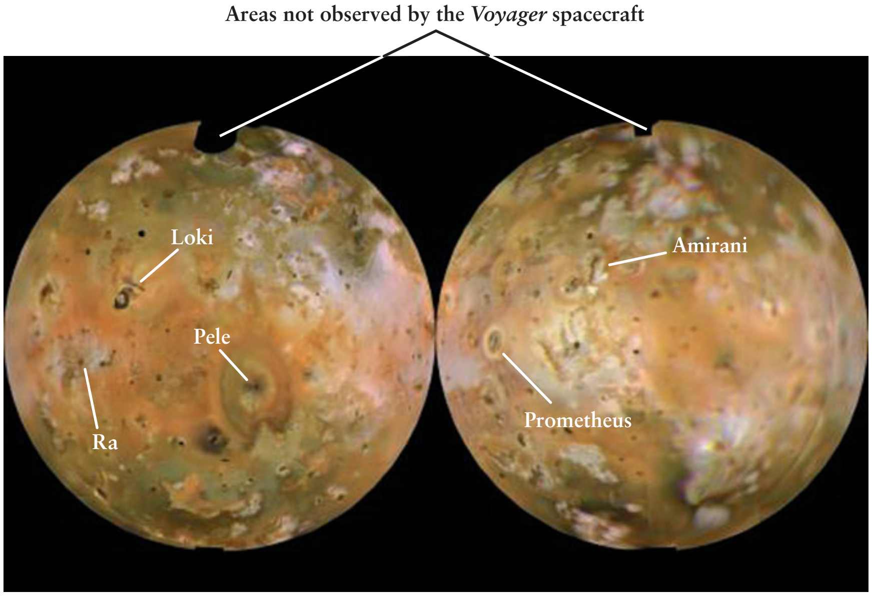
Figure 13-4 RIVUXG

Io These mosaics of Io's two hemispheres were built up from individual *Voyager* images. The extraordinary colors are probably caused by deposits of sulfur and sulfur compounds ejected from Io's numerous volcanoes. The labels show several of these volcanoes, which are named for sun gods and fire gods of different cultures. (NASA/JPL)



around its orbit, Europa and Ganymede exert gravitational tugs on it in a regular, rhythmic fashion, thanks to the 1:2:4 ratio among the orbital periods of these three satellites. These tugs distort Io's orbit into an ellipse, and so its speed varies as it moves around its orbit. Consequently, Io's long axis "nods" back and forth by about ½° as seen from Jupiter. Due to this "nodding," the tidal stresses that Jupiter exerts on Io vary rhythmically as Io moves around its orbit. These varying tidal stresses alternately squeeze and flex Io. Just as a ball of clay or bread dough gets warm as you knead it between your fingers, this squeezing and flexing causes **tidal heating** of Io's interior. (The Earth's Moon also "nods" as it orbits our planet, but experiences almost no tidal heating because the Earth's tidal forces are very weak compared to those of massive Jupiter.)

The theory of tidal forces helped predict Io's volcanic nature

Tidal heating adds energy to Io at a rate of about 10^{14} watts, equivalent to 24 tons of TNT exploding every second. As this energy makes its way to the satellite's surface, it provides about 2.5 watts of power to each square meter of Io's surface. By comparison, the average global heat flow through the Earth's crust is 0.06 watts per square meter. Only in volcanically active areas on Earth do we find heat flows that even come close to Io's average. Thus, Peale, Cassen, and Reynolds predicted "widespread and recurrent surface volcanism" on Io.

Io's Active Volcanoes

No one expected that *Voyager 1* would obtain images of erupting volcanoes on Io. After all, a spacecraft making a single trip past the Earth would be highly unlikely to catch a large volcano actually erupting. But, in fact, *Voyager 1* images of Io revealed *eight* different giant plumes of gas from volcanic eruptions. (Figure 13-5*a* shows two of these.) More recently, the *Galileo* spacecraft, which came as close as 200 km to Io's surface during its eight years in orbit around Jupiter, returned detailed images of several such plumes (Figure 13-5*b*). These observations resoundingly confirm the Peale-Cassen-Reynolds prediction: Io is by far the most volcanic world in the solar system.

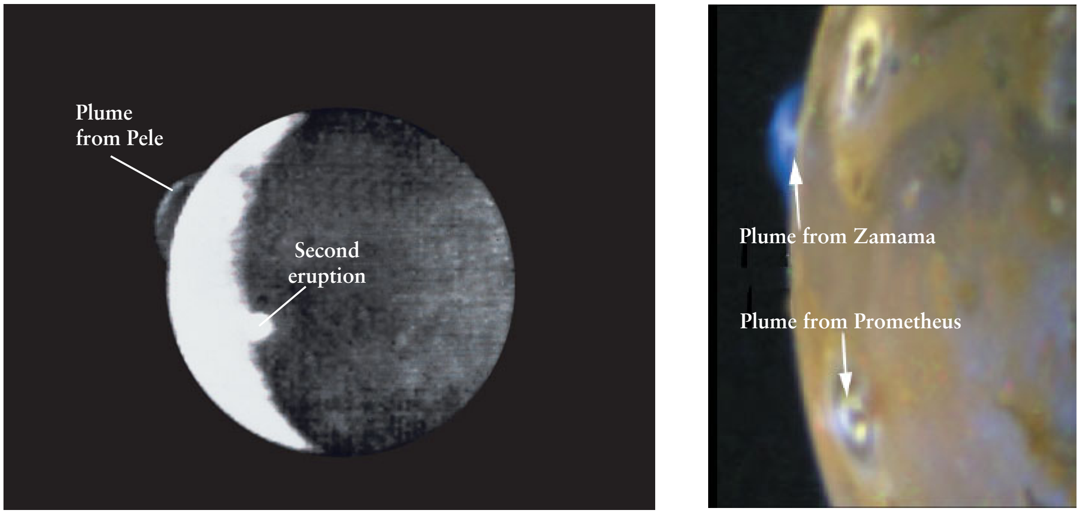
Io's volcanic plumes rise to astonishing heights of 70 to 280 km above Io's surface. To reach such altitudes, the material must emerge from volcanic vents with speeds between 300 and 1000 m/s (1100 to 3600 km/h, or 700 to 2200 mi/h). Even the most violent terrestrial volcanoes have eruption speeds of only around 100 m/s (360 km/h, or 220 mi/h). Scientists thus began to suspect that Io's volcanic activity must be fundamentally different from volcanism here on Earth.

The Nature of Io's Volcanic Eruptions

An important clue about volcanism on Io came from the infrared spectrometers aboard *Voyager 1,* which detected abundant sulfur and sulfur dioxide in Io's volcanic plumes. This led to the idea that the plumes are actually more like geysers than volcanic eruptions. In a geyser on Earth, water seeps down to volcanically heated rocks, where it changes to steam and erupts explosively through a vent. Planetary geologists Susan Kieffer, Eugene Shoemaker, and Bradford Smith suggested that sulfur dioxide rather than water could be the principal propulsive agent driving volcanic plumes on Io. Sulfur dioxide is a solid at the frigid temperatures found on most of Io's surface, but it should be molten at depths of only a few kilometers. Just as the explosive conversion of water into steam produces a geyser on the Earth, the conversion of liquid sulfur dioxide into a high-pressure gas could result in eruption velocities of up to 1000 m/s.

VIDEO 13.2

Io's dramatic coloration (see Figure 13-4) is probably due to sulfur and sulfur dioxide, which are ejected in volcanic plumes and later fall back to the surface. Sulfur is normally bright yellow, which explains the dominant color of Io's surface. But if sulfur is heated and suddenly cooled, as would happen if it were ejected from a volcanic vent and allowed to fall to the surface, it can assume a range of colors from orange



(a) *Voyager 1*, March 1979

(b) *Galileo*, November 1997

Figure 13-5 R I V U X G

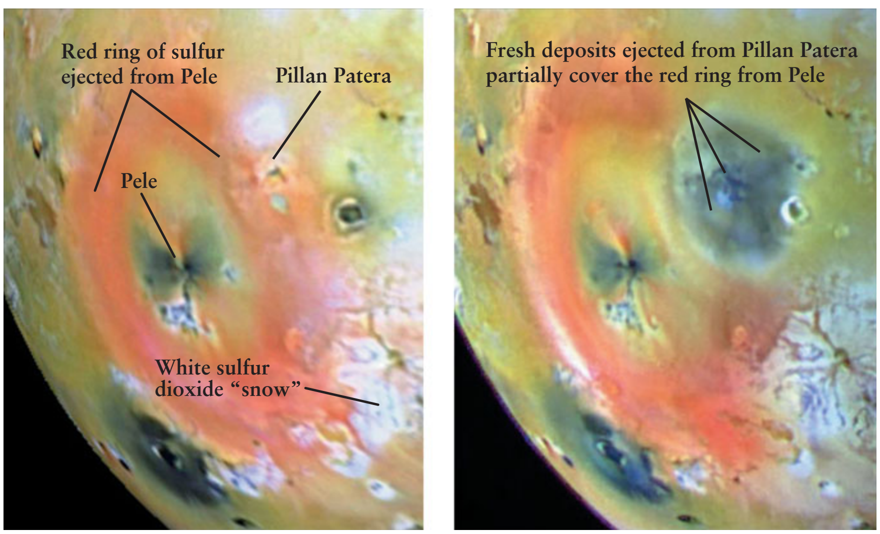
Volcanic Plumes on Io **(a)** Io's volcanic eruptions were first discovered on this *Voyager 1* image. The plume at upper left, which extends 260 km (150 mi) above Io's surface, came from the volcano Pele (see Figure 13-4). **(b)** This *Galileo* image shows plumes from the volcanoes Prometheus (see Figure 13-4) and Zamama. The plumes rise to about 100 km (60 mi) above Io's surface and are about 250 km (150 mi) wide. (a: NASA; b: Planetary Image Research Laboratory/University of Arizona/JPL/NASA)

and red to black. Indeed, these colors are commonly found around active volcanic vents (Figure 13-6). Whitish surface deposits (examine Figure 13-4 and Figure 13-6*a*), by contrast, are probably due to sulfur dioxide (SO_2). Volcanic vents on Earth commonly discharge SO_2 in the form of an acrid gas. But on Io, when hot SO_2 gas is released by an eruption into the cold vacuum of space, it crystallizes into white snowflakes. This sulfur dioxide "snow" then rains back onto Io's surface.

The plumes are not the whole story of volcanism on Io, however. The sources of the plumes appear in *Voyager* images as black spots 10 to 50 km in diameter, which are actually volcanic vents (see Figure 13-4). Many of these black spots, which cover 5% of Io's surface, are surrounded by dark lava flows. Figure 13-7 is a *Galileo* image of one of Io's most active volcanic regions, where in 1999 lava was found spouting to altitudes of thousands of meters along a fissure 40 km (25 mi) in length. The glow from this lava fountain was so intense that it was visible with Earth-based telescopes. Thus, Io has two distinct styles of volcanic activity: unique geyserlike plumes and lava flows that are more dramatic versions of volcanic flows on Earth.

While sulfur and sulfur compounds are the key constituents of Io's volcanic plumes, the lava flows must be made of something else. Infrared measurements from the *Galileo* spacecraft (Figure 13-8) show that lava flows on Io are at temperatures of 1700 to 2000 K (about 1450 to 1750°C, or 2600 to 3150°F). At these temperatures sulfur could not remain molten but would evaporate almost instantly. Io's lavas are also unlikely to have the same chemical composition as typical lavas on Earth, which have temperatures of only 1300 to 1450 K. Instead, lavas on Io are probably **ultramafic lavas.** These are enriched in magnesium and iron, which give the lava a higher melting temperature.

