More Subsystem Design

9

IT is difficult to say which subsystem on board a spacecraft is the most important, or indeed rank them in order of importance. Subsystem engineers invariably say that their own piece of territory on the spacecraft is the most important. Table 7.1 in Chapter 7 lists the subsystems; a serious failure of any one of them will result in the end of the spacecraft mission. So there is some validity to what subsystem engineers say.

This chapter discusses the remainder of the subsystems, focusing on the main factors that drive their design. As a consequence, I should warn you that this chapter is longer than previous ones. We start with the propulsion subsystem simply because it follows logically from the discussion of orbits in Chapters 2, 3, and 4.

The Propulsion Subsystem

As we recall from Chapter 7, the propulsion subsystem's main functions are to provide a capability to transfer the spacecraft between orbits, and to control the mission orbit (see Chapter 3) and the spacecraft attitude (see Chapter 8), using on-board rocket systems. Although this statement seems elaborate, it is reasonable to say that the propulsion subsystem's job is fairly intuitive; that is, the spacecraft has rocket engines on board that are fired to move the vehicle as appropriate to ensure it reaches its intended destination in near-Earth or interplanetary space. There are effectively two aspects to the propulsion subsystem's job: *orbit transfer*, and *orbit and attitude control*.

Orbit Transfer

Some spacecraft can be placed directly into their mission orbit by the launch vehicle, and thus have no need to perform orbit transfers. However, other spacecraft have to transfer between orbits to reach their final destination. This process of orbit transfer usually involves the use of a large rocket engine onboard the spacecraft, and such a system is referred to as *primary*

G. Swinerd, *How Spacecraft Fly: Spaceflight Without Formulae*,
DOI: 10.1007/978-0-387-76572-3_9, © Praxis Publishing, Ltd. 2008

propulsion. If a spacecraft needs such a system, then the mass of the rocket hardware and of the required propellant has a major effect on the overall mass of the spacecraft. A good example of a spacecraft that requires such a primary propulsion system onboard is one of the communication satellites that we talked about in Chapter 2. They can be injected into a low Earth orbit (LEO) by the launch vehicle, but then need to be boosted to the high geostationary Earth orbit (GEO) before they can begin operations. This is usually achieved by using the *Hohmann transfer*. This strategy, which was invented by Walter Hohmann, is illustrated in Figure 9.1. Hohmann was one of those extraordinary individuals you find in the history of spaceflight. During the early 1900s he was a city architect by profession, in Essen, Germany. But in his free time he devoted his energies to thinking about interplanetary spaceflight. He published his work, including the details of his orbital transfer method, in 1925, at a time when the first Earth satellite was still over a quarter of a century away. His transfer method has been used hundreds of times in placing spacecraft into GEO because it does the job using the minimum amount of rocket fuel, which. means that the overall mass of the spacecraft is minimized and the launch costs can be reduced.

Referring to Figure 9.1, if we assume that the spacecraft is placed initially in LEO by the launcher, how does the Hohmann transfer take us to GEO? To do this, the spacecraft's primary engine is fired twice—once at point 1, to boost it into the geostationary transfer orbit (GTO), and then again at point 2 to push it into the final GEO. The elliptical GTO becomes a sort of bridge, spanning the space between the low orbit and the high mission orbit. Focusing on the first engine firing at point 1, we know from our basic orbits in Chapter 2 that the speed of the spacecraft at this point in the circular LEO

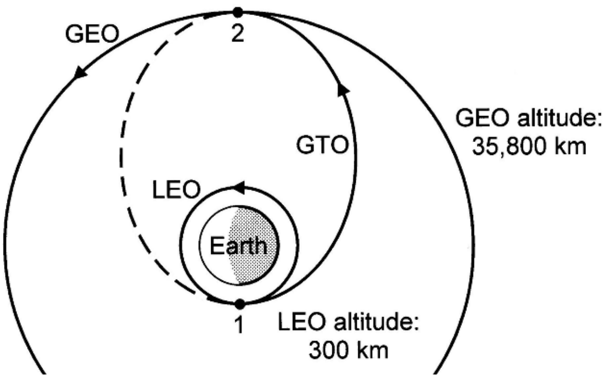


Figure 9.1: The Hohmann transfer between two circular orbits in the same plane. In this case, the spacecraft engine is fired at point 1, to transfer from LEO to GTO. The engine is then fired again at point 2, to take the vehicle into the GEO.

is about 8 km/sec (5 miles/sec). Recalling Newton's cannon on the mountaintop (Chapter 2), we know that the speed at point 1 at the perigee (the low point of an elliptic orbit) of the GTO is higher. This is very much related to our discussion about what happens if we increase the barrel speed of Newton's cannon beyond the circular orbit speed. With some simple calculations we can determine that the speed at point 1 in the GTO is about 10 km/sec (6.2 miles/sec). Thus if the rocket engine firing at point 1 increases the speed of the spacecraft by 2 km/sec (1.2 mile/sec), then the vehicle will transfer from the LEO to the GTO. A similar calculation shows that the speed increase required at point 2 is about 1.5 km/sec (0.9 miles/sec) to transfer between the GTO and the GEO.

The speed change produced by a rocket burn is referred to as a ΔV (pronounced "delta vee"), and one of the prime jobs of the mission analysis team is to calculate the total mission ΔV required to take the spacecraft from launch pad to final mission orbit. Mission analysts spend a lot of time calculating ΔVs because it is directly related to the amount of rocket fuel required. The first person to realize this was Konstantin Tsiolkovsky, another spaceflight visionary who worked as a high school mathematics and science teacher in Russia around the turn of the 20th century. He published his *rocket equation* in 1903, in what was arguably the first mathematical treatise on rocket science. This equation calculates the rocket fuel mass for a particular orbit transfer directly from the corresponding ΔV, which is an important calculation for the spacecraft designer. Mission analysts should also try to minimize the ΔV, because minimum ΔV means minimum fuel mass; in turn, minimum fuel mass means maximum payload mass onboard the spacecraft, which means that the overall effectiveness of the spacecraft in achieving its objective is enhanced. You may recall a similar argument in Chapter 5 when we discussed maximizing the payload of a launch vehicle.

We can use Tsiolkovsky's rocket equation to estimate the rocket fuel mass required for the transfer from LEO to GEO. The total ΔV of about 3.5 km/sec (2.2 miles/sec) requires that about 70% of the initial mass of the spacecraft in LEO needs to be propellant. This leaves only 30% of the initial mass as hardware (payload and subsystems), which poses a problem for the designer. To help overcome this, it is common practice for the launcher to inject the spacecraft directly into the GTO, so that only the 1.5 km/sec ΔV at point 2 is handled by the spacecraft's own primary propulsion. In this case, only about 40% of the initial mass needs to be rocket fuel, giving the designer a much more reasonable 60% of hardware mass to play with.

In our discussion of the Hohmann transfer, we used the LEO to GEO journey as an example. But this type of transfer can be used to move a spacecraft between any two circular orbits that share the same plane, for example, between two LEOs a few hundreds of kilometers apart, where the ΔV might be a few hundred meters per second. The original application that Hohmann had in mind was the transfer between the orbits of two planets, say Earth to Jupiter, where the ΔV might be on the order of 10 km/sec (6.2 miles/sec).

Orbit and Attitude Control

As we saw in Chapter 8, the attitude control subsystem borrows the services of the propulsion subsystem to do the job of attitude control. The small thrusters, fired in opposed pairs (see Figure 8.4), are used as control torquers. These small rocket engines, grouped in clusters, around the spacecraft, are referred to as the spacecraft's *secondary propulsion system.*

The orbit control function, which we discussed in Chapter 3, is also handled by the spacecraft's secondary propulsion. We found that when the spacecraft is launched into the ideal mission orbit, it does not stay there, unfortunately. Perturbing forces due to drag, the Sun's, Moon's and Earth's gravity, and light pressure cause small changes in the orbit, which must be controlled if the mission orbit is to be maintained. To correct the orbit, small changes in the spacecraft's orbital speed are required, and we have seen that this can be achieved by firing small rocket engines onboard the spacecraft. Two small thrusters are fired, not in opposite directions to produce a torque, but this time in the same direction to produce a small ΔV to correct (or control) the orbit, as shown in Figure 9.2.

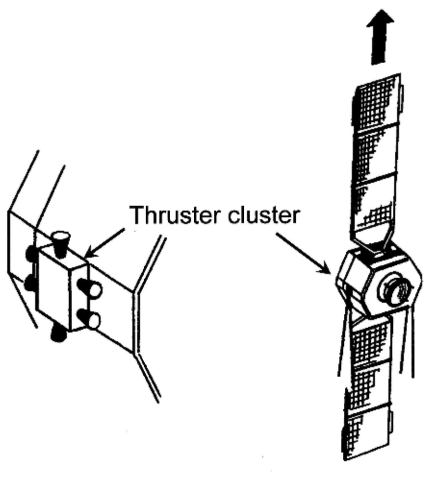


Figure 9.2: Two thrusters are fired in the same direction, to produce a small change in orbital speed to control the orbit against perturbations.

Spacecraft Propulsion Technology

Currently there are two main types of onboard spacecraft propulsion systems: chemical and electrical. Chemical systems, such as the large launcher rocket engines we discussed in Chapter 5, use the chemical energy obtained by burning a propellant/oxidizer mix. This produces a hot and energetic gas, which is then exhausted through a rocket nozzle to produce thrust. Electrical systems, on the other hand, use electrical power to accelerate the propellant gas out of the rocket nozzle. High exhaust speeds can be obtained in this way, but the amount of mass per second that can be accelerated is generally small, unless a large amount of electrical power is used. As a consequence, the thrust of electrical systems is small (usually small fractions of a Newton). This chapter focuses on the more commonly used chemical propulsion systems. Big chemical rocket engines on spacecraft are either solid propellant systems or liquid bi-propellant systems.

