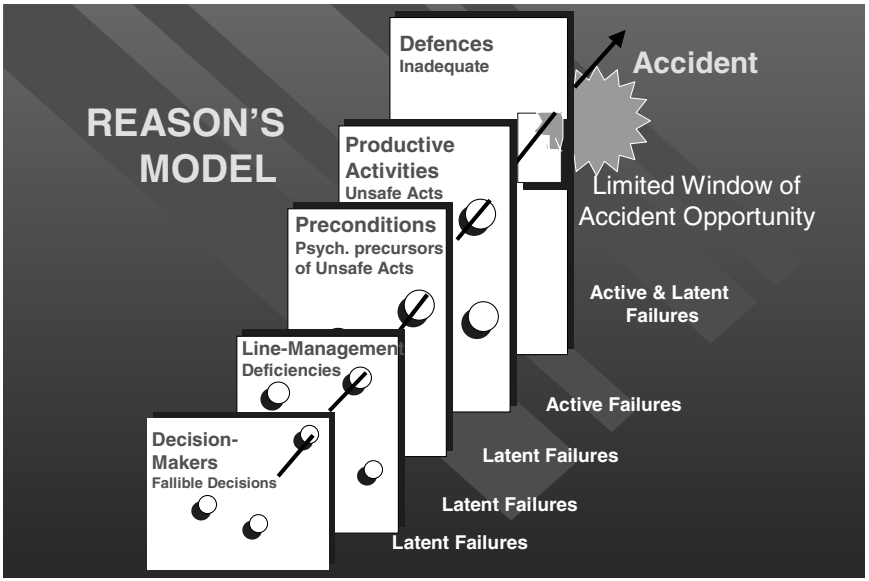


FIGURE 17.7 Reason's model (from Reason 1990).

Dr. Reason believes that aviation mishaps and other mishaps involving high-technology industries are rarely caused exclusively by mistakes or failures on the part of front-line operators. Instead, he says, accidents often result from the interaction of a series of flaws known as latent conditions that have been embedded in the system.

Front-line operators are those who are at the ''sharp end of the pyramid,'' according to Dr. Reason, meaning that these people are performing the actual hands-on tasks, as opposed to managers, who are further behind the front-line. Examples of front line operators include pilots, air traffic controllers, and mechanics.

Reason classifies system failures according to how quickly the failure is manifested or made known. An active failure is typically committed by a front-line operator. Its consequences are usually made known soon after the error was committed, such as flight crew forgetting to set wing flaps prior to takeoff or a mechanic failing to replace engine O-rings.

Latent conditions are systemic flaws whose consequences may not surface until long after being introduced into the system. These dormant conditions usually result from decisions, actions, or inactions of those who are far removed from the front line, such as managers or regulatory authorities. Inadequate regulatory oversight of a rapidly expanding new-entrant air carrier is an example of a latent condition. In itself, that condition may not result in an accident; however, when combined with other active failures or latent conditions, the ''window of accident opportunity'' may be opened.

Each layer of the system can introduce flaws. Reason classifies ''decision-makers'' as the layer that is furthest removed from the front line operations. He says this layer contains the ''architects and the high-level managers of the system,'' whose function is to set the system's strategies and philosophies. They are senior-level management or even regulatory officials. ''A large part of their function concerns the allocation of finite resources, such as money, equipment, and time,'' says Dr. Reason. ''Their aim is to deploy these resources to maximize both productivity and cost.'' When these objectives conflict, flawed decisions can result which will be reflected throughout the system. Consider the aviation decision-maker who decides to improve on-time performance at all costs, or to do everything possible to maximize the corporation's financial profits. Taken independently, these decisions (latent conditions) may sound reasonable. However, if proper checks and balances are not in place, these decisions may have consequences that affect system safety at some future point.

Once decision-makers set the strategies for the system, ''line managers'' are responsible for turning the ideas into practices. This is done by implementing standard operating procedures (SOPs), training programs, company directives, and the like. When these practices are implemented in less-than-optimum fashion, latent conditions are created.

''Preconditions'' are the production qualities and problems that reside within the system. Preconditions that affect the system in a positive manner include attributes like good equipment, good employee moral, good training programs, and positively motivated workers. Conversely, unhealthy characteristics such as an unproductive working environment, poorly motivated workforce, and poorly established procedures are examples of latent conditions that can later interact with other failures to threaten system safety.

The act of the front-line operators performing their job functions is what Reason terms ''productive activities.'' An active failure occurs when a front-line operator commits an unsafe act, such as deviation from SOPs or omission of a checklist item.

"Defences" are the system's safety net, and the final system layer described by Reason. Examples of defences include pilot discipline, proper training, checklist usage, warning devices, quality assurance (QA) checks, and other error-trapping techniques. To illustrate, many maintenance functions require a QA check, where someone independently inspects the work of others. If a mechanic failed to install an engine pylon bolt, for instance, the systems defences (QA check) are designed to catch the error before it becomes catastrophic.

Reason theorizes that an "accident trajectory" forms when unsafe acts or active failures interact with systemic latent conditions. When defences are breached, the trajectory opens the window of accident opportunity, enabling an accident.

When investigating mishaps, the investigator should begin with the error(s) of the front-line operator, then work back along the accident trajectory through each layer in the system. As each layer is examined, the focus should be on identifying the latent conditions that could have influenced the actions of the front-line operator to result in an accident or incident.

Remember the Goal. Dr. Reason insists that the purpose of his model is not merely to remove blame from front-line operators and apportion it throughout the entire system. Instead, his intention is to identify systemic flaws so that proper remedial actions can be taken to correct them.

As stated previously, when solving accidents or incidents, the easiest thing for investigators is to find the active failures of the front-line operators (proximate error). The weakness with this approach, however, is that the latent conditions remain undetected. So if the objective of an investigation is, as ICAO says, "*prevention* of accidents and incidents," then the best way to accomplish that objective is by identifying and correcting systemic deficiencies.

17.42 ACCIDENT INVESTIGATION

Jurisdiction of Accident Investigation

Recall that ICAO annexes contain international Standards and Recommended Practices (SARPs). Because ICAO Standards are considered necessary for the safety or regularity of international civil aviation, ICAO member countries are expected to conform to these standards. However, in the event individual countries cannot conform, they are obligated to submit to ICAO a "notification of difference" to explain their areas of deviation. Recommended Practices, on the other hand, are considered desirable to follow, but not mandatory.

Annex 13 (ICAO 2001) specifies that the "State of Occurrence shall initiate an investigation into the circumstances of the accident." Put simply, the country where an accident occurred is responsible for conducting the investigation. However, Annex 13 allows the State of Occurrence to delegate portions of, or the entire investigation to, authorities from the country where the aircraft was registered (State of Registry), the country where the operator was headquartered (State of the Operator), or the country where the aircraft was designed or manufactured (State of Design or the State of Manufacturer).

Example 1: On November 12, 2001, an American Airlines Airbus A300-600 crashed about two minutes after departing New York's John F. Kennedy International Airport. This case involved an accident in the United States with a French-

designed and manufactured aircraft. Who would be responsible for investigating this accident? The State of Occurrence was the United States, and the U.S. investigative agency (NTSB) conducted the investigation. In accordance with Annex 13, France, being the State of Design and Manufacturer, had a right to participate in the investigation through an Accredited Representative designated by the French Bureau Enquetes Accidents (BEA).

Example 2: A Qantas Boeing 747-400 overran a runway upon landing at Bangkok International Airport, Thailand, on September 23, 1999. The State of Occurrence was Thailand, the State of Registry and Operator was Australia, and the State of Design and Manufacturer was the United States. Who was responsible for investigating this accident? As the State of Occurrence, Thailand was responsible, but in accordance with the provisions of Annex 13, it delegated the investigation to Australia (State of Registry and Operator). Could the investigation have been delegated to the U.S. NTSB? Yes, because both the State of Design and State of Manufacturer was the United States (Boeing). To further the example, could the investigation have been delegated to United Kingdom's Aircraft Accident Investigation Branch (AAIB)? No, because the United Kingdom has no direct responsibility. However, if for some reason the UK had been requested to provide search and rescue or some form of investigation support, it would have been entitled to appoint an Accredited Representative who could have participated in the investigation.

When an accident occurs outside the territory of any ICAO member country, Annex 13 says that the State of Registry shall institute and conduct the investigation. As above, the Annex also says that the State of Registry may ''delegate the whole or any part of the investigation to another State by mutual arrangement and consent.''

Example 3: On October 31, 1999, Egyptair Flight 990, a scheduled international flight from New York to Cairo, a Boeing 767-366-ER, crashed in international waters of the Atlantic Ocean about 60 miles south of Nantucket Island, Massachusetts. Under Annex 13, whose responsibility was it to investigate the accident? The aircraft was Egyptian registered, so Egypt had the responsibility to conduct the investigation. However, because the crash occurred just off the coast of the United States, Egypt delegated the investigation to the United States due to that nation's close proximity to the accident site and experience with underwater salvage.

Responsibilities and Structure of an Investigation

Annex 13 emphasizes the importance of allowing an investigation to be conducted with complete independence and not subjected to political pressures. In fact, several countries have structured their accident investigation agencies to be completely autonomous of other government bodies to ensure independence. When the U.S. NTSB was formed in 1966, for example, it was a government agency under the auspices of the U.S. Department of Transportation, the agency that also houses the Federal Aviation Administration (FAA). Eight years later the Independent Safety Board Act of 1974 was passed by the U.S. Congress, reestablishing the NTSB as an independent governmental agency. Moving the NTSB from DOT provided complete freedom for the NTSB to investigate accidents without political constraints. (For an example of why an independent agency is important, see the above case study, Why a Systems Approach Is Important, where the NTSB determined a probable cause as ''failure of the airline industry and the Federal Aviation Administra-

tion to provide flight crews with procedures, requirements, and criteria compatible with departure delays in known icing conditions.'')

17.43 THE ROLE OF AN AERONAUTICAL AND ASTRONAUTICAL ENGINEER IN AN INVESTIGATION

Investigations often use expert resources to facilitate prompt and accurate findings. As an expert in areas such as structures, aerodynamics, or other engineering disciplines, an aeronautical or astronautical engineer may become involved with assisting an investigation.

Some agencies, such as the U.S. NTSB, utilize a party system, whereby party status is extended to organizations familiar with the process, which bring needed technical expertise and value to the investigation. For example, the FAA is always granted party status because the air traffic control system, certification of aircraft and airmen, and enforcement of the Federal Aviation Regulations falls under the FAA's purview. The aircraft and engine manufacturers are typically parties to the investigation because they have the most knowledge about and experience with their products. If the accident involves an airline, that airline will likely be a party, along with unions representing pilots, flight attendants, and mechanics. This is because the NTSB has learned over the years that these different organizations each bring unique knowledge and perspective to the investigative process.

While serving as NTSB chairman, Jim Hall told a group of aviation attorneys, ''The underlying premise of the party system is a strong one—that everyone has an overriding interest in safety and that everyone wants to find out what happened so that steps can be taken to ensure such an accident is not repeated.'' The system works. As Chairman Hall explained, ''The party participants—be they air carriers, manufacturers, pilot organizations, emergency response providers, suppliers or maintenance providers—all have at one time or another provided the NTSB with the technical depth of knowledge we have needed to determine the probable cause of a transportation disaster.''

While other countries may not use a party system, many use a system where outside expertise is used to some degree. For example, the Transportation Safety Board of Canada uses qualified observers who interact with and participate in many phases of the investigation. For this reason, it is felt that the reader should have some general background in the conduct of an investigation.

On the Site

Accident investigation is probably an art as much as a science. Dedicated investigators spend a lifetime accruing techniques and skills of the trade, so it is not possible to explain all techniques in a section of any book. Indeed, volumes have been written to explain various methods of accident investigation. The purpose of this section is to provide a general overview of aircraft accident investigation techniques.

Upon arriving on an accident scene, one of the first things an investigator should do is perform an overall survey (sometimes called a walk-through) of the wreckage site to get a general feel for the situation. Documenting the wreckage by photog-

raphy and precise diagramming should be the next order of priority. General layout of the destroyed aircraft provides important clues as to how and why the pieces ended up where they did. To the experienced investigator, ground scars can tell a fascinating story of the final flight path of how the aircraft initially struck terrain. The distance between propeller slash marks can be measured, and if the prop's RPM is known, the aircraft's ground speed at impact can be determined. Conversely, if ground speed is known, the prop RPM at impact can be calculated.

Trees can provide additional clues. In a heavily wooded area, their damage patterns can indicate the flight path angle of descent and aircraft bank angle. This information, coupled with the general wreckage distribution pattern, begins to tell the story of the last moments of flight. Generally, wreckage that resulted from a steep angle of impact remains relatively close together. Pieces of wing or tail sections found far from the main wreckage may provide clues to an in-flight breakup.

A day or two after the accident, dead leaves on an otherwise green tree may be an indication of fuel spilling onto the leaves during impact. If there is a question of whether or not fuel was onboard at the time of impact, this may provide a good clue. Questions about fuel type can be answered by having the leaves lab tested to check their flash point and vapor pressure. Each fuel type has its unique signature for these items.

All of these are only general rules of thumb, however, and to ensure validity they must be carefully weighed with other evidence. ''Until you get the big picture, a lot of things can be masked. Get the complete picture and don't jump to conclusions,'' warns an NTSB instructor.

In the case of a suspected in-flight breakup, expect to find evidence of wing or tail overstressing. With positive-g overstressing, the wings are forced upward relative to the fuselage. The upper spar cap may show signs of compression stress due to this leveraged upward bending. Likewise, the skin of the wing's upper surface may be wrinkled and the lower spar cap may show signs of tensile stress.

Ice that accumulates on wing and tail leading edges during flight can have a devastating effect on aircraft performance. It is usually difficult to find physical evidence of in-flight icing in wreckage, because it likely will have melted by the time the investigator arrives on scene. But if these surfaces are not damaged by fire, careful inspection can still provide clues. If the leading edge of an airfoil is clean of dirt, grime, and debris and the rest of the wing has telltale signs of impact dirt and soot, this could indicate that ice was covering the wing's leading edge at impact, then melted off.

Propeller damage may support evidence of an engine developing power. Significant degrees of propeller twisting, prop leading edge damage, and scratch marks across the propeller's surface can be indicators of engine power. Again, there are few hard and fast rules, so the investigator must weigh all of the evidence together.

Aircraft collisions, whether in mid-air or on the ground, will leave telltale paint transfer and scratch marks on one or both aircraft. Measuring the angles of these marks relative to the aircraft's longitudinal axis enables basic mathematical functions to be used to determine collision angle. This information can be critical in determining whether the pilots could have seen each other or whether obstructions inside or outside the aircraft limited their visibility.

When wreckage is burned, there is always the issue of whether the fire was the cause or the result of the accident. With an in-flight fire, the aircraft may have dropped debris in the final moments of flight. Therefore, when in-flight fire is suspected, it is important to search the areas along the flight path for supporting evidence. Once located, aircraft parts should be pieced together to compare burn

marks. Finding one part scorched but an adjoining piece untouched by fire suggests the obvious: the parts separated before fire eruption.

Even with charred or fragmented cockpit instruments, it is sometimes possible to determine the indicator's position at impact. The needle of a conventional (analog) instrument may momentarily strike the instrument face at impact, then spring back rapidly and strike the backside of the glass cover. This spring action of the needle may leave a slight impression in the instrument's face, or it may leave a subtle deposit of luminous paint on the instrument's glass. These trademarks may be detected with an ultraviolet light.

Digital aircraft utilizing "glass cockpit" technology present additional challenges and opportunities for the investigator. No longer can wrecked instruments be coaxed into sharing their secrets—only computer chips remain. But those circuits sometimes hold a wealth of information in their nonvolatile memories.

By looking at a light bulb's filament, it may be possible to determine if the bulb was illuminated at impact. The filament of an illuminated bulb is warm and relatively flexible. At impact this filament tends to stretch and become elongated. Conversely, an unlit bulb's filament will be relatively cold and brittle and may shatter upon impact.

Behind the Scenes of an Investigation

When the on-site investigation is completed, there is still potential evidence to be gathered. ATC radio transmissions and radar presentations are recorded, and those tapes are available to the investigation if requested. Weather reports need to be collected, and maintenance records should be carefully reviewed.

Aircraft components may need to be sent off-site for examination. A portion of the wing spar could be sent to a metallurgy lab to determine failure mode. The power plant may need tear-down and inspection by the manufacturer. For aircraft so equipped, the cockpit voice recorder (CVR) and flight data recorder (FDR) must be sent to an investigative agency lab for readout. Not all countries are equipped to do these readouts, so it is common to send them to the U.S. NTSB, Australian Transportation Safety Board, the TSB of Canada, the U.K. AAIB, or the French BEA.

Witness interviews should be conducted as soon as possible. If too much time elapses, the witnesses' thoughts may be affected by media reports and other sources. Additionally, with the passage of time the witness can be affected by "closure theory," where the witness's mind subconsciously tries to fill in the blanks to help reconcile the traumatic event that he or she witnessed. For instance, witnesses of in-flight breakups often report hearing an explosion, followed by the wings falling off. In reality, this loud popping sound is typically caused by the sudden release of the wing's energy when the wing spar snaps. However, the witness's mind subconsciously links the loud popping noise and the wing separation with an in-flight explosion. The person's mind may then quietly add other elements associated with explosions, such as fire and smoke.

Determination of Findings and Recommendations

After all of the effort has been expended to investigate a mishap properly, the most important facet is to determine findings accurately and issue recommendations to

prevent a similar mishap. Remember, this is the sole reason that the investigation was conducted. Unless this effort is properly undertaken the investigative effort is futile. Annex 13 should be consulted, as it contains specific guidelines for specific formats to be used.

17.44 SUMMARY

The goal of incident and accident investigation is to identify each of the factors that enabled the mishap and then develop ways to correct the problems. Investigators must realize there are often several factors that combine and lead to a mishap. Do not fall into the trap of thinking that once you have found the active error of a front-line operator, the accident is solved. According to Captain Dan Maurino of ICAO, "The discovery of human error should be considered as the starting point of the investigation, and not the ending point that has punctuated so many previous investigations." That is sound advice for any investigation.

REFERENCES

Boeing Commercial Airplane Company. 2001. *Statistical Summary of Commercial Jet Airplane Accidents, Worldwide Operation, 1959–2000,* Boeing Commercial Airplane Company, June.

International Civil Aviation Organization (ICAO). 1993. *Human Factors Digest Number 7: Investigation of Human Factors in Accidents and Incidents,* ICAO, Montreal.

International Civil Aviation Organization (ICAO). 2001. *International Standards and Recommended Practices: Aircraft Accident and Incident Investigation. Annex 13 to the Convention on International Civil Aviation,* ICAO, Montreal.

Reason. J. 1990. *Human Error,* Cambridge University Press, Cambridge.

U.S. Federal Aviation Administration (FAA). 1995. *Federal Aviation Administration Human Factors Team Report on: The Interfaces Between Flightcrews and Modern Flight Deck Systems,* June 18.

U.S. National Transportation Safety Board (NTSB). 1979. *Allegheny Airlines, Inc. Nord 262, Mohawk/Frakes 298, N29824,* Clarksburgh, WV, February.

U.S. National Transportation Safety Board (NTSB), 1993. *Takeoff Stall in Icing Conditions. USAir Flight 405, Fokker F28, N485US, LaGuardia, NY. March 22, 1979.* Aircraft Accident Report, NTSB/AAR-93-02, February.

FURTHER READINGS

International Civil Aviation Organization (ICAO). *Accident Investigation Manual,* ICAO, Montreal (1970).

International Civil Aviation Organization (ICAO). *Accident Prevention Manual,* ICAO, Montreal (1984).

Maurino, D., Reason, J., Johnston, N., and Lee, R., *Beyond Aviation Human Factors,* Ashgate, Brookfield, NH (1995).

Sumwalt, R. L., "Mishaps and Human Factors," *ISASI Forum,* April–June 1998.

Sumwalt, R. L., ''The Quest for Aviation Safety's Holy Grail: Finding Underlying Causes of Accidents and Incidents, or, If You Really Want to Improve Aviation Safety, You Must First Identify Systemic Problems,'' in *Proceedings of Workshop on Accident and Incident Human Factors,* FAA, June 1995.

U.S. Department of Transportation (DOT), *Aircraft Accident Investigation Procedures and Techniques,* DOT (1978).

Walters, J. M., and Sumwalt, R. L., *Aircraft Accident Analysis: Final Reports,* McGraw-Hill, New York (2000).

Wood, R. and Sweginnis, R., *Aircraft Accident Investigation,* Endeavor Books, Caspter, WY (1996).

SECTION 18
AIRCRAFT MAINTENANCE

Section Editor: James T. Garris

NOTE

This section refers to commercial transport aircraft (as governed by FAR Part 25) and not to general aviation aircraft (as governed by FAR Part 23).

The articles in this section describe the regulatory, economic, and operating environment associated with the maintenance of commercial transport aircraft.

PART 1

THE ECONOMICS OF MAINTENANCE

William Tsai

18.1 AIRPLANE MAINTENANCE

An airplane is delivered from the manufacturer to the airline in an airworthy condition with all airplane systems in a fully functional state. Because an airplane is continuously utilized for revenue flights, it is physically subjected to structural fatigue induced by vibration, aerodynamic loads, flight maneuver loads, engine thrust loads, takeoff loads, landing loads, ground loads (taxi and towing), pressurization cycles, and thermal cycles (Figure 18.1). The operating environment can cause physical damage to the airplane, such as corrosion resulting from water, saltwater, or chemicals. Natural phenomena such as hailstorms, lightning strikes, sandstorms, and volcanic ash can cause physical damage to the airplane. Man-made objects such as ground equipment can also cause damage to the airplane. Objects such as rocks or bolts can cause foreign object damage (FOD). Even birds in flight can extensively damage an airplane. However, the most prevalent causes for the physical deterioration of the airplane structure are fatigue and corrosion. Through normal continuous usage, airplane systems will experience degradation and eventual reduced performance. Airplane maintenance is essential to ensure full functionality of airplane systems and the airworthiness of the airplane.

18.2 DESIGN SERVICE OBJECTIVE (*DSO*)

At a certain point in the ownership of an airplane it will be more economical to purchase a new airplane than to expend resources for escalating annual maintenance cost. The design service objective for an airplane is the number of flight hours, flight cycles, or years established as a design goal by the airplane manufacturer based on the expected utilization at the time of original airplane design and certification. The design service objective is the expected product life duration, which provides a common design basis for the durability of the airplane, systems and components. Airplane design is a trade-off between the weight and the performance of the airplane to find the optimum for airframe durability and operating economics. The application of new lightweight cost-effective materials, in-service experience, improved stress analysis, and improved testing techniques in new design improvements have helped to reduce the weight of the airplane. To achieve an even longer design life, a more durable airplane structure could be designed, but that new design would be burdened with increased weight, thereby penalizing the payload and range of the airplane. The design service objective for various Boeing airplane products is shown in Table 18.1. The design service objective is defined as the number of flight cycles, flight hours, or years that an airplane can be operated in which the primary aircraft structure is crack-free despite structural fatigue or service loads. Operating an airplane beyond the design service objective imposes escalating maintenance cost, which rapidly becomes cost prohibitive. As the airplane continues to

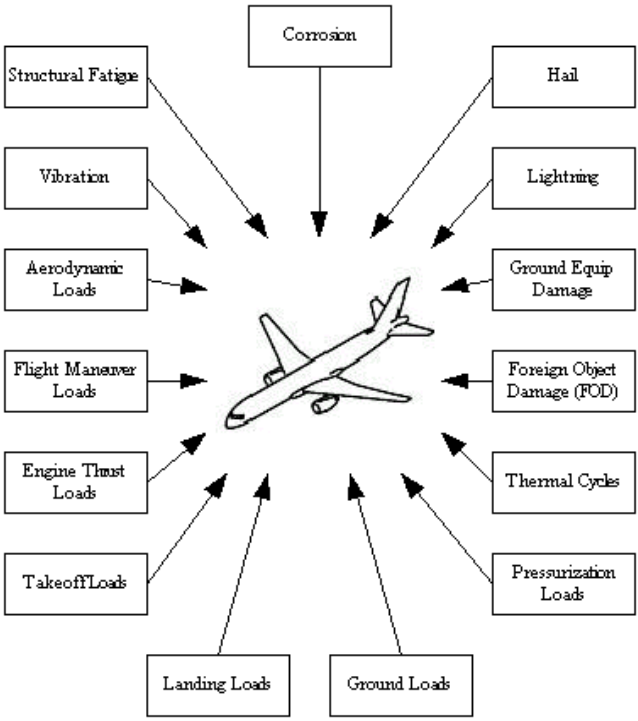


FIGURE 18.1 Physical effects on airplane structure.