Solidified ultramafic lavas are found on Earth, but primarily in lava beds that formed billions of years ago when the Earth's interior was much hotter than today. The presence of molten



(a)

(b)

WEB LINK 13.5

Figure 13-6 R I V U X G

Rapid Changes on Io **(a)** This *Galileo* image shows a portion of Io's southern hemisphere. Sulfur ejected from the active volcanic vent Pele produced a red ring that resembles paint sprayed from a can. **(b)** A few months later, a volcanic plume erupted from a vent called Pillan Patera, spraying dark material over an area some 400 km (250 mi) in diameter (about the size of Arizona). Other than Earth, Io is the only world in the solar system that shows such noticeable changes over short time spans. (NASA/JPL)

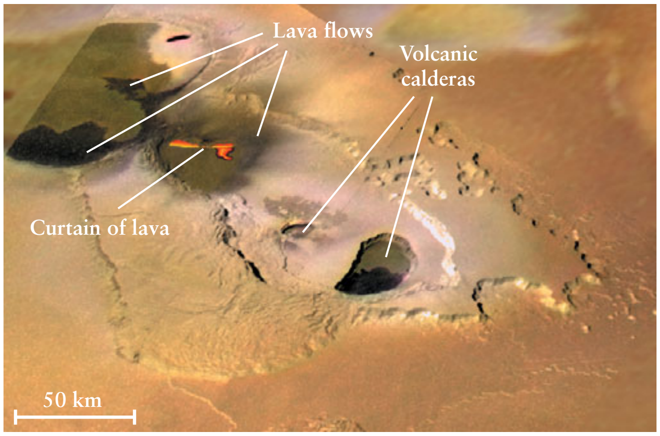


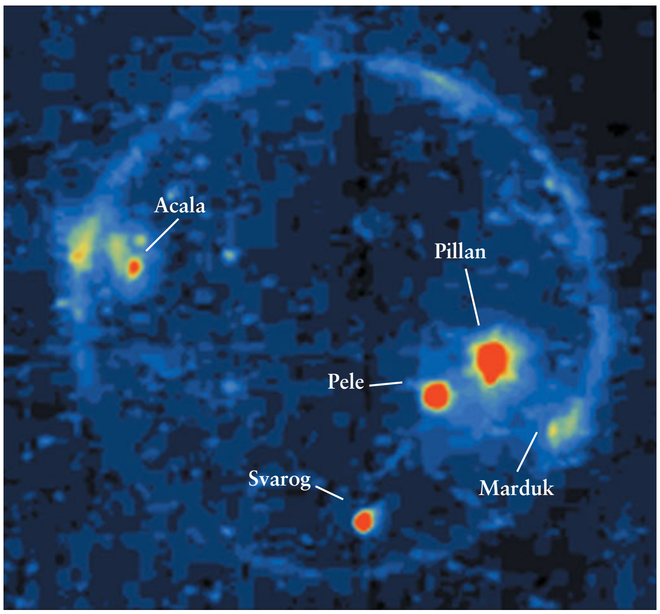
Figure 13-7 R I V U X G

Io's Lava Flows and a Curtain of Fire This mosaic of *Galileo* spacecraft images shows a nested chain of volcanic calderas (pits) with black lava flows. The red streak at upper left is a fissure in the surface about 40 km (25 mi) long, through which lava is erupting upward in a vertical sheet or curtain. The erupting lava reaches the amazing height of 1500 m (5000 ft). (University of Arizona/JPL/NASA)

ultramafic lava on Io suggests that its interior, too, is substantially hotter than that of the present-day Earth. One model proposes that Io has a 100-km (60-mi) thick crust floating atop a worldwide ocean of liquid magma 800 km (500 mi) deep. (As we saw in Section 9-2, the mantle that lies underneath the Earth's crust is a plastic solid rather than a true liquid.) The global extent of this "magma ocean" would explain why volcanic activity is found at all points on Io's surface, rather than being concentrated in pockets as on Earth (see Section 9-3). It could also explain the origin of Io's mountains, which reach heights up to 10 km (30,000 ft). In this model, these mountains are blocks of Io's crust that have tilted before sinking into the depths of the magma ocean.

Io's worldwide volcanic activity is remarkably persistent. When *Voyager 2* flew through the Jovian system in July 1979, four months after *Voyager 1*, almost all of the volcanic plumes seen by *Voyager 1* were still active. And when *Galileo* first viewed Io in 1996, about half of the volcanoes seen by the *Voyager* spacecraft were still ejecting material, while other, new volcanoes had become active.

Io may have as many as 300 active volcanoes, each of which ejects an estimated 10,000 tons of material per second. Altogether, volcanism on Io may eject as much as 10^{13} tons of matter each year. This is sufficient to cover the satellite's entire surface to a depth of 1 meter in a century, or to cover an area of 1000 square kilometers in a few weeks (see Figure 13-6). Thanks to this continual "repaving" of the surface, there are probably no long-lived features on Io, and any impact craters are quickly obliterated.



WEB LINK 13.6

Figure 13-8 R I V U X G

Io's Glowing Volcanoes This false-color *Galileo* image (made while Io was in the darkness of Jupiter's shadow) shows the infrared and visible glow from several volcanic regions, including Pele and Pillan (see Figure 13-6). These regions are substantially hotter than typical lava flows on Earth. (Planetary Image Research Laboratory/University of Arizona/JPL/NASA)

13-5 Jupiter's magnetic field makes electric currents flow through Io

Most of the material ejected from Io's volcanoes falls back onto the satellite's surface. But some material goes on a remarkable journey into space, by virtue of Io's location deep within Jupiter's magnetosphere.

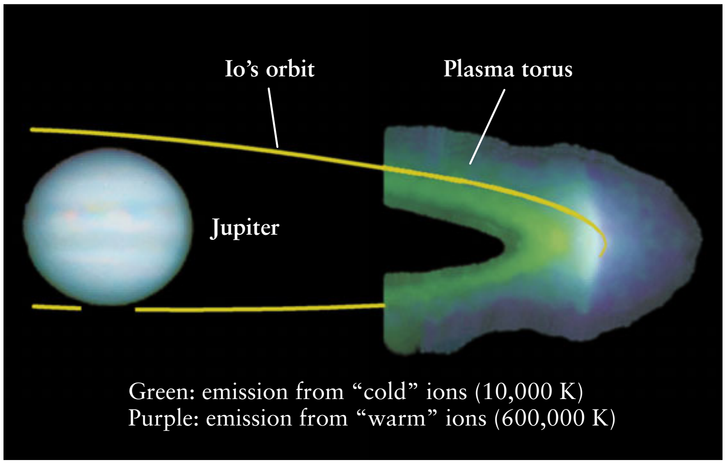
The Io Torus

The Jovian magnetosphere contains charged particles, some of which collide with Io and its volcanic plumes. The impact of these collisions knocks ions out of the plumes and off the surface, and these ions become part of Jupiter's magnetosphere. Indeed, material from Io is the main source for Jupiter's magnetosphere as a whole. Some of the ejected material forms the **Io torus**, a huge doughnut-shaped ring of ionized gas, or plasma (see Section 12-7), that circles Jupiter at the distance of Io's orbit. The plasma's glow can be detected from Earth (Figure 13-9*a*).

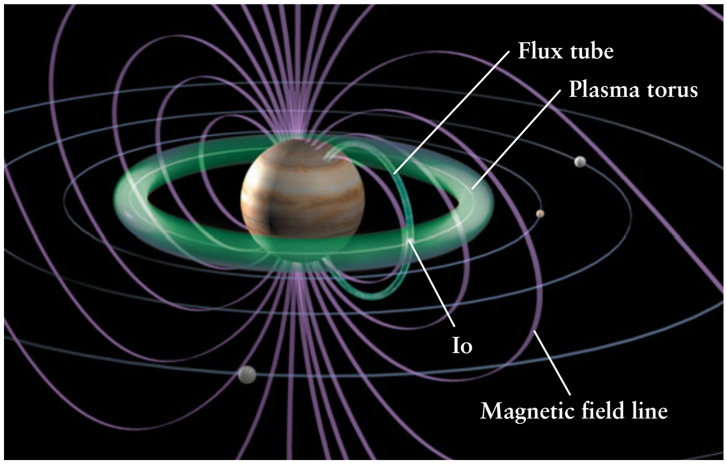
Jupiter and Io together act like an immense electric generator

Jupiter's magnetic field has other remarkable effects on Io. As Jupiter rotates, its magnetic field rotates with it, and this field sweeps over Io at high speed. This generates a voltage of 400,000 volts across the satellite, which causes 5 million amperes of electric current to flow through Io.

ANALOGY Whenever you go shopping with a credit card, you use the same physical principle that generates an electric current within Io. The credit card's number is imprinted as a magnetic code in the dark stripe on the back of the card. The salesperson who swipes your card through the card reader is actually moving the card's magnetic field past a coil of wire within the reader. This generates within the coil an electric current that carries the same coded information about your credit card number. That



(a)



(b)

VIDEO 13.3

Figure 13-9 R I V U X G

The Io Torus (a) An Earth-based telescope recorded this false-color infrared image of Io's plasma torus. Because of Jupiter's glare, only the outer edge of the torus could be photographed; an artist added the visible-light picture of Jupiter and the yellow line indicating the rest of the torus. **(b)** A flux tube (a region of concentrated magnetic field) carries plasma between Jupiter and Io, forming an immense electric circuit. (a: Courtesy of J. Trauger; b: Alfred T. Kamajian and Torrence V. Johnson, "The Galileo Mission to Jupiter and Its Moons," *Scientific American*, February 2000, p. 44)

information is transmitted to the credit card company, which (it is hoped) approves your purchase. An electric generator at a power plant works in the same way. A coil of wire is moved through a strong magnetic field, which makes current flow in the coil. This current is delivered to transmission lines and eventually to your home.

In fact, a current flows not only through Io but also through the sea of charged particles in Jupiter's magnetosphere and through Jupiter's atmosphere. This forms a gigantic, oval-shaped electric circuit that connects Io and Jupiter in the same way that wires connect the battery and the lightbulb in a flashlight (Figure 13-9*b*). Jupiter's aurora, which we discussed in Section 12-7, is particularly strong at the locations where this current strikes the upper atmosphere of Jupiter.

Part of the electric current that flows between Io and Jupiter is made up of electrons that spiral around Jupiter's magnetic field lines. As they spiral, the electrons act like miniature antennas and emit radio waves. This is the source of Jupiter's bursts of decametric radio radiation, which we discussed in Section 12-7.

Probing Io's Interior

The electric current within and near Io also creates a weak magnetic field, which was first detected by *Voyager 1*. Remarkably, measurements made by *Galileo* during its close approaches to Io suggest that the magnetic field near Io may be too strong—comparable to that of Mercury, although weaker than that of Earth—to be generated by Io's electric current alone. This has led scientists to speculate that Io generates its own magnetic field through the motions of molten material in its interior, just as the Earth does (see Section 7-7 and Section 9-4). If these speculations are correct, Io is the smallest world in the solar system to generate its own magnetic field.

If Io has undergone enough tidal heating to melt its interior, chemical differentiation must have taken place and the satellite should have a dense core. (We described chemical differentiation in Box 7-1 and Section 8-4.) To test for this, scientists measured how Io's gravity deflected the trajectory of *Galileo* as the spacecraft flew past. From these measurements, they could determine not only Io's mass but also the satellite's oblateness (how much it deviates from a spherical shape because of its rotation). The amount of oblateness indicates the size of the core; the greater the fraction of the satellite's mass contained in its core, the less oblate the satellite will be. The *Galileo* observations suggest that Io has a dense core composed of iron and iron sulfide (FeS), with a radius of about 900 km (about half the satellite's overall radius). Surrounding the core is a mantle of partially molten rock, on top of which is Io's visible crust.

13-6 Europa is covered with a smooth layer of ice that may cover a worldwide ocean

WEB LINK 13.7

Europa, the second of the Galilean satellites, is the smoothest body in the solar system. There are no mountains and no surface features greater than a few hundred meters high (Figure 13-10). There are almost no craters, indicating a young surface that has been reprocessed by geologic activity. The dominant surface feature is a worldwide network of stripes and cracks (Figure 13-11). Like Io, Europa is an exception to the general rule that a small world should be cratered and geologically dead (see Section 7-6).

Europa's Surface: Icy but Active

Neither *Voyager 1* nor *Voyager 2* flew very near Europa, so most of our knowledge of this satellite comes from the close passes

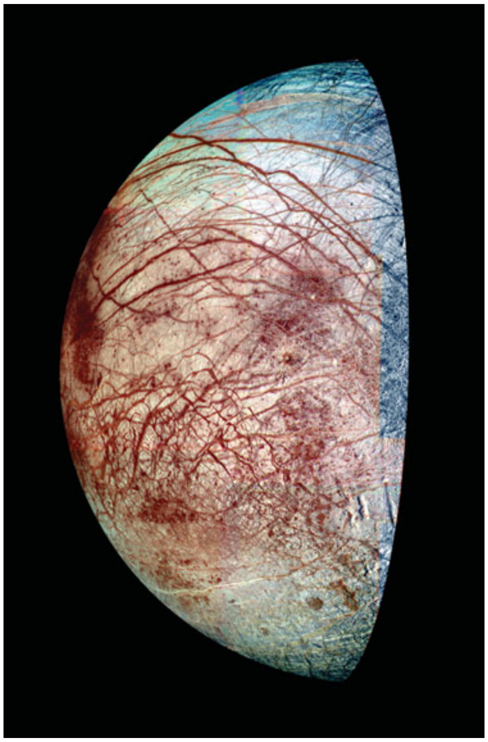


Figure 13-10 R I V U X G

Europa Dark lines crisscross Europa's smooth, icy surface in this false-color composite of visible and infrared images from *Galileo.* These are fractures in Europa's crust that can be as much as 20 to 40 km (12 to 25 mi) wide. Only a few impact craters are visible on Europa, which indicates that this satellite has a very young surface on which all older craters have been erased. (NASA/JPL/University of Arizona)

made by *Galileo*. (Both Figure 13-10 and Figure 13-11 are *Galileo* images.) But even before these spacecraft visited Jupiter, spectroscopic observations from Earth indicated that Europa's surface is almost pure frozen water (see Figure 7-4). This was confirmed by instruments on board *Galileo,* which showed that Europa's infrared spectrum is a close match to that of a thin layer of fine-grained water ice frost on top of a surface of pure water ice. (The brown areas in Figure 13-10 show where the icy surface contains deposits of rocky material from meteoritic impacts, from Europa's interior, or from a combination of these sources.)

Geology on Europa is based on solid ice and liquid water, not rock and molten magma

The purity of Europa's ice suggests that water is somehow brought upward from the moon's interior to the surface, where it solidifies to make a fresh, smooth layer of ice. Indeed, some *Galileo* images show what appear to be lava flows on Europa's surface, although the "lava" in this case is mostly ice. This idea helps to explain why Europa has very few craters (any old ones have simply been covered up) and why its surface is so smooth. Europa's surface may thus represent a water-and-ice version of plate tectonics.

Although Europa's surface is almost pure water ice, keep in mind that Europa is *not* merely a giant ice ball. The satellite's density shows that rocky material makes up about 85 to 90% of Europa's mass. Hence, only a small fraction of the mass, about 10 to 15%, is water ice. Because the surface is icy, we can conclude that the rocky material is found within Europa's interior.