Primary Propulsion

Solid Propellant Systems

A typical *solid propellant* main engine on a spacecraft looks like the system shown in Figure 9.3. In concept, it is a simple device, composed of a tank containing solid fuel, an engine nozzle, and some kind of *pyrotechnic device* to light it. Just like the large space shuttle solid propellant booster rockets (it may be helpful to recall the material associated with Figure 5.2 in Chapter 5), once lit, it will burn until the propellant is exhausted, and then it becomes

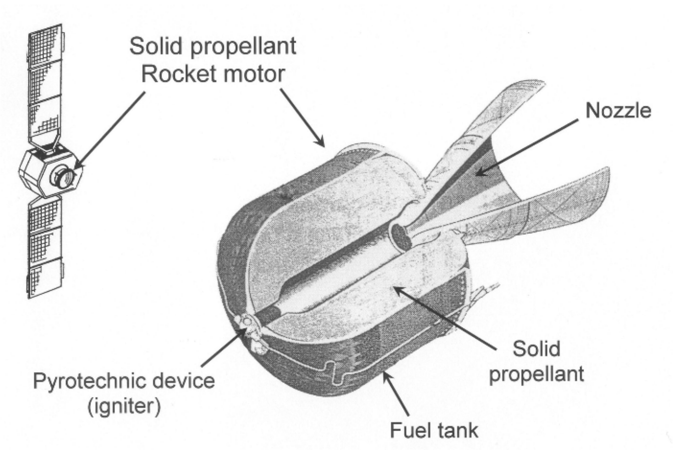
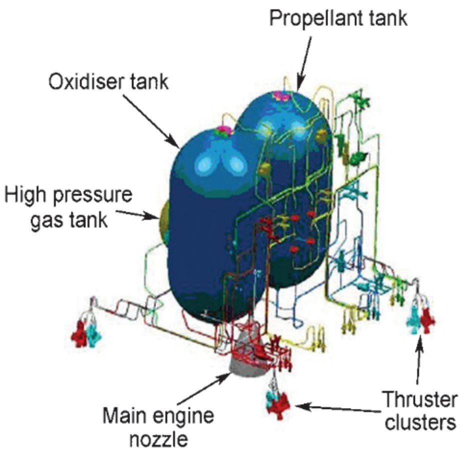


Figure 9.3: Diagram of a typical solid propellant primary engine.

inert and plays no further role in the spacecraft's mission. The engine burn is usually of short duration (a few tens of seconds) and high thrust (tens of thousands of Newtons), producing high accelerations—typically a 0 to 60 mph in 1 second kind of performance! In our orbit transfer example from LEO to GEO in Figure 9.1, the engine burn at point 2 has been performed many times with this kind of rocket engine. To get the right ΔV at this point, the mission analysis team has to do careful calculations, using the rocket equation, to determine how much propellant is needed. This precise amount is then loaded into the motor. Despite the simplicity of the system, one major disadvantage of using a solid propellant motor is the one-shot characteristic; multiple rocket engine burns are not possible.

Liquid Bi-Propellant Systems

Over time, spacecraft missions have become more sophisticated, requiring more than a single primary engine burn. To accommodate this requirement, a *unified liquid bi-propellant system* has become increasingly popular as an alternative to the solid propellant systems. The system is unified in the sense that the spacecraft's primary engine, and the small thrusters that deal with orbit and attitude control are on the same circuit. The big and small rockets onboard the spacecraft share the same fuel supply. The propellant used is in liquid form, and it is a bi-propellant because there are two liquids—a fuel and an oxidizer. The commonly used fuel and oxidizer are monomethylhydrazine and nitrogen tetroxide, respectively. These liquids are *hypergolic,* which basically means that if we mix them together, they spontaneously explode! So to fire the main rocket engine, for example, all we have to do is feed the fuel and oxidizer into the rocket's combustion chamber (see Figure 5.3), and a hot gas is produced explosively, which is then exhausted through the engine nozzle to produce thrust. The same principle is used to operate the small thrusters on board the vehicle. The liquid propellants are held under the pressure of a gas, such as helium, so that the feed system in this case operates by squeezing the liquids down the fuel lines under this gas pressure. Of course, the hypergolic character of the fuel/oxidizer mix means that the plumbing associated with the propulsion system is complex, in order to ensure that the two liquids do not mix before they get to the combustion chambers of the thrusters. The consequences for the spacecraft were this to happen would be catastrophic! Figure 9.4, which shows an example of the plumbing associated with such a system, gives a good idea of this complexity. There is also the issue of the safe handling of the fuel and oxidizer at the launch site, when the spacecraft's fuel tanks are being loaded. These workers have to be equipped with protective, pressurized suits to safeguard them from the effects of an accidental escape of these nasty substances.



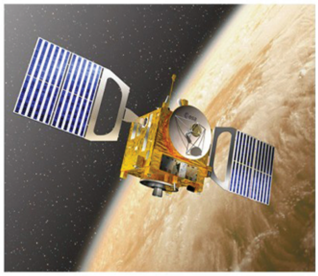


Figure 9.4: A diagram of the unified liquid bi-propellant propulsion system used on the European Space Agency (ESA) Venus Express spacecraft. As well as the main elements that are labeled, there is also an array of valves, sensors, and regulators to manage the hypergolic fuel safely. (Image credits: left—courtesy of EADS Astrium; right—courtesy of ESA.)

The main engine used in unified liquid bi-propellant systems usually has a much lower thrust than solid propellant primary engines. A thrust level of around 400 Newtons is common, giving a 0 to 60 mph in 2 minutes type of performance for an average-sized spacecraft! Consequently, the duration of an engine firing is often much longer, up to 1½ hours, to give enough time for the required ΔV to be achieved.

Secondary Propulsion

We have already discussed the operation of small thrusters on spacecraft equipped with unified liquid bi-propellant systems. However, spacecraft equipped with solid propellant primary engines, have a need for an independent thruster arrangement, and this is usually a *mono-propellant hydrazine* system. "Mono" implies that this involves a single liquid fuel, as opposed to the fuel/oxidizer combination in the bi-propellant system; this single fuel is fed under gas pressure to the thrusters. To fire a particular thruster, the fuel inlet valve is opened, and the mono-propellant hydrazine is squirted through a bed of chemicals in the thruster, causing an exothermic (heat producing) chemical reaction, which breaks down the liquid fuel into hydrogen, nitrogen, and ammonia gases. These hot gases are then exited through the thruster nozzle to produce thrust. Figure 9.5 shows a typical mono-propellant hydrazine thruster, which fits comfortably in the palm of a hand.

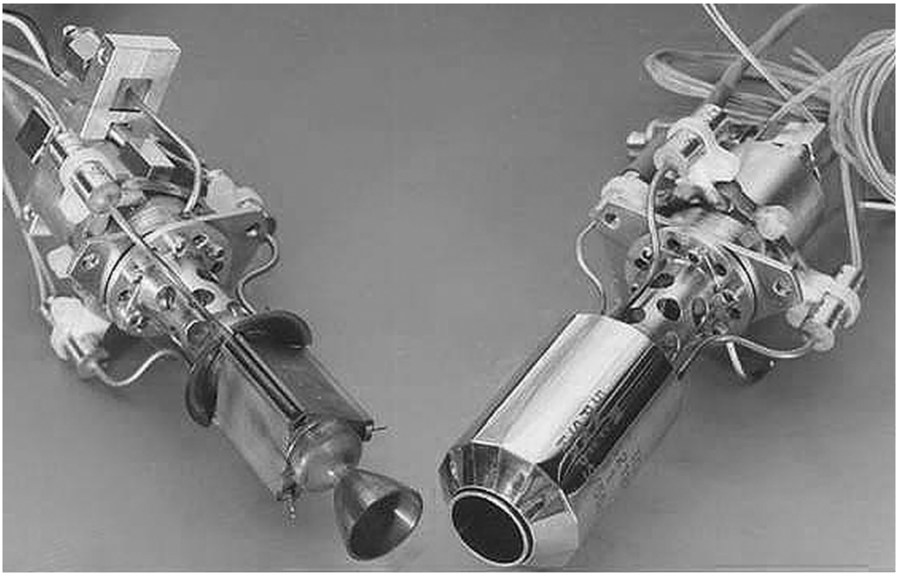


Figure 9.5: An example of a mono-propellant hydrazine thruster, with a thrust level of 5 Newtons. Left: The thruster showing the classic thruster nozzle. Right: The same thruster in its normal operational configuration with collar attached. (Image courtesy of EADS Astrium.)

The Electrical Power Subsystem

As we recall from Table 7.1 in Chapter 7, the main function of the power subsystem is to provide a source of electrical power to support payload and subsystem operation. This job is critical to the overall health of the spacecraft. Just about every type of spacecraft payload and all the subsystems, with the exception of the structure and possibly the thermal control, need a reliable source of electrical power to operate. A failure, or brief interruption, of the power subsystem function can have catastrophic consequences for the spacecraft mission.

In the same way that the spacecraft has primary and secondary propulsion (see the previous section), the spacecraft also has primary and secondary power systems. The *primary power system* is the main source of electrical energy; for example, for Earth-orbiting satellites this is often the conversion of sunlight into electricity using a solar panel (or array). As you have read through the previous chapters in this book, you've seen numerous pictures of spacecraft, and the majority of them are equipped with solar arrays. The *secondary power system* comprises electrical storage devices. In the vast majority of spacecraft, this implies the use of battery technology. However, there are other possibilities, although they are rarely used. For example, a fly wheel can be installed as an alternative electrical storage device. While the spacecraft is in sunlight, solar panel power can be supplied to a torque motor to spin a large wheel. When the spacecraft is in darkness, and the primary power source no longer works, the rotational energy in the wheel can be extracted and converted back into electricity.

A spacecraft is very much like an automobile inasmuch as there is a primary power source combined with a secondary power system. In a car the primary system is the generator that provides electrical power all the time that the engine is turning. The secondary power system is the battery, which allows the car's systems to operate even when the engine is not running. Despite the terminology, the secondary system is very important, in that it provides a means of starting the car and bringing the primary power source on line!

The most common mode of operation for the supply of electrical power on board Earth-orbiting spacecraft involves a solar array/battery combination. While the spacecraft is on the sunny side of the Earth, the solar array operates to produce power for the vehicle's payload and subsystems. At the same time, an extra amount of power is produced by the solar array to charge the battery system. Then, when the spacecraft enters the Earth's shadow on each orbit revolution, the solar array ceases to function and the vehicle's systems are powered by discharging the stored energy in the batteries.

Commonly Used Primary Power Sources

In terms of primary power sources, there are a number of options (including solar arrays), and the main candidates are listed in Table 9.1. Figure 9.6 shows the most suitable primary power source for a spacecraft, given the duration of its mission and its electrical power requirement. From this we can see that, of the six sources listed in Table 9.1, only four are suitable for long-duration space missions, and of these only two—solar arrays and radioisotope thermal generators (RTG)—are commonly used.