TABLE 18.1 Design Service Objective of Airplanes

Airplane model	Design service objective		
	Number of flight cycles	Number of flight hours	Years
B707	20,000	60,000	20
B717	60,000	60,000	20
B727	60,000	50,000	20
B737	75,000	51,000	20
B747	20,000	60,000	20
B757	50,000	50,000	20
B767	50,000	50,000	20
B777	40,000	60,000	20
DC-8	25,000	50,000	20
DC-9	40,000	30,000	20
DC-10-30/40	30,000	60,000	20
MD-11	20,000	60,000	20
MD-80	50,000	50,000	20
MD-90	60,000	90,000	20

Source: Banis et al. 1999.

experience aging effects, the inspection and repair costs associated with structural fatigue and corrosion will continue to escalate with each successive scheduled maintenance check. The wear and tear of flying an airplane 10 to 14 hours every day for many years will accumulate on the airframe. Besides the airframe, airplane systems will age and become more susceptible to failure as the mechanical components, electrical wiring, and electrical connectors are subjected to years of flight environment exposure. Airplane components will become difficult to replace as they become obsolete or as suppliers are no longer available. Structural fatigue, corrosion, airplane utilization, operating conditions, regulatory environment, and business strategy are important determinants for an operator's decision regarding the economics of continued operation of an airplane approaching or exceeding the design service objective.

18.3 AIRPLANE AGING AND MAINTENANCE COST

As an airplane ages, generally the maintenance requirements will increase with time, thereby increasing the annual maintenance cost. There are three distinct phases of maintenance cost induced by airplane aging. The first phase is termed ''new'' and generally lasts about *5* years. In this phase, the maintenance cost is rapidly escalating since at first the new airplane has minimal maintenance requirements. As repairs are needed, the maintenance cost will appear to undergo rapid escalation. The second phase is termed ''mature'' and generally lasts about 10 years. In this phase, the maintenance cost of the airplane has stabilized. The third phase is termed ''old'' and generally lasts about 5 years. In this phase, the maintenance cost is rapidly escalating again as the airplane is reaching the design service objective and the long-term effects of continuous utilization are being realized. Different airlines utilize different business strategies to account for airplane aging. An airline may choose to sell an airplane well before the design service objective while the airplane still retains value, before the eventual escalating maintenance cost. Other airlines will retain the airplane as long as possible by implementing repairs and aging airplane programs in order to preclude the enormous investment cost associated with a new airplane purchase.

18.4 AIRPLANE FUNCTIONALITY

As airplane designs have evolved, more systems have been added to provide more functionality. As airlines and regulatory agencies increased their requirements, the airplane manufacturers added more and more components and systems to the latest airplane designs. The older airplanes had a significantly lower level of avionics units than current airplanes. An early airplane design such as the Boeing 707 contained about 68 avionics units, while a current design such as the Boeing 777 airplane has about 101 avionics units. The trend is toward increased quantity of avionics units to meet the requirements for increased functionality (Figure 18.2). The growth in functionality means more systems and components installed on the airplane, thereby adding to the maintenance cost and inventory requirements. The increased need for airplane functionality has resulted in greater complexity and the potential for more maintenance. Continuous improvement in component reliability

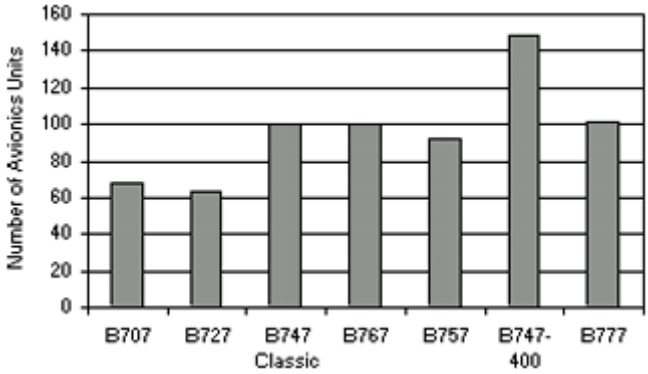


FIGURE 18.2 Avionics units in electrical equipment bays in airplanes.

has helped to counteract the associated cost of increased component quantity driven by increased functionality requirements.

18.5 TOTAL OPERATING COST

The total operating cost for an airplane consists of the cost categories, such as flying operations, aircraft and traffic servicing, promotion and sales, maintenance, transport related, passenger service, depreciation and amortization, and general and administrative. A percentage breakdown of the total operating costs for the Majors during a typical year is shown in Figure 18.3. The Majors are defined as certificated

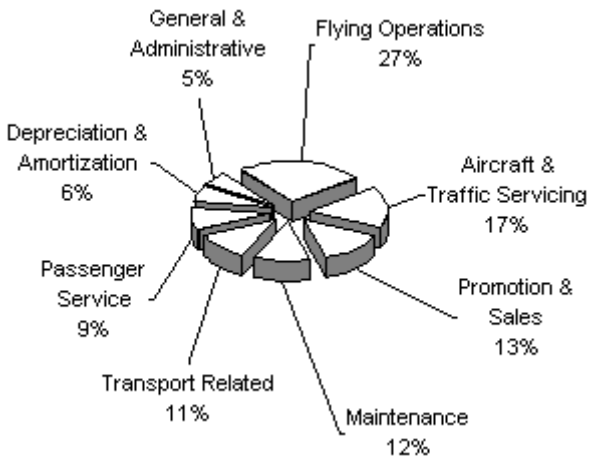


FIGURE 18.3 Percentage breakdown of total operating cost (from U.S. DOT 1999/1998).

air carriers in the United States with annual operating revenues of more than U.S. $1,000,000,000. The flying operations cost is the flight crew cost, fuel, oil, insurance, tax, rental and other expenses related to in-flight operation of the airplane. The aircraft and traffic servicing cost is the handling and servicing of the airplane while it is on the ground, including the ground personnel cost and baggage handling personnel cost. The cost of controlling in-flight movement of the airplane, scheduling and preparation of crews for flight assignments, and any other expense incurred on the ground is also included in the aircraft and traffic servicing cost. The promotion and sales cost is the selling, advertising, marketing, public relations, tariff development, flight schedule development, and seat reservation system involved with the promotion of air carrier to the flying public. The maintenance cost is the cost incurred in terms of personnel, parts, materials, and equipment in order to maintain and support the airplane. The transport-related cost is any other expenses associated with airplane operations not accounted for in the existing categories. The passenger service Cost is the cabin crew cost, the passenger food cost, and other costs associated with ensuring the safety, comfort, and convenience of the passenger during in-flight operations or during a flight interruption. The depreciation and amortization cost is the depreciation and amortization of owned or leased equipment and property such as airplanes, flight equipment, ground equipment and ground property. The general and administrative cost is the cost of corporate-level activities such as financial accounting, purchasing, legal representation, corporate administration, and other general corporate-level activity not attributable to any other function.

18.6 DIRECT OPERATING COST

The direct operating cost is a subset of the total operating cost. It consists of the flying operations, aircraft and traffic servicing, maintenance, and passenger service costs. The direct operating cost is directly attributed to an airplane flight. For example, if the flight schedule is changed to eliminate a specific flight, the direct operating cost will also be eliminated. Excluded from the direct operating cost are fixed costs that are not directly attributed to a specific flight. These excluded costs are infrastructure and corporate costs such as general and administrative, depreciation and amortization, promotion and sales, transport related, and insurance. The direct operating cost is the cost of the flight crew, cabin crew, fuel, maintenance, catering, baggage handling, servicing, air traffic control, and landing fees directly required to support an airplane flight. The direct operating costs of various airplanes are shown in Table 18.2.

18.7 MAINTENANCE COST

The maintenance cost is a subset of the direct operating cost. A well-established industry metric for airplane maintenance cost is the cost per flight hour with the cost in the units of U.S. dollars (USD). The maintenance cost per flight hour is the monetary expenditure for airplane maintenance required for keeping the airplane in an airworthy, flying condition on a flight-hour basis. The maintenance cost can range from $515 USD per flight hour for a two-engine narrow body airplane to

TABLE 18.2 Direct Operating Cost of Airplanes

Aircraft model	Number of seats	Direct operating cost (USD per flight hour)
B747-100	462	$7,224
B747-400	375	$6,455
B747-200/300	369	$7,207
L-1011-100/200	320	$4,891
DC-10-10	297	$5,703
DC-10-40	285	$4,547
B777	274	$3,804
MD-11	259	$6,539
DC-10-30	250	$5,972
L-1011-500	244	$3,855
A300-600	228	$4,783
B767-300ER	211	$3,383
B757-200	186	$2,623
B767-200ER	180	$3,168
MD-90	150	$3,976
B727-200	150	$2,567
B737-800	149	$1,665
A320-100/200	148	$2,227
B737-400	142	$2,095
MD-80	139	$2,139
B737-300	133	$1,878
DC-9-50	125	$2,137
A319	124	$2,254
B717-200	119	$2,571
B737-100/200	113	$2,121
DC-9-40	111	$1,598
B737-500	110	$1,881
DC-9-30	101	$1,897
F-100	97	$2,081
DC-9-10	69	$1,870

Source: ATA 200.

$1,699 USD per flight hour for a four-engine wide-body airplane as shown in Table 18.3. The maintenance cost is the total labor and material required to perform airframe maintenance, engine maintenance, and component maintenance (Figure 18.4). Airframe maintenance can be further separated into line and base maintenance. Similarly, engine maintenance can be separated into line maintenance and engine shop. Line maintenance is the maintenance activity that occurs when the airplane is on the ramp or at the gate. Base maintenance is the maintenance activity that occurs when the airplane is in the hangar or parked on the tarmac in a non-revenue capacity. Component maintenance is the maintenance activity that occurs in the component shop. Labor consists of the direct labor and the maintenance burden associated with the direct labor. The maintenance burden, also known as overhead, consists of benefits, such as vacation, sick leave, and medical, dental, and vision care. Overhead also consists of unallocated labor and expenses such as engineering staff, management, computing support, technical data staff, security,

TABLE 18.3 Maintenance Cost of Airplanes

Airplane	Maintenance cost (USD per flight hour)
Two-engine narrow-body jet	$515
Two-engine wide-body jet	$780
Three-engine narrow-body jet	$712
Three-engine wide-body jet	$1,459
Four-engine narrow-body jet	$990
Four-engine wide-body jet	$1,699

Source: U.S. DOT, FAA.

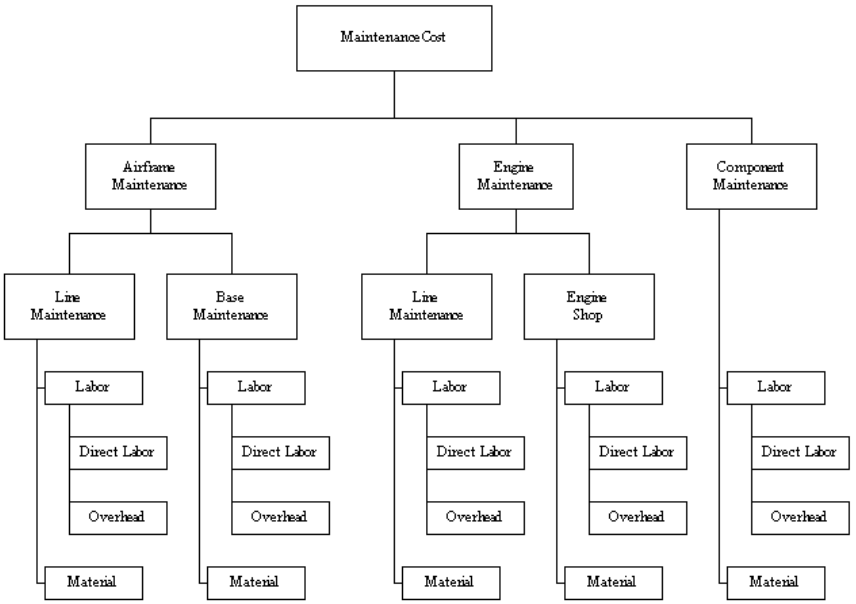


FIGURE 18.4 Elements of maintenance cost.

facilities, utilities, rental, and training. Direct labor can be described as labor that can be charged to a specific maintenance task. Overhead cannot be charged to a specific maintenance task but is rather an activity needed to support the direct labor.

18.8 SCHEDULED AND UNSCHEDULED MAINTENANCE

Scheduled maintenance is preventative maintenance performed by an airline per an established plan to preclude further degradation of airplane systems and the poten-

tial loss of functionality. Scheduled maintenance is a planned activity occurring at specific flight hours, flight cycles, or calendar days. Unscheduled maintenance is reactive or corrective maintenance performed in response to an existing or imminent failure of an airplane system or component. Unscheduled maintenance is maintenance performed by an airline without prior planning. While scheduled maintenance has already been factored into the airline flight schedule, unscheduled maintenance, by its unpredictable nature, has the potential to disrupt the flight schedule. Because it is unplanned, an unscheduled maintenance task can consume more time and therefore be more costly than a comparable scheduled maintenance task. The preplanning to ensure the availability of manpower, parts, materials, and equipment for scheduled tasks makes scheduled maintenance more efficient and cost effective than unscheduled maintenance.

18.9 MAINTENANCE COST COMPARISON

Comparing the maintenance cost of different airlines is a difficult undertaking since no airline is exactly like any other airline. No airline has the same airplane fleet, route structure, regional market, workforce, labor agreements, governing laws, aviation regulations, business strategy, product standards, or operating conditions. A comparison of annual maintenance cost is made even more difficult because the maintenance cost tends to vary with the maintenance cycle of the individual airplanes within an airline's fleet. For example, if airline ABC has more airplanes undergoing heavy checks in a particular year, then, even though the efficiencies and cost structures may be similar, airline ABC will have a higher maintenance cost than airline XYZ for that year. Generally, maintenance cost comparison over the length of a maintenance cycle would be more representative than a maintenance cost focused on one specific year.

18.10 TURNAROUND TIME (TAT)

An important indicator of the efficiency of a maintenance operation is the turnaround time (TAT) required to perform the repair and maintenance of airplane systems and components. TAT is the cycle time in calendar days or man hours to complete a maintenance task. It is fundamental to the efficiency of an operator. Good planning and processes are the critical determinants of TAT ensuring the availability of parts, materials, equipment, and facilities as well as technician experience and training. Efficient TAT practiced in line or hangar maintenance reduces the downtime of the airplane. Efficient TAT in the shop minimizes the amount of inventory required to support maintenance operations.

18.11 PASSENGER YIELD

A measure of the economic performance of an airline is the passenger yield. Passenger yield is passenger revenue/passenger-mile, which is defined as the revenue collected from the passenger divided by the number of miles that the revenue pas-

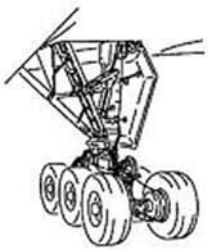
senger was transported. In a recent typical year, the passenger yield was 14.18 cents per mile (U.S. DOT, FAA). Passenger yield is the remuneration received by an airline for transporting a passenger a unit distance and is a relative indicator of economic performance.

18.12 DISCRETIONARY MAINTENANCE

Maintenance cost will vary between operators because of differences in labor cost, maintenance burden, productivity, airplane utilization, fleet age, operating environment, and regulatory requirements. The airline business strategy will affect the maintenance cost by imposing nontechnical maintenance requirements such as passenger seating upgrades and in-flight entertainment systems. Most cabin upgrades are discretionary and primarily do not affect the airworthiness of the airplane. However, some airlines and use cabin upgrades to differentiate themselves from the competition, thereby making cabin upgrades a key for increasing revenue generation. Another example of discretionary maintenance is the number of man hours expended on airplane cleaning and painting, which is more related to customer satisfaction than to a technical maintenance requirement. An airline with a primary focus on high-end passenger service, as compared to an airline with a primary focus on economy passenger service, will gladly incur the additional maintenance cost to meet the primary goal of high-end passenger service as established by business strategy.

18.13 HIGH COST OF AIRPLANE COMPONENTS

From 1960 to 1998, the annual production of commercial transport aircraft ranged from a low of 233 airplanes to a high of 589 (NTS 1999). Comparatively, during the same time frame, passenger car sales ranged from a low of 6.6 million cars to a high of 11.0 million (NTS 1999). The relatively low number of airplanes produced means that the components installed on those airplanes were also produced in low quantities. Airplane components have rigorous aerospace specifications for mission performance, operational reliability, dimensional control, and material quality thereby imposing a high production cost (Figure 18.5). In addition to high relia-



B77 Main Landing Gear

Rigorous Aerospace Specification

- Mission Performance
- High Reliability
- Material Quality
- Precision Tolerances
- High Strength
- Low Weight
- Survive Flight Conditions
- Survive Ground Environment

FIGURE 18.5 Rigorous aerospace specifications.

bility, airplane components must have high strength, low weight, and be able to survive flight conditions and ground environments that present extreme ranges in temperature, pressure, and moisture. Although avionics components are located in pressurized areas of the airplane and therefore protected against extreme conditions, they must be protected from lightning strike, electromagnetic interference (EMI), and high-intensity radio frequency (HIRF). The total effect of these numerous different requirements is to increase the cost of each airplane component. Ranging from high-strength bolts, state-of-the-art avionics to high-tech engine materials, airplane components are uniquely specialized for aviation usage. The extremely low production quantities, the inherent complexity, and the quality standards of airplane components impose a high unit cost. Compliance with regulatory agency requirements for documentation and certification of airplane components adds additional cost to each component.

18.14 COMPONENT RELIABILITY

The mean time between unscheduled removals (MTBUR) of a component is the average number of flight hours accumulated on the component before an unscheduled removal of the component occurs. MTBUR is a key indicator of component reliability and the economic effectiveness of components installed on the airplane. The higher the component MTBUR values, the less often maintenance is being performed on the airplane and therefore the lower the maintenance cost. Component reliability is fundamental to the operating economics of an airplane, with a direct correlation to schedule reliability and maintenance cost.

18.15 SCHEDULE RELIABILITY

Current-technology airplanes have a low incidence of technical airplane problems affecting the flight schedule. A delayed flight, a canceled flight, an air-turnback, or a diversion as a result of a technical maintenance problem does not occur very often. (Note: In the aviation industry, a delay caused by a technical maintenance problem must exceed 15 minutes from the planned revenue departure time to be counted as a delay.) The reliability of the latest mechanical and avionics components installed on airplanes has improved as a result of constant technical upgrades to eliminate or reduce the frequency of failure modes. Consequently, the technology incorporated into airplanes has continued to improve the reliability and safety of the airplane. The Boeing 777 airplane has a schedule reliability of 99.12%, as shown in Table 18.4. This means that a Boeing 777 airplane will be delayed less than 1% of the time as a result of technical maintenance problem. This schedule reliability number relates to delays caused by a technical maintenance problem and does not include delays due to weather, air traffic control, catering, fueling, baggage handling or passenger loading. Maintainability has been designed into the latest airplane models to allow for quicker troubleshooting of technical problems and easier access to airplane components. The maintainability designed into the latest airplane models allows maintenance technicians to rapidly resolve technical problems and to successfully dispatch airplanes thereby achieving higher schedule reliability.

TABLE 18.4 Schedule Reliability of Airplanes

Airplane model	Schedule reliability[a]
B737-300/400/500	99.23%
B737-600/700/800	99.09%
B747-100/200/300	97.43%
B747-400	98.56%
B757	98.87%
B767	98.69%
B777	99.12%
MD-11	97.85%

Source: Boeing.

[a]These data apply to a recent one-year period and are for reference only.

18.16 FIX OR FLY?

As the maintenance crew readies an airplane for dispatch, they may encounter an airplane defect that requires them to make the decision of whether to "fix or fly." If the airplane defect is not flight critical, the maintenance crew may defer the item for later rectification and proceed to dispatch (fly) the airplane. The minimum equipment list (MEL) provides relief for the operator by specifying what equipment has to be operational and what equipment can be inoperative for dispatch of the airplane. The built-in redundancy designed into the airplane allows dispatch with some selected equipment inoperative. The current geographical location of the airplane with a technical maintenance problem is an important determinant of whether the airplane is dispatched. If the technical maintenance problem occurs while the airplane is at the main base, the airline may choose not to dispatch the airplane since the main base has all of the spare parts, ground equipment, test equipment, and maintenance technicians available. This is preferable to dispatching the airplane on a long international trip to locations that may not have the same level of resources. If the airplane is located at a remote location, the airline may prefer to dispatch the airplane to the main base, where more resources are available to rectify the problem. If the technical maintenance problem is flight critical, the maintenance crew will fix the airplane, which may consequently delay or cancel the airplane flight.

18.17 ECONOMICS OF A MAINTENANCE EVENT

The concept of scheduled maintenance is to perform maintenance prior to component or system failure in order to increase the predictability of the flight schedule. Unwarranted maintenance cost reduction such as reducing the availability of frequently needed spare parts, equipment, or manpower can have the unintended effect of increasing the unpredictability of the flight operations. "Airplane on-ground" (AOG) means that a technical maintenance problem is delaying or possibly canceling an airplane flight. An AOG status for an airplane is a serious economic

situation for any airline because it represents an enormous investment not being utilized and is compounded by the loss in revenue. When the purchase price of an airplane can be anywhere from $30 million to $200 million USD, an AOG situation represents a major underutilization of assets. To return an airplane into service for revenue generation, an airline will focus maximum resources to resolve the technical maintenance problem affecting the airplane. Predictability of the flight schedule is fundamental to efficient and economical airline operations. Any unpredictability in terms of delayed or canceled flights due to sudden technical maintenance problems will negatively affect the economics of airline operations. Besides the loss of goodwill on the affected passengers, a canceled flight translates into lost revenue for the airline. Unpredictable flight operations caused by the lack of proper airplane maintenance, parts, materials, equipment, or skilled labor can incur a significant cost to the airline in terms of lost revenue. The lost revenue can easily outweigh the gains of cost-reduction efforts in maintenance operations such as manpower or inventory reduction. A canceled flight can have severe repercussions on flight operations as the airline attempts to position and move airplanes from other locations in order to compensate. For a small or medium airline, one flight cancellation can affect flight operations for an entire week as airplanes are repositioned to compensate for the capacity shortage. A major airline may incur the additional expense of excess capacity by having a spare or standby airplane to serve as a substitute for an airplane with a maintenance problem. Airlines can also compensate by pulling an airplane out of hangar maintenance earlier than planned. The estimated economic consequence of a delayed flight, a canceled flight, and an engine in-flight shutdown are respectively shown in Table 18.5. Although the economic consequences of an engine in-flight shutdown are extremely significant, an engine in-flight shutdown is an infrequent event occurring on the magnitude of about 1 in 100,000 flights for the Boeing 777 airplane engines (Pandey 2000).

18.18 MAINTENANCE RISK

Hypothetically, what if a maintenance technician notices a main landing gear tire with a significant amount of wear? It might be a judgment call whether the tire could withstand a few more flight cycles without consequence. But is it worth putting a $30 to $200 million USD airplane at risk for the sake of saving the cost of a tire change? Postponing, avoiding, or eliminating a maintenance activity in order to reduce maintenance cost must be weighed against the potential risk to the investment, airplane, and passenger safety. In the aviation industry, some airlines arc very conservative and have actually exceeded regulatory agency requirements and manufacturers' recommendations in order to reduce the maintenance risk as-

TABLE 18.5 Economic Consequence of Maintenance Event

Type of event	Cost
Flight delay	$10,000 USD per hour
Flight cancellation	$50,000 USD per event
In-flight shutdown (IFSD) of engine	$500,000 USD per event

Source: Rankin et al. 1998.

sociated with a specific recurring technical maintenance problem. Inversely, no airline can afford to overmaintain an airplane by performing unnecessary maintenance. Each airline has its own unique operating environment from which it must determine the appropriate level of maintenance for an acceptable level of maintenance risk and cost.