Europa is too small to have retained much of the internal heat that it had when it first formed. But there must be internal heat nonetheless to power the geologic processes that erase craters and bring fresh water to Europa's surface. What keeps Europa's interior warm? The most likely answer, just as for Io, is heating by Jupiter's tidal forces. The rhythmic gravitational tugs exerted by Io and Ganymede on Europa deform its orbit into an ellipse; the varying speed of Europa as it goes around its orbit causes Europa to "nod" back and forth, causing variations in tidal stresses that make Europa flex. But because it is farther from Jupiter, tidal effects on Europa are only about one-fourth as strong as those on Io, which may explain why no ongoing volcanic activity has yet been seen on Europa.

Some features on Europa's surface, such as the fracture patterns shown in Figure 13-11, may be the direct result of the crust being stretched and compressed by tidal flexing. Figure 13-12 shows other features, such as networks of ridges and a young, very smooth circular area, that were probably caused by the internal heat that tidal flexing generates. The rich variety of terrain depicted in Figure 13-12, with stress ridges going in every direction, shows that Europa has a complex geologic history.

Among the unique structures found on Europa's surface are **ice rafts.** The area shown in Figure 13-13*a* was apparently subjected to folding, producing the same kind of linear features as those in Figure 13-12. But a later tectonic disturbance broke the surface into small chunks of crust a few kilometers across, which then "rafted" into new positions. A similar sort of rafting hap-

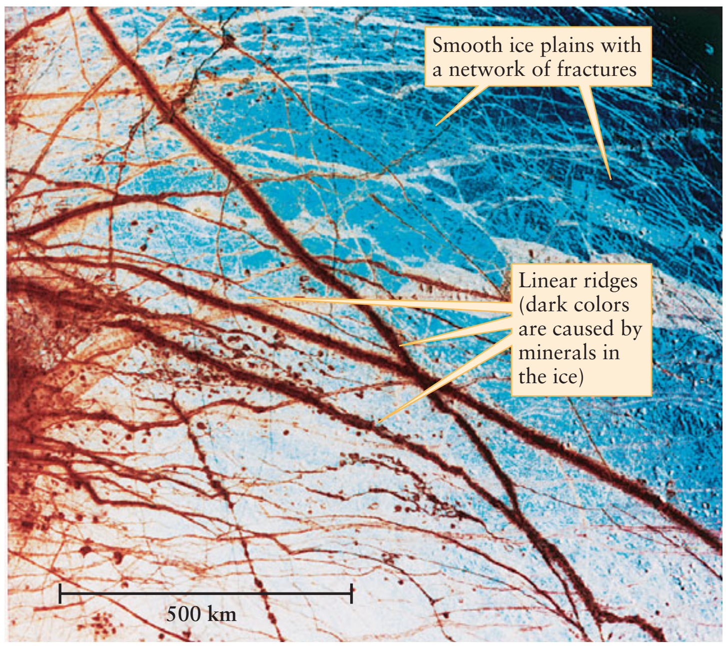


Figure 13-11 R I V U X G

Europa's Fractured Crust False colors in this *Galileo* composite image emphasize the difference between the linear ridges and the surrounding plains. The smooth ice plains (shown in blue) are the basic terrain found on Io. (NASA/JPL)

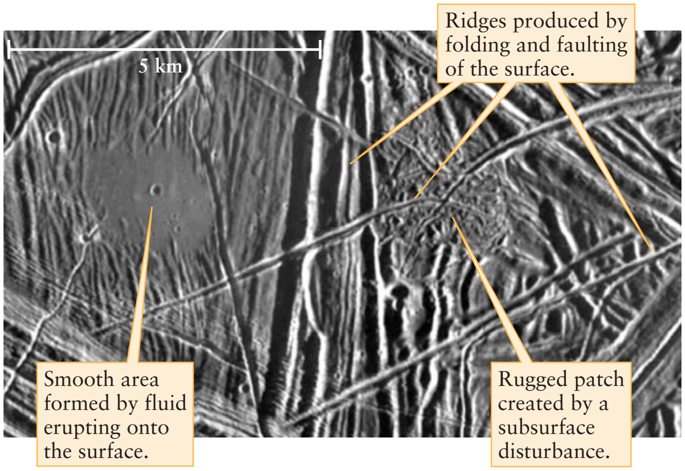


Figure 13-12 R I V U X G

Europa in Close-up This high-resolution *Galileo* image shows a network of overlapping ridges, part of which has been erased to leave a smooth area and part of which has been jumbled into a rugged patch of terrain. Europa's interior must be warm enough to power this complex geologic activity. (NASA/JPL)

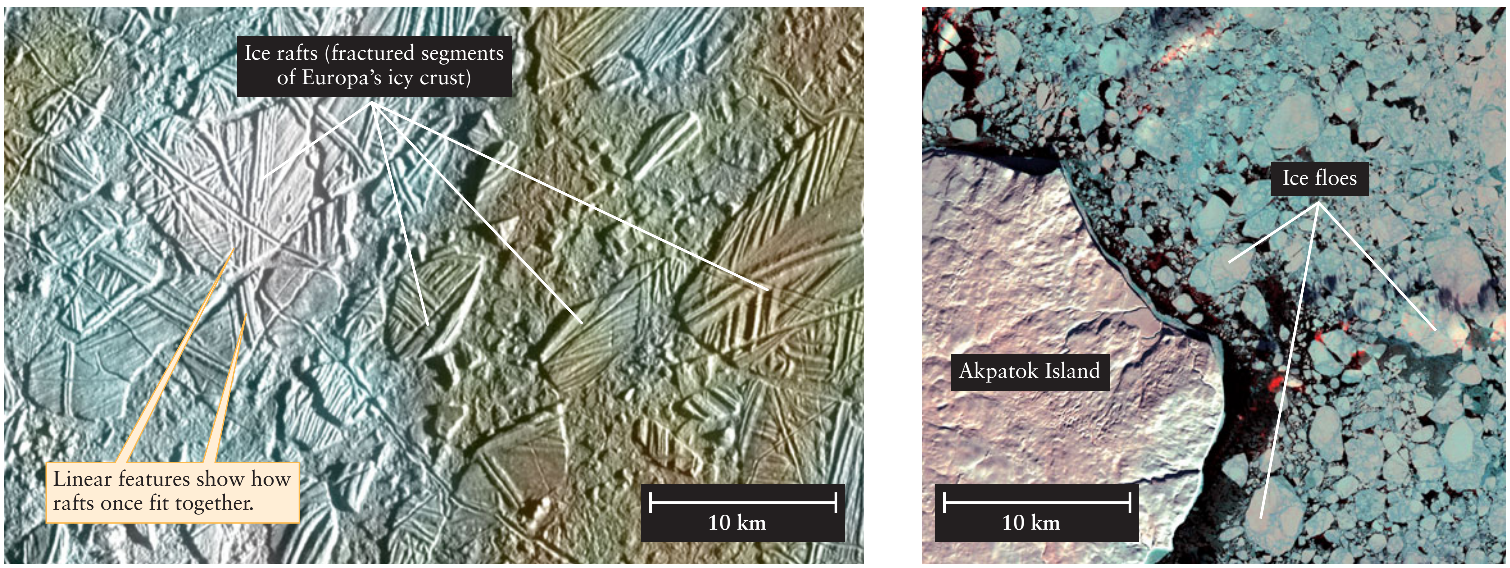
pens in the Earth's Arctic Ocean every spring, when the winter's accumulation of surface ice breaks up into drifting ice floes (Figure 13-13*b*). The existence of such structures on Europa strongly suggests that there is a subsurface layer of liquid water or soft ice over which the ice rafts can slide with little resistance.

An Underground Ocean?

The ridges, faults, ice rafts, and other features strongly suggest that Europa has substantial amounts of internal heat. This heat could prevent water from freezing beneath Europa's surface, creating a worldwide ocean beneath the crust. If this picture is correct, geologic processes on Europa may be an exotic version of those on Earth, with the roles of solid rock and molten magma being played by solid ice and liquid water.

A future spacecraft may use radar to penetrate through Europa's icy crust to search for definitive proof of liquid water beneath the surface. But magnetic field measurements made by the *Galileo* spacecraft have already provided some key evidence favoring this picture. Unlike the Earth or Jupiter, Europa does not seem to create a steady magnetic field of its own. But as Europa moves through Jupiter's intense magnetic field, electric currents are induced within the satellite's interior, just as they are within Io (see Section 13-5), and these currents generate a weak but measurable field. (The strength and direction of this induced field vary as Europa moves through different parts of the Jovian magnetosphere, which would not be the case if Europa generated its field by itself.)

To explain these observations, there must be an electrically conducting fluid beneath Europa's crust—a perfect description of an underground ocean of water with dissolved minerals. (Pure water is a very poor conductor of electricity, so other substances must be present to make the water conducting.) If some of this water should penetrate upward to Europa's surface through cracks in the crust, it would vaporize and spread the dissolved minerals across the terrain. This could explain the reddish-brown colors in Figure 13-13*a*.



(a) Ice rafts on Europa

(b) Ice floes on Earth

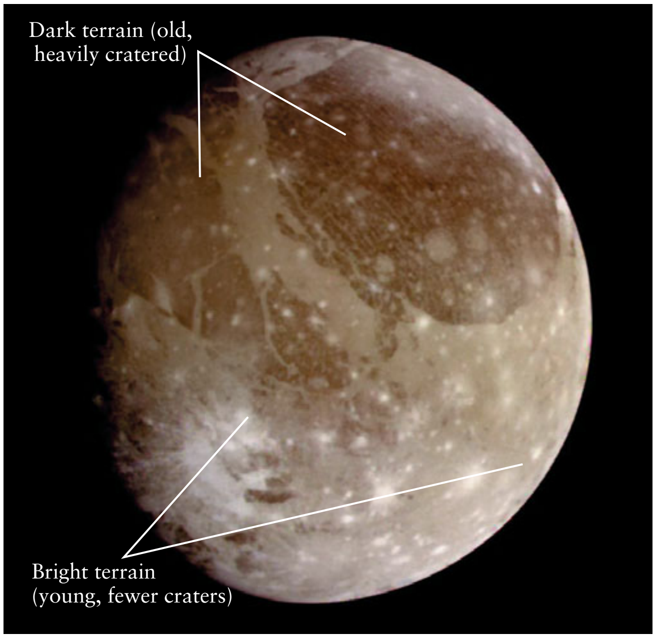
Figure 13-13 R I V U X G

Moving Ice on Europa and Earth **(a)** Some time after a series of ridges formed in this region of Europa's surface, the icy crust broke into "rafts" that were moved around by an underlying liquid or plastic layer. The colors in this *Galileo* image may be due to minerals that were released from beneath the surface after the crust broke apart. **(b)** Europa's ice rafts are analogous to ice floes created when pack ice breaks up, as in this spacecraft view of part of the Canadian arctic. (a: NASA/JPL; b: USGS and NASA)

By combining measurements of Europa's induced magnetic field, gravitational pull, and oblateness, scientists conclude that Europa's outermost 100 to 200 km are ice and water. (It is not clear how much of this is liquid and how much solid.) Within this outer shell is a rocky mantle surrounding a metallic core some 600 km (400 mi) in radius.

Remarkably, Europa also has an extremely thin atmosphere of oxygen. (By Earth standards, this atmosphere would qualify as a near-vacuum.) We saw in Section 9-5 that oxygen in the Earth's atmosphere is produced by plants through photosynthesis. But Europa's oxygen atmosphere is probably the result of ions from Jupiter's magnetosphere striking the satellite's icy surface. These collisions break apart water molecules, liberating atoms of hydrogen (which escape into space) and oxygen.

The existence of a warm, subsurface ocean on Europa, if proved, would make Europa the only world in the solar system other than the Earth on which there is liquid water. This would have dramatic implications. On Earth, water and warmth are essentials for the existence of life. Perhaps single-celled organisms have evolved in the water beneath Europa's crust, where they would use dissolved minerals and organic compounds as food sources. In light of this possibility, NASA has taken steps to prevent biological contamination of Europa. At the end of the *Galileo* mission in 2003, the spacecraft (which may have carried traces of organisms from Earth) was sent to burn up in Jupiter's atmosphere, rather than remaining in orbit where it might someday crash into Europa. An appropriately sterilized spacecraft may one day visit Europa and search for evidence of life within this exotic satellite.



VIDEO 13.4

Figure 13-14 R I V U X G

Ganymede Two distinct types of terrain—one dark, heavily cratered, and hence old, the other bright, less cratered, and hence younger—are visible in this *Galileo* image of the hemisphere of Ganymede that always faces away from Jupiter. Craters in general appear bright, suggesting that the impacts that formed the craters excavated the surface to reveal ice underneath. (NASA/JPL)

13-7 Liquid water may also lie beneath the cratered surfaces of Ganymede and Callisto

Both Ganymede and Callisto show evidence of past geologic activity

Unlike Io and Europa, the two outer Galilean satellites have cratered surfaces. In this respect, Ganymede and Callisto bear a superficial resemblance to our own Moon (see the illustration with Table 13-1). But unlike the Moon's craters, the craters on both Ganymede and Callisto are made primarily of ice rather than rock, and Ganymede has a number of surface features that indicate a geologically active past.