Solar Arrays

Recall that in Chapter 6 we (figuratively) went into the garden and presented a square meter of area to the Sun, and found that the Sun's power falling this was roughly 1.4 kW, neglecting the losses that occur due to passage through the atmosphere. In Earth orbit, this *power flux* is essentially a free resource that is just too good to miss, so spacecraft are usually equipped with solar arrays designed to convert some of this solar power into electricity. Most

Table 9.1: Primary electrical power sources used on spacecraft

Type	Usage and principle of operation
Primary batteries	Primary (unrechargeable) batteries are used for short duration missions, for example, to power a launch vehicle during the few minutes' climb to orbit.
Fuel cells	Fuel cells are essentially chemical engines that produce electrical power, with water as a by-product. This makes them particularly suitable for manned missions. The duration of their operation is limited, however, by the requirement to fuel the chemical reaction continuously with oxygen and hydrogen. Fuel cells provide the primary power source for the space shuttle.
Solar arrays	Solar arrays operate by converting sunlight into electrical power. With the intensity of sunlight at Earth (about 1400 W per square meter), solar arrays on Earth-orbiting spacecraft provide about 100 W of useful electrical power for every square meter of solar array surface area. So, although solar arrays are commonly used, they are inefficient (see text for more detail).
Solar dynamic devices	A solar concentrator, such as a large parabolic mirror, is used to focus the power of the Sun to heat a working fluid, such as water. The high-pressure steam produced can then be used to drive a turbine generator—the dynamic part—to produce electrical power. Solar dynamic devices are more efficient than solar arrays, but are generally much heavier. As such, they are only considered for use on large spacecraft, such as space stations. Overall, they are rarely used.
Radioisotope thermal generators (RTGs)	The heat energy obtained from a radioactive material (such as an isotope of plutonium) is converted into electrical energy. RTGs are widely used on spacecraft that operate a long way away from the Sun, where sunlight is so feeble as to make the use of solar arrays impractical. Each RTG is cylindrical in shape, typically 1 m in length and 30 cm in diameter (see Figure 9.8). Such a device is around 40 kg in mass, and produces about 200 W of useful electrical power. A quick calculation (200 W divided by 40 kg) gives a rough estimate of the amount of power produced per kilogram—approximately 5 W. So if large amounts of power are required for payload and subsystem operation, then the mass of the power supply can be significant. There is also political opposition to the

	use of RTGs by the "Green" lobby, which fears the effects of the dispersal of radioactive material in the event of a launch failure.
Nuclear reactors	These are scaled-down versions of the nuclear reactor systems found in terrestrial power stations. They are used only for applications requiring large amounts of power—on the order of hundreds to thousands of kilowatts. To date, they have been used rarely, and then only on military space systems such as large active radars for surveillance.

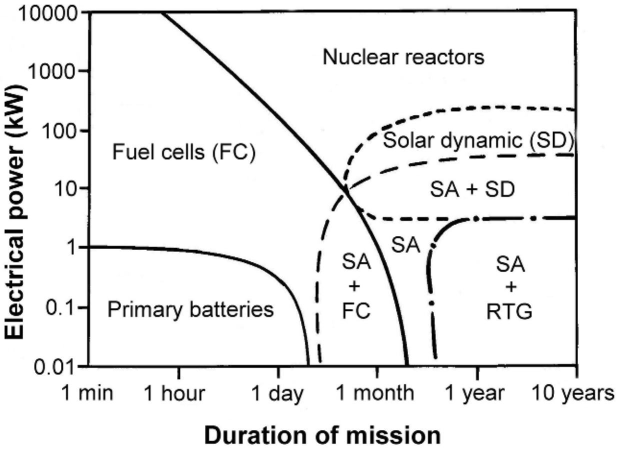


Figure 9.6: The most suitable primary power source is shown, depending on the level of electrical power required by the spacecraft and the duration of its mission. Note: SA, solar array; RTG, radioisotope thermal generator.

solar arrays used for space applications are made out of the semiconductor material silicon, and they have an efficiency of around 10%. This means that of the 1400 W of solar power falling on each square meter of array, only one tenth of this (140 W) is produced as useful electrical power to run the spacecraft's systems.

Unfortunately, there are a number of other factors that have to be taken into account, that further reduce the solar array's efficiency. For example, for best performance the array surface needs to be presented to the Sun so that the Sun's rays fall at right angles to its surface. This requires the array to be accurately pointed by the spacecraft's attitude control subsystem (ACS) (see Chapter 8), and this may be done only to a certain tolerance—say, to within 5 degrees of the ideal. Such a *pointing error* will reduce the array's efficiency.

Another factor is temperature. Arrays get hot while in the Sun, typically up to around 50°C (in Earth orbit), and the hotter they get the less efficient they are; for example, silicon arrays can lose 10% of their electrical performance with a 25°C increase in temperature. A third issue is that the spacecraft's solar array panels cannot be covered completely in useful silicon material. Each panel is usually composed of smaller silicon cells, and the required electrical interconnection between these means that about 10% of the panel area is dead with respect to power generation. The final factor is damage to the array caused by particle radiation (see Chapter 6). After 10 years of operation in GEO, the electrical output from a solar array can be reduced by around 30% of its beginning of life performance.

All these factors vary according to the spacecraft's mission and orbit, but in Earth orbit a useful rule of thumb is to expect about 100 W of useful electricity—enough to run an old-fashioned domestic light bulb—for every square meter of silicon solar array on the spacecraft. Many modern spacecraft, particularly communication satellites, can have power requirements up to 10 kW, which means that an lot of solar array area is needed. Accommodating such large arrays can have a significant impact on the overall configuration of the spacecraft.

The above discussion applies to Earth-orbiting spacecraft, which are effectively at a distance from the Sun of 1 astronomical unit (AU). As we have seen, the Sun's intensity at 1 AU is about 1.4 kW per square meter, but this intensity varies with distance from the Sun. In fact it decreases in proportion to the inverse square of the distance; Chapter 1 discussed the inverse square law in the context of Newton's law of gravity. Basically, the idea is that at twice the distance from the Sun at 2 AU, the Sun's intensity will fall to a quarter ($1/2^2$) of what it was at Earth. At 3 AU it will decrease to a ninth ($1/3^2$), and so on. So, for example, if our mission is to take us to Saturn, which is about 10 AU from the Sun, the solar power flux there is about 1400 W divided by 100 (10^2)—about 14 W per square meter. On top of this, there are all the various inefficiencies of the solar array in converting this power flux into electricity, so we can see that at this kind of distance the use of solar arrays to generate electrical power becomes impracticable.

What is the maximum distance from the Sun where solar arrays can still be used effectively to generate spacecraft power? This is difficult to answer, but I would suggest 5 AU (the distance of Jupiter from the Sun) as the outer boundary. A quick calculation tells us that at 5 AU the solar intensity is 1400 W divided by 25 (5^2), which is about 56 W per square meter. Taking account of the efficiency of the array, we might expect about 6 W of electrical power per square meter of array, which still sounds marginal. But one saving grace here is the temperature of the array, which out in the cold reaches of Jupiter's orbit will typically be less than –100°C. As mentioned above, solar arrays become less efficient as their temperature rises, but then by the same token they become more efficient as their temperature drops. Consequently, the array can achieve a useful level of power generation even at these distances. Interestingly, at the time of this writing (2007), the European Space Agency (ESA) spacecraft Rosetta is on its way to rendezvous with a comet in 2014, and the interception will take place at a distance of about 5¼ AU from the Sun where the solar intensity is about 50 W per square meter. The Europeans are generally averse to using RTGs, on environmental grounds, and so the spacecraft is equipped with solar arrays for power generation. However, to generate the required 395 W of electricity to run the spacecraft—equivalent to four light bulbs—64 square meters of solar array are needed! A quick calculation, 395 W divided by 64 m^2, gives about 6.2 W of electrical power per square meter of array. The overall array efficiency (6.2 W/m^2 of electricity divided by 50 W/m^2 of solar intensity) is about 12%, which is just about feasible. It is at these great distances from the Sun where it is advantageous to think about the use of RTGs.

Radioisotope Thermal Generators (RTGs)

To overcome the problems of power generation at large distances from the Sun, the use of RTGs provides a solution. These have been used on many spacecraft (for example, Pioneer, Voyager, Ulysses, Galileo, and Cassini), which have ventured to distant parts of the solar system, at and beyond Jupiter's orbit. The idea is that they take their own energy source with them, in the form of a radioactive material. The radioactivity heats the material, and this heat is converted to electricity through the *thermoelectric effect.* This was discovered, it is said accidentally, by Thomas Seebeck in 1821, and you may have done experiments in school science lessons to demonstrate the effect, using a *thermocouple.* Figure 9.7 illustrates a simple thermocouple: two wires made of dissimilar metals A and B are formed into a circuit, with some kind of meter to measure electrical current. If junction 1, where the two metals are joined, is heated, and junction 2 is cooled, an electrical current is generated in the circuit. The heat differential produces electricity. The RTG is a more sophisticated device, but this simple thought experiment demonstrates the underlying principle of its operation. Essentially, the temperature difference between the hot core of the radioactive material and the cold of space is exploited to produce electrical power for the spacecraft. The typical size, mass, and power output of a RTG are given in Table 9.1, and Figure 9.8 shows what they look like.

There are disadvantages to using RTGs, some of which we have already mentioned. Given that a typical RTG contains a radioactive material such as

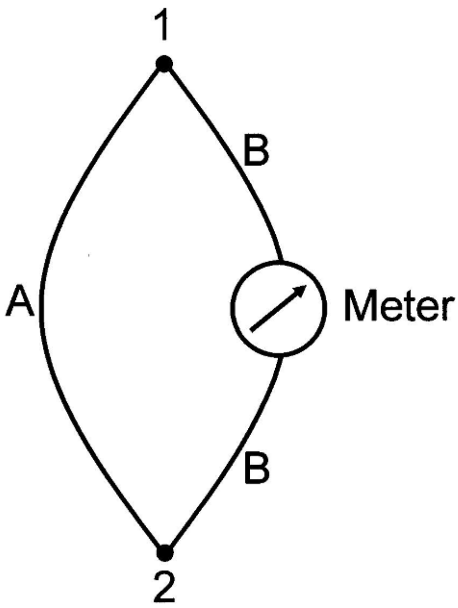


Figure 9.7: A simple thermocouple circuit composed of wires of different metals A and B, and a meter to measure electrical current. When junction 1 is heated and junction 2 cooled, an electrical current will flow in the circuit.