18.19 AIRPLANE LEASE (DRY LEASE OR WET LEASE?)

An airplane can either be purchased from the airplane manufacturer or leased from a lessor. Leasing can be categorized as either a capital lease or an operating lease. In a capital lease, sometimes referred to as a finance lease, the airplane appears as an asset on the airline financial statement, while in an operating lease the airplane does not appear as an asset. At the end of a capital lease, an airline can have the option to purchase the airplane. A capital lease is used because of potential tax depreciation advantages for the lessee and the lessor. A capital lease is typically a long-term arrangement, while an operating lease that is typically a short-term arrangement. Operating leases can be further subdivided either into a ''dry leases'' and ''wet leases'' (Figure 18.6). In a dry lease, only the airplane is provided. In a wet lease, the airline is provided with the airplane, flight crew, maintenance, ground staff, ground equipment, insurance and fuel. (A wet lease is sometimes referred to as an ACMI, for ''airplane, crew, maintenance, and insurance.'') A wet lease is used to accommodate a short-term requirement for additional lift capacity. A wet lease is suited to a small airline because the airline can achieve greater lift capacity

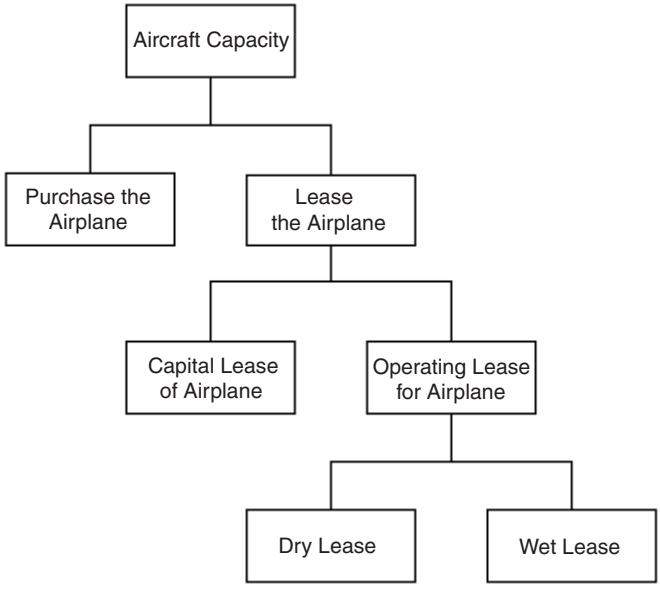


FIGURE 18.6 Aircraft capacity through purchase or lease.

without large capital expenditure. A wet lease allows an airline to begin immediate operations without having to obtain an air carrier operating certificate. Business strategy, internal finances, operating environment, and market conditions will determine whether an airline purchases or leases an airplane.

18.20 IN-HOUSE OR OUTSOURCE?

An airline can choose either choose to perform maintenance in-house or to have a subcontractor perform maintenance. An airline with a small fleet of airplanes can find it more economical to utilize a subcontractor than to establish the maintenance capability internally and incur the substantial investment cost. An advantage of utilizing a maintenance subcontractor is that the airline avoids the maintenance burden or overhead associated with an in-house maintenance workforce. As a business strategy, an airline may choose to utilize subcontract maintenance in order to focus resources on core competencies such as flight operations, marketing, and customer service. Airlines specializing in charter operations may find it to be more economical to have a subcontractor perform maintenance during the off-season as opposed to having an idle in-house workforce during the charter season. Inversely, airlines will choose to have the in-house maintenance capability in order to maintain independence, control, and flexibility. For instance, an airline may want to modify an airplane fleet in order to support a new marketing campaign or comply with a new safety requirement. With in-house maintenance capability, the required airplane modifications would proceed immediately as directed by the airline. With outsource maintenance, the required modifications would be delayed until the subcontractor had sufficient availability of resources. Airlines with in-house maintenance capability believe they have a competitive advantage that provides the assurance of workmanship integrity and product quality. Airlines that utilize outsource maintenance must maintain close surveillance on the billing charges from the subcontractors, be cognizant of market competitive prices for outsource maintenance, and establish controls to monitor product quality. Economies of scale allow a large operator to have lower unit cost than a small operator. Consequently, a small operator with a low volume of work may find it more economical to outsource the work than to perform the work in-house.

18.21 POWER-BY-THE-HOUR (PBH)

"Power-by-the-hour" is a contractual outsourcing agreement in which an airline pays a fixed amount of money based on flight hour utilization to a maintenance subcontractor for the performance of required maintenance. The power-by-the-hour concept can be applied to components, airframe, and engines. Power-by-the-hour is more apt to be utilized by smaller airlines and most likely for engine maintenance. A power-by-the-hour engine maintenance agreement allows a small airline to avoid the expense of an engine shop, engine test cell, engine spares, equipment, and specially trained personnel while allowing it to benefit from the economies of scale gained by a maintenance subcontractor performing engine maintenance for many airlines. Airlines will use power-by-the-hour when their workloads do not provide justification for in-house capability.

18.22 LINE REPLACEABLE UNIT (LRU) AND SHOP REPLACEABLE UNIT (SRU)

Due to the limited amount of time available for the maintenance crew to work on an airplane during an airplane turnaround while it is at the gate or ramp, the airplane manufacturers have designed selected parts to be line replaceable units (LRUs). LRUs by definition are designed such that the unit can be easily removed from the airplane and replaced within the time-limited constraints of line maintenance. In the airplane design, the airplane parts categorized as LRUs have the provision for easy access and removal with the knowledge that at some point in time the LRU may eventually have to be replaced for maintenance. Instead of being embedded into the airplane structure, a mechanical LRU is designed with attach points for easy removal and installation. Instead of being hardwired to the airplane, an avionics LRU is designed with electrical connectors for easy replacement of the LRU. When the LRU reaches the component shop or engine shop, maintenance personnel must evaluate the LRU and decide whether to repair or replace the LRU. Even though the component may be of high economic value, it may be severely damaged. If the LRU has a level of damage or deterioration making it uneconomical to repair, it is judged to be "beyond economical repair" (BER) and will be scrapped. Contained within the LRU there can be lower-level items called shop replaceable units (SRUs). An SRU is an item that should be replaced within a shop environment. An example is a circuit card of an avionics box that for reasons of access and cleanliness is more suited for replacement in the shop environment than during line maintenance on the airplane. Another example is a flap assembly which has high economic value, categorized as repairable and therefore an LRU. A bushing within the flap assembly would be categorized as an SRU. The bushing would be checked to determine whether it should be replaced or repaired in the shop.

18.23 AIRPLANE PARTS

The Boeing 747-400 airplane consists of about 6 million parts, of which half are fasteners. Included are 171 miles (274 km) of wiring and 5 miles (8 km) of tubing. To support a Boeing 777 fleet of 10 airplanes with an inventory fill rate of 95%, the airplane manufacturer recommends an initial spares inventory of about 4,500 parts, based on parts most likely to be required by the airline. The value of such a spares inventory of airframe and engine parts can be equal to about 6% of the purchase price of the airplane fleet. With the overwhelming number of parts used to support airplane maintenance operations, the airline industry developed a concept of part categorization for planning and procurement. Airplane materials and parts can be categorized as consumable, expendable, life-limited, repairable, or rotable (Table 18.6). Consumables and expendables are of low economic value, while repairables and rotables are of high economic value. Life-limited parts can be of either high or low economic value. Repairables and rotables have authorized repair procedures, while consumables, expendables, and life-limited parts do not have authorized repair procedures. An airline will closely monitor the maintenance activity level of repairables and rotables because of their high economic value and critical function for normal airplane operation. An airline will ensure an adequate inventory of the repairables and rotables to preclude an airplane on-ground situation.

TABLE 18.6 Airplane Part Categories

Item	Economic value	Action	Life	Quantity[a]
Consumable	Low	Discard and replace	One-time usage	5,000 max
Expendable	Low	Discard and replace	One-time usage	3,000 to 4,500
Life-limited	Low or high	Discard and replace	Chronological time or flight cycle limits	125 est.
Repairable	High	Repair	Less than life of the airplane	1,500 to 2,100
Rotable	High	Repair	Life of the airplane	

[a]Rough estimate of item quantity on a typical airplane (for reference only).

Life-limited parts are closely monitored to ensure regulatory compliance and avoid potential safety problems.

Consumable

A consumable is a bulk-like material that is a one-time usage item without the possibility of repair. Examples of consumables are abrasive, adhesive, solvent, protective coating, chemical, compound, primer, enamel, dye, hydraulic fluid, oil, grease, sealant, patch, tape, splice, shim, and lockwire.

Expendable

An expendable is a part that, upon failure or damage, is discarded. Repair of an expendable is not economical because a repair would exceed the low cost of the item and because restoration of full functionality and the original specification would be questionable. Examples of expendables are bolt, screw, fastener, washer, nut, pin, rivet, nutplate, spring, bushing, bearing, fitting, bracket, duct, clamp, check valve, relief valve, ball valve, control cable, filter, electrical switch, pressure switch, limit switch, sensor, relay, capacitor, diode, resistor, circuitbreaker, and light bulb. Also considered expendable are parts such as cotter pin, seal, gasket, and O-ring that for reasons of good maintenance practice should be discarded after one-time usage.

Life-Limited Part (LLP)

A life-limited part is a part on the airplane that must be discarded and replaced before a specific number of years or flight cycles. A service life has been established for life-limited parts because they are critical parts that deteriorate with usage, deteriorate with chronological time, serve an emergency function, or operate in a hostile environment. For example, the cost of planned component replacement precludes the unacceptable consequence of an inoperative battery in an emergency situation or the sudden in-flight failure of an engine fan disk. Examples of life-

limited parts are underwater locator beacon battery, emergency flashlight battery, engine fire bottle squib cartridge, escape slide/raft, passenger chemical oxygen generators, engine fan disk, engine fan shaft, high-pressure compressor stage disk and shaft, low-pressure turbine shaft, and high-pressure turbine disk. The life-limited parts of an engine are extremely expensive and dominate the life-limited part replacement costs at an airline. Life-limited parts should not be repaired.

Repairable

A repairable is a part of high economic value that can be economically repaired to full functionality but with limitations as to how many times the part can be repaired. Usually there are damage limits governing the feasibility of repairing the part. The part can be repaired provided that the damage has not exceeded the allowable damage limits. If the part has exceeded the allowable damage limits, the part should be scrapped. Examples of repairables are floor panel, radome, trailing edge wedge, fairing panel, fuselage skin and engine low pressure compressor (LPC) first stage fan blade. If the damage limits for the LPC first stage fan blade are exceeded, the fan blade must be removed, scrapped, and replaced with a new fan blade. If the damage to the LPC first stage fan blade is within the allowable damage limits, the fan blade can be repaired by blending. Some airlines or manufacturers may group repairables and rotables together without distinction between the two categories.

Rotable

A rotable is a part of high economic value that is typically repaired instead of replaced. As dictated by maintenance requirements, rotables are removed from the airplane, repaired in the shop, and stored in the inventory stock. In essence, rotables are parts that are cycled through the airplane, shop, and inventory. Rotables are exclusively categorized as LRUs. Examples of rotables are rudder trim actuator, fuel boost pump, landing gear, autopilot flight director computer (AFDC), HF communication transceiver, integrated drive generator (IDG), engine-driven pump (EDP), air-driven pump (ADP), air data inertial reference unit (ADIRU), ground proximity warning computer, instrument landing system (ILS) receiver, traffic collision avoidance system (TCAS) computer, and auxiliary power unit (APU). Rotables do not have an established life and can be economically repaired to full functionality for the life of the airplane.

18.24 SUMMARY

On a macro level, there are many choices and options for airplane maintenance, based on business strategy, internal finances, airplane fleet size, skill level, and labor rates. Among some of the options discussed are whether the airplane is purchased or leased and whether maintenance is performed in-house or outsourced. On a micro level, the maintenance crew must make day-to-day decisions regarding the dispatch of the airplane during line maintenance. To offset the high initial investment cost, airplanes are designed for long design service life to allow economic operations

over a long time period. Meanwhile, in-production and in-service airplanes are continuously improved by incorporating the latest modifications, consisting of design improvements, new technologies, and new materials. The inherent reliability of the airplane, systems, and components directly affect the maintenance cost. The reliability of airplane components is continuously upgraded to ensure maximum schedule reliability, thereby improving airplane-operating economics. The low production quantity of airplane components, combined with rigorous aerospace standards, high material cost, and labor-intensive processes, makes airplane maintenance a costly business. History has shown that an accident or incident caused by deficient maintenance can seriously damage an airline's reputation and negate all of the marketing expenditures to create a positive consumer image of the airline. In a safety-sensitive industry, some airlines have taken many years to recover or never recovered from a notorious accident or incident. An airplane should be neither overmaintained nor undermaintained. Overmaintaining an airplane can have significant consequences for operating economics, while undermaintaining an airplane can have serious safety implications. The economics of maintenance is a technology-dependent, labor-intensive, capital-intensive domain tightly linked to the requirements of aviation safety, regulation compliance, and market conditions. Airplane maintenance should be viewed not as a cost incurred to comply with regulatory agency requirements but as a necessary and critical function to ensure the highest level of safety for the flying public.

REFERENCES

Air Transport Association (ATA) 2000. Annual Report 2000.

Banis, D., Marccau, J. A., and Mohaghegh, M. 1999. "Design for Corrosion Control," *AERO,* no. 7, http://www.boeing.com/commercial/aeromagazine/aero_07.corrosion.html.

The Boeing Company. BCAG Reliability & Maintainability Engineering, 08-01-1999 through 07-31-2000, Schedule Reliability Summary by Aircraft Series.

National Transportation Statistics (NTS) 1999. Bureau of Transportation Statistics, U.S. Department of Transportation.

Pandey, M. R., 2000. "ETOPS." Presented by The Boeing Company at Appraiser's Forum.

Rankin, W. L., Hibit, R., Allen, J., and Sargent, R. 1998. "Development and Evaluation of the Maintenance Error Decision Aid (MEDA) Process," *International Journal of Industrial Ergonomics,* vol. 16.

U.S. Department of Transportation (DOT), Bureau of Transportation Statistics 1999/1998. Air Carrier Financial Statistics Quarterly, Fourth Quarter, December.

U.S. Department of Transportation (DOT), Federal Aviation Administration. Bureau of Transportation Statistics (BTS) Form 41 data.

PART 2

NATIONAL AND INTERNATIONAL REGULATIONS

Jeremy F. Haines

18.25 BACKGROUND

Aviation is one of the most heavily regulated industries today. This stems from the very nature of this type of transportation system, which tends to be very unforgiving of any shortcoming. As a result, the traveling public has looked to governments to ensure that a safe and reliable service is provided. While the industry was developing in the early years of the last century, regulations were progressively put into place relating to design standards, maintenance requirements, and operations. Much of this regulation was driven by lessons learned from experience. Unfortunately, much of this experience manifested itself as catastrophic accident. It has been the mandate of the government and its regulatory authority to investigate all accidents and ensure any necessary corrective action is put into place. This includes not only current designs and operating airplanes but also regulatory requirements for new designs.

This process has developed at a national level in all countries that manufacture airplanes or operate significantly sized commercial airplane fleets. These national requirements tend to be very similar in nature since the sources for the requirements are mainly the same and there has been considerable cooperation between the regulatory authorities. In addition, the U.S. Federal Aviation Regulations (FARs) have been used as a basis for regulation for many countries. Some countries have actually adopted the FARs. Other countries with fewer resources have negotiated with larger countries to provide regulatory oversight.

International airplane travel has forced countries to recognize each other's requirements and to provide the stimulus for harmonization. This is very evident in Europe, where the majority of regulatory authorities have formed a Joint Airworthiness Authority (JAA). At present 20 countries are full members of the JAA and a further 12 are candidate members. The objective is to develop and adopt Joint Aviation Requirements (JARs) in the fields of aircraft design and manufacture, aircraft operations and maintenance, and the licensing of aviation personnel. The JAA establishes procedures for joint certification of products and services and, where it is considered appropriate, performs joint certification.

Since 1947, the International Civil Aviation Organization (ICAO) has existed as a specialized agency of the United Nations. The ICAO is not a regulatory body, but contracting states collaborate to secure the highest practicable degree of uniformity in regulations, standards, procedures, and licensing. These international standards are adopted as annexes to the ICAO Convention. From a maintenance perspective, the following are relevant:

Annex 1, Personnel Licensing

Annex 6, Operation of Aircraft

Annex 8, Airworthiness of Aircraft

Annex 13, Aircraft Accident Inquiry

Harmonization of requirements and the ICAO Convention enable regulatory authorities to accept the airworthiness of airplanes flying in from other states, including both standards and workmanship. This has facilitated not only the international operations but also airplane marketing, sales, and leases throughout the world.

18.26 FEDERAL AVIATION REGULATIONS (FAR)

Since the processes that have been established worldwide are similar to or based on the U.S. model, the remainder of this Part will describe how the FARs impact on maintenance. The U.S. Congress established that the promotion of aviation safety and the development of civil aeronautics were in the public interest. As a result, the Federal Aviation Administration (FAA) was created as the regulatory authority to set up procedures and standards for the design, manufacture, and operation of airplanes. The standards for design and operation are issued as regulations. These can apply to both manufacturers and operators. Approvals by the regulatory authority showing compliance with these standards are shown by means of certificates. The FAA issues its standards through the Federal Aviation Regulations. The Code of Federal Regulation is divided into 50 Titles, which represent broad areas subject to federal regulation. The Aviation Regulations are found in the Code of Federal Regulations Title 14, Chapter I, Parts 1 through 199. The major Parts from Chapter I of the Code that have an impact on Maintenance are as follows:

FAR Part 21, Certification Procedures for Products and Parts

FAR Part 25, Airworthiness Standards: Transport Category Airplanes

FAR Part 33, Aircraft Engines

FAR Part 34, Fuel Venting and Exhaust Emission

FAR Part 36, Noise Standards

FAR Part 39, Airworthiness Directives

FAR Part 43, Maintenance, Preventive Maintenance, Rebuilding and Alteration

FAR Part 121, Certification and Operations

FAR Part 129, Foreign Air Carriers and Foreign Operators of US-Registered Aircraft

FAR Part 135, Air Taxi Operators and Commercial Operators

Parts 25, 33, 34, and 36 contain the design standards that are used as the basis of certification. Parts 39, 43, 121, 129, and 135 are applicable to operating airplanes.

FAR Part 21 specifies the Certification Procedures for Products and Parts. This FAR contains the general certification under which the FAA approves a design and a manufacturer may subsequently produce a product. The subparts covering each certificate are as follows:

Part 21B, Type Certificate

Part 21E, Supplemental Type Certificate

Part 21G, Production Certificate

Part 21H, Airworthiness Certificate

Part 21K, Part Manufacture Approval

Part 21O, Technical Standard Order

Standards for Safety in Design are contained in Part 25, which prescribes the airworthiness standards for the issue of type certificates, and changes to those certificates, for transport category airplanes. Mandatory changes to the certified design are covered by Part 39, Airworthiness Directives, which prescribes inspections and conditions and limitations associated with certified aeronautical products that affect safety. Part 25.1529 (Appendix H) contains the requirement for Instructions for Continued Airworthiness. This means that a manufacturer must support a new design by Maintenance Manuals including an approved Maintenance Program.

The standards in the FARs that relate to how an operator must perform maintenance are found in Part 43 and Part 121.

Part 43, Maintenance, Preventive Maintenance, Rebuilding and Alteration, requires that methods, techniques and practices acceptable to the administrator be used for inspection and repair of civil aircraft.

Part 121, Certification and Operations—Domestic, Flag and Supplemental Air Carriers and Commercial Operators of Large Aircraft, governs the operations of most U.S. carriers engaged in air transportation. Other Parts address commuter, air taxi, agricultural, and foreign operators.

Part 121, Subpart L, Maintenance, Preventive Maintenance and Alterations, includes Parts 121.361 through 121.380a.

The significant Parts of Subpart L are detailed as follows:

Part 121.365, Maintenance, preventive maintenance and alteration organization, requires a certificate holder to have an adequate maintenance and inspection organization.

Part 121.367, Maintenance, preventive maintenance and alterations programs, requires a certificate holder to have a program to ensure maintenance is performed in accordance with the manual by competent personnel using adequate facilities and equipment and that each aircraft released to service is airworthy.

Part 121.380, Maintenance recording requirements, requires a certificate holder to keep records of airworthiness releases and current status of all maintenance items and life limits, including airworthiness directives.

Part 121.379, Authority to perform and approve maintenance, preventative maintenance and alterations, requires certificate holders to use approved data for major repairs and alterations.

Part 121.703, Mechanical Reliability Report, requires a certificate holder to report to the FAA the occurrence or detection of failures, malfunctions, or defects.

Part 121.705, Mechanical Interruption Report, requires a certificate holder to provide to the FAA a summary report of certain occurrences, such as flight interruptions and premature engine removals.

Part 121.373, Continuing Analysis and Surveillance Program, requires a certificate holder to establish and maintain a system for the continuing analysis and surveillance of the performance and effectiveness of its maintenance program.

The FAA issues Advisory Circulars (ACs) to inform the aviation public in a systematic way of nonregulatory material providing guidance, policy, and infor-

mation. These ACs provide a generally accepted means, but not the only means, of compliance with the FARs. Operators can deviate from an AC with the consent of FAA. A Master Index lists all effective ACs. There are several Advisory Circulars that have an impact on Maintenance Programs, as follows

AC 120-16C, Continuous Airworthiness Maintenance Programs, provides information and guidance on continuous airworthiness maintenance programs which gives the airline the privilege and responsibility for aircraft maintenance. There are five elements of an airworthiness maintenance program:

Responsibility for airworthiness

Maintenance and inspection organization

Performance and approval of maintenance and alterations

Arrangements for maintenance and alterations

Continuous analysis and surveillance

AC 25-19, Certification Maintenance Requirements, provides guidance on the selection and documentation and control of certification maintenance requirements.

AC 43-9C, Maintenance Records, describes methods, procedures and practices determined to be acceptable for showing compliance with the general aviation maintenance record making and record-keeping requirements of Parts 43 and 91. This material is issued for guidance and outlines several methods of compliance with the regulations.

AC 43-12A, Preventative Maintenance, provides information concerning preventive maintenance, who may perform it, the standards of performance applicable to it, authority for approving aircraft for return to service, and the applicable recording requirements.

AC 43.13-1B, Acceptable methods, Techniques and Practices—Aircraft Inspection and Repair, contains methods, techniques, and practices acceptable to the Administrator for inspection and repair of nonpressurized areas of civil aircraft only when there are no manufacturer repair or maintenance instructions. This is a good general source of information used by maintenance to satisfy FAR Part 43. This FAR Part 43 identifies the documentation required (sign-off) to return an aircraft back to service and also defines what is a major and minor repair or alteration.

AC 120-17A, Maintenance Control by Reliability Methods, provides information and guidance materials which may be used to design or develop maintenance reliability programs utilizing reliability control methods. This will satisfy the FAR Part 121.373 requirement for each certificate holder to establish and maintain a system for the continuing analysis and surveillance of the performance and effectiveness of its maintenance program.

AC 120-42A, Extended Range Operation with Two-Engine Airplanes (ETOPS), states an acceptable means, but not the only means, for obtaining approval under FAR Section 121.161 for two-engine airplanes to operate over a route that contains a point farther than one hour flying time at the normal one-engine inoperative cruise speed from an adequate airport.

AC 121-1A, Standards Operations Specifications Aircraft Maintenance Handbook, provides procedures acceptable to the FAA that may be used by operators when establishing inspection intervals and overhaul times.

AC 121-22A, Maintenance Review Board, provides guidelines that may be used by industry during the development and revision of the initial minimum scheduled maintenance/inspection requirements for derivative or newly type certificated transport category aircraft and powerplants for submittal to the FAA for approval. These

initial minimum scheduled maintenance/inspection requirements are referred to in this AC as the Maintenance Review Board Report (MRBR). The requirements, after approval by the FAA, become the framework around which each air carrier develops its own individual maintenance program and which the manufacturer uses in the manuals to satisfy the requirement for instructions for continuing airworthiness.

AC 129-4, Maintenance Programs for Foreign Operators of US-Registered Aircraft under FAR Part 129, provides information and guidance about acceptable maintenance programs for U.S.-registered aircraft subject to FAR Part 129.

Operations specifications are based on FAR Part 121.25. FAA regulations require that specific operational authorizations and limitations applicable to an air operator certificate be issued in conjunction with the issue of the air operations specifications. These are to supplement the general requirements of the basic certificate and to list authorizations and limitations not specifically covered by other FAA regulations. Part 121.25 lists the contents of the operating certificate and the operations specification. The operations specification documents the complete maintenance program for the operator. This combined issuance of the air operator certificate and the operations specification constitute the FAA approval of the unique airline operation.

Variables affecting carrier operations are:

Aircraft types
Operating environment
Operator capability
Level of experience
Facilities
Organizational structure

Manufacturers and operators are required to report certain types of service problems in accordance with FAR Parts 21.3, Reporting of failures, malfunctions, and defects, 121.703, and 135.415, Service Difficulty Reports. The FAA, if it determines that a reported service problem constitutes an unsafe condition, can mandate corrective action. This mandatory corrective action is accomplished by issuing an airworthiness directive. Airworthiness directives are rules and must be complied with as specified in each airworthiness directive. The manufacturer is required by Part 21 to develop any service instructions (inspections, repairs, or modifications) that are necessary and that will be approved by the FAA. In the United States, an airworthiness directive is the only means by which owners and operators (airlines) can be forced to accomplish a manufacturer service bulletin. Airworthiness directives are the end result of the Continued Airworthiness Program performed by the FAA.