The Two Faces of Ganymede's Surface

WEB LINK 13.10

Ganymede is the largest satellite in our solar system and is even larger than the planet Mercury. Like our Moon, Ganymede has two distinct kinds of terrain, called **dark terrain** and **bright terrain** (Figure 13-14). On the Moon, the dark maria are younger than the light-colored lunar highlands (see Section 10-1). On Ganymede, by contrast, the dark terrain is older, as indicated by its higher density of craters. The bright terrain is much less cratered and therefore younger. Because ice is more reflective than rock, even Ganymede's dark terrain is substantially brighter than the lunar surface.

Voyager and *Galileo* images show noticeable differences between young and old impact craters on Ganymede. The youngest craters are surrounded by bright rays of freshly exposed ice that has been excavated by the impact (see the lower left of Figure 13-14). Older craters are darker, perhaps because of chemical processes triggered by exposure to sunlight.

The close-up *Galileo* image in Figure 13-15 shows the contrast between dark and bright terrain. The area on the left-hand side of this image is strewn with large and small craters, reinforcing the idea that the dark terrain is billions of years old. The dark terrain is also marked by long, deep furrows. These are deformations of Ganymede's crust that have partially erased some of the oldest craters. Since other craters lie on top of the furrows, the stresses that created the furrows must have occurred long ago. Since geologic activity is a result of a planet or satellite's internal heat, this shows that Ganymede must have had a warm interior in the distant past.

Linear features are also seen in the bright terrain on the right in Figure 13-15. The *Voyager* missions discovered that the bright terrain is marked by long grooves, some of which extend for hundreds of kilometers and are as much as a kilometer deep. Images from *Galileo* show that some adjacent grooves run in different directions, indicating that the bright terrain has been subjected to a variety of tectonic stresses. *Galileo* also found a number of small craters overlying the grooves, some of which are visible in Figure 13-15. The density of craters shows that the bright terrain, while younger than the dark terrain, is at least a billion years old.

After the *Voyager* missions, it was thought that the bright terrain represented fresh ice that had flooded through cracks in

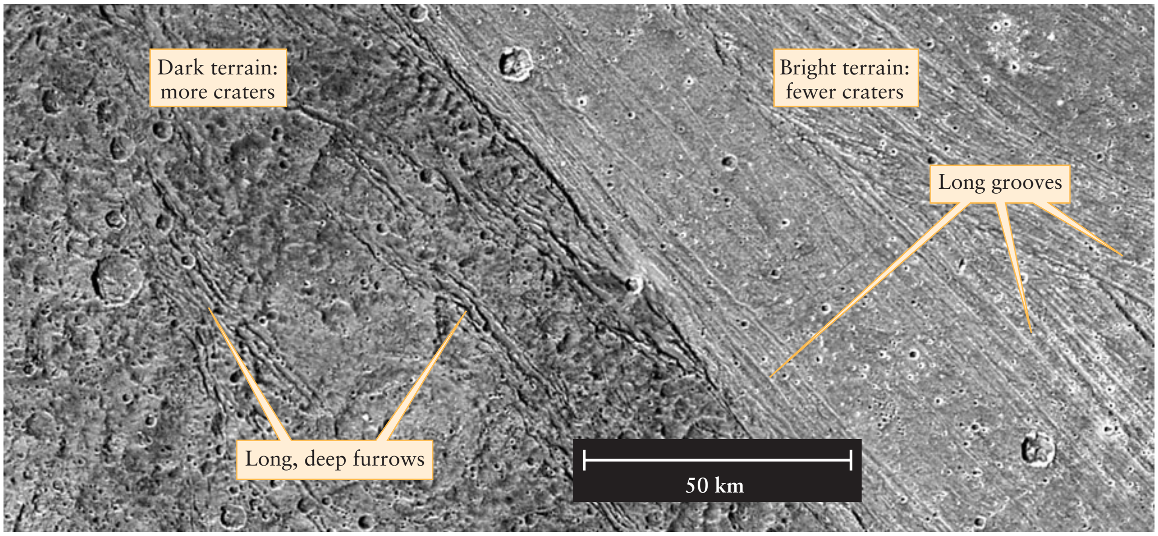


Figure 13-15 R I V U X G

Evidence for Geologic Activity on Ganymede Stresses in Ganymede's crust have produced linear features that extend for hundreds of kilometers across both the dark and bright terrain. These features must be quite ancient, since numerous impact craters have formed on top of them. (NASA)

Ganymede's crust, analogous to the maria on our Moon (see Section 10-1). However, high-resolution *Galileo* images do not show the kinds of surface features that would be expected if this terrain was volcanic in origin. Yet there are several low-lying areas within the bright terrain that appear to have been flooded by a watery fluid that subsequently froze. These low-lying areas may be *rift valleys* where the crust has been pulled apart, like Valles Marineris on Mars (see Section 11-5). In order to flood the floors of these valleys, there must have been liquid water beneath Ganymede's surface when the valleys formed. This is additional evidence that Ganymede must have had a warm interior in the past, and that this internal warmth persisted at least until the era when the bright terrain formed a billion years ago. This is rather surprising, because Ganymede is too small to have retained much of the heat left over from its formation or from radioactive elements, and it orbits too far from Jupiter to have any substantial tidal heating.

Ganymede's Interior and Bizarre Magnetic Field

An equally great surprise from *Galileo* observations was the discovery that Ganymede has its own magnetic field, and that this field is twice as strong as that of Mercury. This field is sufficiently strong to trap charged particles, giving Ganymede its own "mini-magnetosphere" within Jupiter's much larger magnetosphere. The presence of a magnetic field shows that electrically conducting material must be in motion within Ganymede to act as a dynamo, which means that the satellite must have substantial internal heat even today. A warm interior causes differentiation, and *Galileo* measurements indeed showed that Ganymede is highly differentiated. It has a metallic core some 500 km in radius, surrounded by a rocky mantle and by an outer shell of ice some 800 km thick.

An even more surprising result is that Ganymede's magnetic field varies somewhat as it orbits Jupiter, and that these variations are correlated with Ganymede's location within Jupiter's magnetic field. One explanation for this variation is that Ganymede has a layer of salty, electrically conducting liquid water some 170 km beneath its surface. As the satellite moves through Jupiter's magnetic field, electric currents are induced in this layer just as happens within Europa (see Section 13-6), and these currents generate a weak field that adds to the stronger field produced by Ganymede's dynamo.

A possible explanation for Ganymede's internal heat is that gravitational forces from the other Galilean satellites may have affected its orbit. While Ganymede's present-day orbit around Jupiter is quite circular, calculations show that the orbit could have been more elliptical in the past. Just as for Io and Europa, such an elliptical orbit would have given rise to substantial tidal heating of Ganymede's interior. It may have retained a substantial amount of that heat down to the present day.

Callisto's Perplexing Surface

WEB LINK 13.11

In marked contrast to Io, Europa, and Ganymede is Callisto, Jupiter's outermost Galilean satellite. Images from *Voyager* showed that Callisto has numerous impact craters scattered over an icy crust (Figure 13-16). Callisto's ice is not as reflective as that on Europa or Ganymede, however; the surface appears to be covered with some sort of dark mineral deposit.

CAUTION! Figure 13-16, as well as the images that accompany Table 13-1, shows that Callisto is the darkest of the Galilean satellites. Nonetheless, its surface is actually more than twice as reflective as that of our own Moon. Callisto's surface may be made of "dirty" ice, but even dirty ice reflects more light than the gray rock found on the Moon. If you could somehow replace our own Moon with Callisto, a full moon as seen from Earth would be 4 times brighter than it is now (thanks also in part to Callisto's larger size).

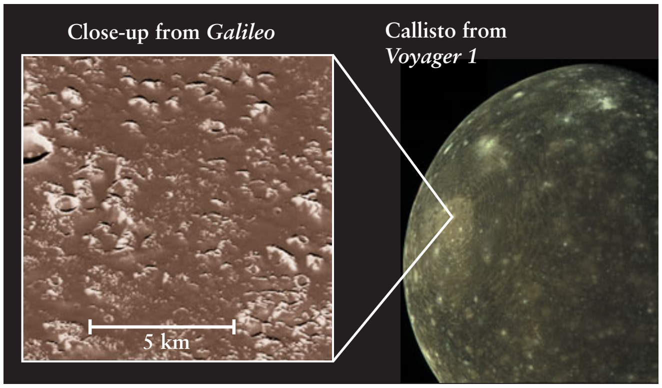


Figure 13-16 R I V U X G

Callisto Numerous craters pockmark Callisto's icy surface. The inset shows a portion of a huge impact basin called Valhalla. Most of the very smallest craters in this close-up view have been completely obliterated, and the terrain between the craters is blanketed by dark, dusty material. (Left: Arizona State University/JPL/NASA; right: JPL/NASA)

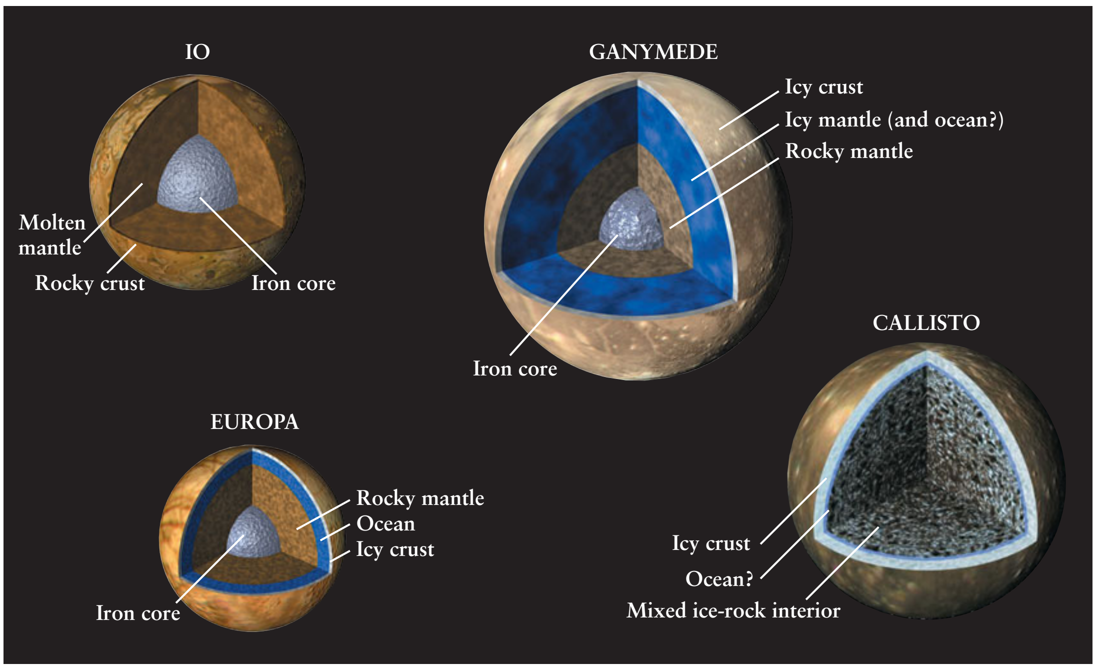
The *Voyager* images led scientists to conclude that Callisto is a dead world, with an ancient surface that has never been reshaped by geologic processes. There is no simple relationship between Callisto's orbital period and those of the other Galilean satellites, so it would seem that Callisto never experienced any tidal heating and hence was unable to power any geologic processes at its surface. But once the *Galileo* spacecraft went into orbit around Jupiter and began returning high-resolution images, scientists realized that Callisto's nature is not so simple. In fact, Callisto proves to be the most perplexing of the Galilean satellites.

One curious discovery from *Galileo* is that while Callisto has numerous large craters, there are very few with diameters less than 1 km. Such small craters are abundant on Ganymede, which has probably had the same impact history as Callisto. This implies that most of Callisto's small craters have been eroded away. How this could have happened is not known. Another surprising result is that Callisto's surface is covered by a blanket of dark, dusty material (see the inset in Figure 13-16). Where this material came from, and how it came to be distributed across the satellite's surface, is not understood.

Callisto's Interior: Hot or Cold?

Callisto's most unexpected feature is that, like Europa and Ganymede, it appears to have an induced magnetic field that varies as the satellite moves through different parts of Jupiter's magnetosphere. Callisto must have a layer of electrically conducting material in order to generate such a field. One model suggests that this material is in the form of a subsurface ocean like Europa's but only about 10 km (6 mi) deep. However, without tidal heating, Callisto's interior is probably so cold that a layer of liquid water would quickly freeze. One proposed explanation is that Callisto's ocean is a mixture of water and a form of "antifreeze." In cold climates on Earth, antifreeze is added to the coolant in automobile radiators to lower the liquid's freezing temperature and prevent it from solidifying. Ammonia in Callisto's proposed subsurface ocean could play the same role.

The presence of a liquid layer beneath Callisto's surface is difficult to understand. Even with ammonia acting as an antifreeze, this layer cannot be too cold or it would solidify. This argues for a relatively *warm* interior. But *Galileo* measurements suggest that Callisto's interior is not highly differentiated. This



INTERACTIVE EXERCISE 13.1

Figure 13-17

Interiors of the Galilean Satellites These cross-sectional diagrams show the probable internal structures of the four Galilean satellites of Jupiter, based on information from the *Galileo* mission. (NASA/JPL) implies that the interior is *cold,* and may never have been heated enough to undergo complete chemical differentiation. The full story of Callisto's interior, it would seem, has yet to be told.

Although Ganymede and Callisto have many differences, one feature they share is an extremely thin and tenuous atmosphere. The molecules that make up Ganymede's atmosphere are oxygen (O_2) and ozone (O_3). These are thought to be released from gas bubbles trapped in Ganymede's surface ice. On Callisto, by contrast, the atmosphere is composed of carbon dioxide (CO_2). This gas may have evaporated from Callisto's poles, where temperatures are low enough to form CO_2 ice.