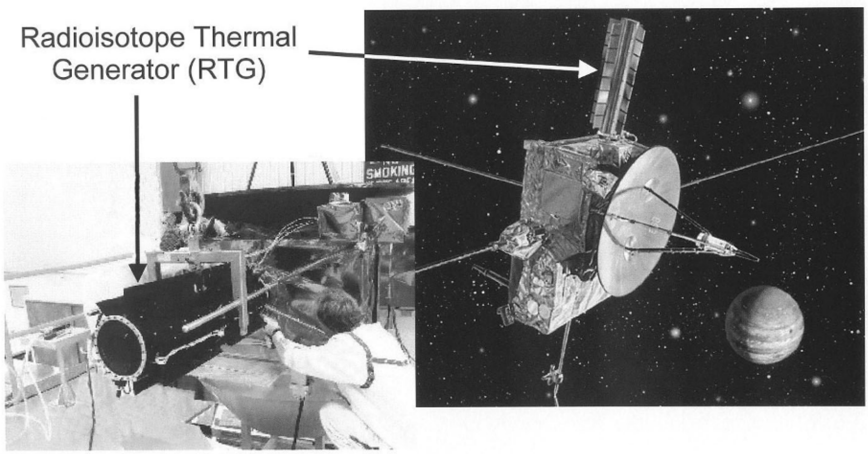


Figure 9.8: The RTG on the ESA/National Aeronautics and Space Administration (NASA) spacecraft Ulysses, shown during spacecraft assembly (left) and in flight (right). (Image credits: left—NASA/Jet Propulsion Laboratory (JPL)—Caltech; right—ESA.)

plutonium, there are safety concerns related to the radiation hazard RTGs pose when the spacecraft is being assembled. This radiation is also potentially damaging to onboard electronics, and so RTGs need to be positioned away from radiation-sensitive equipment. Also, the heat source needs to be intense to produce the necessary electrical performance. As a consequence, RTGs become hot and produce about 10 times more heat power than electrical power. For a typical device, each RTG installed on the spacecraft radiates about 2 kW of heat, which can create problems with the spacecraft's thermal control and during the launch when the spacecraft is confined in a limited volume inside the launcher fairing.

Typical Power System Operation

Despite the diversity of possible electrical power sources (see Table 9.1 and Figure 9.6), the vast majority of spacecraft operate on a solar array/battery combination. This is particularly true for Earth-orbiting spacecraft, which we shall focus on here. Figure 9.9 is a simplified diagram showing how this typical arrangement works. The main feature is the connection of the spacecraft's electrical loads across the solar array (via points 6 and 7), so that the loads are supplied by the array while the spacecraft is in sunlight. However, there is a complication. We know from the above discussion about solar arrays that if a specified area of array is presented to the Sun, it will produce a specified amount of electrical power. But we also know that the spacecraft loads vary; for example, payload instruments will be switched on and off at various times, and subsystem elements such as reaction wheels, or the communications equipment, will run when they are required. So, on the one hand, the solar array output is constant, but then on the other, the

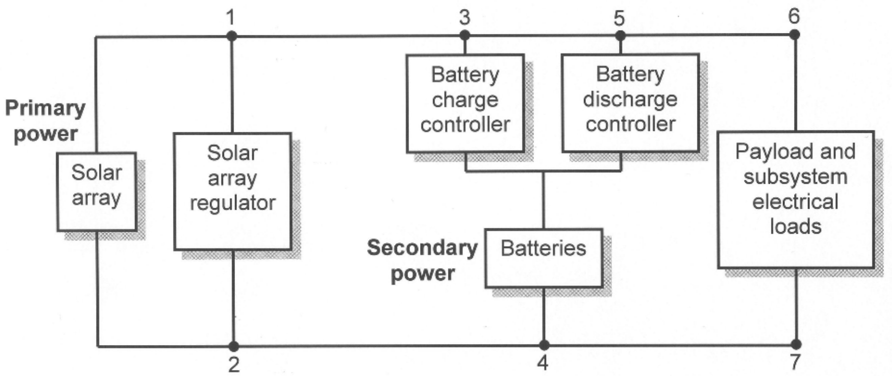


Figure 9.9: A simplified block diagram showing a typical electrical power distribution arrangement for a spacecraft with a solar array/battery power system.

electrical loads it needs to supply are continuously varying. To address this mismatch, a *solar array regulator* is fixed across the solar array (at points 1 and 2). This can be a simple device that takes the excess power not required by the spacecraft loads and dissipates it externally as heat—a bit like an electric bar fire. At the other end of the scale, it can be a sophisticated device, controlled by the onboard computer, which switches patches of solar array area in and out to ensure that the array output matches the loads at any particular time.

While all this is going on, the battery system is also connected across the solar array (via points 3 and 4), so that it can be charged up while the spacecraft is in sunlight; it becomes just another element of the spacecraft's electrical loads. Then, when the spacecraft enters Earth's shadow on each orbit, the electricity for the payload and subsystems can be supplied by the stored energy in the batteries (through points 4 and 5). This process of charging and discharging the batteries needs to be carefully controlled, to ensure that the batteries last long enough to do their job throughout the entire spacecraft mission. The main factors that govern the lifetime of the batteries are the total number of charge/discharge operations needed during the spacecraft's mission, and the amount of stored energy that is taken out of the battery on each such operation. For a typical Earth-orbiting spacecraft, this charge/discharge operation usually occurs once per orbit, and so for a spacecraft in a LEO there are typically about 5000 charge/discharge cycles per year. The *charge/discharge controllers* (points 3 and 5) are essential components to make sure the batteries do not die prematurely. If you think again about the battery in a car, the battery charge and discharge process is not usually controlled in any way; in fact the driver takes this role. Unlike the battery system in a spacecraft, which can last for 10 or 15 years as a consequence of careful control, the battery in a car can expire at the most inopportune moment before its mission (the car's lifetime) is completed!

This section has discussed how the power subsystem achieves the vital role of keeping all the electrical spacecraft systems alive. We now turn to the basics behind spacecraft communications.

The Communications Subsystem

As we recall from Chapter 7, the prime functions of the communications subsystem are to provide a communications link with the ground, to downlink payload data and telemetry, and to uplink commands to control the spacecraft. This is a fairly intuitive task involving the requirement for two-way communication between the spacecraft and the ground. Any activity in orbit, such as taking measurements of the space environment or imaging a galaxy, requires a means of communicating the data to the ground, otherwise there's no point to doing it!

Communications Frequencies

Information on a satellite communications link is carried by electromagnetic (EM) waves; Figure 6.2 in Chapter 6 illustrated the various parts of the EM spectrum. As a consequence, the speed of communication is the speed of light, which is around 300,000 km per sec (186,000 miles per second), so that communication with spacecraft in LEO is effectively instantaneous. However, for a communication satellite in GEO, the altitude of the satellite is around 38,000 km, so that EM waves take just over a tenth of a second to travel from the ground to the spacecraft. This may not seem a lot, but bear in mind that for me (in the United Kingdom) to hold a telephone conversation with someone in the United States requires four such trips for the EM waves; my voice needs to travel up to the satellite, and then down to a ground station in the U.S. My friend's response then needs to make the same return trip, requiring about half a second of travel time. If I talk with someone in Australia, the communications route may involve more than one satellite, so the travel time can be large enough to produce awkward pauses that can disrupt the flow of a conversation. Of course, for interplanetary spacecraft the distances are such that a two-way conversation is not an option; for example, the travel time one-way for EM waves to a spacecraft at Saturn is at least an hour and a quarter.

The wavelength of EM radiation used for satellite communications is between 2 and 30 cm, which is referred to as the microwave part of the EM spectrum (see Figure 6.2). This is also the part of the spectrum used by microwave ovens to heat up dinner in the evening. This type of cooker heats food by bombarding it with EM radiation with a wavelength of typically about 12 cm. Fortunately, the microwave beam from a large satellite communications dish antenna at a ground station is well focused along the axis of the dish and is usually pointed skywards, so it is not a health hazard. Figure 9.10 reviews how the wavelength of EM radiation is defined, but it also shows some other important features. The intensity of the radiation is governed by the *amplitude* of the wave—or the wave height. In terms of the visible part of the spectrum, for example, a bright light has a larger amplitude than a dim one. The figure also shows how the phase of an EM wave is defined. This is measured in degrees, and indicates where you are on the wave along its wavelength; for example, the leading edge of the wave is defined to be at 0-degree phase, the first crest is at 90 degrees, the trough at 270 degrees, and so on. We will see later why this feature of the wave is important.

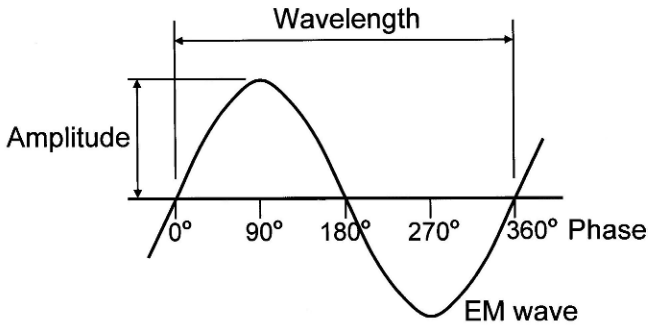


Figure 9.10: The wavelength, amplitude, and phase of an EM wave (see text).

Another aspect of communications is that it is common to talk about the part of the spectrum in terms of *frequency*, rather than wavelength. For every wavelength of EM radiation, there is a corresponding frequency. Generally speaking, short wavelength radiation has a high frequency and long wavelength radiation has a low frequency. This focus on frequency can be seen by looking at a domestic FM stereo radio. A radio station may be listed on the tuning dial as having a frequency of, say, 100 MHz, where the "M" is an abbreviation for "Mega," meaning a million, so we have a frequency of 100 million Hz. "Hz" is short for Hertz, which means cycles per second, named in honor of Heinrich Hertz, a German physicist who made important contributions to electromagnetism. Thus the frequency of the radio station is 100 million cycles per second but maybe this still doesn't mean much to you. Another way of thinking about this is that the wavelength of this signal is such that 100 million wavelengths pass the radio every second traveling at the speed of light. Given that the speed of EM waves is constant, for this to happen a simple calculation shows that the wavelength of a 100-MHz signal must be 3 m.

Satellite communications occur at a higher frequency, and so have a shorter wavelength. Typically the frequencies used are between 1 and 15 GHz, where the "G" stands for "Giga," meaning 1000 million. So the frequencies used are between 1000 million and 15,000 million cycles per second, which (as mentioned above) correspond to a wavelength range of 30 cm to 2 cm, respectively. Why are these particular frequencies used? This choice is dictated by the physics of the atmosphere. For a ground station to talk to a spacecraft, the EM waves must necessarily pass through Earth's atmosphere. However, for frequencies less than about 1 GHz, the energy of the radiation is sapped by charged particles, such as electrons, in the ionosphere. The ionosphere is a region of Earth's atmosphere at heights greater than about 80 km where the atmosphere's molecules of oxygen, nitrogen, etc. are stripped of their electrons by the Sun's ultraviolet radiation. For frequencies higher than 15 GHz, the radiation is absorbed by molecules of water vapor and oxygen in the lower part of the atmosphere. Thus the 1- to 15-GHz frequency range provides a convenient "window" through which the ground station can talk to the spacecraft, and vice versa.

Digital Communications

Another feature of satellite communications is that they are predominantly digital. In terms of satellite communications, *digital* means that effectively all the information in the communications link is converted into a string zeros and ones, each "0" and "1" being referred to as a bit (binary digit) of information. This *binary digital language* is the same as that used by your desktop computer to perform its routine internal operations and to communicate with other computers across worldwide digital networks. In recent years, the use of this digital technique has increased, being applied to television, radio, photography, music, and more. In these areas, and in space communications, the advantage of digital methods is that they produce a better quality of sound or picture, as it is generally easier to distinguish a digital signal from the various sources of interference which compete with it. It is amazing to think how such huge communications and consumer electronics industries can be built on the use of a language that is fundamentally made up of just zeros and ones!