18.27 FAA CERTIFICATION

The FAA Aircraft Certification Office (ACO) evaluates design for compliance with regulations and special conditions and issues design approvals by type certificate (TC), supplemental type certificate (STC), technical standards orders (TSO), and engineering approvals for parts manufacturing and repairs (see Figure 18.7). The ACO monitors service difficulties reporting and ensures continued airworthiness by issuing airworthiness directives. This office has the prime responsibility for proposing the need for regulation change.

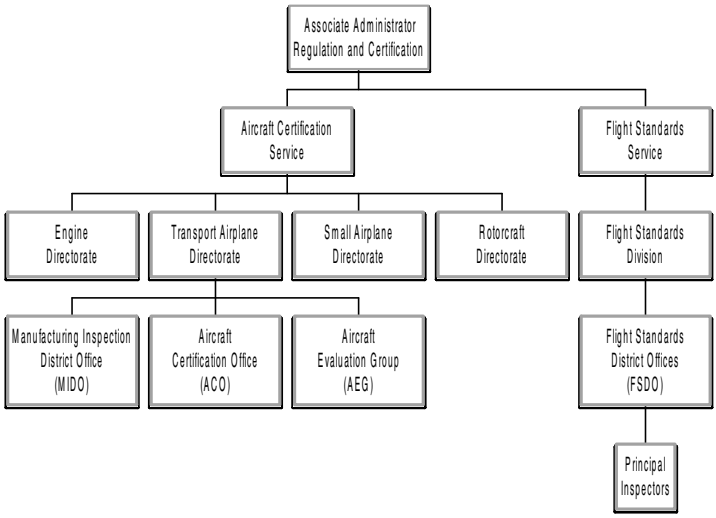


FIGURE 18.7 Federal aviation airworthiness and carrier.

The FAA Manufacturing Inspection District Office (MIDO) approves and surveys production approvals by issuing production certificates and part manufacturing approvals and confirming first article products and witnessing tests. MIDO issues airworthiness and export certificates.

The FAA Aircraft Evaluation Group (AEG) provides coordination between certification and flight standards. In particular, the AEG manages the Maintenance Review Board (MRB) for the FAA and provides the communication link between the Flight Standards District Office (FSDO), the Principal Maintenance Inspectors (PMI), the Principal Operations Inspectors (POI), and the Aircraft Certification Office (ACO).

An aeronautical product must conform to its type design, supplemental type design, and any applicable airworthiness directives and be in a condition for safe operation. The word ''airworthiness'' is used throughout the regulations without being defined. A clear understanding of its meaning is essential. Two conditions must be met for an airplane to be considered airworthy:

1. An aeronautical product is airworthy when it conforms to the regulations under which it has been certified. This means that the product has to comply with the documentation that demonstrates that the design meets the regulations. This documentation has been submitted to the FAA as the basis for certification approval. The airplane configuration and the components installed must be consistent with the drawings, specifications, and other data that are part of the type certificate.
2. An aeronautical product must conform to its type design, supplemental type design, and any applicable airworthiness directives, and be in a condition for safe operation. This means that an airplane continues to be airworthy when it continues to meet the design (or approved design changes) and is being inspected in accordance with, and meeting the requirements of, the approved manuals. The condition for safe operation refers to the condition of the airplane relative to wear and deterioration

Regulatory approvals of airplane certification are shown by means of certificates. These certificates cover the manufacture and delivery of each airplane. The operator also requires certificates to show continuing airworthiness. The airplane requires three certificates:

1. The Type Certificate (FAR 21 Subpart B) is issued by the FAA Aircraft Certification Office (ACO) to approve each specific airplane type design as conforming to the Airworthiness Standards.

Type certification requires that the configuration and features of the type design comply with the requirement of the FARs that are applicable. This compliance must be substantiated by test or analysis and presented as documentation to support the type approval application. It establishes the airworthiness of the design by showing conformity to "the basis of certification." The basis of certification (FAR PART 25, 33, 34, and 36) includes:

- Configuration and features of the product
- Structural strength of the product (dimensions, materials, and processes)
- Life limits, airworthiness limitations, and certification maintenance requirements
- Operating limitation in the Airplane Flight Manual (noise, fuel venting, and exhaust emissions).

The fatigue evaluation of structure required by Part 25.571 for certification can result in mandatory maintenance requirements. These can take the form of life limits for the safe life structure and supplemental fatigue-related inspections (airworthiness limitations, AWL) for the principal structural elements for which the baseline inspection program does not satisfy the damage tolerance requirements.

System design and analysis required by Part 25.1309 for certification can result in mandatory maintenance requirements. Nowadays design certification is much dependent on reliability substantiation by analysis and, where necessary, by appropriate ground, flight, or simulation tests. Failures are classified according to their severity, that is, the failure effect on the airplane and its occupants, both direct and consequential. To demonstrate analytically that catastrophic and hazardous events will be extremely improbable or remote, a requirement for maintenance may be included in the calculation of probability. Any such maintenance then becomes mandatory as part of certification and is known as a certification maintenance requirement (CMR).

2. The Production Certificate (FAR 21 Subpart G) is issued by FAA Manufacturing Inspection District Office to the manufacturer as approval of the manufacturing and quality system for building airplanes. It certifies that the manufacturer has established and can maintain a quality control system so that each product will meet the design provisions of the pertinent type certificate.

To obtain a production certificate, it is necessary that a manufacturer satisfy a FAA inspection that there are adequate facilities, quality control procedures, and a drawing system that will enable manufacture of repeat airplanes that conform to the type certificate requirements. The production certificate shows that the manufacturer has approved facilities and approved quality control procedures to meet the design provisions of the type certificate. The production certificate authorizes the manufacturer to produce the particular type airplane.

3. The Airworthiness Certificate (FAR 21 Subpart H) issued by the FAA Manufacturing Inspection District Office as approval to deliver and operate an airplane that has been built and tested in compliance with production and type certification. It certifies that, as of the date of issuance, the aircraft to which the certificate has been issued has been inspected and found to conform to its type certificate and to be in a condition for safe operation. Note that this certificate is also used by the operator.

With a type approval certificate and a production certificate, a manufacturer is authorized to produce airplanes and apply for an airworthiness certificate for each copy to show that it conforms to the type design and is in a condition for safe operation. The airworthiness certificate confirms that the aircraft is airworthy, that is, conforms to its type certificate, and is in a condition for safe operation. As long as the airworthiness certificate remains valid, it permits the operation of the airplane within the limitations and conditions provided.

The operator requires two certificates:

1. The airworthiness certificate (FAR 21 Subpart H) is issued by the FAA Manufacturing Inspection District Office.

This Certificate was originally issued to the manufacturer and remains valid as long as the airplane is being maintained and inspected in accordance with the requirements of the approved manuals and regulations.

2. The operating certificate (FAR 121) is issued by the FAA Flight Standards District Office (FSDO). This is the approval that a carrier meets the requirements for the operation for which he has applied and may operate while those conditions apply. It authorizes an air carrier to engage in scheduled air transportation.

The airline must define the type of operation, the airplane type operated, the operating environment, and the operator capability. The Operations Specification matches Part 121 to the unique operation of the airline and documents the complete maintenance program for the operator. It is necessary for maintenance organizations to be cognizant of applicable national and international regulations. Failure to comply can result in disruption to operations and degradation in safety. This description of regulation and certification procedures demonstrates the complexity and extent of the necessary regulation. These or similar requirements are in use throughout the world and are in place to promote aviation safety.

PART 3

MAINTENANCE PROGRAMS

Adel S. Zeki

18.28 AIRPLANE MAINTENANCE

One of the least desirable things that can happen to an airline or operator is to have an airplane on the ground not generating revenue, knowing very well the only time this airplane is useful is when it is in the air transporting payload. Airline management recognizes, however, that all is not revenue and that a carefully planned and executed maintenance program is vital to operational reliability, airplane performance, and continuing airworthiness of the airplane.

Maintenance requirements for any specific airline are dictated by a number of factors, such as type of airplane, fleet size, route structure, and flying schedules. To a large degree, the type of equipment establishes maintenance frequency cycles, while the size of the fleet determines quantitative maintenance cycle loads. Route structure and flight schedules influence the location and number of stations that must possess the capability of performing the work.

The airplane manufacturer, throughout the design and manufacture phase of the airplane, develops basic maintenance and planning data. These data are usually divided into work packages containing all tasks and work items to be accomplished within specified airplane operation time periods (intervals) or calendar times. These time intervals may be progressively increased as experience is gained on the airplane, as design improvements keep pace with the state of the art, and as advanced techniques are developed in flight and maintenance operations. The work packages are based on airplane flying times and designated with letters, as in the following example. "A" check accomplished at 200 flight hours, 2A at 400 flight hours, and 4A at 800 flight hours. "C" check accomplished at 3,200 flight hours, 2C at 6,400 flight hours, and 4C at 12,800 flight hours. "D" check at 25,000 or 5 years, whichever comes first. These packages increase in scope and detail progressively from the "A" through the "D" checks.

The "A" check usually consists of a visual inspection of the interior and exterior areas of the airplane. This includes inspection of certain specified equipment for general condition, servicing systems and equipment with air, fluids, etc., operational check of electronics systems and flight controls.

The "C" check includes all items contained in "A" and multiple "A" checks plus some additional work items:

1. Visual interior inspection of areas such as equipment compartments and wing bays, empennage, wheel wells, nacelles, and engines
2. Visual exterior inspection of fuselage, wings and control surfaces, and empennage
3. Operational checks of control surfaces, instruments, and essential flight equipment
4. Equipment checks such as operational checks of systems and components
5. Servicing such as lubrication, cleaning filters, screens, etc., replenish air, fluids, etc.

6. Accomplishing nonroutine work (unscheduled work) as discovered by routine inspection
7. Fixing faults and problems reported by pilots in the airplane technical log book

The "D" checks include all items contained in the "A" and "C" checks and their multiples, plus structural inspections and equipment tests as follows:

1. Structural inspection of the control surfaces, landing gear and supporting structure, doors, hatches, fuselage, nacelles, horizontal and vertical stabilizers, and wings
2. Inspection and refurbishing of areas such as control cabin, passenger cabin, cargo compartments galleys, and toilets
3. Detailed visual inspection and operational checks of all systems
4. Hard time and life-limited component changes as scheduled
5. Structural, systems, and component repairs and replacement as discovered by routine inspection
6. Servicing such as cleaning filters, screens, lubrication, and replenishment of air, fluids, etc.
7. Engine run-up if applicable
8. Test flight if required
9. Fixing flight test faults as reported after test

Airline maintenance philosophy, practices, and procedures vary widely. In fact, because the operational requirements are unique to each airline situation, no two airlines operate under identical maintenance plans.

Because of these situations, the maintenance planning data (MPD) document, which is developed by the airplane manufacturer, is rearranged by the individual airlines to suit the unique operational and technical situations encountered by the airline and also to suit its maintenance philosophy.

18.29 MAINTENANCE PROGRAM REQUIREMENTS AND THE HISTORICAL DEVELOPMENT OF AIRCRAFT MAINTENANCE THEORIES

If we take a brief look at the historical development of aircraft maintenance theories, we find that during the early 1930s the industry believed that mechanical parts wear out with time, wear causes failure, and failure degrades safety. This belief led to the periodic disassembly of everything from components and engines to structures and radio sets. With this philosophy the operating time of the airplane controlled all activities in airplane maintenance. *Hard Time,* also called scheduled overhauls, was the only recognized primary maintenance process.

After the Second World War, a second primary maintenance process was recognized. It was called *on condition* maintenance because it permitted the use of periodic condition checks instead of scheduled overhauls. This process was mainly applied to items known to wear, such as brakes and tires.

Neither of the two processes fully accommodated the improvements in the design and materials of the new components and parts which were introduced on jet airplanes during the 1960s. The industry analyzed the behavior of hundreds of components during millions of flight hours and finally discovered that the reliability of complex assembles as opposed to single-cell items does not decrease with increasing age. Preventive maintenance processes like on condition and hard time therefore could not be used as a means to ensure their continued reliability. The proper way to maintain these complex assemblies turned out to be a third primary maintenance process called *condition monitoring*.

Condition monitoring is not a preventive maintenance process. It allows malfunctions to occur and relies upon the analysis of data generated by such malfunctions for taking appropriate actions. In the early 1970s, when wide-body jets were introduced, airplane maintenance theory had arrived at its current form. It states:

1. Mechanical parts wear out over a wide range of time.
2. Part reliability is a function of its inherent design. Most designs include redundancy that prevents failures from having an effect on safety.
3. Good and adequate maintenance ensures that a part is as reliable as its initial design.
4. Overmaintenance does not improve reliability. However, less than adequate maintenance can degrade reliability.
5. Engineering modification is required to improve inherent reliability.
6. There are only three primary maintenance processes: hard time, on condition, and condition monitoring.
7. The function, failure mode, and failure effect of a unit can be used in a logical manner to arrive at the minimum preventive maintenance requirements to protect inherent reliability.

18.30 MAINTENANCE STEERING GROUP—1 (MSG-1) DECISION LOGIC

In July 1968, representatives of various airlines developed Handbook MSG-1, *Maintenance Evaluation and Program Development,* which included decision logic and interairline/manufacturer procedures for developing the maintenance program for the new Boeing 747 airplane. Subsequently, it was decided that experience gained on this project should be applied to update the decision logic and delete certain 747 detailed procedural information so that a universal document could be made applicable for later new type aircraft. This was done and resulted in the document titled *Airline/Manufacturer Maintenance Program Planning Document,* MSG-2, which is a decision logic used to develop scheduled maintenance programs for the aircraft of the 1970s.

18.31 MAINTENANCE STEERING GROUP—2 (MSG-2) DECISION LOGIC

Airline/Program Planning Document MSG-2, dated March 25, 1970, shows in detail how the maintenance theory is being used to determine the essential scheduled

maintenance requirements for a new airplane. Airplane systems, structures, and engines are put through the MSG-2 decision logic to find which tasks should be done for safety and economic considerations. These tasks are intended to prevent adverse effects on operating safety, ensure the availability of hidden functions, and restore inherent system reliability. The final result is one list of preventive maintenance tasks divided into hard time and on condition tasks, and another list of those items that require no scheduled maintenance and are therefore included in the condition monitoring process.

18.32 MAINTENANCE STEERING GROUP—3 (MSG-3) DECISION LOGIC

In 1979, a decade after the publication of MSG-2, experience and events indicated that an update of the MSG analysis procedure was both timely and appropriate in order for the document to be suitable to develop maintenance programs for new airplanes, systems, and power plants.

An Air Transport Association (ATA) Task force reviewed MSG-2 and identified various areas that were likely candidates for improvement. Some of these areas were the rigor of the decision logic, the clarity of the distinction between economics and safety, and the adequacy of treatment of hidden functional failures. Additionally, the development of new-generation airplanes provided a focus, as well as motivation, for an evolutionary advancement in the development of the MSG analysis concept. New regulations, which had had an effect on maintenance programs, had been adopted and therefore needed to be reflected in MSG analysis procedures. Against this background, Air Transport Association airlines decided that a revision to existing MSG-2 procedures was both timely and appropriate. The active participation of members of the aviation industry generated the MSG-3 document. As a result, there were a number of differences between MSG-2 and MSG-3, which appeared both in the organizational and presentation of the material as well as in the detailed procedural content. However, MSG-3 did not constitute a fundamental departure from the previous version. It was built upon the existing framework of MSG-2, which had been validated by 10 years of reliable airplane maintenance operations.

Airline/Manufacturer Maintenance Program Planning Document MSG-3, dated October 1980, was first revised in 1983 to become MSG-3 Revision 1 and then in 1993 to become MSG-3 Revision 2. This latest revision is the present-day standard to determine the essential scheduled maintenance requirements for new airplanes. MSG-3 includes detailed decision logic for assigning maintenance tasks and task intervals.

Quite often, airlines operate a fleet of airplanes that are divided between both MSG-2 and MSG-3 decision logic programs. Generally, all new airplanes manufactured today will follow the MSG-3 concept, while older airplanes continue with the older philosophy of MSG-2.

18.33 DEVELOPING THE MAINTENANCE PROGRAM PROPOSAL

Upon completion of the MSG-3 process, the participating airlines and the airplane manufacturer meet in committee to arrive at a proposed maintenance program which

is submitted to the Federal Aviation Administration (FAA) Maintenance Review Board (MRB). After reviewing the proposed program, the MRB holds further conferences with the committee to discuss work requirements and time intervals. Upon final FAA approval, each airline operating the new airplane uses the FAA-approved program as the basis for its initial maintenance program. The new MRB requirements and other requirements recommended by the manufacturer's maintenance planning data (MPD) document must be tailored by the individual airlines to suit the operational and environmental conditions under which the airline is operating. This involves adjustment of task content and check intervals to account for the influence of climatic conditions such as temperature, humidity, salt air, smog, etc., and the operating environment which influence

18.34 OPERATIONS SPECIFICATIONS

The civil aviation authorities (the FAA in the United States) require that specific operational authorization and limitations applicable to an Air Operator Certificate (AOC) be issued for the new airplane. These specific operational authorizations are referred to as operations specifications and are intended to supplement the general requirement of the basic certificate. They also list authorizations and limitations not specifically covered by the civil authority regulations. The operations specifications are divided into separate parts as follows:

Part A, General

Part B, En route authorization

Part C, Airports authorizations and limitations

Part D, Maintenance

Part E, Weight and balance

Part F, Interchange of equipment operations

Part G, Aircraft leasing operations

Part D, Maintenance, is where the civil authority, i.e., the FAA, approves the airline maintenance program. This part provides detailed maintenance-related authorizations and limitations for a particular operator that are not specifically described by the regular civil authority regulations, such as time limitations for inspections and overhaul which vary with the type of aircraft and maintenance program followed. When specified, these limitations become as binding on the operator as other regulations.

Part D of the Operations Specifications is divided into two categories of material. One specifies the inspection, check, and overhaul time limits for airframes and power plants tasks, and the other covers a number of maintenance-related authorization pages unique to the particular characteristics of the proposed operations. Examples of such authorizations are the short-term escalation of check intervals, ferry flight continuing authorization, and reliability program authorizations.

Maintenance Program and Related Maintenance Practices

Just as maintenance programs vary among operators using similar airplanes, so do the practices and procedures applied in carrying out these programs. These practices

are heavily influenced by airline maintenance policy, management decision, economic situation, age of airplanes, etc. One airline will wash the airplane at each "C" check and paint it at every other "D" check, while another will wash only when necessary and paint annually. Domestic airlines, due to route structure, will usually be able to bring all airplanes to the maintenance base every other night or twice a week, while another, operating internationally, may return certain airplanes to the maintenance base only once in a two-month period. These variables in the opportunity to perform maintenance obviously influence maintenance practices, procedures and cost. Maintenance scheduling is much simpler for the operator who can return his airplanes home frequently. In this case, maintenance workloads are more easily controlled at a constant level and maintenance manpower requirements are therefore more stable.

Some airlines perform maintenance under severe conditions, while others provide elaborate facilities and equipment. Some operators have fitted steel docks equipped with electrical outlets and floodlights, water and air, benches with vises, and even tool boards. After the airplane has been positioned precisely by placing the wheels over index markings on the hangar floor, these docks are moved into positions around and under the airplane on steel rails embedded in the floor. Operators having such support equipment usually also have elaborate tail docks of three, four, or five levels (some with elevators), each similarly equipped and moving on positioning rails. At the other extreme is the operator whose maintenance support equipment consists of step ladders, with platform steps which can be raised and lowered hydraulically, platforms on steel tubing frames that may be mechanically raised or lowered, "cherry pickers" to perform work on the vertical tail, and a variety of scaffolding. These contribute to congestion in the working area. However, most operators will provide maintenance equipment somewhere between these two extremes.

Maintenance facilities vary with the prevailing climate. Operators at bases in colder regions and extremely warm areas equip hangars with doors, heat, and air conditioning, while in tropical areas hangar doors, heat, and air conditioning may be omitted.

The composition of maintenance crews varies as widely as do equipment and facilities. The well-equipped facility usually permits more efficient maintenance than the congested shop. An effective maintenance crew in a well-equipped facility will average 20 to 30 men, while the congested facility may require 30 to 45 men to perform the same maintenance check in the same elapsed time as that required by the more efficient shop.

Scheduling for a Maintenance Event (Check)

Maintenance work packages are assembled in advance by production planning and control and are based on the scheduled work to be performed during the specific check on the individual airplane. Engineering orders, service bulletins, and modifications are added to scheduled work items as applicable to either airplane or equipment. Items which may have been deferred from the previous scheduled check(s) are also included. This work package is delivered by production scheduling to the hangar supervisor or the maintenance control center, where it is analyzed for scope and man hours required for accomplishment within the available airplane downtime. Maintenance shifts and crew strength within each shift is adjusted accordingly. Specialized shops are notified of known support requirements and the

inventory control department is advised of known requirements for components, parts, and materials.

The Maintenance Day

The maintenance day normally begins at midnight, although this may vary with arrival schedules of airplanes requiring maintenance. After airplane arrival, time is allowed to offload passengers, baggage, and other cargo. Work assignment for first shift crews will consume 15 to 20 minutes. The airplane is then positioned in its designated place in the hangar and the docks or work stands are moved into place, toolboxes are located convenient to the work area, and the check is begun.

In the meantime, the last day's flight logs are reviewed and snags written by the pilot or copilot are added to the work package. During this time the inspectors perform their zonal inspection tasks and each discrepancy discovered is added to the work package. These additions are assigned for correction to lead men of appropriate teams. Meanwhile, supplies are delivered to the maintenance area, including those items of parts and materials for which they were previously alerted, plus other required items as needed. This action will continue as long as there is a requirement.

Some maintenance organizations provide mobile tool cribs which are moved to the site, while others require special tools to be withdrawn from a centrally located tool crib. In most cases, the mechanic obtains needed tools by surrendering a tool check which bears his name, Many operators now attempt to issue all special tools required for a specific maintenance task to shift supervisor, who, if the task is not completed, passes accountability on to the supervisor of the succeeding shift This procedure is most effective when mobile tool cribs or portable tool boards are used, as accountability can be transferred more rapidly.

The first shift, which usually begins from 12:01 a.m. to 8:30 a.m., is usually the strongest in manpower, while the second shift, which begins from 8:01 a.m. To 4: 30 p.m. is the next strongest. Assuming that the end of the second shift will complete the majority of maintenance tasks, the third shift, from 4:01 p.m. to midnight, has the least manpower. This shift also overlaps the second shift by 30 minutes for pickup of unfinished items and cleaning of small jobs remaining to be accomplished.

The foregoing description of shift schedules and shift strength is merely representative, as circumstances change at each base and for each operator the requirement may change to fit those circumstances.

PART 4

AIRCRAFT STRUCTURAL CORROSION

Aydin Akdeniz

18.35 INTRODUCTION

Metals in nature exist as compounds, such as oxides, sulfites, and nitrates, which are fairly stable. Most pure metals and alloys are not stable and as soon as they are formed tend to revert back to their natural states by combining with elements in the atmosphere. This chemical reaction of metals reverting back to their natural states, such as oxides, is known as a corrosion reaction.

Two types of corrosion can be considered:

1. Corrosion involving a liquid phase when conductive solution is present
2. Dry corrosion involving metal/gas or metal/vapor reaction

18.36 THERMODYNAMICS OF CORROSION

As with all natural processes, corrosion can be observed from a thermodynamic and kinetic viewpoint.

$$G = H - TS$$

where G = Gibbs free energy (cal/mol)
H = enthalpy (heat energy) (cal/mol)
T = absolute temperature
S = entropy

For a given process in nature to be spontaneous, ΔG must be negative. A given metal's electrode potential can be calculated when concentrations (i.e., activities) are known.

ΔG of a chemical reaction is in equilibrium:

$$aA + bB \rightleftharpoons cC + dD \tag{18.1}$$

$$\Delta G = \Delta G + RT\ln \frac{[C]^c[D]^d}{[A]^a[B]^b} \tag{18.2}$$

where R = gas constant
[ ] = concentration of species

$$\Delta G = n\mathcal{F}\varepsilon \Rightarrow \text{electrochemical reaction is in equilibrium} \tag{18.3}$$

where $\mathcal{F}$ = Faraday constant = 96,000 coulombs/equivalent weight
ε = the potential (volts) of the reaction for oxidation.