The diagrams in Figure 13-17 summarize the probable interior structures of the Galilean satellites. They demonstrate the remarkable variety of these terrestrial worlds.

13-8 Titan has a thick atmosphere and streams carved by liquid hydrocarbons

WEB LINK 13.12

Unlike Jupiter, Saturn has only one large satellite that is comparable in size to our own Moon. This satellite, Titan, was discovered by Christiaan Huygens in 1665. Titan's diameter of 5150 km makes it second in size among satellites to Ganymede (Section 13.7). Like Ganymede, Titan has a low density that suggests it is made of a mixture and ice and rock. (See Table 7-2 for a comparison between Titan, the Moon, and the Galilean satellites.) But unlike Ganymede or any other satellite in the solar system, Titan has a thick atmosphere with a unique chemical composition.

Titan's Atmosphere

By the early 1900s, scientists had begun to suspect that Titan (Figure 13-18) might have an atmosphere, because it is cool enough and massive enough to retain heavy gases. (Box 7-2 explains the criteria for a planet or satellite to retain an atmosphere.) These suspicions were confirmed in 1944, when Gerard Kuiper discovered the spectral lines of methane in sunlight reflected from Titan. (We discussed Titan's spectrum and its interpretation in Section 7-3.)

Because of its atmosphere, Titan was a primary target for the *Voyager* missions. To the chagrin of mission scientists and engineers, however, *Voyager 1* spent hour after precious hour of its limited, preprogrammed observation time sending back featureless images like the one at the center of Figure 13-18. Titan's atmosphere is so thick—about 200 km (120 mi) deep, or about 10 times the depth of the Earth's troposphere—that it blocked *Voyager*'s view of the surface. The haze surrounding Titan is so dense that little sunlight penetrates to the ground; noon on Titan is only about 1/1000 as bright as noon on Earth, though still 350 times brighter than night on Earth under a full moon.

By examining how the radio signal from *Voyager 1* was affected by passing through Titan's atmosphere, scientists inferred that the atmospheric pressure at Titan's surface is 50% greater than the atmospheric pressure at sea level on the Earth. Titan's surface gravity is weaker than that of Earth, so to produce this pressure Titan must have considerably more gas in its atmosphere than Earth. Indeed, calculations show that about 10 times more gas (by mass) lies above a square meter of Titan's surface than above a square meter of Earth.

Voyager data show that more than 95% of Titan's atmosphere is nitrogen. Most of this nitrogen probably came from ammonia (NH_3), a compound of nitrogen and hydrogen that is quite common in the outer solar system. Ammonia is easily broken into hydrogen and nitrogen atoms by the Sun's ultraviolet radiation. Titan's gravity is too weak to retain hydrogen atoms (see Box 7-2), so these atoms escape into space and leave the more massive nitrogen atoms behind.

The remainder of Titan's atmosphere is predominantly methane (CH_4), which is the principal component of the "natural gas" used on Earth as fuel. The interaction of methane with ultraviolet light from the Sun produces a variety of other carbon-hydrogen compounds, or **hydrocarbons.** For example, the *Voyager* infrared spectrometers detected small amounts of ethane

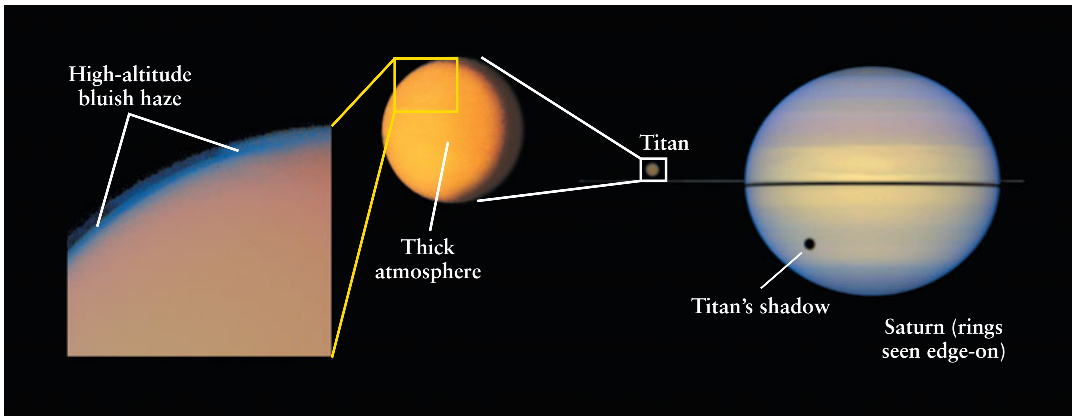


Figure 13-18 R I V U X G

Titan The Hubble Space Telescope image at far right shows Titan and the shadow it casts on Saturn's clouds. The *Voyager* image in the middle shows how Titan's atmosphere gives it a nearly featureless appearance. At far left is a color-enhanced *Voyager* image that shows a haze extending well above the satellite's visible edge, or limb. (Left: JPL/NASA; center: NASA; right: Erich Karkoschka, LPL/STScI/NASA)

(C_2H_6), acetylene (C_2H_2), ethylene (C_2H_4), and propane (C_3H_8) in Titan's atmosphere.

On Earth, water can be a gas (like water vapor in the atmosphere), a liquid (as in the oceans), or a solid (as in the polar ice caps). On Titan, the atmospheric pressure is so high and the surface temperature is so low—about 95 K (–178°C, or –288°F)—that any water is frozen solid. But conditions on Titan are just right for methane and ethane to exist as gases, liquids, or solids. This raises the tantalizing possibility that Titan could have methane rain and lakes of liquid ethane.

A faint mist of methane rain falls continually on Titan

Nitrogen in Titan's atmosphere can combine with airborne hydrocarbons to produce other compounds, such as hydrogen cyanide (HCN). Hydrogen cyanide, along with other molecules, can join together in long, repeating molecular chains to form substances called **polymers.** Droplets of some polymers remain suspended in the atmosphere, forming an **aerosol.** (Common aerosols on the Earth include fog, mist, and paint sprayed from a can.) The polymers in Titan's airborne aerosol are thought to be responsible for the reddish-brown color seen in Figure 13-18.

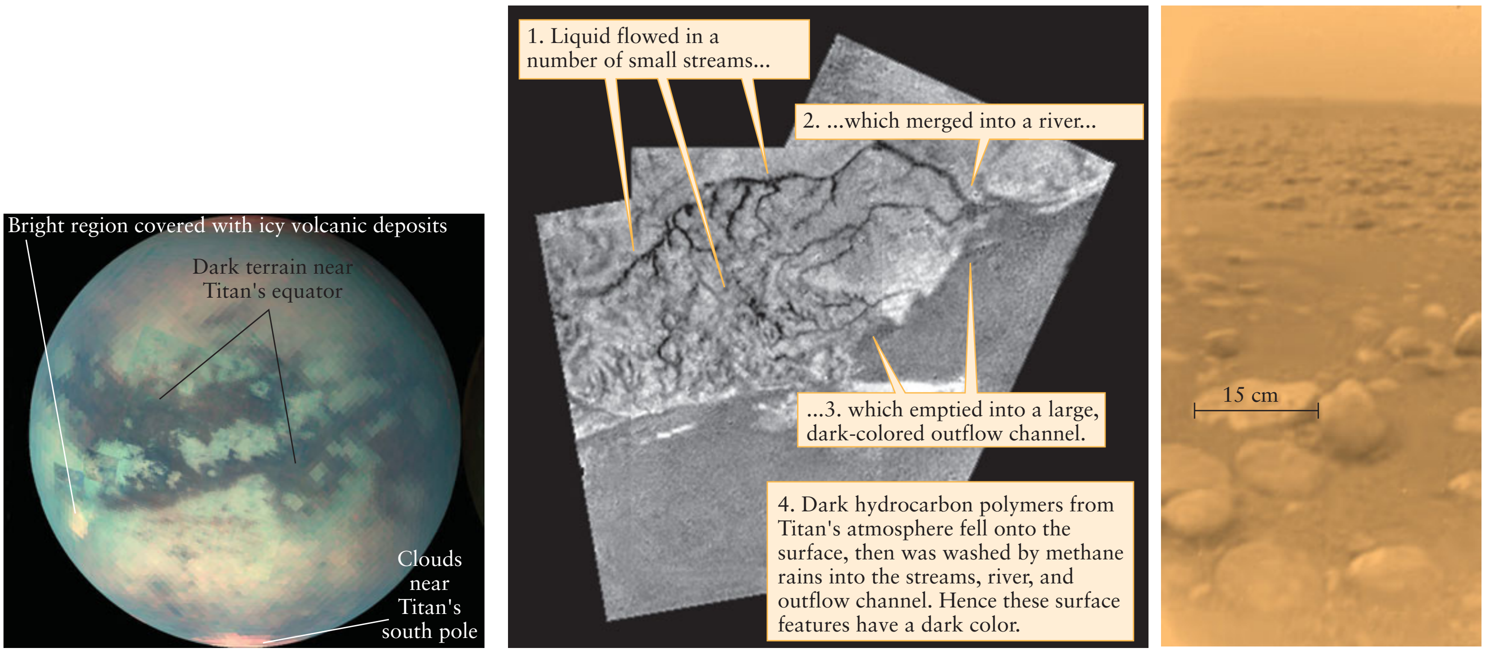
Exploring Titan with *Cassini* and *Huygens*

WEB LINK 13.13

The designers of the *Cassini* mission to Saturn made the exploration of Titan a central goal. During its primary mission in Saturn orbit from 2004 to 2008, the *Cassini* spacecraft will make 44 close flybys of Titan. Titan's atmosphere is more transparent to infrared wavelengths than to visible light, so *Cassini* has used its infrared telescope to obtain the first detailed images of the satellite's surface (**Figure 13-19*a***). Most of the surface is light-colored, but a swath of dark terrain extends around most of Titan's equator.

Cassini has used its on-board radar to map the dark terrain. (Like infrared light, radio waves can pass through Titan's atmosphere.) These images reveal a series of long, parallel lines of sand dunes about 1 to 2 km apart. These are aligned west to east, in the same direction that winds blow at Titan's equator, and so presumably formed by wind action. Unlike Earth sand, which is made of small particles of silicate rock, sand on Titan is probably small particles of water ice combined with polymers that have fallen from Titan's atmosphere. These polymers give the dark terrain its color.

The infrared telescope aboard *Cassini* also detected a surface feature that appears to be a volcano. The presence of volcanism on Titan can explain why there is so much methane in the satellite's atmosphere. Methane molecules are broken apart by ultraviolet photons from the Sun, and the hydrogen atoms escape into space. This process would destroy all of Titan's methane within 10 million years unless a fresh supply is added to the atmosphere by volcanic activity. Volcanic activity can also explain the presence of isolated very bright patches on Titan's surface (see Figure 13-19*a*). A volcanic eruption would outgas not just methane but



(a) R I V U X G (b) R I V U X G (c) R I V U X G

VIDEO 13.5 VIDEO 13.6

Figure 13-19

Beneath Titan's Clouds (a) This mosaic of Titan was constructed from *Cassini* images made at visible and infrared wavelengths during a close flyby in December 2005. **(b)** The *Huygens* lander recorded this view of features on Titan carved by flowing liquid. It is too cold on Titan for water to be a liquid, so these streams and river must have carried liquid methane or ethane. **(c)** This nearly true-color *Huygens* image shows the view from Titan's surface. The 15-cm (6-in.) wide "rock" is about 85 cm (33 in.) from the camera. Unlike Earth rocks, Titan "rocks" are chunks of water ice. (NASA/JPL/ESA/University of Arizona) also water vapor and carbon dioxide. These gases would quickly freeze onto the surface, making a patch of highly reflective ice.

The most remarkable images of Titan's surface have come from *Huygens,* a small lander that *Cassini* carried on its journey from Earth and released before entering Saturn orbit. Named for the discoverer of Titan, *Huygens* entered Titan's atmosphere on January 14, 2005. The lander took $2\frac{1}{2}$ hours to descend to the surface under a parachute, during which time it made detailed images of the terrain (see Figure 13-19*b*). After touchdown, *Huygens* continued to return data for another 70 minutes (see Figure 13-19*c*) before its batteries succumbed to Titan's low temperatures.

Huygens images such as Figure 13-19*b* show that liquids have indeed flowed on Titan like water on Earth. Yet none of the images made by the lander during its descent or after touchdown showed any evidence of standing liquid. There was nonetheless evidence of recent methane rainfall at the *Huygens* landing site. A lander instrument designed to measure atmospheric composition was equipped with a heater for its gas inlet. During the three minutes it took to warm this inlet, the measured methane concentration jumped by 30% and then remained steady. The explanation is that the ground was soaked with liquid methane that had fallen as rain and which evaporated as it was warmed by the heater. *Huygens* measurements indicate that the average annual methane rainfall on Titan is about 5 cm (2 in.), about the same as the amount of water rain that falls each year at Death Valley in California. The rivers and outflow channels that *Huygens* observed during its descent may date from an earlier period in Titan's history when methane was more abundant and rainfall was more intense.

Before the *Cassini* mission scientists expected to find lakes of liquid ethane across Titan's surface. As of this writing (2006) the only location where *Cassini* appears to have located such lakes is near Titan's north pole (**Figure 13-20**). It is currently winter in the northern hemisphere, so temperatures are lowest there. Scientists also expected to find extensive clouds of ethane in Titan's atmosphere; these, too, have only been found near the north pole. Over the course of a Titan year (which, like that of its parent planet Saturn, lasts 29.37 years) the supply of liquid ethane may migrate with the seasons from pole to pole, much as frozen carbon dioxide migrates seasonally on Mars (see Section 11-6).