Satellite Telephone Communications

To describe how satellite communications work, let's look in more detail at the process of making an intercontinental telephone call using a GEO communication satellite system. Just as it was in the 1870s when Alexander Graham Bell first invented it, the telephone receiver is an *analogue* device. It operates using continuously varying physical quantities, such as electrical current, without a single digital bit 0 or 1 to be seen anywhere.

In fact, making a telephone call many years ago was a completely analogue process. In this process, the voice produces pressure waves in the air—sound waves—that impinge on a small circular metal disk in the telephone mouthpiece. These waves then cause the disk (sometimes called a diaphragm) to vibrate in sympathy with the voice. Attached to the diaphragm is a lightweight coil of wire, which is positioned adjacent to a permanent magnet. As the wire coil moves up and down with the diaphragm in the magnetic field, a current is induced in the wire (see Chapter 6 for a brief explanation of electromagnetic induction), and this current can be thought of as an electrical representation of the information contained in the voice, that is, speech. This electrical version of the voice then propagates

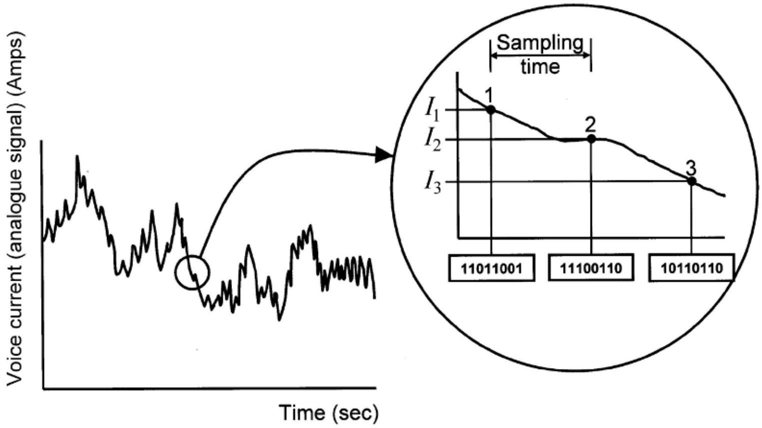


Figure 9.11: A rapidly varying and complex analogue signal is represented on the left. A small section of this is blown up on the right, with the time axis stretched, showing the process of conversion of the analogue signal to a digital one.

down the cable to the destination telephone handset, where the voice current is passed through the diaphragm wire coil of its earpiece. The combination of the voice current in the coil and a magnet in the earpiece causes the diaphragm to move in sympathy with the current. This movement in turn produces sound waves in the air, re-creating an understandable representation of the voice for the recipient of the telephone call.

Because satellite communication is a digital process, at some point the analogue signal—in this case the voice current—needs to be converted into a sequence of zeros and ones to produce a digital signal. This process of converting an analogue signal into a digital one is referred to as *digital encoding* (the method of encoding described below is not currently the most commonly used technique, but it is probably one of the easiest to understand). How can we convert a rapidly varying voice current into a string of 0 and 1 values suitable for transmission to a satellite? The left-hand side of Figure 9.11 is a representation of the analogue signal produced by a telephone, that is, the variation of electrical voice current produced by the telephone mouthpiece. In general, this is a complex and rapidly varying signal, and at any precise moment in time it will have a particular numerical value. The first step, then, is to convert decimal numbers, representing the value of the current at a particular time, into binary numbers, which are simply a sequence of zeros and ones. Table 9.2 shows how the first eight numbers, including zero, can be written in binary digital language. To do so, we need a string of three bits (zeros or ones). The number of zeros and ones in the string required to represent a particular decimal number can be

Table 9.2: The first eight decimal numbers (including zero) represented in binary as strings of three zeros or ones

Decimal number	Binary number		
	2^0	2^1	2^2
	1	2	4
0	0	0	0
1	1	0	0
2	0	1	0
3	1	1	0
4	0	0	1
5	1	0	1
6	0	1	1
7	1	1	1

calculated by thinking about powers of 2. Since 8 is 2^3 ($2 \times 2 \times 2$), we require a string of three 0 and 1 values. Continuing in this way, the first 16 (2^4) decimal numbers require a string of four bits, the first 32 (2^5) require a string of five bits, the first 64 (2^6) require six bits, and so on. This is great for people like myself who have quite a few years behind them—instead of having to put fifty something candles on your birthday cake, you only need six candles to write your age in binary. Of course, you do need candles of two colours to distinguish the "0" and the "1". With these kind of jokes on offer you can see that, in my household, birthdays are a whole lot of fun!

As shown in Table 9.2, in binary digital language each bit represents a power of 2, so under the heading "Binary number" we have powers of 2 (2^0, 2^1, 2^2), which take the values of 1, 2, and 4, respectively. We do not often come across a number to the power 0, but any number to the power 0 (e.g., 2^0) takes the value 1. So for each decimal number, we put a 1 where the power of 2 contributes and a 0 where it does not. Here are a couple of examples: on Table 9.2, the decimal number 5 can be written in binary digital language as 1 0 1, because 5 in simple arithmetic is $(\mathbf{1} \times 1) + (\mathbf{0} \times 2) + (\mathbf{1} \times 4)$. Similarly, 6 is 0 1 1, because 6 is $(\mathbf{0} \times 1) + (\mathbf{1} \times 2) + (\mathbf{1} \times 4)$. Thus decimal numbers can be represented as strings of zeros and ones. This strange binary language is what a desktop computer is using all the time to perform its mathematical routines.

In Figure 9.11, the analogue signal—the voice current—is shown on the left-hand side, and we need to convert this into a long sequence of zeros and ones suitable for transmission to the satellite. To show how this is done, we have magnified a small part of the analogue signal on the right-hand side of the figure. At a particular moment, at point 1, the voice current takes a particular value I_1, and this value can be converted into binary digital language. It is common to convert it into a string of eight bits, which means that the voice current value at this point can be assigned to any one of 256 (2^8) possible levels. Then a small fraction of a second later, called the *sampling time*, the value I_2 of the voice current is measured again, and it too is converted to a string of 8 bits. This process continues throughout the telephone conversation, converting the voice current into a long string of 0 and 1 values, as shown in the boxes beneath each point. It is usual to sample the voice current about 8000 times a second (so that the sampling time is 0.000125 second). This short sampling time is needed so that all of the complex and rapidly varying detail of the original voice current is captured in the digital signal. Using this method, the data rate of one digital telephone voice is around 8 $\times$ 8000 or 64,000 bits per second—8 bits produced 8000 times each second. This is referred to as 64 kbps, where “k” stands for kilo, meaning a thousand. Computer-savvy readers will be familiar with data rates; for example, a broadband Internet connection might be, say, 5 Mbps, where the “M” stands for Mega, meaning a million.

There is still another step in the process of transmitting a voice to the satellite. The digital bit stream representing a voice now needs to be somehow put onto an EM wave so that it can be transmitted from the ground station antenna to the spacecraft. This step in the process is referred to as *modulation* (Fig. 9.12). Modulation entails putting the information

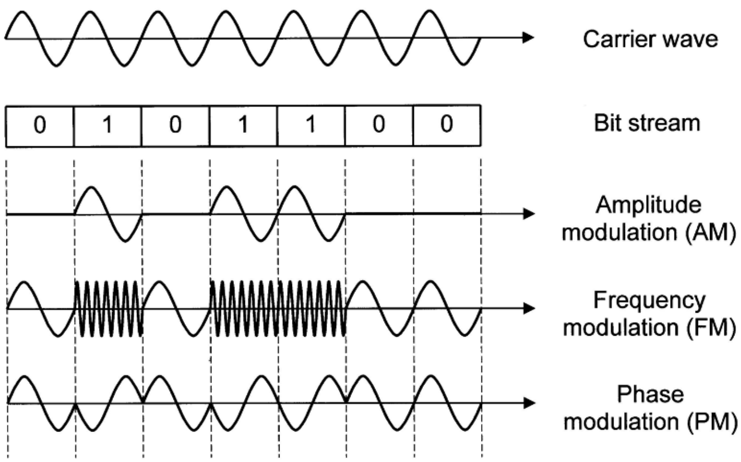


Figure 9.12: Three types of modulation, AM, FM, and PM, are used to put the digital bit stream onto the carrier wave.

contained in the digital bit stream onto a *carrier wave*, which can then be transmitted to the satellite. The carrier wave's job is simply to carry the information from ground to spacecraft. Initially, a carrier wave is an EM wave with a single frequency, as shown at the top of the figure. The next layer of the diagram shows the digital bit stream that needs to be carried by the wave, and there are three main ways of doing this. The first is *amplitude modulation* (AM), which is shown in the third layer in the figure. The amplitude of the wave is varied depending on the value of the bit; in this case, the amplitude is zero when a 0 bit is being carried, and nonzero when a 1 bit is transmitted. For *frequency modulation* (FM), as shown in the fourth layer of the diagram, the wave frequency is varied. When a 0 is carried, the frequency is low (long wavelength), and when a 1 is carried the frequency is high (short wavelength). Most conventional radios have AM and FM bands on their tuning dial. The third type, *phase modulation* (PM), is shown by the bottom line of the diagram. In this case, every time the value of a bit changes from a 0 to a 1, or vice versa, the phase of the carrier wave is changed by 180 degrees (see information on phase in Figure 9.10). In digital space communications, this is the most commonly used form of modulation.

When the carrier wave is received at the destination ground station, the process described above needs to be reversed so that the person at the other end of the phone can hear what the speaker said. The bit stream needs to be recovered by a process of *demodulation* of the carrier wave, and the analogue signal—the voice current—needs to be recovered by *decoding* the digital bit stream. It does seem to be a complex process, but it is all done routinely everyday, without anyone noticing!

It is also important to realize that there are thousands of these telephone conversations going on at the same moment, all of them sharing the same uplink to the spacecraft. To prevent one telephone conversation from bumping into another, each has a separate carrier wave at slightly different frequencies. So typically the ground station transmits a great wodge of carrier waves with frequencies ranging from, say, 6 to 6.5 GHz. On receipt of the uplink signal, the job of the satellite is to amplify the signal, change its frequency to, say, 4 to 4.5 GHz, and then retransmit it on the downlink to the destination ground station. The amplification is necessary as the signal strength after the uplink journey is small, about 10^{-8} W (or 0.000 000 01 W), a tiny fraction of the output of a domestic light bulb! The change in frequency is needed because the spacecraft usually uses the same antenna for both uplink and downlink, and the frequency shift prevents the information in the two links from becoming scrambled.