Combining Eqs. (18.2) and (18.3) above:

$$\varepsilon = \mathring{\varepsilon} + \frac{2.3RT}{n\mathcal{F}} \log_{10} \frac{[C]^c\,[D]^d}{[A]^a\,[B]^b} = \text{electrode potential (Nernst equation)} \quad (18.4)$$

where $\mathring{\varepsilon}$ = metals standard oxidation potential

$$\varepsilon = \mathring{\varepsilon} + \frac{0.059}{n} \log \frac{[\text{oxidized}]}{[\text{reduced}]} \quad (18.5)$$

A metal's electrode potential indicates its tendency to corrode. By convention, the more negative the electrode potential, the more active the metal will be.

As shown in Eq. (18.5), every electrochemical reaction involves oxidation at one electrode and reduction at the other.

Anode: where loss of electrons (oxidation) occurs

Cathode: where gain of electrons (reduction) occurs

By calculation (Nernst equation), the reversible electrode potentials at the anode and cathode, the potential for corrosion (thermodynamics) can be determined (Figure 18.8).

If the reversible potentials (ε_{res}) are linearly plotted on a vertical line, the thermodynamics for a corrosion system can be established. If the equilibrium (reversibility) of each electrode electron flow (current) occurs, this current dictates the rate of corrosion (electrochemical reactions) and is governed by kinetics.

As shown in Figure 18.8, the change in electrode potential occurs when current i is allowed to flow in the system.

$$\text{Corrosion rate} = \frac{I_{corr}}{\mathcal{F}} \quad \text{Faraday's law (uniform corrosion)}$$

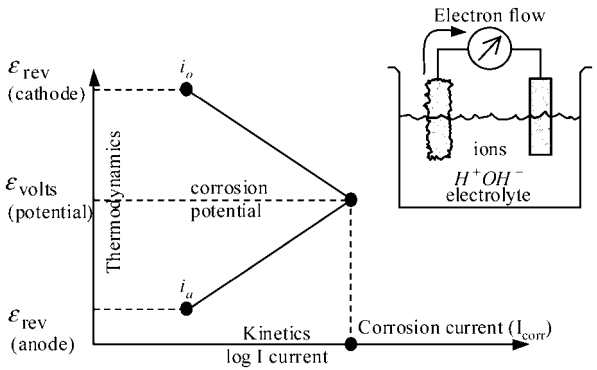


FIGURE 18.8 Current-potential diagram for a typical corrosion reaction.

18.37 AIRCRAFT MATERIALS

Material for aircraft structural application must meet a wide range of design requirements for corrosion prevention:

- Proper choice of environmentally stable materials should be made.
- Parts should be designed to minimize exposure of corrosion-prone end grain boundaries to moisture.
- Parts should be designed to prevent the ingress and accumulation of moisture. Protective treatments or surface coating should be carefully selected to enhance corrosion resistance of components operating in particularly aggressive environments.
- Assembly techniques should be used to avoid damage to protective coatings and seal structure against moisture penetration. Parts should be designed to allow access to all areas of structure for inspection and maintenance.
- Inhibitors and water-displacing fluids may be used for supplementary protection but should not be relied upon as the primary form of corrosion protection.
- Since corrosion is inevitable in the long term, some allowance should be made for the effects of corrosion in the initial stage.

Since the economic life of an aircraft may be limited by corrosion rather than fatigue, it is important that corrosion resistance be given high consideration during the design stage. The aim should be to ensure that the final structure is of optimum configuration, not only for immediate flight performance, but also to allow manufacture, assembly, and in-service inspection and maintenance to be performed easily. It should be performed in a manner that maintains the quality of the structure and its ability to resist corrosion.

Water-displacing compounds (also called corrosion-inhibiting compounds, or CICs) may be useful in providing supplemental protection during aircraft operation. They are applied as fluids to faying surfaces and crevices for displacing moisture.

18.38 TYPES OF AIRCRAFT MATERIAL CORROSION

Corrosion is a serious airworthiness concern and a costly problem on all types of aircraft. Approximately 95% of commercial airplane structural repairs are corrosion related. A greater concern is the aircraft structural integrity throughout its projected life or design service objective (DSO). Environment and inadequate maintenance practices combine to make aircraft materials vulnerable to corrosion as shown in Figure 18.9. Unchecked corrosion can cause structural and system failures and, in extreme instances, loss of the aircraft.

The responsibility for keeping corrosion under control starts with the airplane designer realizing the importance of material selection, drainage material finishes, and proper sealant for structural durability and easy access for proper maintenance. Operators must have trained personnel to clean corrosion and repair structure as appropriate to carry the design and regulatory loads. Regulatory agencies must make sure that the manufacturers and the airplane operators are following rules and regulations, as shown in Figure 18.9.

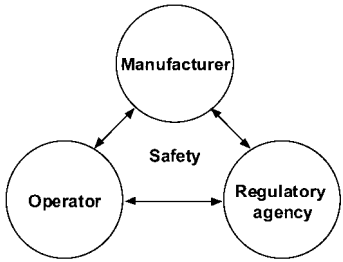


FIGURE 18.9 Industry must work together for successful design and in-service operation of airplanes.

Corrosion Types

The following definitions and illustrations are common types of corrosion found in jet transport aircraft.

Pitting Corrosion

Pitting corrosion is a localized type of corrosion which leads to the formation of deep and narrow cavities in metals. The developing corrosion pits can act as stress risers that could evolve into structural fatigue or stress corrosion cracking, as shown in Figure 18.10.

Crevice Corrosion

Crevice corosion is a form of corrosive attack occurring when a corrosive liquid, such as salt spray, gains access to crevices in or between components. It also occurs

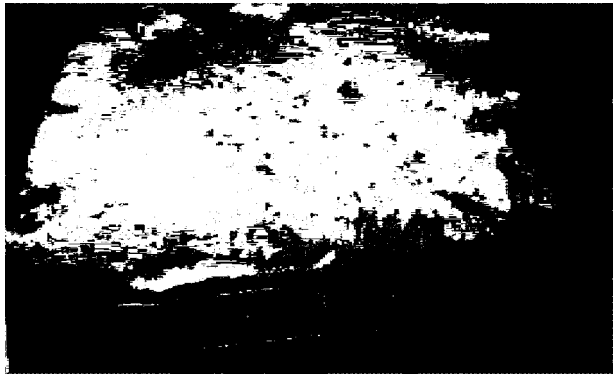


FIGURE 18.10 Pitting corrosion example on aluminum web.

when moisture and salt are present in a joint, creating a very acidic and corrosive condition (see Figures 18.11 and 18.12).

Crevice corrosion has many characteristics in common with pitting corrosion. For example, it is common in passive metals such as stainless steels and aluminum alloys which form protective oxide films, and it is often observed in solutions containing high concentrations of both chloride and hydrogen ions, which promote the breakdown of these films.

Filiform Corrosion

Filiform corrosion is the forming of a network of threadlike filaments of corrosion products on the surface of a metal coated with a paint film. This is a form of crevice

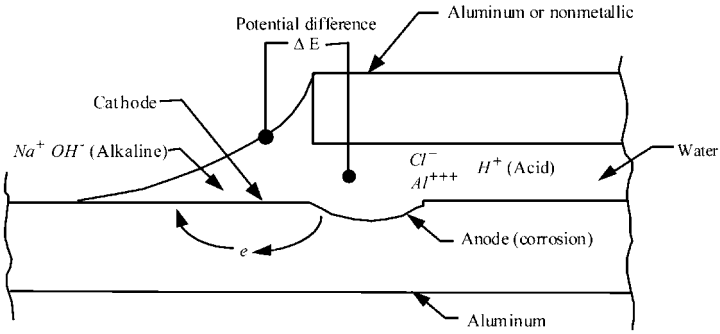


FIGURE 18.11 Schematic of a typical concentration cell (crevice) with salt (NaCl) present.

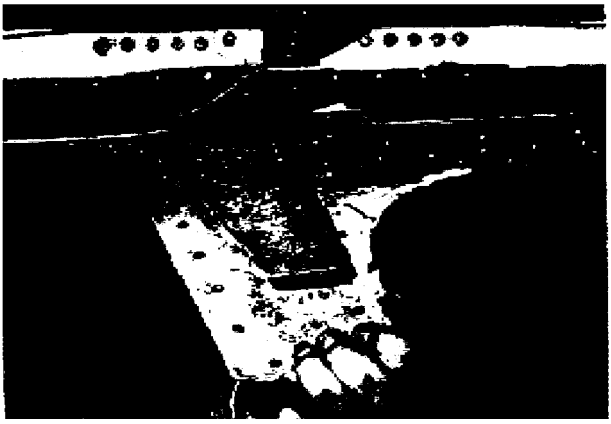


FIGURE 18.12 Results of concentration cell corrosion in a delaminated bond of a fuselage tear strap.

corrosion which begins at a break in the paint film and develops as a filament. It may become quite severe before it is detected, since it will often be hidden by cladding or paint, as shown in Figures 18.13 and 18.14.

Intergranular / Exfoliation Corrosion

Intergranular (exfoliation) corrosion is a highly localized form of dissolution which affects the grain boundary regions in a polycrystalline metal. The corrosive attack can produce a network of corrosion or cracking on the metal surface, occasionally dislodging whole grains, or it may penetrate deeply into the metal, leaving behind very little visible evidence of the damage, as shown in Figures 18.15 and 18.16.

Exfoliation corrosion is a form of intergranular corrosion that follows the elongated grain paths created by the manufacturing processes of rolling, extruding, or

FIGURE 18.13 Corrosion around fastener heads.

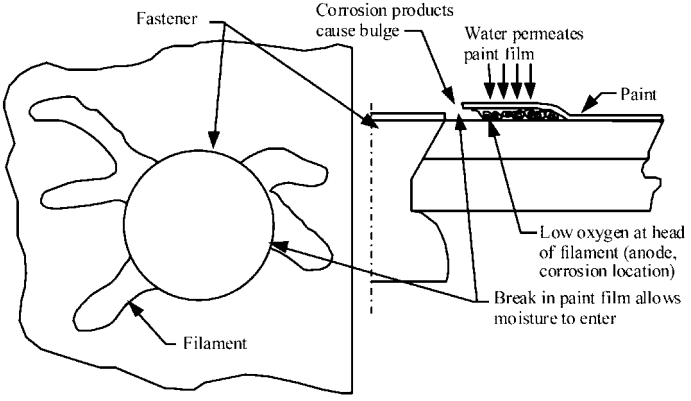


FIGURE 18.14 Schematic of filiform corrosion.

FIGURE 18.15 Exfoliation corrosion on wing trailing edge lower chord beam.

FIGURE 18.16 Magnified cross-section of exfoliation corrosion.

forging. Such corrosion can cause rapid material loss because of the localized intensity of the corrosion.

Stress Corrosion Cracking

Stress corrosion cracking is a form of intergranular corrosion that generally follows a single plane in the metallic grain structure in aluminum alloys. Environment and stress, either applied or residual, must be present, with the resulting crack generally propagating at a rapid rate, as shown in Figure 18.17. Some materials susceptible to stress corrosion cracking are the 7000 series aluminum alloys in the T6 condition and high-strength alloy steels (tensile strengths >200ksi).

General Surface Corrosion

When observed with the unaided eye, this type of corrosion has no distinctive appearance, such as exfoliation or pitting, but is relatively uniform, as shown in

FIGURE 18.17 Stress corrosion cracking shown by the arrows on a extruded 7079-T6 aluminum angle.

Figure 18.18. While the usual rates of such corrosion are low, resulting in slow loss of structure, it can develop into other corrosion forms with higher rates (e.g., pitting or exfoliation) if left unattended.

Fretting Corrosion

Wear from differential movement between two surfaces can repeatedly remove the protective oxide, exposing fresh metallic surfaces. In addition, metal particles may be dislodged by local welding or shear action. These particles corrode rapidly because of the large surface area exposed, adding to the corrosion and oxidation products and thereby compounding the wear. Pits generated by such fretting can become stress risers and eventually lead to fatigue cracking. A classic example of fretting corrosion is shown in Figure 18.19. An incipient fatigue crack starting at a pit is evident in the lower photo.

Dissimilar Metal / Bimetallic Corrosion

This corrosion is generated when a galvanic cell is formed by the contact of dissimilar metals in the presence of an electrolyte. The active anode corrodes. Figure

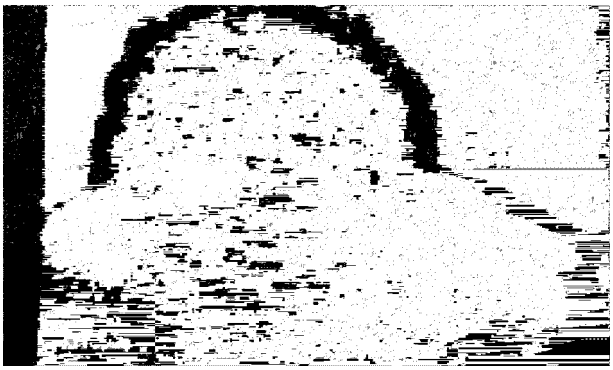


FIGURE 18.18 General surface corrosion on wing skin.

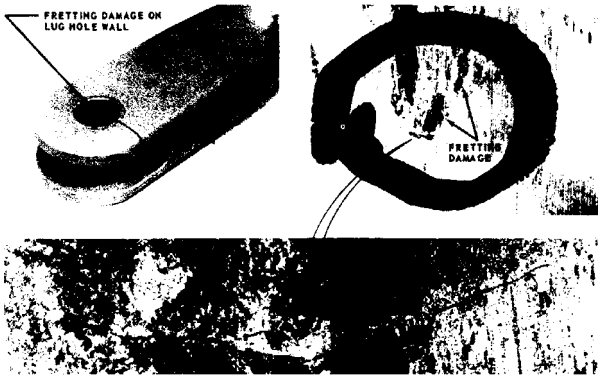


FIGURE 18.19 Fretting corrosion in the bore of a 7075-T6 aluminum lug.

18.20 shows a classic case of dissimilar corrosion where accelerated corrosion of the aluminum lug (anode) occurs due to contact with the Al-Ni-bronze bushing (cathode). In this case, the aluminum is corroding faster than it would if not contacted by the Al-Ni-bronze bushings.

As a result of corrosion on airplane structures, all aircraft must be inspected for signs of corrosion and condition of protective coatings during scheduled inspections. These must be carried out more frequently when the aircraft carries corrosive cargo and aircraft is operated under extreme conditions of humidity and temperature. Usually, inspections for corrosion and prevention are carried out with general inspections for other forms of damage, such as fatigue, and mandated by regulatory agencies.

FIGURE 18.20 Dissimilar metal corrosion between steel fasteners and aluminum stringer on airplane structure.

Since corrosion reduces the structural strength of aircraft, inspection methods and intervals should be designed to detect structural deterioration at early stages before the load-carrying capabilities of principal structures fall below the minimum regulatory requirements.

FURTHER READING

Aircraft Engineering and Aerospace Technology, vol. 68, no. 3 (1996).

Akdeniz, A., "Airplane Structural Maintenance and Corrosion Prevention" (1996).

Akdeniz, A. and Das, G., "Influence of Undetected Subsurface Corrosion on Structural Airworthiness of Aging Jet Transports," paper presented at NASA/FAA/DoD Aging Airplane conference (1997).

Boeing Commercial Airplane Group, Service Engineering, *1998 Boeing Structures Seminar Book.*

Hoeppner, D. W., "Corrosion of Aircraft," UCLA extension.

U.S. Department of Commerce, *AGARD Corrosion Handbook,* NTIS, Washington, DC.

PART 5

AEROSPACE PAINTS AND PROTECTIVE COATINGS

Lori Straus

Aircraft have long used organic and inorganic coatings as a means of protecting the structure. Even in the days of wood and fabric, coatings were used to provide moisture resistance, fungal resistance, and cosmetic appearance. Although the chemistry of the coatings has changed, as have the construction materials that they are protecting, organic and inorganic coatings still provide the majority of protection and appearance qualities on modern aircraft. Coatings used on aircraft fall into two major categories: organic and inorganic. Inorganic coatings or finishes are used primarily for corrosion protection and wear resistance. There are some limited applications for inorganic coatings as decorative finishes. Organic finishes or coatings are used for the purposes of corrosion protection, fluid resistance, wear protection, UV protection, and decoration.

18.39 INORGANIC COATINGS

Inorganic coatings and processes include anodizing, chemical conversion coatings, plating, and thermal spray coatings. In general, anodizing is used as a base for paint on aluminum, titanium, and magnesium. It is also used for some light-wear applications. Chemical conversion coatings are also used as a base for paint on aluminum and magnesium. Metallic plating is used to prevent wear, provide base metal corrosion protection of low-alloy steels, and provide dissimilar metal protection (galvanic coupling). Thermal spray coatings are primarily used to prevent wear but can also provide base metal corrosion protection of low-alloy steels.

18.40 ANODIZING

Anodizing is an electrolytic process by which a controlled-thickness oxide coating is produced. In this process the part to be anodized is the anode in an electrolytic cell and current is applied for a specified time to form the coating.

$$2\text{Al part} + 3\text{H}_2\text{O} \rightarrow \text{Al}_2\text{O}_3 \text{ anodize layer} + 6\text{H}^+ + 6\text{e}^-$$

Anodizing is most commonly used on aluminum, although it can be used on titanium and magnesium as well. This controlled oxide resists further corrosion unless the oxide layer is damaged, at which time it is possible for corrosion to occur. Historically, these coatings have been sealed in deionized water (maximum corrosion resistance, poor paint adhesion, no color change) or a dilute chromate solution (some corrosion resistance, good paint adhesion, yellow/green color change). The sealing process hydrates the porous anodize layer. The most common solution used for anodizing of aluminum sheet and wrought products for corrosion resistance and

paint adhesion is chromic acid. Chromic acid anodizing produces a thinner coating than sulfuric acid anodizing, resulting in a smaller fatigue effect. Sulfuric acid anodizing is commonly used on castings and parts which require a colored, or dyed, seal. These parts include fasteners and decorative trim. It should be noted that although castings may be anodized, it can be quite difficult due to the inherent porosity of the casting itself and the likelihood of solution entrapment. Phosphoric acid anodizing and unsealed chromic acid anodizing are used for parts which are to be structurally bonded (metalbond).

In an effort to reduce chrome waste streams associated with the anodizing process, many parts have been converted to nonchromate anodizing solutions such as boric acid-sulfuric acid in lieu of chromic acid anodizing. This change in chemistry has eliminated one source of Cr+6 in the manufacture of aircraft. Another change to the anodizing process has been to eliminate the postanodize sealing process and allow the parts to remain unsealed if they are to be subsequently painted. Although this reduces the inherent corrosion resistance of the anodize layer, maximum primer adhesion to the substrate is gained.

Anodizing can also be used as a means of creating a wear-resistant surface on most 2000, 6000, and 7000 series aluminum alloys. These thick sulfuric acid coatings provide good light wear and corrosion resistance. However, the coating is brittle and will dramatically reduce the fatigue life of the material. Hard anodizing is commonly used on actuators and other hydraulic or pneumatic parts.

During the anodizing process, the coating is formed from the base metal, and there is a net dimensional growth of approximately one-half the coating thickness. It may be necessary to take this growth into account when designing parts, especially parts which are to be hard anodized. It should also be noted that fatigue resistance decreases with an increase in anodize layer thickness. The following are common thickness ranges for the varying types of anodize:

Chromic acid or boric-sulfuric acid	0.00004–0.0001 in.
Sulfuric acid	0.0001–0.001 in.
Hard anodize	0.0018–0.0035 in.

18.41 CHEMICAL CONVERSION COATINGS

Chemical conversion coatings are applied as a nonelectrolytic immersion process during which an oxide, chromate, or phosphate film is formed.

$$2Al \text{ part} + H_2CrO_4 + 6HCl \rightarrow Al_2O_3 \text{ film} + 2CrCl_3 + 5H_2O$$

$$2Cr \text{ part} + 6HNO_3 \rightarrow Cr_2O_3 \text{ film} + 6NO_2 + 3H_2O$$

As with anodizing, chemical conversion coatings are used as a base for paint adhesion and corrosion protection. As a stand-alone finish, the level of corrosion protection is significantly reduced as compared to anodizing. Unlike an anodized surface, a chemically conversion coated surface conducts electricity, because chemical conversion coatings are many times thinner than anodize coatings. For aluminum parts which are required to be electrically conductive and will receive no subsequent finishes, the electrically conductive class of conversion coating is used to impart some degree of corrosion protection. Unlike anodize coatings, chemical conversion coatings do not cause an appreciable change in part dimension.

The generic class of chemical conversion coatings also includes phosphate coatings for steels, phosphate-fluoride coatings for titanium, and passivation of stainless steels. Both of the phosphate coatings act as bases for subsequent applications of paint or dry film lubricants. Passivation thickens the natural passive layer present on stainless steels and also removes any free iron that may have been deposited on the parts during machining operations. It is necessary to remove this free iron to prevent surface rust from occurring on stainless steels.

Because the chemical conversion process can be performed manually, it is common for aircraft operators to repair anodize coatings with conversion coatings.

18.42 PLATING

There are three main applications for plating on commercial aircraft: to provide base metal corrosion protection (sacrificial plating), to prevent galvanic corrosion, and to provide wear resistance. Electrolytic cadmium plating is commonly used on stainless and alloy steels, nickel alloys, and copper alloys to make them galvanically compatible with aluminum or other dissimilar metals or to act as a sacrificial coating. Chrome plating is used on the same alloys to provide a very hard, wear-resistant surface. Whereas in anodizing the part is the anode, during plating the part becomes the cathode.

$$Cd^{+2} + 2e^{-} \rightarrow \text{Cd plating on part}$$

In addition to providing galvanic protection, cadmium plating provides corrosion protection to alloy steels, which, if left unplated, are highly prone to oxidation. Cadmium plating almost always receives a postplate chromate treatment to impart it with improved corrosion resistance as well as to provide an excellent base for paint. Because of its excellent antigalling properties, it is commonly used on threaded parts. In the case of low-alloy steels which are heat treated above 220 ksi, it is necessary to plate parts in a low-hydrogen embrittling (LHE) cadmium bath to avoid hydrogen embrittlement of the base material. Because of its embrittling characteristics, cadmium plating should never be used on titanium. Like chrome, cadmium has been a target for elimination due to environmental concerns. In many applications, cadmium plating has been replaced by zinc-nickel plating or aluminum coatings such as ion vapor deposited (IVD) aluminum.

Chrome plating is used on parts which are subject to extreme wear, such as landing gear actuators and axles. Because of its internal plating stresses, chrome plate forms minute, irregular cracks. Improper plating or abusive postplate grinding can cause these cracks to link up and form a network of larger cracks that penetrate the plating layer to the substrate. In order to prevent crevice corrosion from occurring at the base of these through-cracks, it is common practice to apply a wipe-on/wipe-off layer of chromated epoxy primer. Grinding or polishing is almost always required after chrome plating to ensure a smooth surface.

Although not used as often in the manufacture of commercial aircraft parts, nickel plating is commonly used in the repair of previously chrome-plated parts. If parts become damaged or corroded during service, it is common to blend out the damage or corrosion mechanically, fill the blended-out cavity with nickel, and then chrome-plate over the top of the nickel. Nickel is very useful as a repair material because of its ability to be plated to much greater thicknesses than chrome. Brush, stylus, or contact plating is another method which is commonly used in maintenance

but not in the manufacture of aircraft. This method allows small localized areas to be plated without immersion of the entire part. Due to the lack of plating tanks at maintenance bases, airlines commonly use the brush plating method to repair small areas.

To avoid the risk of hydrogen embrittlement in susceptible alloys, it is necessary to bake the parts at an elevated temperature after plating. In some cases, alloys more susceptible to hydrogen embrittlement, such as high-strength alloy steels, require preplate thermal stress relief in addition to postplate thermal baking.

Plating to aluminum and titanium alloys, especially in important wear and other load-carrying applications, must be done with extra precautions because of the difficulty of achieving good adhesion to these alloys. These metals oxidize so rapidly that it is extremely difficult to provide an oxide-free surface upon which to plate. Because of this rapidly forming oxide, poor adhesion is often the result.

Another method of plating, electroless plating, does not require an external source of electricity. The metal is deposited by chemical reaction onto the substrate surface. This method is very sensitive to substrate surface cleanliness and deposition conditions and must be used with extra precautions when critical applications are involved.

Typical plating thicknesses are as follows:

Cadmium plating	0.0003–0.0005 in.
Zinc-nickel plating	0.0003–0.0005 in.
LHE cadmium plating	0.0005–0.0008 in.
Chrome plating	0.003–0.005 in.
Nickel plating	0.003–0.040 in.
Ion vapor deposited aluminum	0.001 in. minimum

18.43 THERMAL SPRAY COATINGS

Thermal spray coating involves heating a material, in powder or wire form, to a molten or semimolten state. The material is propelled using a stream of hot gas to deposit it, creating a surface coating on a given substrate. These molten particles hit the surface of the part with a great force that causes them to splatter onto the surface and on top of one another to form a coating. The coating material may consist of a single element but is often an alloy or composite. Tungsten carbides, aluminum, cobalt alloys, and copper-nickel-indium are common materials that are applied by thermal spray.