Impact craters are also conspicuously absent on Titan: Only a handful of craters have been observed by *Cassini*. The majority of craters may have been eroded by wind or rainfall, or covered by material ejected from volcanoes.

Titan's Geologic History

We have seen that Titan must have been volcanically active to provide its atmosphere with methane. But what has powered the volcanic activity? Titan is too small to have retained the internal heat from its formation. Furthermore, although Titan's rotation is synchronous like that of Io, its orbit around Saturn is so circular that it cannot be heated by tidal flexing in the way that Io is (see Section 13-4).

One explanation has been proposed by planetary scientists Gabriel Tobie, Jonathan Lunine, and Christophe Sotin. In their model, Titan developed a solid rocky core, a mantle of liquid water, and an icy crust soon after it formed some 4.5 billion years ago. The internal heat left over from formation facilitated volcanic activity that outgassed Titan's original complement of atmospheric methane molecules. Some of this methane was reabsorbed into the interior; the rest was destroyed by ultraviolet light from the Sun within the first billion years. From 3.5 billion to 2 billion years ago, Titan's atmosphere was essentially free of methane.

Over time, the rocky core increased in temperature due to the decay of radioactive elements like uranium. (These elements are concentrated in Titan's core to a greater extent than on Earth, which has rocky substances throughout its volume.) About 2 billion years ago, the core became warm enough to partially melt, and convection began in the melted portion. This allowed heat to flow from the core into the mantle, triggering a second episode of methane outgassing. This, too, eventually came to an end, and methane once again disappeared from Titan's atmosphere.

The third and final release of methane began about 500 million years ago, as the last of Titan's internal heat from radioactivity caused convection to take place in the satellite's ice crust. The amount of methane released during the current episode is enough to sustain the gas in the atmosphere, but not enough to provide extensive lakes of methane or ethane.

Within the next several hundred million years, methane outgassing on Titan will cease for good. In time, the hydrocarbon haze will disappear, the skies will clear, and the surface of Titan will be revealed to whatever outside observers may gaze upon it in that distant era.

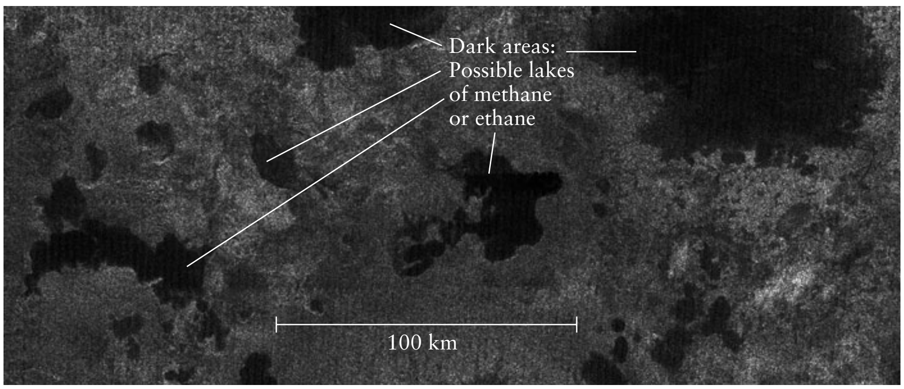


Figure 13-20 R I V U X G

Hydrocarbon Seas on Titan? *Cassini* used its onboard radar to record this view of a region near Titan's north pole. The dark regions show where the surface does not reflect radio waves at all, which would be the case if these regions were lakes filled with liquid methane or liquid ethane. In support of this idea, many of the dark regions appear to have liquid-carved inflow or outflow channels. (NASA/JPL)

13-9 Jupiter has dozens of small satellites that have different origins

WEB LINK 13.14

Besides the Galilean satellites, Jupiter has at least 59 other small satellites. Each is named for a mythological character associated with Jupiter. Appendix 3 gives details of some of the satellites' orbits. It is helpful to sort Jupiter's 63 known satellites (as of 2006) into three groups according to their sizes and orbits.

Four planet-sized satellites, Io, Europa, Ganymede, and Callisto, can be thought of as four of the terrestrial worlds of the solar system. All are comparable in size to the planet Mercury (see Table 13-1). The semimajor axes of their orbits range from 5.9 to 26 times the radius of Jupiter. (By comparison, the distance from the Earth to the Moon is about 60 times the Earth's radius.)

Four small inner satellites, Metis, Adrastea, Amalthea, and Thebe, have orbital semimajor axes between 1.8 and 3.1 times the radius of Jupiter. As Figure 13-21 shows, these satellites have irregular shapes, much like a potato (or an asteroid). The largest of the inner satellites, Amalthea, measures about 270 by 150 km (170 by 95 mi)—about 10 times larger than the moons of Mars, but only about one-tenth the size of the Galilean satellites. Discovered in 1892, Amalthea has a distinct reddish color due to sulfur that Jupiter's magnetosphere removed from Io's volcanic plumes, then deposited onto Amalthea's surface. The three other inner satellites are even smaller than Amalthea. They were first discovered in images from *Voyager 1* and *Voyager 2*.

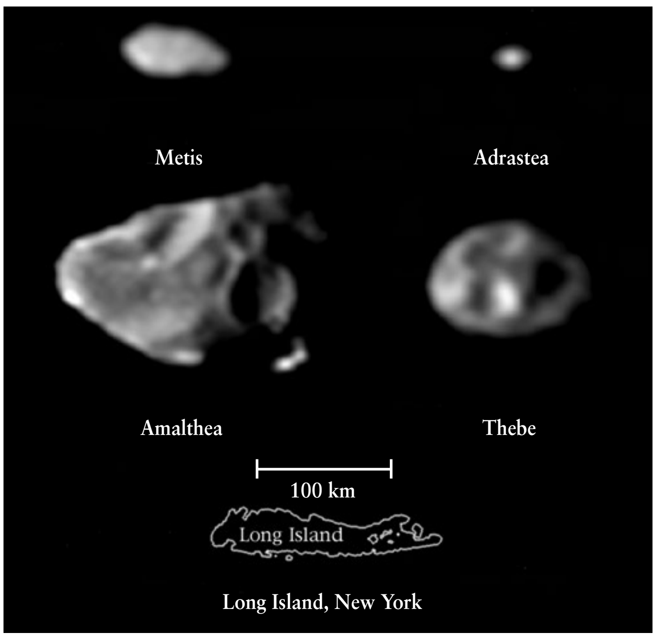


Figure 13-21 R I V U X G

Jupiter's Four Inner Satellites First observed in 1892, Amalthea was the last satellite of any planet to be discovered without the aid of photography. Metis, Adrastea, and Thebe, by contrast, are so small that they were not discovered until 1979. These images from the *Galileo* spacecraft are the first to show these three satellites as anything more than points of light. (NASA/JPL; Cornell University)

Fifty-five small outer satellites have diameters from 184 km for the largest to about 1 km for the smallest. All orbit far beyond Callisto, with semimajor axes that range from about 100 to 360 times the radius of Jupiter. (Twenty-three of these satellites were discovered in 2003 alone, and even more may be discovered as telescope technology continues to evolve.)

The Galilean satellites and the four inner satellites all orbit Jupiter in the plane of the planet's equator. (The same is true of Jupiter's rings, discussed in Section 12-9.) Furthermore, these are all **prograde orbits,** which means that these objects orbit Jupiter in the same direction as Jupiter's rotation. This is what we would expect if these satellites formed from the same primordial, rotating cloud as Jupiter itself. (In a similar way, the planets all orbit the Sun in nearly the same plane and in the same direction as the Sun's rotation. As we saw in Section 8-3, this reinforces the idea that the Sun and planets formed from the same solar nebula.)

Only a few of Jupiter's satellites orbit the planet in the same direction as Jupiter rotates

In contrast, the 55 outer satellites all circle Jupiter along orbits that are inclined at steep angles to the planet's equatorial plane. These satellites are thought not to have formed along with Jupiter; rather, they are probably wayward asteroids captured by Jupiter's powerful gravitational field. One piece of evidence for this is that 48 of the outer satellites are in **retrograde orbits:** They orbit Jupiter in a direction opposite the planet's rotation. Calculations show that it is easier for Jupiter to capture a satellite into a retrograde orbit than a prograde one.

13-10 The icy surfaces of Saturn's six moderate-sized moons provide clues to their histories

Saturn, too, has an extensive collection of satellites: 56 were known as of 2006. As we did for Jupiter's satellites in Section 13-9, we can subdivide the satellites of Saturn into three categories based on their sizes and their orbits. (Appendix 3 has information about the orbits of all of Saturn's satellites.) We will see that while Saturn does not have a set of large moons like the Galilean satellites of Jupiter, it does have a remarkably diverse collection of moderate-sized satellites.

Saturn's Retinue of Moons

One planet-sized satellite, Titan, is actually one of the terrestrial worlds of our solar system. With its diameter of 5150 km, Titan is intermediate in size between Mercury and Mars. Its prograde orbit has a semimajor axis of just over 20 times the radius of Saturn and lies in nearly the same plane as Saturn's equator and rings.

Six moderate-sized satellites were all discovered before 1800. These six satellites have properties and diameters that lead us to group them in pairs: Mimas and Enceladus (400 to 500 km in diameter), Tethys and Dione (roughly 1000 km in diameter), and Rhea and Iapetus (about 1500 km in diameter). The semimajor axes of their orbits range from 3.1 to 59 times the radius of

Saturn. Like Titan, all of their orbits are prograde and close to the plane of Saturn's equator.

Forty-nine small satellites have sizes from 3 to 266 km. Some are in prograde orbits as small as 2.2 times the radius of Saturn, and lie within or close to the rings (see Figure 12-25). These may be jagged fragments of ice and rock produced by impacts and collisions. The shepherd satellites that shape Saturn's F ring (see Figure 12-24) may be such fragments. Others are in retrograde orbits with semimajor axes as large as 380 times Saturn's radius, and may be captured asteroids.

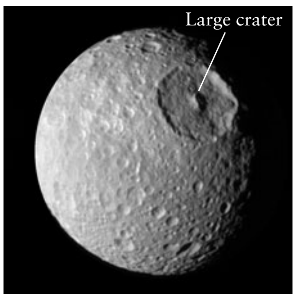
Six Strange Satellites

The most curious and mysterious of these worlds may be Saturn's six moderate-sized satellites (Figure 13-22). These six satellites have very low average densities of less than 1400 kg/m^3. This implies that unlike the satellites of the inner solar system or the Galilean satellites of Jupiter, they are composed primarily of ices (mostly frozen water and ammonia) with very little rock. The motions of these mid-sized satellites are not surprising: All six orbit the planet in the plane of Saturn's equator and in the direction of the planet's rotation, and all six exhibit synchronous rotation. What is surprising is the curious variety of surface features that the *Voyager* and *Cassini* spacecraft discovered on these satellites.

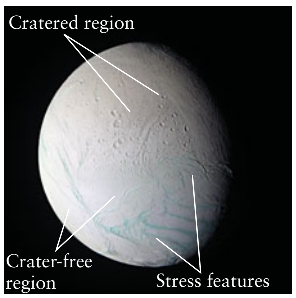
Mimas, the innermost of the six middle-sized moons, has the heavily cratered surface we would expect on such a small world. Its most remarkable feature is a crater so large (Figure 13-22*a*) that the impact that formed it must have come close to shattering Mimas into fragments.