The Communications Subsystem

The focus of our discussion so far has been on telephone communications, but space communication is not just about this. For example, images taken by interplanetary spacecraft or Earth-observation satellites are commonly conveyed by space communications links, and this is again a digital process. The cameras on spacecraft use digital photography, so the pictures already come in a handy digital form, allowing the same type of method that we described above to transmit them from the spacecraft to the desktop computers on the ground.

The design of the spacecraft's communications subsystem is strongly influenced by the ground communications systems that it will talk to during the mission. Figure 9.13 shows typical spacecraft and ground communications antennas, used by GEO communication satellites. The main job of the communications subsystem engineer is to determine the size of the spacecraft antenna and the amount of power it needs to radiate to achieve a good-quality link. The physical characteristics of the overall (spacecraft and ground) system play a major part in this design process, and these include the ground-to-spacecraft range, the frequencies used, the size and power of ground antennas, and the losses caused by atmospheric absorption.

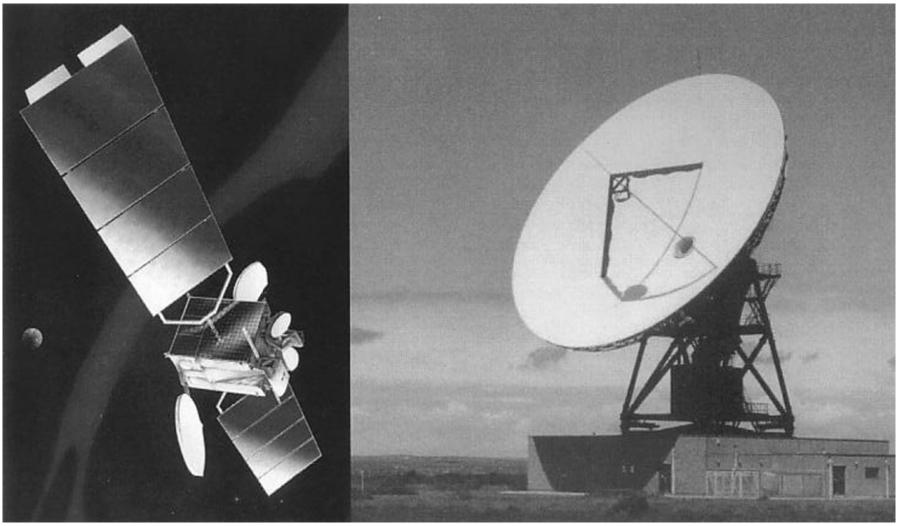


Figure 9.13: Spacecraft and ground communications antennas are shown, typical of those used for a GEO communication satellite mission. The ground station picture was taken at the Goonhilly Down ground station during a family holiday in Cornwall. (Spacecraft image courtesy of Lockheed Martin.)

The two most important attributes of the spacecraft communications subsystem are *radiated power* and *gain*. If you use a mobile phone, you have some idea of what the radiated power of a communications system is all about; they have a little scale indicator on the screen to tell the user if there is a signal. If there is one, then somewhere nearby there is a mobile phone mast that has been designed to have sufficient radiated power to cover a particular area so that calls can be made and received. Further away from the mast, the signal strength will drop, so another mast has to provide ground coverage area there so that the service can be maintained. In the same way, an important measure of the effectiveness of a spacecraft's communications subsystem is the amount of radiated power it generates. However, a number of watts of radiated power is not the end of the story; the effectiveness of the system can be further enhanced if the spacecraft's communications antenna (usually dish-shaped) has high gain. The simple rule is that a large dish has high gain, whereas a small dish has low gain. We can understand the idea of gain if we consider the small light bulb inside a flashlight. The amount of radiated power it generates, in terms of the amount of light it gives off, is small, and certainly much smaller than that of a standard domestic light bulb. If we take the small bulb out of the flashlight and connect it to a battery, it provides insufficient light in a darkened room. However, if we put the bulb back into the flashlight and turn it on, the flashlight's dish-shaped reflector focuses the light from the small bulb into a beam, which can be dazzling if pointed directly into your eyes. The flashlight's reflector has a measure of gain, effectively increasing the radiated power of the bulb along the axis of the flashlight where the beam is concentrated.

This is very much like a dish antenna on a spacecraft. The radiated microwave power generated by the spacecraft is focused into a beam by the dish, and this beam can then be pointed at a receiving dish on the ground. This has the effect of increasing the received power on the ground. To achieve the level of received power on the ground required to ensure a good quality of link, spacecraft designers have a choice: they can achieve the required level either by having a small amount of radiated power and a large gain (a big dish), or by having lots of radiated power and a small gain (a small dish). This *power-gain tradeoff* is one of the main design issues for the spacecraft communications subsystem engineer. This issue affects spacecraft design, in particular that of interplanetary spacecraft that travel to distant parts of the solar system. For example, probes such as Cassini/Huygens (see Fig. 7.6 in Chapter 7), and the New Horizons spacecraft (Fig. 9.14) recently launched to investigate Pluto, look almost like flying communications dishes. At great distances from the Sun, generating large amounts of electrical power onboard the spacecraft is difficult, so the

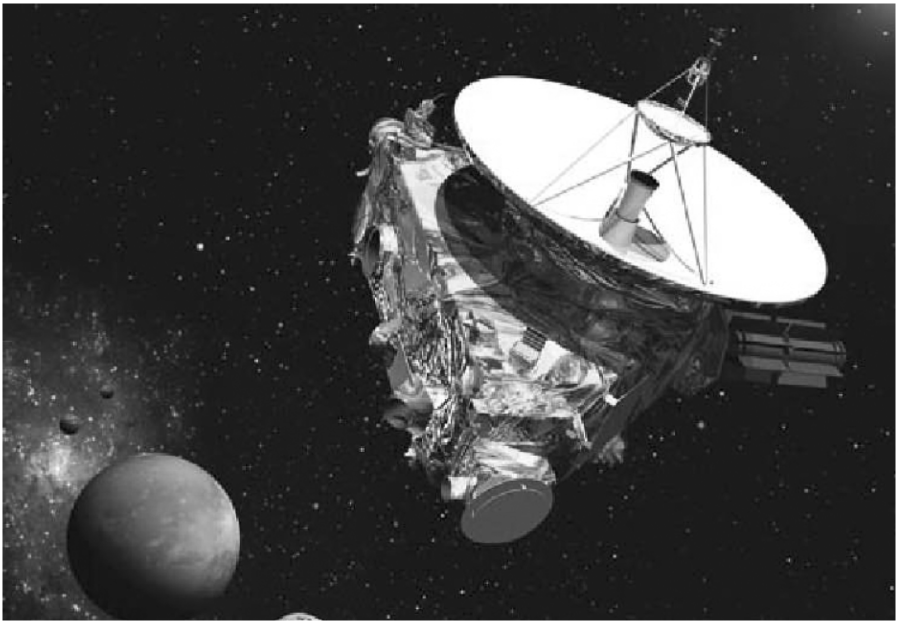


Figure 9.14: An artist's impression of the New Horizons spacecraft at Pluto. (Image courtesy of NASA.)

communications subsystem is likely to have low radiated power. To achieve the link quality required to return science data across such great distances, these spacecraft require large, high gain antennas.

Let's move on now and briefly discuss the subsystem that is tasked with the job of moving all the binary digits around the spacecraft.

The On-Board Data Handling (OBDH) Subsystem

As we recall from Chapter 7, the main functions of the OBDH subsystem are to provide storage and processing of payload and other data, and to allow the exchange of data between subsystem elements. As this brief description suggests, the subsystem is made up mostly of computers, their peripherals, and software. The OBDH subsystem is distributed throughout the vehicle, to make possible the data links between all the spacecraft subsystems needed for successful operation, such as the ACS control loop we discussed in Chapter 8 (see Figure 8.1), and the transfer of payload data from the payload equipment to the communications subsystem ready for downlink to the ground. Although the OBDH subsystem is a virtual one, in the sense that it consists mostly of computer processors and programs, nevertheless it is perhaps the most real part of the spacecraft for the operators as it is the part that they talk to. This interaction with the ground takes place in two directions—through the uplink, which is mostly *command*, and through the downlink, which is mostly *payload data* and *telemetry*.

The Command Function

It is through the command uplink that the ground operators talk to the spacecraft to get it to do something. Commands are received, interpreted, and distributed by the OBDH subsystem. The commanded tasks may be as simple as, say, switching a battery heater on, or may involve a more complicated job, such as repointing a space observatory to a new place in the sky, or commanding an Earth-observation satellite to image a particular ground target of interest. This command function is important, as it allows the ground team to control all aspects of the spacecraft's operation. As such, it must be done in a reliable way; for example, the uplinked commands must be verified, to ensure that they have been received correctly. The OBDH subsystem must then provide confirmation that the uplinked command instructions have been carried out correctly, which is usually done through the use of telemetry (see below). All this sounds rather tedious, but some very expensive spacecraft missions have been lost simply due to incorrect or unverified commands being executed on board the spacecraft.

The Payload Data and Telemetry Functions

Perhaps the most important role of the OBDH is to ensure that data generated by the payload are transferred to the communications subsystem ready for downlinking to the ground. In some cases, these payload data need to be processed by the OBDH subsystem computer(s), and this processing involves things like *storage*, *error detection and correction* and *compression* (see below). For some spacecraft, such as communication satellites or those with imaging payloads, the volume of data produced by the payload can be large. For LEO spacecraft, sometimes a ground station is not in view as the payload data are being generated, and so the data must be stored on board while waiting for an opportunity to downlink to a ground station. These storage devices are part of the OBDH subsystem, and in the past tape recorders were commonly used. More recently, *solid-state memories*, like those in a desktop computer, have been used instead. They provide huge amounts of data storage, up to hundreds of Gbits (where the "G" stands for "Giga," meaning 1000 million), but they are prone to radiation-induced errors, as mentioned in Chapter 6. To minimize the effects of such data errors, the memory is continuously scanned by error detection and correction computer programs, as part of the OBDH subsystem function. In some cases, the data volume generated by the payload is too large for the downlink communications to handle, and so the data need to be compressed using OBDH software. The raw payload data can be compressed (or reduced in volume) in such a way as not to compromise the value of its content too much. This is done by eliminating redundant or duplicated content, removing unwanted information, or reducing the resolution of digital imagery.

The other main job of the OBDH subsystem is to generate telemetry, which we discussed briefly in Chapter 7. Various sensors are placed around the spacecraft, to monitor the health and operating status of the onboard equipment. They check the temperatures of electronic equipment, the pressures in fuel tanks, voltages and currents in power supplies, and so on. The operational status of equipment is also monitored in terms of whether items are switched on or off. All these accumulated data are converted into a digital bit stream, which is then downlinked with a data rate of a few kbps (kilobits per second) to the ground for display on computer monitors in the operations room. In this way, any problems that occur on the spacecraft can be quickly spotted and corrected by the operations team.