Thermal-sprayed coatings can be a cost-effective way to add superior performance qualities to a given substrate. Different types of coatings can be sprayed to obtain the optimum coating properties to resist wear, corrosion, and fatigue and increase hardness and lubricity.

The coating materials can be applied using several different processes. Thermal coating methods utilize fuel combustion, plasma spray, and electric arc delivery systems. Coatings can be applied under standard atmospheric conditions or in specialized, highly controlled atmospheric environments. Coatings can be applied manually or with automated robotics.

Many industries use thermal-sprayed coatings to extend product life, increase performance, and reduce production and maintenance costs. Thermal-sprayed coat-

ings can be a cost-effective means of protecting substrate surfaces from wear or corrosion. Other primary uses of thermally sprayed coatings include dimensional restoration, maintenance of precise clearances, and modification of thermal and electrical properties. Thermal spray coatings are also being utilized to replace chrome plating due to environmental and health hazards associated with the chrome-plating processes.

18.44 ORGANIC FINISHES

Organic finishes are used to provide environmental protection as well as for decorative appearance. Included in this category are paints—primers, topcoats, sealants, and corrosion-inhibiting compounds. Primers are used to provide corrosion protection and as a base for subsequent topcoatings and sealants. Topcoats are used to protect the underlying primer from UV degradation and for decorative purposes. Sealants are most commonly used to prevent moisture ingression into crevices and joints, to provide pressure tightness for the fuselage, to provide aerosmoothing for the exterior, and to provide fuel tightness for the fuel tanks. Corrosion-inhibiting compounds are applied as water-displacing materials.

18.45 PRIMERS

Corrosion-inhibiting primers are applied to provide corrosion resistance to the underlying base materials and effect adhesion of subsequent coatings. The majority of primers used on aircraft contain chromates as corrosion-inhibiting agents. The exception to this is those primers used on nonmetallic parts. Chemically the primers tend to be either epoxy-amines or urethane (modified) compatible epoxies. These chromated primers are formulated for general use, for fuel resistance, or for exterior aerodynamic applications. The primers used in fuel tanks are also formulated to resist microbial corrosion, while the primers used on the exterior of the airplane fuselage are formulated to resist filiform corrosion. The characteristic which sets aerospace primers apart from primers used in other industries is their requirement to be resistantant to hydraulic fluids, the most aggressive of which is a phosphate ester used on all commercial jet aircraft. In addition to hydraulic fluids, the coatings must be resistant to solvents, oils, fuel, lubricants, and maintenance chemicals. As with the inorganic coatings, there is environmental pressure to remove heavy metals, especially chromates, from the organic coatings. Corrosion-inhibiting chrome-free primers are being developed, but they have many hurdles to overcome before they will be deemed equivalent to the existing chromated primers.

18.46 TOPCOATS

Aerospace topcoats generally fall into two chemical families: epoxies and polyurethanes. In general, epoxies are used on the interior structure of the airplane and polyurethanes are used on the exterior areas of the airplane. On the interior of the airplane, topcoats are used for aesthetic purposes where structural components may

be visible to passengers, such as the area around entryways and hatches. On the exterior of the airplane, topcoats are used to provide resistance to fluids, both operational and maintenance, UV protection for primers and nonprimed composite substrates, and as the decorative paint scheme for the carrier (including maintenance, registry, and warning markings). For both primers and topcoats, it is necessary for the paint manufacturers to balance the mutually exclusive requirements of chemical resistance and flexibility.

18.47 SPECIALTY COATINGS

Many other coatings are used on the airplane which serve very specific purposes. Coatings used within the passenger cabin of the airplane have much more stringent antistain requirements than do the other decorative paints used on the aircraft. In addition, these interior coatings may have additional requirements that must be met concerning toxicity and smoke generation in the event of an on-board fire. Specialty coatings are used on the flight deck that have very low gloss to avoid the possibility of the pilot's sight being impaired by glare off of his or her instrument panel. On the exterior of the airplane are coatings designed to prevent static build-up on composite parts, provide resistance to rain erosion on composite parts, provide resistance to elevated temperatures, and provide an anti-skid emergency walkway on the upper surface of the wing.

18.48 SEALANTS

The majority of the sealants that are used on the aircraft are polysulfide materials. These sealants are used to seal crevices between parts to prevent moisture ingression and direct fluids to specific locations within the airplane where they can be drained overboard. Sealants are also used to create a pressure-tight vessel within the main cabin of the airplane and for aerosmoothness on the exterior of the airplane. These sealants may be chromated or nonchromated and may or may not contain microballoons to decrease their density. The second-most common use of sealants is within the fuel tank. The fuel tank's structural joints and fasteners are sealed to prevent fuel from moving from tank to tank or leaking out of the tanks. Earlier aircraft used rubber bladders to contain the fuel, and many military aircraft still use this method. The sealants used within the fuel tank are nonchromated and also may or may not contain microballoons to decrease density. High-temperature silicone sealants are used in areas which will see temperatures greater than 200°F. Food-grade silicone sealants are used in areas which may contact passenger foods or beverages, such as the galley areas.

18.49 CORROSION INHIBITING COMPOUNDS

Corrosion-inhibiting compounds (CICs) are petroleum-based compounds dispersed in a solvent. These compounds fall into two major categories: water displacing and barrier. Water-displacing materials are sprayed onto structure to penetrate faying

surfaces and wick into small crevices, thereby displacing any water which may be there and preventing any subsequent water from entering. The barrier-type materials are applied to greater film thickness than the water-displacing materials, providing additional corrosion protection to open surfaces yet still maintaining good penetrating properties. The barrier-type materials are applied in those areas where water is more likely to sit for long periods of time without draining away.

18.50 CHANGES DUE TO ENVIRONMENTAL REGULATIONS

Environmental regulations have had an increasing impact on aircraft finishes over the past 10 years. The first regulations that went into effect brought about the elimination or reduction of materials such as lead and asbestos. Limitations on chromium and cadmium have been increasing in recent years. In the solvents and organic coatings areas, broad restrictions have been placed on materials which contain smog formers and ozone depleters. Many of the solvents that were previously used or are used today fall into this category. Many surface-preparation materials—cleaners, degreasers—have been switched over to water-based materials rather than solvent-based materials. In addition, the smog-forming volatile organic compound (VOC) level has been severely limited in the primers and topcoats that we use today. There is no evidence that this trend will be reversed. Quite to the contrary, the restrictions are likely to become more severe. Many of the new ''environmentally friendly'' materials are much less forgiving when it comes to application characteristics, and care must be taken when using them to ensure equivalent performance as compared to their conventional counterparts.

18.51 MAINTENANCE

The most important factor when it comes to finishes and maintenance is that the majority of the finishes that are on the airplane are there to prevent corrosion. If the finishes are damaged or worn, they must be replaced to ensure the integrity of the aircraft. CICs are especially effective when reapplied periodically in service.

PART 6

AIRFRAME MAINTENANCE

Aydin Akdeniz

GLOSSARY

Accidental damage (AD): Physical deterioration of an item caused by contact or impact with an object or influence that is not a part of the aircraft or by human error during manufacturing, operation of the aircraft, or maintenance practices.

Airworthiness limitations: A section of the instruction for *Continued Airworthiness* that contains each mandatory replacement time, structural inspection interval, and related structural inspection task. This section may also be used to define a threshold for the fatigue-related inspections and the need to control corrosion to level 1 or better. The information contained in the airworthiness limitations section may be changed to reflect service and/or test experience or new analysis methods.

Corrosion prevention and control program (CPCP): A program of maintenance tasks implemented at a threshold designed to control aircraft structure to corrosion level 1 or better (corrosion within permissible limits of structural damage blendout).

Damage tolerant: A qualification standard for aircraft structure. An item is considered damage-tolerant if it can sustain damage and the remaining structure can withstand reasonable loads without structural failure or excessive structural deformation until the damage is detected.

Environmental deterioration (ED): Physical deterioration of an item's strength or its resistance to failure as a result of a chemical interaction with the climate or environment.

Repeat Interval: The interval, expressed in flight cycles, flight hours, and/or calendar time, between successive accomplishments of a specific maintenance task.

Scheduled maintenance check: Any maintenance opportunity that is prepackaged and accomplished on a regular basis.

Principal structural element (PSE): Any detail, element, or assembly, which contributes significantly to carrying flight, ground, pressure, or control loads and whose failure could affect the structural integrity necessary for the safety of the aircraft.

Structural assembly: One or more structural elements which together provide a basic structural function.

Structural detail: The lowest functional level in an aircraft structure. A discrete region or area of a structural element, or a boundary intersection of two or more elements.

Structural element: Two or more structural details which together form an identified manufacturer's assembly part.

Structural Function: The mode of action of aircraft structure. It includes acceptance and transfer of specified loads in items (details/elements/assemblies) and provides consistently adequate aircraft response and flight characteristics.

Tasks—Maintenance: A set of actions required to achieve a desired outcome which restores an item to, or maintains an item in, serviceable condition, including inspection and deterioration of condition.

Inspection—detailed: An intensive visual examination of a specific structural area, system, installation, or assembly to detect damage, failure, or irregularity. Available lighting is normally supplemented with a direct source of good lighting at an intensity deemed appropriate by the inspector. Inspection aids such as mirrors, magnifying lenses, etc. may be used. Surface cleaning and elaborate access procedures may be required.

Inspection—general visual (surveillance): A visual examination of an interior or exterior area, installation, or assembly to detect obvious damage, failure, or irregularity. This level of inspection is made under normally available lighting conditions such as daylight, hangar lighting, flashlight, or droplight and may require removal or opening of access panels or doors. Stands, ladders, or platforms may be required to gain proximity to the area being checked.

Inspection—special detailed: An intensive examination of a specific item(s), installation, or assembly to detect damage, failure or irregularity. The examination is likely to make extensive use of specialized inspection techniques and/or equipment. Intricate cleaning and substantial access or disassembly procedure may be required.

Design service objective (DSO): A concept established to minimize structural maintenance cost. Minimum period of service during which the primary structure of the airplane is designed to be essentially free of detectable fatigue cracks with a high degree of reliability and confidence.

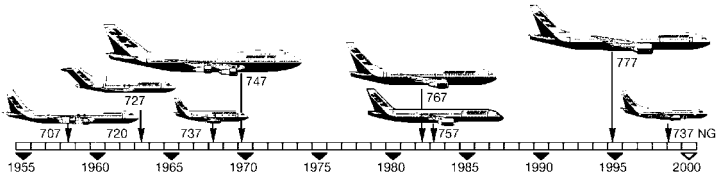
18.52 INTRODUCTION

Jet transport aircraft have demonstrated many years of safe and reliable service. Airline and manufacturer experience in developing scheduled maintenance requirements has shown that the use of a logical decision process produces more efficient programs. The primary objective of a structural inspection program is to maintain airworthiness throughout the life of a given aircraft fleet in an economical manner. The structural maintenance program is based on an assessment of structural design information, fatigue and damage tolerance evaluations, and service experience with similar structure and pertinent test results.

Economic and market conditions have resulted in the use of commercial jet transport aircraft well beyond their original economic design service objective (DSO), as shown in Figure 18.21. Increase in the active service life of the commercial fleet is due to low fuel costs and increasing fleet-replacement costs. Supplemental inspections, based on the damage-tolerant philosophy, provide continuing airworthiness until the economics of repair and modifications dictate retirement of an aircraft from service. Safety is not compromised if the mandated safety inspections and scheduled aircraft maintenance procedures are carried out. The air transportation industry consensus is that older jet transport aircraft will continue to be in service despite the anticipated substantial increases in the required structural repair and modification.

18.53 MAINTAINING STRUCTURAL SAFETY

Criteria and procedures employed in commercial aircraft design for the past three decades have resulted in long-life, damage-tolerant structures with a credible safety



Model	Number of airplanes		Minimum design service objectives*		High-time airplanes**	
	Total	Active	Flights	Hours	Flights	Hours
707	734	58	20,000	60,000	36,400	91,900
720	153	0	30,000	60,000	45,000	69,300
727	1,821	1,179	60,000	50,000	77,500	80,300
737	2,941	2,691	75,000	51,000	92,800	84,800
747	1,125	942	20,000+	60,000	33,900	106,400
757	778	766	50,000+	50,000	27,300	52,900
767	677	667	50,000+	50,000	33,000	61,700
777	104	104	44,000	60,000	4,800	10,600

* At this point in the airplane lifetime, some initial fatigue might occur.
**Different airplanes are in high flights and high hours.
+ Some derivatives have special design objectives.

FIGURE 18.21 Basic commercial jet fleet summary (January 1998) for Boeing Heritage Airplanes.

record, as shown in Figure 18.22. This achievement is the result of diligent attention-to-detail design, manufacturing, maintenance, and inspection procedures.The design concepts, supported by testing, have worked well due to the system that is used to ensure the commercial fleets are kept flying safely throughout their operational life. The major participants in this system are:

- The airworthiness authorities, who establish rules and regulations, approve the designs, and oversee the airline maintenance performance
- The manufacturers, who design, build, and support the aircraft in service
- The airline companies, who operate, inspect, and maintain the fleet

Continuing aircraft safety depends on diligent performance of and participation by all the participants in the system.

Commercial jet aircraft structures in the 1950s and 1960s were designed using the fail-safe design concept. This concept ensures the ability of a critical component

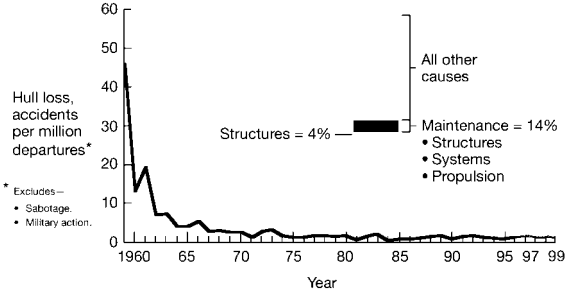


FIGURE 18.22 Safety record—worldwide commercial jet fleet (year-end 1999).

to maintain design capability and functionality with the complete failure of a single element or an obvious partial failure of a single element. Experience has shown that the fail-safe design philosophy has generally been effective in allowing sufficient opportunities for timely detection of structural damage. However, failure modes were not always predicted with sufficient accuracy to ensure that structural failures would be obvious and within safety limits. Structural failures could progress in unanticipated ways, and a number of older aircraft were found to have unexpected defects.

By the early 1970s, inspection planning was recognized as a safety issue. It was clear that the airline operators were expected to find structural cracks that were often difficult to detect. This prompted the industry and airworthiness authorities to focus more on the adequacy of inspection programs for the timely detection of structural discrepancies before they become critical. Combined industry and airworthiness authority activities in the 1970s resulted in changes to the regulatory requirements. These requirements reflected state-of-the-art developments in fracture mechanics and embraced damage tolerance as a design philosophy. In addition to residual strength evaluations, damage growth and inspection requirements with consideration for damage at multiple sites were incorporated in 1978 in Federal Aviation Regulation, (FAR)/AC 25.571, Amendment 45 for new aircraft, and in Civil Aviation Authority, (CAA) Notice 89 and AC 91-56 for development of Supplemental Structural Inspection Program (SSIP) for aircraft certified prior to 1978.

18.54 MAINTAINING AIRCRAFT CERTIFIED DAMAGE TOLERANT USING MSG-3 PROCESS

For aircraft models certified after 1978, maintenance for damage tolerance was developed using the Maintenance Steering Group—3 (MSG-3) process. The MSG-3 guidelines, written by the Air Transport Association (ATA), provide procedures for developing initial maintenance programs for three principal sources of aircraft structural damage: fatigue, corrosion, and accidental damage, as shown in Figure 18.23.

The MSG-3 guidelines provide information for a structural rating system development to be used by the industry as a means of determining inspection require-

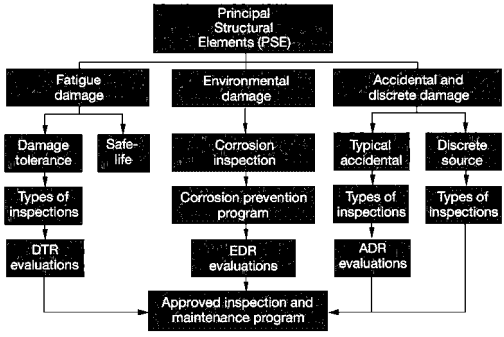


FIGURE 18.23 MSG-3 process for developing maintenance for principal damage sources.

ments. The rating systems are based on past practices and manufacturer/operator experiences with a similar structure. The rating systems used by manufacturers are accidental damage rating (ADR), environmental deterioration rating (EDR) for corrosion, and damage tolerance rating (DTR) for fatigue-damage detection assessment. The assessment of structure for selecting a maintenance task also includes the following:

1. The susceptibility of the structure to each source of deterioration
2. The consequences of structural deterioration to ensure continued airworthiness:
   - Effect on aircraft, e.g., loss of function or reduction of residual strength
   - Multiple-site or multiple-element fatigue damage
   - The effect on aircraft flight or response characteristics caused by the interaction of structural damage for systems or powerplant items failure
   - In-flight loss of structural items
3. The applicability and effectiveness of various methods to prevent, control, or detect structural deterioration, taking into account inspection thresholds and repeat intervals

The initial structural maintenance plan for a new model is directed toward detecting corrosion and accidental damage. Based on manufacturer and operator experience, ADR and EDR for each principal structures element (PSE) are established and initial inspections developed. Simultaneously, a feasibility study for fatigue-damage detection is made to determine if a practical program can be constructed using the DTR system.

As the fleet matures, the risk of fatigue increases and the initial inspection program is reassessed using the DTR system. If the initial program is inadequate for finding fatigue cracks in significant structures, additional or supplemental inspections are required and provided by a supplemental structural inspection program.

18.55 STRUCTURAL MAINTENANCE PROGRAM DEVELOPMENT

As part of the structural maintenance program development, it is necessary to rate each principal structural element. In terms of susceptibility (likelihood of damage) and detectability (timely detection of damage). The rating system accounts for the susceptibility of the PSE to the likely source of damage and type of deterioration of the PSE to the damage source. Differences between metallic and nonmetallic portions of the PSE must be taken into account.

The rating systems are compatible and allow comparative assessments for each PSE group. Emphasis is placed on rating each PSE in relation to other PSEs in the same inspection area, leading to increased emphasis for the foremost PSE. Manufacturer and operator experience is a key ingredient for these evaluations.

Rating systems for fatigue damage should incorporate results from the manufacturer's residual strength and crack growth evaluations. The applicability and effectiveness of various inspection methods, detectable damage sizes, and access requirements are key ingredients for these evaluations.

18.56 RATING ENVIRONMENTAL DETERIORATION (EDR)

Environmental deterioration rating systems allow for evaluations of susceptibility to and timely detection of corrosion and stress corrosion. Susceptibility to corrosion is assessed on the basis of probable exposure to an adverse environment and adequacy of the protective system. For example:

1. Exposure to a deteriorating environment such as cabin condensation, galley spillage, toilet spillage, cleaning fluids, etc.
2. Contact between dissimilar materials (potential for galvanic activity)
3. Breakdown of surface-protection systems (e.g., deterioration of paint, primer, bonding, sealant, corrosion-inhibiting compounds, and cladding systems), with the resulting corrosion of metallic materials or fluid incursion into permeable, nonmetallic materials, etc.

Material characteristics, coupled with the likelihood of sustained tensile stress, are used to assess susceptibility to stress corrosion. Timely detection is determined by sensitivity to relative size of damage and visibility of the PSE for inspection. Rating system evaluations are made taking into account the requirement for each operator to control the aircraft structure at corrosion level 1 or better, as shown in Figure 18.24.

18.57 RATING FATIGUE DAMAGE (FD)

The rating system must lead to an inspection program that provides a high probability of detecting fatigue damage in the fleet before such damage reduces any aircraft's residual strength below permissible levels. To achieve this, the rating system considers the following:

1. Residual strength, including the effects of multiple-site fatigue damage, where appropriate (critical).
2. Crack growth rate, including effects of multiple-site or multiple-element fatigue damage, where appropriate.

Corrosion Severity Relative to SRM Allowable Blend-out Limit		Corrosion Level for Given Extent and Number of inspection Intervals to Reach Given Severity			
		LOCAL		WIDESPREAD	
		Single Interval	Multiple Interval	Single Interval	Multiple Interval
Well Below		LEVEL1	LEVEL1	LEVEL1	LEVEL1
Approaching		LEVEL1	LEVEL1	LEVEL2	LEVEL1
At Limit		LEVEL1	LEVEL1	LEVEL2	LEVEL1
Above Limit	(1)	LEVEL2	LEVEL1*	LEVEL2	LEVEL1*
	(2)	LEVEL3	LEVEL1**	LEVEL3	LEVEL1**

(1) Corrosion is not an urgent airworthiness concern.

(2) Corrosion is an airworthiness concern.

* Operator experience over several years has demonstrated only light corrosion between successive inspections, but latest inspection and cumulative blend-out exceeds manufacturer's allowable limit.

** Highly unlikely event following multiple applications of preventive measures and corrosion-inhibiting compounds.

SRM Manufacturer's approved Structural Repair Manual.

FIGURE 18.24 Effects of corrosion extent and severity on corrosion level.

3. Damage-detection period, which corresponds to the interval for the fatigue damage to grow from the threshold of detection (detectable) to the limiting size defined by 1 above. This period varies according to the inspection method used and may be influenced by structural parts of processes, e.g., sealant obscuring part of the damage.
4. Detection standards for applicable inspection methods. Estimated detectable crack lengths can be used for the fatigue damage. Detection evaluations are required as part of aircraft type certification.
5. Applicable inspection levels and methods (e.g., visual and NDI), directions (e.g., external and internal), and repeat intervals

18.58 IMPACT OF COMBINED FATIGUE AND CORROSION DAMAGE

A combination of severe corrosion and fatigue damage can significantly affect damage-tolerance capability by reducing residual strength, increasing multiple-site damage, accelerating crack growth rates, and assumes that the structure has an increased initial damage size. Distribution of corrosion is unpredictable in the fleet and significantly reduces the safe damage-detection period for fatigue damage. An example illustrating the potential effect of fatigue damage occurring in the presence of severe corrosion is shown in Figure 18.25. Airplane structural integrity can be significantly affected by the reduced number of opportunities to detect damage before it exceeds safe allowable limits.

Figure 18.26 shows a typical fleet damage rate assumed for aging fleet assessments in the early 1980s when SSIPs for fatigue cracking were developed. The fleet damage rate for corrosion was assumed to be fairly constant. However, surveys and events in the late 1980s have indicated an increasing number of corrosion problems in aging aircraft.

Figure 18.27 shows an observed fleet damage rate during the same period as Figure 18.26. Figure 18.27 indicates a significant increase in the fleet damage rate

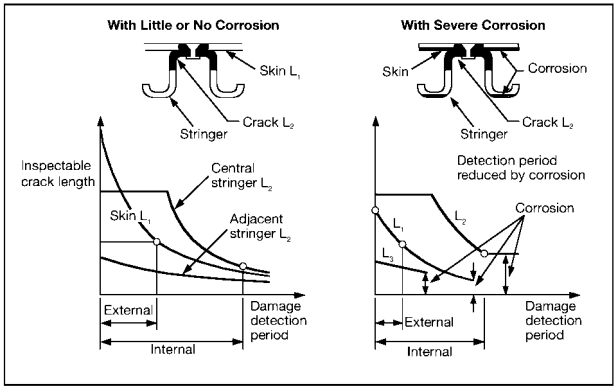


FIGURE 18.25 Reduction in damage detection period due to severe corrosion.

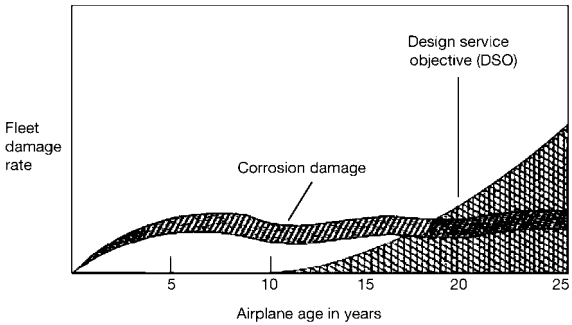


FIGURE 18.26 Assumed fleet damage rate.