In contrast to Mimas's heavily cratered surface, its neighbor and near twin, Enceladus, has extensive crater-free regions (Figure 13-22*b*). Some form of geologic activity has resurfaced these areas within the past 100 million years, and may still be under way. Geologic activity may also explain why Enceladus has an amazingly high albedo of 0.95, making it the most reflective large object in our solar system. Some process has freed Enceladus's icy surface of rock and dust, whose presence on the other icy satellites is the likeliest explanation for their lower albedos. Most remarkably, *Cassini* discovered that Enceladus is

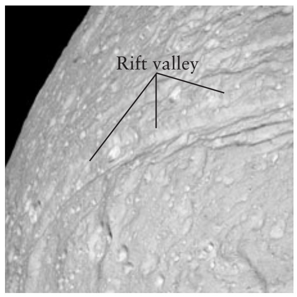
Although just 500 km across, Enceladus has powerful volcanoes at its south pole



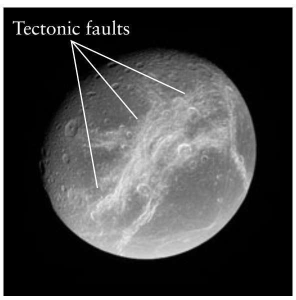
(a) Mimas
(diameter 392 km)



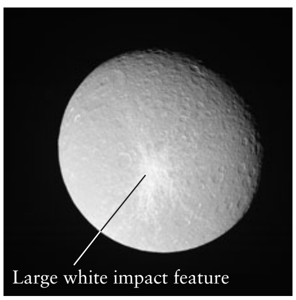
(b) Enceladus
(diameter 500 km)



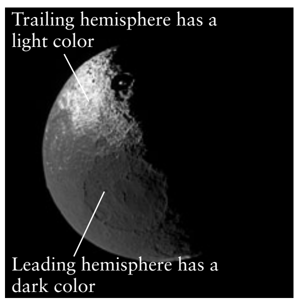
(c) Tethys
(diameter 1060 km)



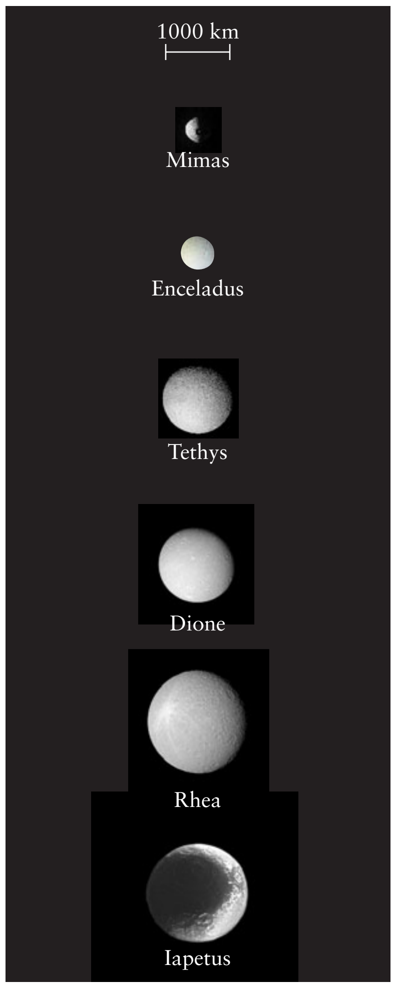
(d) Dione
(diameter 1120 km)



(e) Rhea
(diameter 1530 km)



(f) Iapetus
(diameter 1460 km)

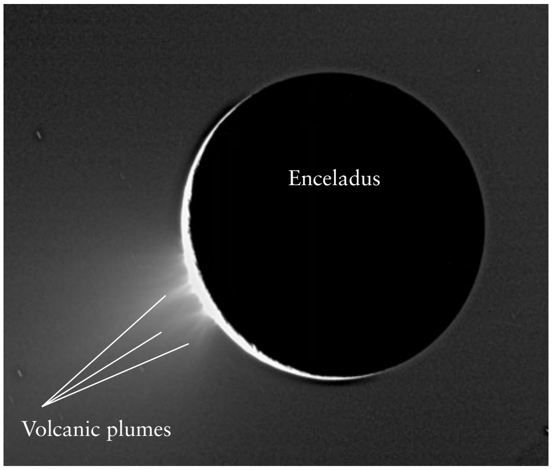


(g) Satellites to scale

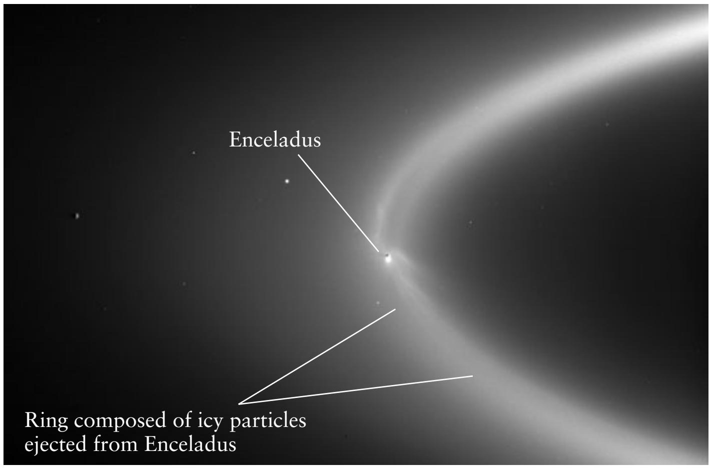
WEB LINK 13.15

Figure 13-22 R I V U X G

Saturn's Six Mid-Sized Satellites **(a)–(f)** These images from *Cassini* show the variety of surface features found on Saturn's six moderate-sized satellites. The satellites depicted in parts (a)–(f) are *not* shown to scale. **(g)** These images show the relative sizes of Saturn's moderate-sized satellites. (NASA/JPL/Space Science Institute)



(a)



(b)

Figure 13-23 R I V U X G

Volcanoes on Enceladus (a) This *Cassini* image shows plumes of icy particles being ejected from ice volcanoes near the south pole of Enceladus. The faint plumes are best seen when they are backlit by the Sun, so Enceladus appears as a crescent. **(b)** A wide-angle backlit view from *Cassini* shows how particles ejected from Enceladus become part of Saturn's rings. This structure is actually part of the E ring, within which Enceladus orbits (see Figure 12-21). (NASA/JPL/Space Science Institute)

volcanically active, though with ice taking the place of lava. Figure 13-23*a* shows several distinct plumes ejecting icy particles from near the satellite's south pole, the same region where the surface is laced with bluish stress cracks (see Figure 13-22*b*). Some of the particles are ejected with such force that they escape Enceladus altogether and go into orbit around Saturn, forming a tenuous ring (Figure 13-23*b*).

What could be powering the geologic activity on Enceladus? One guess is tidal heating, which provides the internal heat for the inner Galilean satellites of Jupiter (see Section 13-4). The satellite Dione orbits Saturn in 65.7 hours, almost exactly twice Enceladus's orbital period of 32.9 hours, and the 1:2 ratio of orbital periods should set Enceladus up to be caught in a tidal tug-of-war between Saturn and Dione. Calculations suggest that the amount of tidal heating produced in this way is modest, but should be enough to melt the ices of Enceladus at least part of the time. Curiously, there is also a 1:2 ratio between the orbital periods of Mimas and Saturn's moon Tethys, but the cratered surface of Mimas shows no sign that this moon was ever tidally heated.

Tethys and Dione are the next-largest pair of Saturn's six moderate-sized satellites. Most of Tethys, like Mimas, is heavily cratered. But Tethys also has an immense rift valley called Ithaca Chasma that extends for 2000 km, or about $^3\!/_4$ of the way around Tethys (Figure 13-22*c*). This rift may have formed when the water inside Tethys cooled and solidified, or it may be the result of a shock wave that traveled through Tethys after an immense asteroid impact. (A large impact crater lies on the side of Tethys opposite Ithaca Chasma.) Dione, slightly larger than Tethys, has a strange surface dichotomy: The surfaces of its leading hemisphere (that is, the hemisphere facing toward its direction of orbital motion) and trailing hemisphere are quite different. Dione's leading hemisphere is heavily cratered, but its trailing hemisphere is covered by a network of bright stress cracks caused by some sort of tectonic activity (Figure 13-22*d*).

Because Tethys and Dione have about double the diameter of Enceladus, the energy to power their geologic activity might have come from within. These larger satellites retain heat more effectively, including heat from the decay of radioactive materials in each satellite's interior. Such materials are scarce within the icy satellites of Saturn, but only relatively little heat is needed to power ice volcanism.

Rhea and Iapetus are the largest and most distant of Saturn's moderate-sized moons. Rhea is heavily cratered, including a bright, fresh-looking crater on its leading hemisphere (Figure 13-22*e*). This may be the result of a more recent impact than those that formed most of Rhea's craters.

Iapetus has been known to be unusual ever since its discovery in 1671: Unlike the other satellites, its brightness varies greatly as it orbits Saturn. Images from *Cassini* show extreme differences between the leading and trailing hemispheres of Iapetus (Figure 13-22*f*). Most of the leading hemisphere is as black as asphalt (albedo = 0.05), but the trailing hemisphere is highly reflective (albedo = 0.50), like the surfaces of the other moderate-sized moons.

What could account for the two-sided nature of Iapetus? One possibility is that the dark material covering Iapetus's leading hemisphere may have come from Phoebe, a smaller satellite that orbits about $3\frac{1}{2}$ times farther from Saturn than Iapetus. Phoebe moves in a retrograde direction along an orbit tilted well away from the plane of Saturn's equator. This motion leads scientists to suspect that Phoebe is a captured asteroid, and its dark surface does indeed resemble the surfaces of a particular class of carbon-rich asteroids. Bits and pieces of this dark charcoal-like material drifting toward Saturn may have been swept up onto the leading hemisphere of Iapetus. However, a sticking point with this scenario is that the dark material on Iapetus is reddish in color, while Phoebe's surface is not. An alternative explanation is that a large portion of the leading hemisphere of Iapetus was coated with methane ice, which darkened when exposed to solar radiation.

Key Words

aerosol, p. 344
bright terrain (Ganymede), p. 340
dark terrain (Ganymede), p. 340
Galilean satellites, p. 330
hydrocarbon, p. 343
ice rafts (Europa), p. 338
Io torus, p. 336
occultation, p. 330
polymer, p. 344
prograde orbit, p. 346
retrograde orbit, p. 346
tidal heating, p. 334
ultramafic lava, p. 335

Key Ideas

INTERACTIVE EXERCISE 13.2

Nature of the Galilean Satellites: The four Galilean satellites orbit Jupiter in the plane of its equator. All are in synchronous rotation.

• The orbital periods of the three innermost Galilean satellites, Io, Europa, and Ganymede, are in the ratio 1:2:4.

• The two innermost Galilean satellites, Io and Europa, have roughly the same size and density as our Moon. They are composed principally of rocky material. The two outermost Galilean satellites, Ganymede and Callisto, are roughly the size of Mercury. Lower in density than either the Moon or Mercury, they are made of roughly equal parts ice and rock.

• The Galilean satellites probably formed in a similar fashion to our solar system but on a smaller scale.

Io: Io is covered with a colorful layer of sulfur compounds deposited by frequent explosive eruptions from volcanic vents. These eruptions resemble terrestrial geysers.

• The energy to heat Io's interior and produce the satellite's volcanic activity comes from tidal forces that flex the satellite. This tidal flexing is aided by the 1:2:4 ratio of orbital periods among the inner three Galilean satellites.

• The Io torus is a ring of electrically charged particles circling Jupiter at the distance of Io's orbit. Interactions between this ring and Jupiter's magnetic field produce strong radio emissions. Io may also have a magnetic field of its own.

Europa: While composed primarily of rock, Europa is covered with a smooth layer of water ice.

• The surface has hardly any craters, indicating a geologically active history. Other indications are a worldwide network of long cracks and ice rafts that indicate a subsurface layer of liquid water or soft ice. As for Io, tidal heating is responsible for Europa's internal heat.

• An ocean may lie beneath Europa's frozen surface. Minerals dissolved in this ocean may explain Europa's induced magnetic field.

Ganymede: Two types of terrain are found on the icy surface of Ganymede: areas of dark, ancient, heavily cratered surface and regions of heavily grooved, lighter-colored, younger terrain.

• Ganymede is highly differentiated, and probably has a metallic core. It has a surprisingly strong magnetic field and a magnetosphere of its own.

• While there is at present little tidal heating of Ganymede, it may have been heated in this fashion in the past. An induced magnetic field suggests that it, too, has a layer of liquid water beneath the surface.

Callisto: Callisto has a heavily cratered crust of water ice. The surface shows little sign of geologic activity, because there was never any significant tidal heating of Callisto. However, some unknown processes have erased the smallest craters and blanketed the surface with a dark, dusty substance.

• Magnetic field data seem to suggest that Callisto has a shallow subsurface ocean.

Titan: The largest Saturnian satellite, Titan, is a terrestrial world with a dense nitrogen atmosphere. A variety of hydrocarbons are produced there by the interaction of sunlight with methane. These compounds form an aerosol layer in Titan's atmosphere and fall as a gentle rain on the surface.

• Titan's surface shows that liquid hydrocarbons have flowed over its surface, forming streams, rivers, and outflow channels. Very little of this liquid appears to be present on Titan's surface today.

Other Satellites: As of 2006, Jupiter has a total of 63 known satellites and Saturn has a total of 56.

• In addition to the Galilean satellites, Jupiter has four small inner satellites that lie inside Io's orbit. Like the Galilean satellites, these orbit in the plane of Jupiter's equator. The remaining satellites are small and move in much larger orbits that are noticeably inclined to the plane of Jupiter's equator. Many of these orbit in the direction opposite to Jupiter's rotation.

• In addition to Titan, six moderate-sized moons circle Saturn in regular orbits: Mimas, Enceladus, Tethys, Dione, Rhea, and Iapetus. They are probably composed largely of ice, but their surface features and histories vary significantly. The other, smaller moons include shepherd satellites that control the shapes of Saturn's rings and captured asteroids in large retrograde orbits.

Questions

Review Questions

1. If you observed the Galilean satellites through a telescope for a single night, could you notice their motions around Jupiter?
2. Why can't the Galilean satellites be seen with the naked eye?
3. During the time it takes Ganymede to complete one orbit, how many orbits do Io and Europa complete?
4. In what ways does the system of Galilean satellites resemble our solar system? In what ways is it different?
5. No spacecraft from Earth has ever landed on any of the Galilean satellites. How, then, can we know anything about the chemical compositions of these satellites?
6. In what ways did the formation of the Galilean satellites mimic the formation of the planets? In what ways were the two formation processes different?