The Thermal Control Subsystem

For the majority of spacecraft in orbit, the thermal control subsystem represents only a small percentage of the spacecraft mass, yet remarkably it generally dominates the overall appearance of a spacecraft. It is what you see! To illustrate the point, Figure 9.15 shows Earth-observation spacecraft SPOT 5 prior to launch (see Table 7.4 in Chapter 4). It looks a bit like a chocolate box wrapped with gold and silver foil, and this “foil” is the *multilayered insulation blanket* that I have mentioned briefly in previous chapters. There are also other features visible, such as mirror-like surfaces and white-painted areas and antennas, and these are all characteristic of a typical thermal control subsystem design. The same features can be seen in many of the pictures of spacecraft in earlier chapters, with the thermal design dominating their appearance. In this section, we have a look at why this is.

As we recall from Table 7.1, the main job of the thermal control subsystem is to provide an appropriate thermal environment onboard, to ensure reliable operation of payload and subsystem elements.

Figure 9.15: The SPOT 5 Earth-observation satellite. The thermal control subsystem design governs the way the spacecraft looks. (Image copyright © CNES/Patrick Dumas.)

Table 9.3: Approximate temperature range for reliable operation of equipment

Component	Approximate temperature range (°C)
Batteries	0 to 25
Propellant (e.g., Hydrazine)	10 to 50
Electronic equipment (e.g., computer processors)	–5 to 40
Mechanical bearings (e.g., reaction wheels)	0 to 45

Equipment Reliability

The average spacecraft is crammed with electronic and mechanical equipment so that it can achieve its mission. Most of this equipment, especially the electronics, has been developed from similar domestic and industrial items that were designed to operate down here on the ground. In other words, they have a *design heritage* such that they work best when they operate at room temperature. One main reason for requiring thermal control onboard a spacecraft is to produce a room temperature environment for the various equipments so that they operate reliably for the lifetime of the spacecraft mission. It is the same with home electronics. For example, if we put our TV set in the freezer or in the oven, its useful lifetime will be considerably reduced. Similarly, reliable, long-term performance of most spacecraft components requires them to operate within thermal tolerances. Table 9.3 shows the approximate temperature range for the reliable operation of some items of spacecraft equipment, which the thermal control engineer has to consider when designing the thermal control subsystem. (The definition of room temperature is quite broad.)

Payload Requirements

As we have already mentioned in Chapter 7, certain payloads may have to be maintained within a strict temperature range to work properly, which will affect how the thermal control subsystem is designed (see Figure 7.1). To illustrate, we can use the example of a spacecraft that has a large imaging payload on board, such as an Earth-observation satellite or a space observatory. In both types of spacecraft the imaging equipment is made up of a variety of mirrors and lenses, and their job is to bring the light entering the payload to a focus in the right place so that the image can be recorded in a way similar to that done by a digital camera. To get the image in focus, the lenses, mirrors, and detectors all need to be kept at the right distance from one another. To do this, the optical components are all firmly mounted on a rigid framework, usually referred to as an *optical bench*, within the spacecraft. This framework needs to be a robust structure so that the imager still works after the rough ride to orbit on the launch vehicle (see Chapter 5). But another feature of importance is its thermal design. When in orbit, the spacecraft is exposed to extremes of temperature, and if unprotected from these, the optical bench framework will expand and contract in size in response to changes in temperature. From the point of view of the image quality, this small, relative movement among the mirrors, lenses, and detectors is clearly undesirable. Consequently, it is the job of the thermal control engineer to design the spacecraft so that the payload is isolated from these extremes of temperature, in order to ensure that it works. In large observatories, such as the Hubble Space Telescope (see Table 7.5 for details), this job can be challenging.

The In-Orbit Thermal Environment

To keep the temperature of the spacecraft and its components within the required range, the thermal control engineer has to consider the factors that heat the spacecraft and those that cool it. The engineer strives to achieve a balance, so that the spacecraft does not get too hot or too cold. If we focus on a spacecraft in Earth orbit, we can summarize its thermal environment in terms of heat inputs (factors that heat it) and heat outputs (factors that cool it).

Heat Inputs

The mechanisms that heat the spacecraft are shown in Figure 9.16. As we mentioned previously in Chapter 6, the major input is that due to direct solar radiation, that is, the direct electromagnetic radiation from the Sun. This amounts to about 1.4 kW of thermal power for every square meter of spacecraft area that is presented to the Sun. In addition, the spacecraft receives solar radiation that is reflected from Earth's surface. About one third of all the sunlight that falls on Earth is reflected back out into space, mostly from cloud and ocean surfaces. This is referred to as *Earth albedo radiation,* and this too heats the spacecraft. Another source of heat, referred to as *Earth heat radiation*, is the infrared radiation (see Figure 6.2) produced by Earth simply because it is a warm body. Sometimes we are very much aware of heat (infrared) radiation, such as when we sit in front of a glowing open fire and feel the heat on our faces. However, such heat radiation is produced by all objects to a greater or lesser extent depending on their temperature. This is true as long as the object's temperature is above *absolute zero.* On the Celsius temperature scale, absolute zero occurs at –273° and is so called because it is

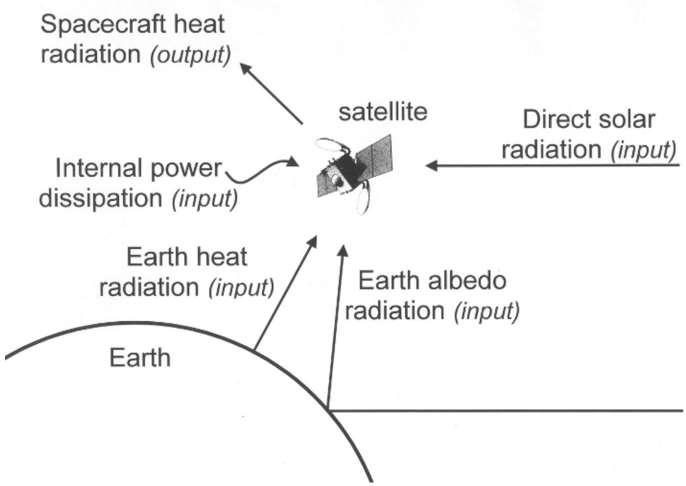


Figure 9.16: A summary of the spacecraft's thermal environment, in terms of things heating it (inputs) and things cooling it (outputs).

the lowest temperature that is physically possible. At –273°C all physical processes stop.

Objects such as Earth and people give off heat radiation simply because their temperature is above absolute zero. Since Earth is an object with a temperature of around 20°C (on average), it produces a heat radiation field that gently warms the spacecraft. Both of these Earth-generated effects, Earth albedo and Earth heat radiation, provide a heat input to the spacecraft that decreases with the altitude of the spacecraft's orbit in accord with an inverse square law (see Chapter 1). The final mechanism that heats the spacecraft is not an external one but rather an internal effect referred to as *internal power dissipation*. Typically a spacecraft is packed with electrical and electronic equipment, and most of it is not very efficient in that a significant percentage (between 10% and 50%) of the electrical power that is fed into the equipment to sustain its operation is wasted in producing heat. This is not just a feature of spacecraft components; the same thing happens in domestic electrical equipment. For example, if we leave the television set on for a couple of hours, and then just place our hand (carefully) by the air vents at the back of the set, we can feel that not all the electricity has been used to produce picture and sound. Some of it is being wastefully dissipated as heat. Internally dissipated power is another significant mechanism that acts to heat up the spacecraft.

Heat Outputs

Unless there is some way of getting rid of the heat, the spacecraft is going to become too hot, and the acceptable temperature ranges shown in Table 9.3 will be exceeded. However, as a warm body, the spacecraft itself will give off its own heat radiation, and the intensity of this radiation will increase as the spacecraft's temperature rises. This spacecraft heat radiation (see Figure 9.16) is the only effective thermal output acting to cool the spacecraft down.

Thermal Equilibrium

From the above discussion about the thermal environment, we can see that at a particular temperature, the heat outputs will match the heat inputs, and a kind of thermal equilibrium will be reached where the temperature of the spacecraft remains more or less constant. This temperature is referred to as the *equilibrium temperature*. It is the job of the thermal control engineer to design the thermal control subsystem to ensure that when thermal equilibrium occurs, the equilibrium temperature is around room temperature. If this can be achieved, then there is a good chance that all of the on-board components will be able to operate reliably for the lifetime of the mission.

Thermal Control Design

How does the thermal control engineer achieve this objective? In the introduction to this section we said that the thermal control subsystem is what you see, and basically this is the clue to how it is done. First we need to discuss the thermal properties of materials, that is, why some surfaces get hot in the Sun, while others stay cool. If we walk barefoot on a beach on a hot summer's day, some surfaces, such as the sand or black tarmac on the seafront promenade, get hot and scold the soles of your feet, whereas others, such as the wooden steps down to the beach, feel quite comfortable.

Along the same theme, I have a friend who used to work as a spacecraft thermal control subsystem engineer, and he had a neighbor who owned a large motor boat. But there was a problem with the boat: parts of the decking were made of stainless steel, and when exposed to the Sun on a hot weekend the temperature of the decking would rise to a level that was hazardous. My friend did a few simple calculations and estimated that the temperature of the stainless steel decking could reach temperatures in excess of 100°C! The solution was an easy one; my friend recommended that the offending parts of the deck be painted white. As we will see below, a white-painted surface is poor at absorbing the Sun's heat but good at giving off heat radiation. As a consequence the temperature of the deck went down to a comfortable room temperature, and the boatman was happy. Obviously, boat builders can learn something from spacecraft engineers!

Some of the spacecraft's surfaces are good at absorbing the Sun's heat, but not good at giving off heat (infrared) radiation when they get hot. An example of this kind of surface is aluminium; it is quite good at absorbing the Sun's heat, but not good at radiating it. An aluminium surface exposed to the Sun in Earth orbit can reach temperatures of around 300° to 400°C. On the other hand, some surfaces are poor at absorbing the Sun's heat, but are good at giving off heat radiation, such as a white-painted surface, which stays cool even when exposed to direct sunlight. If we apply a coat of white paint to our orbiting aluminium surface, its temperature will drop to around, say, 20°C. Obviously, the precise numbers here depend on the orbit, which determines how long the spacecraft is in direct Sun, and how long it is in Earth's shadow. These so-called thermal properties of the surface materials of the spacecraft are used to good effect by the thermal control subsystem engineer to manage the balance between heat inputs and outputs to ensure acceptable local temperatures and an appropriate overall equilibrium temperature—around room temperature—for the spacecraft. This is why spacecraft look the way they do, with various types of surface material to manage this thermal balance.