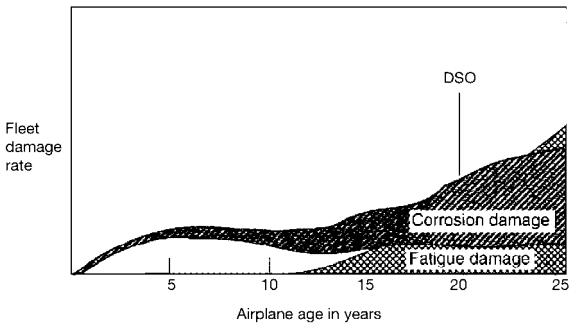


FIGURE 18.27 Observed fleet damage rate.

due to corrosion. Some operators had taken effective action to prevent or control corrosion, but others had allowed corrosion to progress to levels that required major repairs or part replacement. Potential combinations of severe corrosion and fatigue cracking required costly, frequent inspections of all PSEs to assure continuing airworthiness. The way to avoid such impractical restrictions is to prevent or control corrosion, independent of fatigue, prior to any potential interaction with structural fatigue damage.

18.59 RATING ACCIDENTAL DAMAGE (ADR)

Accidental damage rating systems include evaluations of the following:

1. Susceptibility to minor (not obvious) accidental damage based on frequency of exposure to and the location of damage from one or more sources, including:
   - Ground handling equipment
   - Cargo handling equipment
   - Sources resulting from human error during manufacture, maintenance, and/or operation of the aircraft, not included in other damage sources

- Rain, hail, etc.
- Runway debris
- Lightning strike
- Water entrapment

2. Residual strength after accidental damage, normally based on the likely size of damage relative to the critical damage size for the PSE.
3. Timely detection of damage, based on the relative rate of growth after damage is sustained and visibility of the PSE for inspection. Assessments should take into account damage growth associated with nonchemical interaction with an environment, such as disbond or delamination growth associated with a freeze/thaw cycle.

Rating values are assigned to groups of PSEs in the same inspection area on the basis of comparative group assessments.

18.60 USE OF EDR/ADR SYSTEMS

The frequency of the structural inspection for a given area of the aircraft is determined based on EDR and ADR cumulative rating, corrosion-prevention requirements, and operator experience with the existing structural inspection programs for each model.

As shown in Table 18.7, a relationship between EDR/ADR and the inspection interval table is created by the program participants which allows the creation of accurate and realistic inspection frequencies based on fleet experience and design improvements.

Based on the Table 18.7 rating process and Table 18.8 corresponding inspection check frequencies, the resulting structural inspection and structural maintenance requirements for all airplane structures are submitted to the steering committee (manufacturer, operators, regulatory agency) for approval and inclusion in the Maintenance Review Board (MRB) report proposal. Once the MRB is finalized, it is submitted to the FAA for approval.

TABLE 18.7 Relationship between EDR/ADR and Inspection Intervals

EDR/ADR cumulative rating	Ground rules for establishing structural inspection intervals
1	Lowest practical inspection interval for a few critical SSIs in the considered zone
2 3 4	Service-proven acceptable intervals for external or internal surveillance of the considered zone (baseine interval).
5 6	Interval longer than baseline intervals may be considered by the SWG on a case-by-case basis.
7 8	Maximum interval considered adequate for detecting damage from unforeseen causes. Intervals are based on economic considerations for accessibility and repair.

TABLE 18.8 737-300 Fuselage Inspection Intervals Corresponding to Cumulative EDR/ADR

EDR and ADR rating	External inspection [1]	Internal Inspection [2]
1	2A	A
2	5A	5A
3	C	C
4	2C	2C
5	2C	4C
6 or greater	2C	Age exploration 4C interval 1.5 or 1/10 of fleet

A-check frequency: 300 flight cycles
C-check frequency: 1 months or 3,000 flight cycles (whichever comes first) [3]

[1] In addition, based on past practice, critical structure will also be inspected at A-check intervals.

[2] Except, based on past practice for similar structure, age exploration at 4C intervals for 1/5 of the fleet may be utilized.

[3] One flight cycle is equivalent to one takeoff and landing.

FURTHER READING

Akdeniz, A., ''Corrosion Prevention and Control Programs,'' *Airliner Magazine* (March 1996).

Akdeniz, A. and Das, G., ''Influence of Undetected Subsurface Corrosion on Structural Airworthiness of Aging Jet Transports,'' paper presented at 2d NASA/FAA/DoD Aging Airplane conference (1998).

Federal Aviation Administration (FAA), *Maintenance Review board (MRB) Procedures,* FAA Advisory Circular No. 121-22A.

Gornson, U. G.., ''Damage Tolerance—Facts and Fiction,'' paper presented at 17th Symposium of International Committee on Aeronautical Fatigue, Sweden (June 1993).

Maintenance Steering Group—3 Task Force, Air Transport Associaiton of America, *Airline/Manufacturer Maintenance Program Development Document* (*MSG-3*) *Rev. 2.*

PART 7

ENGINE MAINTENANCE

Navaal Ramdin

The majority of commercial airplane engines are dependent on an on-condition maintenance program. Engines are removed off-wing if they fail, exceed inspection limits, engine parameters degrade to limit performance, or the life-limited items become due for replacement. In conjunction with other analytic tools, two types of maintenance actions can provide the necessary details to assess the engine condition. The first consists of the routine scheduled maintenance inspections of the various engine components, assemblies, and systems. These actions are carried out with the engine on-wing and in accordance with current inspection criteria. The purpose of this activity is to ensure continued airworthiness. These periodic inspections are usually based on engine cycles or operating hours. In addition, a number of analytic tools and techniques can be used to assess the health of an engine. These include engine conditioning monitoring, spectrographic oil analysis process (SOAP), and borescope inspections in combination with other types of inspection applications. The second classification addresses damaged or malfunctioning parts which are fixed or replaced as they occur. All maintenance actions fall under the following two categories:

Line maintenance includes tasks that can normally be performed with the engine installed on the airplane. Line maintenance is further defined under two subcategories: scheduled and unscheduled maintenance. Scheduled maintenance tasks are defined in an airline's operational authorization under part D of the airline's operations specification. Typical engine scheduled maintenance tasks include periodic visual and detailed inspections, borescope inspections, oil and fuel filter inspections and replacement, and igniter and magnetic chip inspections. These types of inspections are usually done on-wing. Unscheduled maintenance activity covers items that may be related to component failures, as well as events such as bird ingestion, lightning strikes, and heavy landings.

Heavy maintenance includes those tasks that can be accomplished on an uninstalled engine within a shop environment. It also includes replacement of bearings and other significant parts which are accessible without complete disassembly of the modular sections and repairs. Heavy maintenance also includes the complete disassembly, replacement of parts, rebuilding of parts, and reassembly of each module of the engine. In addition, the engine contains a number of accessories that are also checked at shop visit. Testing of engines after overhaul ensures that the engine meets required certification standards. Maintenance and repair of the engine will often employ a variety of processes that are required to ensure the engine operates within operating limits. These processes include electron-beam and steady state plasma arc welding, diffusion bonding, inertia bonding, superalloy casting, chemical milling laser drilling, broaching, and plasma spraying. Following are some of the types of tasks that are performed during shop visit:

- Visual and detailed inspections
- Cleaning (chemical and abrasive)
- Crack detection and dimensional inspection
- Static and dynamic balancing

- Parts machining for fit
- Welding and bonding
- Pressure checking of bearing compartments

A full discussion of the different types of inspection techniques that are carried out on turbine engines is beyond the scope of this handbook. The inspections outlined in the remainder of this part provide some examples of the analytic and inspection techniques frequently used to assess repair and overhaul of engines.

18.61 ENGINE CONDITION MONITORING (ECM)

The key driver in establishing an engine conditioning monitoring program is being able to detect trends, variations, or fluctuations in critical engine parameters before they cause problems. These changes can affect the safe, continued operation of the engine and airplane. Specific readings that indicate engine performance on-wing and on the ground should be recorded and monitored since they are helpful in detecting engine performance degradation patterns that can affect engine operation. These readings can be used to provide early indication of engine degradation and in some instances may provide an indication on the specific module where degradation may be occurring. This is particularly helpful from an operator perspective since it allows the engine to remain on-wing as long as economically possible without compromising safety and continued airworthiness of the airplane. Impending problems can be detected, observed, and economically corrected. However, this means that engine displays and instruments must be closely monitored and properly interpreted. This type of information is useful in reducing engine-related delays, cancellations, and in-flight shutdowns. Strict surveillance of instrument readings for the purpose of maintenance is called *engine condition monitoring* (ECM). ECM can provide real-time data on critical engine operating parameters.

Instrument readings through flight deck indication, airplane on-board maintenance systems, and data from electronic engine control (EEC) or full authority digital electronic control (FADEC) boxes can be used to make a trend analysis for individual engines. Comparing on-wing engine performance data with data for a new, or newly overhauled, engine provides an indication of variation or changes that may be occurring on the engine when compared to known standards. In order that a comparison be made to engine test cell data, the on-wing data must be converted to standard atmospheric conditions (at sea level) and then compared to statistically compiled data for new or newly overhauled engines. This is just one method of comparing on-wing data to data collected for new engines through other means such as flight test. These comparison methods are dependent on the specific engine manufacturer. In addition to engine indication parameters, ground indicators also provide data on internal degradation. Inspection and servicing tasks such as oil-consumption level checks, borescope, oil filter, fuel filter, and magnetic chip detector inspections may also provide crucial information in confirming ECM trends. These types of inspections may be used to assess degradation levels internal to the engine and enable flight and maintenance crews to detect engine performance deterioration. Causes could include contamination of oil or fuel, engine accessory degradation, module degradation, deteriorated or damaged compressor blades or vanes, or some other unsatisfactory condition within the engine.

For ECM and engine condition trend monitoring, engine instruments may be divided into two categories: engine operating limits and mechanical condition. The instruments used in assessing engine operating limits or parameters can provide information about engine condition and performance. The performance parameter instrumentation includes readings such as internal engine pressures, engine pressure ratio (EPR) (also known as turbine discharge pressure), fuel flow, exhaust gas temperature (EGT), and rpm. In dual axial-flow compressor engines, both the low- and high-pressure compressor rpm (N1 and N2) are usually instrumented. Instrumentation to monitor the parameters of mechanical condition, such as oil temperature and pressure, compressor rpm, and engine vibration, make up the second group. There are operating limits for these parameters. Operation of the engine above these limits in some instances can result in engine failure. As a rule, these limits should not be exceeded. Both groups of instruments can be used to check and monitor the mechanical condition of the engine and can provide indication of engine degradation or impending failure. Operation outside of specific engine limits will result in indication through aural or visual messages in the flight deck and would be evident to the flight crew. These types of problems would have to be rectified on landing and could result in delays or cancellation of flights. Early indication of these types of problems can be identified through ECM.

18.62 ENGINE CONDITION MONITORING (ECM) PROGRAM

On modern airplanes, engine performance data are automatically recorded by airplane computer systems. FADEC or EEC boxes can also be downloaded to retrieve engine-specific data from flights. With the development of wireless technology, pertinent engine parameter information can now be sent directly to locations that collect, observe, and interpret the data as the airplane is in flight. This type of data is useful in the development of trend analysis models for an engine. This is done by application of an ECM computer program to the data collected. This information is best handled by utilizing programs that can convert and normalize the engine performance data to sea-level conditions, making it comparable to engine test cell data. A graphical chart can be produced that shows the engine data for a particular flight segment or period of time. This is particularly useful in looking at long-term trends such as increases in thrust specific fuel consumption and engine gas temperature, which are critical factors in determining engine removal. Analysis of the ECM trend plots by a performance engineer can help operators identify performance degradation trends or malfunctions in the installed engine and engine components before failures occur. Indication of system problems or trends helps the operator assess performance retention as a function of time. Use of ECM in this fashion provides a competitive cost advantage to operators by improving engine reliability and scheduling maintenance when it is convenient to the operator. This is of great economic benefit to operators since it significantly helps to reduce delays, cancellations, and in-flight shutdowns, thereby improving airplane dispatch reliability.

18.63 METHODS OF INSPECTION

Inspecting the engine for deterioration or damage is vital in maintaining engine airworthiness. Many of the inspection techniques are nondestructive because of the

high replacement cost of turbine parts. Various techniques, such as borescope inspections, oil analysis, and chip detector inspections, are easily accomplished on-wing. In addition to these inspections, a number of other inspections involving more qualitative analysis techniques are done at shop visit. Inspection of individual disassembled modules including parts in engine overhaul shop requires sophisticated means of detecting degradation. The following techniques cover a sample of the different types of methods that are currently used to assess the condition of modern engines, during both line and heavy maintenance:

- Borescope
- Spectrographic oil analysis program
- Ferrography
- Fluorescent inspection
- Dye penetrant inspection
- Magnetic testing
- Eddy current inspection
- Radiography
- Ultrasonic inspection
- Electron beam welding
- Steady state plasma arc welding
- Diffusion bonding
- Inertia bonding

Borescope Inspection

Internal inspection of an engine to determine condition is of great importance in assessing the condition of engine components such as blades and vanes. Borescopes are available in rigid or flexible types (fiberoptic). The rigid borescope is limited in terms of the viewing area of the engine. It is useful in inspecting areas that are clearly visible within a specific location, but it does not work well in areas that require different views, such as the front or back of a blade or vane, because of the difficulty associated with moving the probe into ideal viewing locations. The fiberoptic types utilize flexible glass fiber to aid the visual internal inspection of engine sections without the need for major disassembly of the engine or removal of the engine from the airplane. This type of probe used on the borescope can be used to bend around corners to provide better viewing angles internal to the engine with the engine on-wing. Borescopes are also used to inspect engines or modules that have been removed from the wing during shop visits. The borescope consists of an eyepiece connected to either a long, rigid tube or a flexible tube. The flexible tube contains a fiberoptic cable with a light source that carries light to the area being observed and provides visual images. This is particularly useful in areas where viewing access requires going around corners. Borescopes employ a light source and lenses. The mechanic inserts the borescope through openings in the engine for inspection. The openings may be the entrance to the compressor or the turbine, or may be special ports that are designed particularly for borescope inspection. They are located along the engine core to provide access to the major hot and cold sections. The flexible scope allows the mechanic a wider view area of the

location to be inspected by making adjustments to the borescope through controls. Attaching a video recorder monitor to the eyepiece of the borescope can provide an enhanced view of the area being inspected. Use of a monitor can greatly improve the resolution of the inspection area, aid in detecting irregularities, and provide visual documentation as to engine condition.

Metallic Particle Detection in Engine Oil

Contamination is the major cause of bearing and other rotating surface damage or degradation. Detection of wear patterns and establishing degradation trends can be accomplished to predict when critical rotating components may need replacement. Lubrication oil systems have incorporated into them filters and magnetic chip detectors that collect larger engine-related particles in the oil system. Wear from surfaces such as bearings and shafts will produce particles that oil filters and chip detectors may not capture. Accumulation of particles of this size also provides the contamination that leads to degradation. Metal analysis is helpful in predicting failures, but a series of complimentary tests may be necessary to confirm degradation trends. Many techniques are employed by the industry to analyze contamination in engine oils. Some of the most commonly used analysis techniques on commercial turbines include spectrography and ferrographic analysis.

Spectrographic Oil Analysis (SOAP). The basic principle behind SOAP is the relationship between the rate of wear and contamination of the oil. The greater the contamination of the oil, the greater the wear on critical rotating surfaces. Spectrography is helpful in identifying areas of wear and predicting wear patterns, which is helpful in determining the on-wing life of components. Spectrography involves using high energy to excite the metallic atoms in the oil sample. This high energy energizes the metal atoms and causes the excitation and liberation of electrons. When the electrons return to a less excited state, they produce light representative of the different metallic elements that are in suspension in the oil sample. Photo detectors in the spectrograph detect the light. The process is capable of measuring quantities in the range of parts per million. The process is usually able to detect abnormal increases in metal concentrations that exist in the oil. These data, when plotted over time, can provide useful information regarding engine wear. Spectrographic analysis is most useful in detecting particles below 15 μm.

Ferrography. Ferrography is a laboratory technique used to separate ferromagnetic particles on an inclined glass plate by means of a powerful magnetic field. The particles are then viewed under a microscope. They can be analyzed to determine the type of wear patterns encountered and probable areas that may be affected. This technique is helpful in the area of bearing condition monitoring since it is able to identify bearing fragments. Its use is limited to laboratory analysis because of the lack of portability. Analytical ferrography is not a quantitative technique and does not assess the presence of nonmagnetic particles, such as aluminum, brass, copper, and chromium. As a result, it is of limited use in assessing engine metallic particles.

18.64 CLEANING

The engine is disassembled either horizontally or vertically into main subassemblies or modules and then further disassembled into individual parts. Before inspection,

these parts are subjected to cleaning agents such as organic solvents, acids, and other chemical cleaners, including electrolytic cleaning solutions. Rotating parts such as disks are exposed to electrolytic cleaning to aid in inspection of cracks or surface defects.

18.65 CRACK DETECTION

Detection of cracks that are not visible to the naked eye is extremely important in static or rotating engine components. Cracks or anomalies that reside in rotating components could result in catastrophic failures, since these parts are subjected to very large centripetal forces and usually operate in high-temperature and pressure environments. Rotating components such as turbine disks and shafts are difficult to inspect visually for small cracks; these cracks can propagate as the engine accumulates cycles and can lead to component failure and in some instances loss of the engine. Small cracks are usually detected using many different methods. The following inspection techniques are just some of the more commonly used methods; this is not an all-inclusive list since technology changes are constantly improving detection methodology.

Fluorescent Inspection

Fluorescent testing can be defined as a procedure to detect surface related discontinuities in nonporous engineering materials. It is capable of identifying cracks that are undetectable through normal visual inspections. Fluorescence is based on the principle that whenever ultraviolet light falls on the fluorescent chemical the light is absorbed and is reemitted as visible light. The fluorescent liquid used should easily penetrate the cracks. When a fluorescent chemical is applied to a part, it will penetrate the surface of the part if cracks exist. The excess fluorescent compound is then removed. The part is then exposed to ultraviolet light, which illuminates the areas into which the chemical has been drawn and exposes the surface flaws.

Dye Penetrant Inspection

This type of inspection penetrates cracks or pores that exist in the metal part surface. A penetrating dye is induced to enter cracks or pores in the part surface. When this is completed, the part surface is then washed and a developer fluid is applied. The developer fluid contains a white absorbent. Capillary action draws the developer to the crack or pore surface and the stain then indicates the crack location.

Magnetic Testing

Magnetic particle testing is a nondestructive method of testing for surface defects in ferromagnetic materials. Magnetic particle inspections are used to reveal the existence and extent of cracks and some types of flaws in the surface of a part. This type of inspection is applicable only to parts that can be magnetized (metals or alloys containing iron). The part is first magnetized to create a suitable magnetic flux within the part. Magnetic particles in dry powder can then be applied to the

part, or the part can be immersed in a fluid containing magnetic or ferrous particles. Cracks or flaws in the part will create discontinuities in the magnetic field, which polarize the field at the cracks and draw the magnetic particles around the cracks. In some instances the immersion fluid may contain fluorescent particles that accumulate and indicate anomalies in the metal surface. The use of ultraviolet light makes the crack visible. Parts that have been magnetized for crack testing must be demagnetized after inspection.

Eddy Current inspection

Eddy current inspection is well suited for detecting cracks caused by stress or fatigue. When an AC current flows in a coil in close proximity to a conducting surface, the magnetic field of the coil will induce circulating (eddy) currents in that metal surface. The induced eddy currents in the metal surface run in concentric circles. A crack in the surface of the metal will interrupt or reduce the eddy current flow, thus decreasing the loading on the coil and increasing its effective impedance. This is the basis of eddy current testing. By monitoring the voltage across the coil in such an arrangement, we can detect changes in the material of interest. This type of testing works well when the cracks interrupt the surface eddy currents. If, however, the cracks lie parallel to the eddy current path, these cracks may not be detected. The eddy currents generated in the region of a defect differ from the currents in an unflawed area. The magnitude of the difference is used to determine the size of surface flaws. This method works best on surfaces that have relatively smooth surface finishes.

Radiography

Radiography is a useful nondestructive method to detect internal flaws. It can be used on a wide range of materials, metallic and nonmetallic. In essence, a radiograph is the passage of electromagnetic radiation through an object onto a film. X-rays and gamma rays are used as the source to provide the electromagnetic radiation required to penetrate different materials. Thickness and density of the material dictate the amount of radiation that passes through the substance. The part to be examined will absorb some of the radiation, and the remaining radiation will produce an image of the defect on the photographic film. Gamma ray radiography is usually used for detection of internal flaws of engine components that require higher energy levels.

Ultrasonic Inspection

Ultrasonic inspection is effective at detecting flaws both external and internal. This application is particularly useful in assessing critical rotating engine structure for internal material defects. Ultrasonic is concerned with the evaluation of material properties and conditions by probing the material with high-frequency sound waves. Pulses of ultrasonic energy are radiated into the material and subsequently detected using specially designed transducers. As the sound pulses travel into and through the material, they are altered due to attenuation, reflection, and scattering. Most often, ultrasonics is applied to detect thickness and search for flaws in metals—for

example, cracks, voids, porosity, and delaminating (e.g., fan blades and disk.). However, ultrasonics can also ascertain residual stress within metal structures and analyze surface characteristics. Whenever the configuration of the object under test permits, a two- or three-dimensional image of the interior of the object can be made showing reflections of the sound. The latest technology for generating and detecting ultrasonic signals is laser based.

18.66 SHOP PROCESSES

Welding

Welding processes are used extensively in the fabrication and repair of gas turbine components. Examples of commonly use techniques are electron beam welding, tungsten inert gas welding, and steady state plasma arc welding.

Electron Beam Welding (EBW). Electron beam welding utilizes the kinetic energy of a stream of high-velocity electrons to provide the heat required for welding. The process is performed in a vacuum chamber to minimize the effect of gases that produce impurities and distort the beam. It produces chemically pure welds without contamination of the work piece. EBW provides consistent weld size and high joint strength. Welding of thick to thin material and dissimilar metals is easily accomplished. Materials considered difficult to weld by conventional processes may often be welded using EBW since it is able to achieve deep penetration and narrow welds, while providing minimum distortion to the parts. The narrow heat-affected zone allows joining of parts in the finished or near finished shape. In addition, heat-sensitive parts benefit from the low overall heat input. Many steps in the fabrication of case rotating assemblies, fan exit cases, and gearshafts would be extremely costly and time consuming without electron beam welding.

Tungsten Inert Gas (TIG) Welding. TIG welding, also known as gas tungsten arc welding (GTAW), is extensively used in the manufacture and repair of gas turbine parts when high-quality, economical precision welding is required. An arc is formed between a nonconsumable tungsten electrode and the metal being welded. A shielding gas (most commonly argon) is fed through the torch to shield the electrode and molten weld pool. It protects the heated weld zone, molten metal, and electrode from the environment, which could lead to contamination of the weld joint. TIG welding benefits include cost, superior-quality welds, and welds that can be made with or without filler metal. Gas tungsten arc welding is the most popular method for welding aluminum, stainless steels, and nickel-base alloys. This method is employed extensively in the overhaul of gas turbine parts.

Steady State Plasma Arc Welding. Steady state plasma arc welding is produced when ionized gases pass through a restricted metal nozzle. These gases, propelled at extremely high velocities, produce extremely high temperatures as they pass through a restricting metal nozzle. The arc produced when directed to a metal surface will melt the surface, and the filler metal used to make the weld much faster than using TIG welding. The arc generated by this process is stable and can be focused on the workpiece better than other types of arc welding processes. The high heat transfer rate allows for faster, deeper penetration of the metal surface,

which is advantageous when welding thin materials. Another method of utilizing plasma welding is the keyhole method, which requires the plasma to melt through the material. Surface tension forces the molten base metal to flow around the keyhole to form the weld. This is useful only in applications where the plasma can penetrate the joint and is limited to welding thickness up to half an inch, depending on the type of metal used. The process is superior to most other welding techniques for parts with complex geometry.

Bonding

Diffusion Bonding. Diffusion bonding is a method used to join metal and nonmetallic materials. The technique is based on the atomic diffusion of elements at the joining surface. Bonding occurs when metal pieces are heated close to the material melting point and a large pressure is applied that forces the pieces together. Diffusion involves the movement of mass at the atomic level through the lattice of a crystalline solid. Other factors, such as thermal conductivity, thermal expansion, and bonding environment, also affect the bonding process, particularly at high bonding temperature. The process is carried out in a hot press vacuum environment under extremely high temperatures and pressures. The process is used extensively to produce joints that have high strength and good-quality finishes which require no additional work. Diffusion bonding is extensively used to produce a wide variety of parts requiring bonding of dissimilar metals and alloys on turbine engines.