In Figure 9.15 we can see a white-painted communications dish to ensure it stays cool in direct sun light. However, the main feature of the thermal design in this case is the extensive use of thermal blanket to insulate the spacecraft from the direct heating effect of solar radiation. This often gives the spacecraft its characteristic appearance of being wrapped in gold or silver foil. The blanket, sometimes referred to as *multilayered insulation* (MLI), is composed of multiple layers of a thin plastic film with a metallic coating of aluminium, silver, or gold. Each individual layer is similar to the survival blankets handed out at the end of marathons to keep the runners warm. Figure 9.17 shows a small section of MLI, which is made up of around 25 such layers. Each layer is perforated with tiny holes, to ensure that the air trapped in the blanket on the ground can escape easily when the spacecraft reaches the vacuum of orbit. Each layer of the MLI is separated from the next by a thin sheet of nylon bridal veil. In orbit, with a vacuum between each separate layer, the effectiveness of the insulation provided by the blanket is maximized.

However, if we were to wrap the spacecraft completely in multilayered insulation, the heat dissipated by the electrical equipment on board would not be able to escape easily, and the interior of the spacecraft would become hot. So sections of MLI in Figure 9.15 are cut away, and instead radiator surfaces are installed, which are usually mirror-like surfaces that are designed to be poor absorbers of the Sun's heat but good at giving off heat (infrared) radiation. As a result they remain cool, even in direct Sun, and encourage the escape of internally dissipated heat. Usually, the electrical

Figure 9.17: A section of multilayered insulation, with one layer folded back to reveal the nylon bridal veil spacer. Each layer is perforated with small holes, spaced at approximately 1-cm intervals, to allow trapped air to escape from the layers.

Figure 9.18: A lesson from nature in thermal control: polar bears displaying their insulation blankets and thermal radiator surfaces.

devices that produce the most heat are mounted on the interior surfaces of these radiators to keep them cool.

Like many aspects of engineering, this arrangement is a lesson from nature, and is similar to the way a polar bear, for example, manages its thermal control! A polar bear (Fig. 9.18), like a spacecraft, has to survive in a hostile thermal environment. To keep warm in the cold polar regions, it needs good insulation in the form of its fur coat. On the other hand, it does need some effective radiator surfaces, so that it does not overheat when it exerts itself physically or when summer comes. So it has radiators too, such as the pads on its feet, its shiny black nose, and its tongue, to allow its internal heat to be radiated away. Like the spacecraft, it must have enough insulation to keep warm when its environment gets cold, but enough radiator area to cool down when the environment gets too hot.

The Structure Subsystem

The function of the spacecraft structure is to provide a rigid framework to carry all the various payload and subsystem equipment. As we recall from Table 7.1, its function is to provide structural support for all payload and subsystem hardware in all predicted environments (especially the harsh launch vehicle environment). The important phrase here is "in all predicted environments." When structure subsystem engineers begins the process of designing a new spacecraft, the first thing they look at is the most severe of these environments—that of the launch vehicle. In Chapter 5, we saw how harsh this is, with high levels of acceleration, vibration, shock, and noise. The launch vehicle agencies provide detailed information on all these aspects of the ride to orbit, and the main job of the structure subsystem engineer is to produce a structural design that will survive the journey from launch pad to orbit.

Design Requirements

The following is a list of the important aspects that the structure subsystem engineer has to consider, many of which are related to the fundamental requirement to survive launch:

- Low mass. Although there is a need to build a robust structure, nevertheless the structure subsystem engineer must also make every effort to minimize mass. As we said in Chapter 5, the cost of the launch rises steeply as the mass of the spacecraft increases, so this requirement becomes a critical issue about limiting the overall cost of the spacecraft project, as launch is usually a large percentage of this cost.

- Strength and stiffness. The structure must be strong and stiff enough to withstand launch and on-orbit loads without distortion. Unacceptable levels of distortion would compromise the pointing of payload instruments, such as imaging cameras or telescopes, and the pointing of subsystem elements like communications dishes and attitude sensors.
- Environmental protection. The structure must also provide an appropriate level of protection against environmental aspects (see Chapter 6), such as radiation, and shielding against impacts of orbital debris and micrometeors.
- Launch vehicle interface. The manner in which the spacecraft is attached to the launch vehicle also affects the spacecraft's overall design. The attachment must be done in such a way as to ensure a robust connection with the launcher during the ascent to orbit, but it must also be able to reliably release the spacecraft on command once in orbit. The position of this interface within the spacecraft also governs how the launch loads are distributed throughout the spacecraft structure.

Materials

So what kinds of material are used to satisfy the demanding requirements related to having a robust structure, but nevertheless one that is of low mass? A material that is commonly used in spacecraft construction currently is aluminium honeycomb panel. It is composed of a sheet of aluminium honeycomb, bonded between two thin sheets (skins) of aluminium, as shown in Figure 9.19. The honeycomb is basically identical to bees'

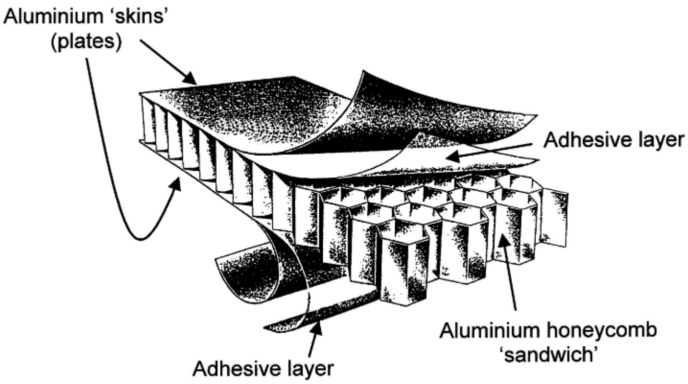


Figure 9.19: Aluminium honeycomb panel is a commonly used material in spacecraft manufacture. Thin aluminium skins are bonded onto a sandwich filling of aluminium honeycomb, making a very light, stiff panel.

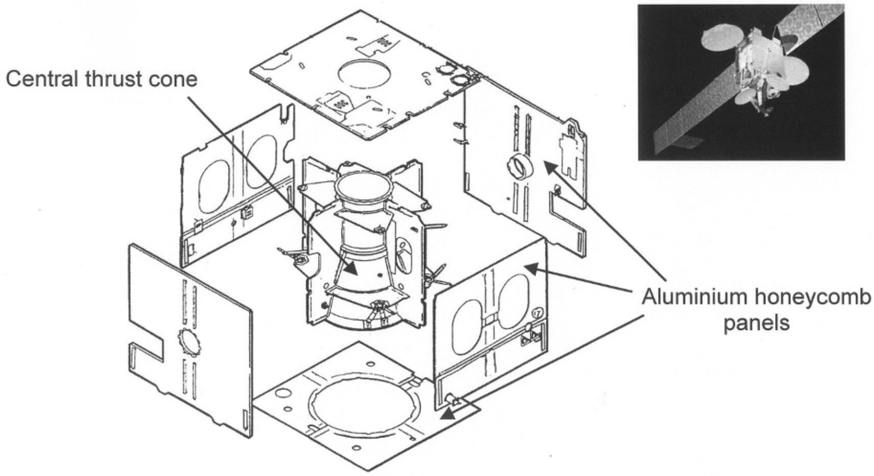


Figure 9.20: An example of a spacecraft structure, based on the Eurostar® configuration. The figure shows the central box-like structure, made up of a number of honeycomb panels. The thrust cone is the housing for the spacecraft's primary propulsion, and is also the position of the interface with the launch vehicle. (Images courtesy of EADS Astrium.)

honeycomb, but made of aluminium. A sheet of this seems fragile and floppy when handled, and can be easily crushed with a pair of pliers. However, when the aluminium skins are glued either side, making a honeycomb sandwich, the resulting panel material is very light and stiff, providing an ideal material for spacecraft manufacture.

An example of the use of this material is given in Figure 9.20, which shows the basic structure of the central box-like body of a communications satellite. The central thrust cone is the housing for the spacecraft's primary propulsion system, and also the location of the interface with the launcher. As such, this element takes most of the loads during the launch and when the main rocket engine is fired on orbit.

Summary

Figure 9.21 is an exploded view of a spacecraft that displays many of the subsystem elements that were discussed here, thus providing a brief summary of the contents of the last three chapters.

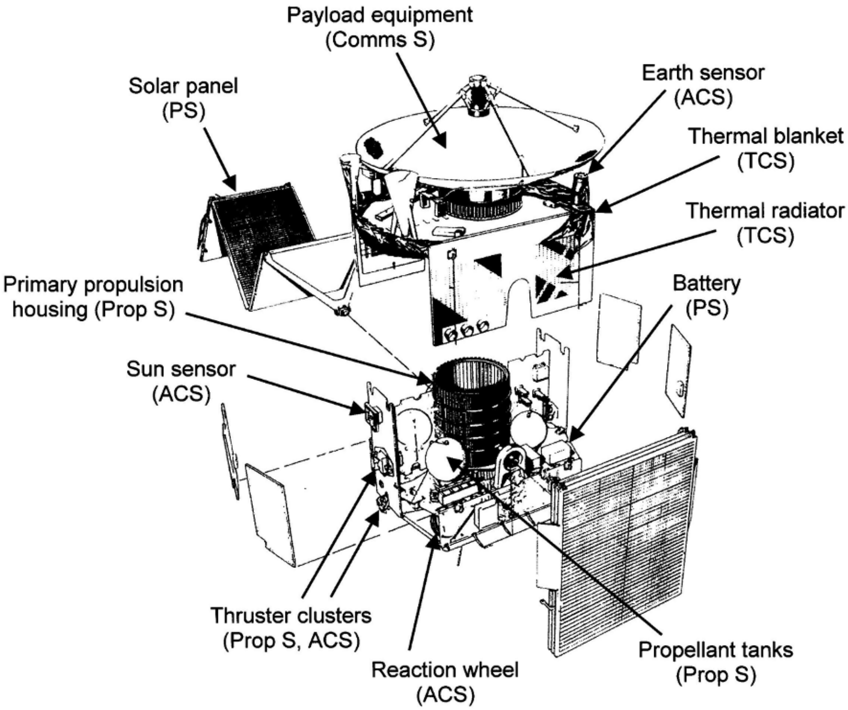


Figure 9.21: An exploded view of a spacecraft showing the various elements of the subsystem design. The example shown is that of a communications satellite, so in this case the payload is identified with the communications subsystem.
Key: ACS, attitude control subsystem; Comms S, communications subsystem; PS, power subsystem; Prop S, propulsion subsystem; TCS, thermal control subsystem. (Backdrop image courtesy of ESA.)