Inertia Bonding. Inertia bonding is used to produce complex parts that would be difficult to machine to achieve the dimensional tolerances needed. The process uses heat and pressure to join metal parts. This is achieved by attaching a part to a flywheel and imparting a large rotational force to the part. This rotational energy is converted to kinetic energy because of the acceleration of the flywheel. The rotating part is then forced against the surface of the stationary part, resulting in frictional heat and pressure that join (inertia bond) the pieces. The resulting frictional heat and pressure join the parts without melting and provide material properties equal to or better than those of the base metals. The process is used extensively to produce compressor disk and shaft assemblies that require little balancing.

18.67 BALANCING

Because of the large rotational speeds of gas turbines, small imbalances are magnified to produce vibration and stresses that increase as the square of the rotational speed. Two principal methods are employed to balance rotating assemblies on turbines. The first is static balancing, done on a single plane, and the second is dynamic balancing, done on two planes. Dynamic balance is most useful for applications such as compressor and turbine rotor assemblies since imbalance may be present at many positions along the axial length of the assembly.

18.68 TESTING

On completion of an assembly for a new or overhauled engine, testing is required. Testing is essential to ensure product reliability, which requires rigorous checks of engines and engine components to meet specific performance requirements. Inside a custom-designed test cell, each engine's performance is closely monitored with specialized equipment that documents performance over the engine's entire operating curve. During testing, records are made of temperature, pressure, vibration, rotor speed, fuel and oil consumption, and engine thrust at different power lever settings. These data are then analyzed to ensure that performance meets the required operating standards. Performance of major engine components is also thoroughly tested. Fuel pumps, fuel controls, governors, actuators, bearings, heat exchangers, and filters are tested separately on test stands provided for this purpose.

18.69 SUMMARY

In-flight shutdown, delay, and cancellation rates continue to improve based on current industry practices. Manufacturing and technology advances, combined with relevant changes in engine maintenance practices, will continue to improve safety and reliability and provide economic gains. Changes in route structures for twins and the advent of long-range operations (LROPs) will continue to expand and challenge current industry standards. It will be imperative for the industry also to improve on the technology currently used to monitor and detect engine degradation that could impact these types of operations. Real-time analysis of engine data during flight is attainable and could provide valuable information on engine condition to engineers monitoring these critical engine parameters. Industry practices, combined with improvements in engine design based on technology improvements and service experience, will continue to improve levels of engine safety and reliability and bring about unprecedented economic improvements.

PART 8

MAINTENANCE HUMAN FACTORS

Bryan P. Kesterson

William L. Rankin

Steven L. Sogg

18.70 PREFACE

Maintenance human factors is defined broadly in this Part to include the application of human factors principles and technologies to all aspects of aircraft maintenance. The term will include the applications of human factors to the aircraft and its operations during its entire life cycle. Maintenance human factors starts with the design of the aircraft by ensuring that the aircraft components and systems are designed for easy maintenance. The design is viewed from the aspect of physical access as well as the on-board maintenance information systems, such as the built-in test equipment (BITE). Maintenance human factors also includes the design of maintenance support products and processes, including maintenance publications (e.g., maintenance manuals, fault-isolation manuals, and illustrated parts catalogs), ground support equipment, and work area design. Finally, maintenance human factors is also applied in day-to-day aircraft maintenance operations, including programs for error management and human factors awareness.

18.71 DEFINITION

The terms *human factors, human factors engineering,* and *ergonomics* are considered synonymous; *human factors* will be used for the remainder of this part. According to the definition adopted by the Internaitonal Ergonomics Association (IEA), ''Ergonomics [*Human factors*] is the scientific discipline concerned with the understanding of the interactions among humans and other elements of a system, and the profession that applies theory principles, data and methods to design in order to optimize human well-being and overall system performance. [Practitioners] contribute to the design and evaluation of tasks, jobs, products, organizations, environments and systems in order to make them compatible with the needs, abilities and limitations of people.''

There are many areas of specialization within human factors. The IEA distinguishes among three major areas. ''Physical ergonomics [human factors] is concerned with human anatomical, anthropometric, physiological and biomechanical characteristics as they relate to physical activity. . . . Cognitive ergonomics [human factors] is concerned with mental processes, such as perception, memory, reasoning, and motor response, as they affect interactions among humans and other elements of a system. . . Organizational ergonomics [human factors] is concerned with the optimization of sociotechnical systems, including their organizational structures, policies, and processes.''

Human factors specialists can come from many different disciplines, although over half the membership of the U.S. Human Factors and Ergonomics Society are

experimental psychologists. Another major discipline is industrial engineering. Other human factors specialists come from the fields of medicine, anatomy, biomechanics, sociology, and other engineering fields. Since the 1980s it has become possible to get an M.S. degree in human factors, where the degree is typically awarded by a psychology or industrial engineering department.

18.72 HUMAN FACTORS PRINCIPLES, STANDARDS, AND TOOLS

Human factors specialists, in general, carry out two types of work—experimental and applied. Experimental work includes the application of the principles of experimental design, subject sampling, standardized data collection, questionnaire development, and data analysis, including descriptive statistics (e.g., means, median, mode, and standard deviation), correlational statistics (e.g., Pearson's *r* and canonical correlation), and statistical significance testing (e.g., *t*-test, univariate and multivariate analysis of variance).

Applied work often involves the application of existing Human Factors standards and tools. The tools include things such as job/task analysis, system modeling, human reliability analysis, and human modeling for reach, access and force requirements analysis. Existing standards used by human factors specialists include MIL-STD-1472 *Military Standard Human Engineering Design Criteria;* ANSI 2535.1-5, *Safety Markers, Warnings, and Cautions;* ISO TC 159, *Ergonomics;* OSHA 65, *International Data on Anthropometry;* and NASA 3000, *Man/System Integration Standards.* The typical content of these standards includes areas like visual displays, audible alerts, controls, labeling, environment, design for maintainability, workspace design, small systems and equipment, hazards and safety, and human–computer interaction.

18.73 REGULATORY REQUIREMENTS REGARDING MAINTENANCE HUMAN FACTORS

The International Civil Aviation Organization (ICAO) was the first organization to speak to the application of human factors in aircraft maintenance. ICAO amended Annex 6, Part I, to include requirements for yearly human factors awareness training for mechanics and for the application of human factors in the day-to-day maintenance operations. Although ICAO annexes do not have to be adopted by the various aviation authorities worldwide, many countries aviation authorities follow ICAO's lead in new regulations. Thus, the ICAO amendments were the first step in these countries' aviation authorities' adopting regulations for human factors in aircraft maintenance. This has led to several of the major aviation authorities adopting regulations regarding maintenance human factors training, which is discussed below.

To date, however, there has not been much regulatory activity regarding maintenance human factors and aircraft design (design for maintainability). The European Joint Aviation Authority (JAA) is considering the development of a new Joint Aviation Regulation (JAR) regarding human-centered design. The JAR has been proposed at this time as a Notice of Proposed Amendment (NPA), 25-310. If passed, the new regulation would require that aircraft manufacturers show "by analysis,

substantiated where necessary by test, that as far as reasonably practicable all design precautions have been taken to prevent human errors in production, maintenance and operation'' that could cause a serious accident.

The JAA has amended Joint Aviation Regulations (JARs) 66 and 145. JAR 66 now requires that apprentice-level mechanics be given human factors awareness training. JAR 145 requires yearly continuation training in human factors after certification.

The U.K. Civil Aviation Authority (U.K. CAA) put out Airworthiness Notice No. 71, Issue 1, on March 20, 2000, recommending strongly that maintenance organizations in the United Kingdom implement a maintenance error management system (MEMS). Also, United Kingdom maintenance organizations have to follow the JAR 66/145 human factors awareness training requirements.

Transport Canada (TC) has recently passed a Canadian Aviation Regulation (CAR) regarding maintenance human factors. The new regulations, among other things, require two things related to maintenance human factors. First, there is a requirement for yearly human factors awareness training for mechanics, including information on human performance, factors influencing human error, and methods for error management and error mitigation. Second, there is a requirement for the implementation of a maintenance safety program, which includes a method for investigation maintenance error and a method for mechanics to report any incidents or unsafe conditions with particular emphasis on real or potential maintenance error (CAR 573.09, 706.07).

Finally, the U.S. Federal Aviation Administration (FAA) has put forth a draft Advisory Circular Number 120xx, *Maintenance Resource Management Training*. The purpose of the Advisory Circular is to strongly encourage U.S. aircraft maintenance organizations to provide human factors training to mechanics.

18.74 MAINTENANCE HUMAN FACTORS IN AIRCRAFT/COMPONENT DESIGN

This application of human factors is typically referred to as *design for maintainability*. Designing an aircraft for ease of maintenance is a responsibility of the aircraft, engine, and component manufacturers. The focus of this design activity is to ensure that the aircraft can be effectively maintained by the mechanics in the field. A well-designed system can significantly reduce training time, task time, and errors and can have a significant effect on the profitability of the airplane in service.

Perhaps the three most important words in aviation maintenance are ''gain access to.'' Space and weight are always at a premium, making physical access to equipment difficult. Human factors has the responsibility for ensuring that mechanics of a range of body dimensions can reasonably fit into access areas and do the work required. In today's world of computer aided design (CAD), this is most commonly done by using a human model in the CAD system. The human model can typically represent people of a variety of sizes. The model also allows for tools to be placed in the human model's hand and for the human model to wear cold-weather clothing. Boeing engineers used this sort of system to show that mechanics could not reach the wiring behind new avionics components to be added to the electronics rack in the 737-700. This analysis forced a redesign of the electronics rack (Figure 18.28).

Once the mechanic has accessed a component, the maintenance work must be done. The same type of analysis that can be done to show access can also be used

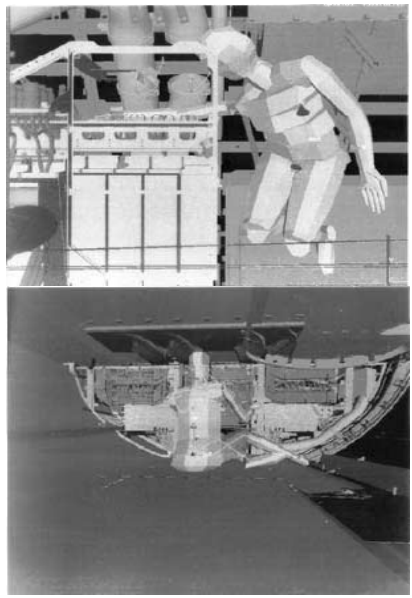


FIGURE 18.28 Human model in computer-aided design system.

to verify that components can be removed and replaced. The analysis determines whether component removal pathways exist and, if they do, whether they can be blocked out to ensure other components are not put into the space. Similarly, the analysis can determine whether the space exists for a tool to be used in the maintenance task. Biomechanics can be used to determine whether mechanics can reasonably be expected to lift the components or whether mechanical lifts will be required.

Like the rest of the world, airplanes are increasingly reliant on computers. The aircraft systems use digital systems to provide the maintainer with information about the health of the systems. This information is used for identifying faults and periodic tests. Consistency and accuracy in the maintainer's interface can make a major difference in the time and errors involved in these sorts of tasks. During the design effort, the same care needs to be put into designing these systems that goes into any other information systems that supports the airplane. For example, the 737-700 stores engine exceedance data in the avionics systems that drive the flight deck displays. The initial menu structure to access the fault data had the mechanic access the display system, then access which one of the redundant units to check, and then allowed the maintainer to get to the exceedance data. Other engine data were found in the engine electronics unit (EEU) which had its own top-level menu pick. A mechanic working looking for an engine exceedance, however, would be unlikely to look at the display systems menu. It more likely, the mechanic would assume exceedance data were with the other engine information and follow the wrong path. The design was changed so that there was a top-level menu item labeled "Engines," under which the mechanic could choose to go to the EEU or to exceedance data.

18.75 MAINTENANCE HUMAN FACTORS IN MAINTENANCE PRODUCT DESIGN

The second area of application for maintenance human factors is with regard to maintenance products, which are commonly supplied by the aircraft, engine, and component manufacturers. This includes the design ground support equipment and maintenance documentation (e.g., maintenance manuals, fault isolation manuals, component repair manuals, and parts catalogs).

Ground support equipment (GSE) is equipment that is needed by mechanics to work on an aircraft. GSE includes work stands, towing equipment, movable stairs, and special tools that are needed to work on aircraft engines, for example. Because this equipment is used by mechanics to carry out a maintenance task, the equipment has to meet human design requirements, which are supplied by the maintenance human factors specialists. These design requirements include:

- Limits on the forces required for GSE assembly and positioning
- Accommodations of safe and efficient tools use
- Accommodations for physical size range of the worker population
- Physical and visual access requirements based on human capabilities and limitations
- Design accommodations and constraints that promote best ergonomic work practices

The second area of application of human factors principles and technologies deals with maintenance documentation. The maintenance engineers writing this documentation use human factors inputs on areas such as:

- Developing the overall format and layout of the document
- How best to convey safety information to the mechanic so that he does not hurt himself/herself during the maintenance task
- The use of simplified English in writing the manuals for use by mechanics for whom English is their second language
- The development and labeling of figures and graphical information
- The verification and validation process for the manuals

If this documentation is to be presented to the mechanic on a computer, then the human factors specialists become involved in the development of the human–computer interface (HCI). Effective presentation of the manual information and associated navigation tools can make a significant difference in the ability of the maintainer. If manuals are provided on computer, an effective user interface will provide navigation tools that reflect the intended task, which is referred to as *task mapping*. Most digital document systems use a *book metaphor,* where the user works from a table of contents similar to one at the front of a paper document. In a comparison between a book metaphor navigation system and a task-mapped navigation system, correct task completion rates went from 13% for book metaphor to 100% for task mapped, and search time went from 27.6 minutes for the book metaphor to 7.4 minutes for task mapped.

18.76 MAINTENANCE HUMAN FACTORS IN AIRCRAFT MAINTENANCE PROGRAM APPLICATIONS

The most common direct application of human factors to maintenance programs is human factors awareness training. Many regulatory agencies currently require this training, both for new staff and as part of ongoing training (see above). Historically, the initial drive for training of this type came out of the flight crew operations side, where human factors training began in the early 1980s. This type of training is called crew resource management (CRM) training. However, in 1988 an Aloha Airlines 737-200 aircraft suffered upper fuselage separation due to cracking. The aircraft had been inspected for cracking only a few weeks earlier, but the cracks had been missed for several reasons to do with human factors issues. This accident was the impetus for the application of human factors in aircraft maintenance and eventually led to the development of maintenance human factors awareness training, also called maintenance resource management (MRM). In addition to the types of things trained in CRM (e.g., communications, decision-making, leadership, and assertiveness), other human factors issues such as perception and human error have been added to the maintenance human factors awareness training curriculum. The following is a list of subject headings for human factors training required for new mechanics under JAR 66:

- General
- Human performance
- Social psychology
- Factors affecting performance
- Physical environment
- Tasks
- Communication
- Human error
- Hazards in the workplace

The human factors awareness or MRM training in the United States is somewhat different. The focus there is more like the focus in CRM, including:

- Situation awareness
- Communications
- Teamwork
- Task allocation
- Decision-making

The other major use of human factors in aircraft maintenance organizations is maintenance error investigation. The Maintenance Error Decision Aid (MEDA) is the most commonly used investigation tool in the industry. The MEDA philosophy is:

- Mechanics do not make errors on purpose.
- Errors are due to contributing factors in the workplace, such as poor lighting, a poorly written maintenance manual, not having a calibrated tool to do a job, and use of wrong parts.
- Most of these factors are under management control and can, therefore, be improved to reduce the likelihood that the next mechanic will make the same error.

Maintenance organizations use MEDA to investigate errors that lead to safety-related or cost-related events, such as flight cancellations, in-flight engine shutdowns, equipment damage, and personal injury.

- Maintenance organizations will also be using human factors principles and technologies to deal with potential safety-related issues, including
- Poor communication between mechanics, especially across shifts
- Fatigue
- Procedural noncompliance
- Shift and task hand-over
- Duplicate inspections
- Technical documentation
- Tool and equipment design
- Workplace/facility environment (e.g., workspace layout and lighting requirements)

18.77 SUMMARY

The application of human factors expertise to aircraft maintenance began in the late 1980s. The original application in aircraft design was in design for maintainability. More recently, human factors principles and technologies have been applied to maintenance products, such as maintenance documentation and ground support equipment. In addition, because of regulation or regulatory authority emphasis, many airlines have started providing human factors awareness training to their mechanics. Airlines have also started to use human factors processes to investigate the causes of maintenance error. Future regulation may require that airlines apply human factors expertise in day-to-day operations.

PART 9

EXTENDED TWIN OPERATIONS (ETOPS)

Harry A. Kinnison

18.78 INTRODUCTION: WHAT IS ETOPS?

The Federal Aviation Administration (FAA) defines ETOPS as ''Extended Range Operations with Two Engine Airplanes''(FAA 1988). The common definition is ''Extended Twin Operations.'' Its meaning is often confused, however. The ''extended range'' we are referring to is not ''down range'' or flight length. Rather, we are referring to the allowable distance the airplane can be from an alternate airport; in other words, a lateral range.

The basic rule for two-engine airplanes comes from FAR 121.161, which states that a two-engine airplane must remain within 60 minutes flying time (with one engine out) of a suitable alternate airport throughout the entire flight. This rule was made in the 1950s for airplanes with reciprocating engines and propellers. When the new generation jets were developed in the early 1980s, operators wanted this 60-minute rule changed so they could fly the new airplanes on their North Atlantic routes.

The FAA published AC 120-42 in 1982, authorizing a deviation from FAR 121.161 by extending the distance to an alternate airport to 120 minutes. This was not a giveaway, however. Approval for extended-range operations was awarded to airlines with specially modified and FAA-approved airplane/engine combinations and with special operations and maintenance procedures in place at the airline. In 1988, the AC was revised to extend the distance to 180 minutes. Deviation for 120 or 180 minutes has since been approved for nearly all two-engine jet airplanes. A more recent ruling by the FAA has extended this deviation to 207 minutes in the North Pacific region for operators of new aircraft specifically designed for ETOPS.

The modifications consisted of engine improvements for greater reliability, modification of the auxiliary power unit (APU) to allow high-altitude start and run capability, and the addition of a hydraulic motor generator (HMG) for back-up power. In the event of an in-flight problem with the engine or engine driven generator (IDG), the APU will be called into use to supply AC power. The HMG is provided for backup in case the APU is inoperative.

18.79 MAINTENANCE PROGRAM REQUIREMENTS FOR ETOPS

The AC identifies a supplemental maintenance program, which consists of a number of specific items needed for ETOPS. Many of these may already be included in an operator's existing maintenance program. Others may require a slight modification of the airline's current program. Each item is discussed briefly below.

An *ETOPS service check* should be performed prior to each ETOPS dispatch. This service check is the approved transit check for the airplane/engine combination plus the action to check oil and calculate oil consumption rate for both engines and the APU.

The *oil consumption monitoring* program is implemented in order to determine the baseline oil consumption rate for all engines and APUs and then monitor these rates for every flight (ETOPS or not). At each check, a comparison is made to determine if the consumption rate is acceptable or if a problem is indicated. The consumption rate is also monitored over time to detect any long-term changes.

Engine condition monitoring (ECM) is used to detect early deterioration of the engines and allow for corrective action before safe operation is affected. Although the ECM program is available to all operators for all engines, its use is optional. For ETOPS operators, however, its use is mandatory. The ECM data are collected each flight (ETOPS or not) after the airplane has reached stable cruise. These data are then entered into the computer program (supplied by the engine manufacturer) and monitored for trends.

The *propulsion system monitoring* program is used to monitor and investigate the causes of in-flight shutdowns (IFSDs). The study and evaluation process may result in additional corrective action or operational restrictions in order to avoid or reduce future IFSDs.

The *resolution of discrepancies* program consists of procedures and instructions in the operator's maintenance program to ensure that prompt corrective action is taken on any problem related to the ETOPS operation and that the corrective action taken has been confirmed successful prior to dispatching an airplane on an ETOPS flight. For cases where maintenance action cannot be verified on the ground, an in-flight check-out must be accomplished. This program must identify procedures used to notify flight crews and dispatchers of any limitations to the flight and of what to look for during flight in order to verify the maintenance action. Verification should be done on the next flight after the maintenance is performed.

The *maintenance of multiple, similar systems* program is implemented in order to avoid duplicate maintenance actions and prevent any inadvertent errors from being repeated on subsequent units. If maintenance must be done on both units of a dual system at the same maintenance visit, precautions must be taken to ensure that both systems are operating correctly before the airplane can be released.

The *ETOPS parts control* program is necessary to identify parts required to maintain the airframe/engine combination in the required ETOPS configuration and must contain procedures to ensure that proper parts are used in maintenance. The applicable configuration, maintenance, and procedures (CMP) document determines the correct configuration for ETOPS through Service Bulletins (SBs) and Service Letters (SLs). These documents will identify any special parts required.

The *APU high-altitude start* program. The APU must be capable of starting and running at altitudes up to 41,000 ft for ETOPS. This program requires that a high-altitude start of the APU be attempted on all non-ETOPS flights and on the return leg of all ETOPS flights. If a 95% success rate is achieved for these tests during the first two months of operation, the test may be run on a sampling basis from that point on.

18.80 ADDITIONAL REQUIREMENTS

Additional requirements have been established by the AC to track ETOPS activities and to document the entire program.

The *ETOPS training program:* The operator's current maintenance training program must be modified to ensure that any mechanic or technician, as well as man-

agement personnel, involved with ETOPS airplane maintenance have been trained on all maintenance aspects of the extended range operations. This training program should include training on the ETOPS philosophy and background; training on the additional equipment on the airframe/engine combination being used for ETOPS; and training on the airline's ETOPS maintenance procedures.

The *ETOPS reliability program,* required by AC 120-42A, is an event-oriented program that concentrates on events that are related to the ETOPS mission. This includes such items as engine in-flight-shutdowns, diversions, and turn-backs; pilot reports (PIREPS) on ETOPS significant systems; and adverse trends noted by the oil consumption and engine conditioning monitoring programs. This reliability program must be in addition to any statistically based reliability program that may already be in place at the airline.

An *ETOPS manual* should be developed as part of the airline's general maintenance manual, or as a stand-alone document, in order to provide a single source of information on how the airline meets the AC 120-42A standards. The ETOPS manual can also be used as a text in the ETOPS training program.

18.81 COMMENTS ABOUT THE ETOPS MAINTENANCE PROGRAM

These ETOPS maintenance requirements do not constitute a significant increase in maintenance activities at the airline. They mostly represent a different approach to maintenance management. Conventional maintenance programs involve *preventive maintenance* techniques, which ensure that the systems and components are restored when deterioration has occurred, and *predictive maintenance* techniques, which allow the operator to use historical or statistical data to predict maintenance problems before they become in-service failures. The two approaches work together to provide scheduled and unscheduled maintenance practices and procedures for an efficient maintenance organization.

In ETOPS, we apply what can be called real-time maintenance techniques. Instead of waiting for a failure to occur or for a failure rate to reach some alert level, the real-time approach to maintenance monitors activities on a continuing basis to determine if any deficiency is developing or if any trend is indicated which will eventually result in an in-service problem.

The ETOPS program does not ask airline personnel to do more maintenance; it only asks them to do maintenance more effectively. It has been stated that the ETOPS maintenance program is what a good maintenance program should be: a more conscientious approach to maintenance, keeping watch to detect problems as they develop instead of waiting for failure to occur. If an airline finds that the ETOPS requirements are a significant increase over their current maintenance program, it may be that the current program is not sufficient.

18.82 ETOPS MAINTENANCE ON NON-ETOPS AIRPLANES

A common question asked by operators is, ''If the ETOPS maintenance program is what a good maintenance program should be, then why can't it be applied to

non-ETOPS airplanes?'' Well, it can be. Several operators who use the same two-engine model airplane in both ETOPS and non-ETOPS service have found that configuring and maintaining all the airplanes to ETOPS standards provides greater flexibility in scheduling as well as improvements in fleet reliability. Some operators only configure and maintain common engine types to ETOPS standards so that these can be interchanged between ETOPS and non-ETOPS airplanes as necessary.

Other operators have found that some elements of the ETOPS maintenance program can also improve operation of three- and four-engine airplanes. Although the modifications done on the early two-engine models have not been done on the threes and fours, the real-time maintenance approach can still be applied with noticeable results.

REFERENCES

Boeing Company. 1998. *ETOPS Guide,* vol. 2, *Maintenance and Program Guidelines.*

Boeing Company. *Configuration, Maintenance and Procedures for ETOPS Operation* [model-specific, various dates].

Federal Aviation Administration (FAA). 1988. *Extended Range Operations with Two Engine Airplanes,* FAA Advisory Circular No., 120-42A.