7. In the classic science-fiction film *2010: The Year We Make Contact,* an alien intelligence causes Jupiter to contract so much that nuclear reactions begin at its center. As a result, Jupiter becomes a star like the Sun. Is this possible in principle? Explain your answer.
8. All of the Galilean satellites orbit Jupiter in the same direction. Furthermore, the planes of their orbits all lie within 0.5° of Jupiter's equatorial plane. Explain why these observations are consistent with the idea that the Galilean satellites formed from a "Jovian nebula."
9. What is the source of energy that powers Io's volcanoes? How is it related to the orbits of Io and the other Galilean satellites?
10. Io has no impact craters on its surface, while our Moon is covered with craters. What is the explanation for this difference?
11. Despite all the gases released from its interior by volcanic activity, Io does not possess a thick atmosphere. Explain why not.
12. Long before the *Voyager* flybys, Earth-based astronomers reported that Io appeared brighter than usual for the few hours after it emerged from Jupiter's shadow. From what we know about the material ejected from Io's volcanoes, suggest an explanation for this brief brightening of Io.
13. How do lavas on Io differ from typical lavas found on Earth? What does this difference tell us about Io's interior?
14. What accounts for the different colors found on Io's surface?
15. What is the Io torus? What is its source?
16. What is the origin of the electric current that flows through Io?
17. How was the *Galileo* spacecraft used to determine the internal structure of Io and the other Galilean satellites?
18. What surface features on Europa provide evidence for geologic activity?
19. What is the evidence for an ocean of liquid water beneath Europa's icy surface? What is the evidence that substances other than water are dissolved in this ocean?
20. What aspects of Europa lead scientists to speculate that life may exist there? What precautions have been taken to prevent contaminating Europa with Earth organisms?
21. Why is ice an important constituent of Ganymede and Callisto, but not of the Earth's Moon?
22. How do scientists know that the dark terrain on Ganymede is younger than the bright terrain?
23. In what ways is Ganymede like our own Moon? In what ways is it different? What are the reasons for the differences?
24. Why were scientists surprised to learn that Ganymede has a magnetic field? What does this field tell us about Ganymede's history?
25. What leads scientists to suspect that there is liquid water beneath Ganymede's surface?
26. Why are numerous impact craters found on Ganymede and Callisto but not on Io or Europa?
27. Describe the surprising aspects of Callisto's surface and interior that were revealed by the *Galileo* spacecraft. Why did these come as a surprise?
28. Compare and contrast the surface features of the four Galilean satellites. Discuss their relative geological activity and the evolution of these four satellites.
29. The larger the orbit of a Galilean satellite, the less geologic activity that satellite has. Explain why.
30. Explain how the 1:2:4 ratio of the orbital periods of Io, Europa, and Ganymede is related to the geologic activity on these satellites.
31. Describe Titan's atmosphere. What effect has the Sun's ultraviolet radiation had on Titan's atmosphere?
32. In what ways do hydrocarbons on Titan behave like water on Earth?
33. Why does the presence of methane in Titan's atmosphere imply that Titan has had recent volcanic activity?
34. What is the evidence that there were liquid hydrocarbons on Titan in the past? That there are liquid hydrocarbons now?
35. How would you account for the existence of the satellites of Jupiter other than the Galilean ones?
36. Which of Saturn's moderate-sized satellites show evidence of geologic activity? What might be the energy source for this activity?
37. Explain why debris from Phoebe would be expected to pile up only on the leading hemisphere of Iapetus. (*Hint:* How do the orbits of these two satellites compare? How does the orbital motion of debris falling slowly inward toward Saturn compare with the orbital motion of Iapetus?)
38. Saturn's equator is tilted by 27° from the ecliptic, while Jupiter's equator is tilted by only 3°. Use these data to explain why we see fewer transits, eclipses, and occultations of Saturn's satellites than of the Galilean satellites.

Advanced Questions

Questions preceded by an asterisk (*) involve topics discussed in Box 1-1 or Box 7-2.

Problem-solving tips and tools

Newton's form of Kepler's third law (Box 4-4) relates the masses of two objects in orbit around each other to the period and size of the orbit. The small-angle formula is discussed in Box 1-1. The best seeing conditions on the Earth give a limiting angular resolution of $\frac{1}{4}$ arcsec. Because the orbits of the Galilean satellites are almost perfect circles, you can easily calculate the orbital speeds of these satellites from the data listed in Table 13-1. Data about Jupiter itself are given in Table 12-1 and data about all of the satellites of Jupiter and Saturn can be found in Appendix 3. For a discussion of escape speed and how planets retain their atmospheres, see Box 7-2.

39. Using the orbital data in Table 13-1, demonstrate that the Galilean satellites obey Kepler's third law.

*40. What is the size of the smallest feature you should be able to see on a Galilean satellite through a large telescope under conditions of excellent seeing when Jupiter is near opposition? How does this compare with the best *Galileo* images, which have resolutions around 25 meters?

41. Using the diameter of Io (3642 km) as a scale, estimate the height to which the plume of Pele rises above the surface of Io in Figure 13-5*a*. (You will need to make measurements on this figure using a ruler.) Compare your answer to the value given in the figure caption.

42. Explain why the volcanic plumes in Figure 13-5*b* have a bluish color. (*Hint:* See Box 5-4.)

43. Jupiter, its magnetic field, and the charged particles that are trapped in the magnetosphere all rotate together once every 10 hours. Io takes 1.77 days to complete one orbit. Using a diagram, explain why particles from Jupiter's magnetosphere hit Io primarily from behind (that is, on the side of Io that trails as it orbits the planet).

44. Assuming material is ejected from Io into Jupiter's magnetosphere at the rate of 1 ton per second (1000 kg/s), how long will it be before Io loses 10% of its mass? How does your answer compare with the age of the solar system?

45. Volcanic eruptions from Io add about 1 ton of material to Jupiter's magnetosphere every second. Based on the information given in Section 13-4, what fraction of the material ejected from Io's volcanoes goes into Jupiter's magnetosphere?

46. How long does it take for Ganymede to entirely enter or entirely leave Jupiter's shadow? Assume that the shadow has a sharp edge.

47. Figure 13-15 shows that Ganymede's dark terrain has both more craters and larger craters than the bright terrain. Explain what this tells you about the sizes of meteoroids present in the solar system in the distant past and in the more recent past.

48. Put a few cubes of ice in a glass, then fill the glass with warm water to make the ice cubes crack. The fracture lines in each cube are clearly visible because they reflect light. Use this observation to suggest why Ganymede's icy bright terrain, which may have been badly fractured by the stresses that produced the grooves (Figure 13-15), is 25% more reflective than the dark terrain.

***49.** Find the escape speed on Titan. What is the limiting molecular weight of gases that could be retained by Titan's gravity? (*Hint:* Use the ideas presented in Box 7-2 and assume an average atmospheric temperature of 95 K.)

50. Many of the gases in the atmosphere of Titan, such as methane, ethane, and acetylene, are highly flammable. Why, then, doesn't Titan's atmosphere catch fire? (*Hint:* What gas in our atmosphere is needed to make wood, coal, or gasoline burn?)

51. Suppose the Earth's Moon were removed and replaced in its orbit by Titan. What changes would you expect to occur in Titan's atmosphere? Would solar eclipses be more or less common as seen from Earth? Explain your answers.

52. At infrared wavelengths, the Hubble Space Telescope can see details on Titan's surface as small as 580 km (360 mi) across. Determine the angular resolution of the Hubble Space Telescope using infrared light. If visible light is used, is the angular resolution better, worse, or the same? Explain your answer.

53. (**a**) To an observer on Enceladus, what is the time interval between successive oppositions of Dione? Explain your answer. (**b**) As seen from Enceladus, what is the angular diameter of Dione at opposition? How does this compare to the angular diameter of the Moon as seen from Earth (about $\frac{1}{2}°$)?

Discussion Questions

54. If you could replace our Moon with Io, and if Io could maintain its present amount of volcanic activity, what changes would this cause in our nighttime sky? Do you think that Io could in fact remain volcanically active in this case? Why or why not?

55. Speculate on the possibility that Europa, Ganymede, or Callisto might harbor some sort of life. Explain your reasoning.

56. Comment on the suggestion that Titan may harbor life-forms.

57. Jupiter's satellite Io and Saturn's satellite Enceladus are both geologically active, and both are in 2-to-1 resonances with other satellites. However, the amount of geologic activity of Enceladus is far less than on Io. Discuss some possible reasons for this difference.

58. Imagine that you are in charge of planning a successor to the *Cassini* spacecraft to further explore the Saturnian system. In your opinion, which satellites in the system should be examined more closely? What data should be collected? What kinds of questions should the new mission attempt to answer?

Web/eBook Questions

59. Various spacecraft missions have been proposed to explore Europa in greater detail. Search the World Wide Web for information about these. How would these missions test for the presence of an ocean beneath Europa's surface?

60. Twenty-one new satellites of Jupiter were discovered in 2003. Search the World Wide Web for information about how these satellites were discovered. How was it determined that these satellites are actually in orbit around Jupiter?

VIDEO 13.4

61. The Surface of Ganymede. Access and view the video "Jupiter's Moon Ganymede" in Chapter 13 of the *Universe* Web site or eBook. Describe the different surface features that you see, and explain how each type of feature was formed.

Activities

Observing Projects

STARRY NIGHT

Observing tips and tools

You can easily find the apparent positions of the Galilean satellites for any date and time using the *Starry Night Enthusiast*™ software on the CD-ROM that accompanies certain printed copies of this textbook. For even more detailed information about satellite positions, consult the "Satellites of Jupiter" section in the *Astronomical Almanac* for the current year. If your goal is to view Saturn's satellites, consult the section entitled "Satellites of Saturn" in the *Astronomical Almanac* for the current year. This includes a diagram showing the orbits of Mimas, Enceladus, Tethys, Dione, Rhea, Titan, and Hyperion. Plan your observing session by looking up the dates and times of the most recent greatest eastern elongations of the various satellites. You will have to convert from universal time (UT), also known as Greenwich Mean Time, to your local time zone. Then, using the tick marks along the orbits in the diagram, estimate the positions of the satellites relative to Saturn at the time you will be at the telescope. Another useful resource is the "Celestial Calendar" section of *Sky & Telescope*. During months when Saturn is visible in the night sky, this section of the magazine includes a chart of Saturn's satellites.

62. Observe Jupiter through a pair of binoculars. Can you see all four Galilean satellites? Make a drawing of what you observe. If you look again after an hour or two, can you see any changes?

63. Observe Jupiter through a small telescope on three or four consecutive nights. Make a drawing each night showing the positions of the Galilean satellites relative to Jupiter. Record the time and date of each observation. Consult the sources listed above in the "Observing tips and tools" to see if you can identify the satellites by name.

64. If you have access to a moderately large telescope, make arrangements to observe several of Saturn's satellites. At the telescope, you should have no trouble identifying Titan. Tethys, Dione, and Rhea are about one-sixth as bright as Titan and should be the next easiest satellites to find. Can you confidently identify any of the other satellites?

STARRY NIGHT

65. Use the *Starry Night Enthusiast™* program to observe the Galilean satellites of Jupiter. Open the **Favourites pane** and click on **Guides > Atlas** to display the entire celestial sphere. Open the **Find** pane and double-click the entry for Jupiter to center this planet in the view. Using the controls at the right-hand end of the toolbar, zoom in to a field of view of approximately **$13' \times 9'$.** Stop time flow and in the toolbar, set the date and time to March 4, 2004, at 12:00:00 A.M to see Jupiter at opposition. Set the **Time Flow Rate** to 20 minutes by clicking on the number in the Time Flow Rate box and setting the value with the keyboard. Then click on the Run Time Forward button (a triangle that points to the right). You will see the four Galilean satellites orbiting Jupiter. (If these moons appear to move too quickly, adjust the Time Flow Rate to 10 minutes.) (**a**) Are all four satellites ever on the same side of Jupiter? (**b**) Observe the satellites passing in front of and behind Jupiter and look for their shadows upon the planet. (Zoom in as needed). Explain how your observations tell you that all four satellites orbit Jupiter in the same direction.

STARRY NIGHT

66. Use the *Starry Night Enthusiast™* program to view Saturn from its satellite Enceladus. First click the **Home** button in the toolbar. Select **Options > Viewing Location . . .** from the menu. In the dialog window that appears at the top of the Viewing Location, set the drop-down box next to the "**View from:**" label to read "**the surface of**" and set the second drop-down box to read "**Enceladus**" (listed under Saturn). In the list of locations that appears under the **List** tab on Enceladus, scroll to "**Ahmad.**" Click on this entry to highlight it and then click the **Set Location** button. To see Saturn from this location on Enceladus, center on Saturn by double-clicking the entry for this planet in the **Find** pane. (**a**) Stop the time flow and set the Time Flow Rate to **1 minute** in the toolbar, then click on the **Run Time Forward** button (a triangle that points to the right). How do the stars appear to move as seen from this location on Enceladus? How does Saturn appear to move? What do these observations tell you about the relationship between the orbital period and rotation period of Enceladus? (**b**) Set the Time Flow Rate to **1 hour.** By stepping forward through time using the rightmost single-step button, determine how much time elapses from when Saturn appears full from this location to when it next appears full. Explain why this is the same as the orbital period of Enceladus, and compare this to the value of the orbital period given in Appendix 3.

STARRY NIGHT

67. Use the *Starry Night Enthusiast™* program to examine the satellites of Saturn. Open the **Favourites** pane and select **Solar System > Saturn.** Remove the astronaut's feet from this view by clicking on **View > Feet.** In this view you can rotate Saturn by placing the mouse cursor over the image of the planet, holding down the mouse button, and moving the mouse (on a two-button mouse, hold down the left mouse button). Use this technique to rotate Saturn so that you are viewing the rings edge-on. Then use the elevation controls in the toolbar (the buttons to the left of the Home button) to move closer to and further away from Saturn. This allows you to identify the satellites of Saturn, since they will appear to move whereas the distant stars will remain stationary as you move with respect to the planet. Alternately, you can reduce the confusion in identifying these moons by clicking on **View > Stars > Stars** to remove the stars from the view. You should be able to see at least eight satellites. Which satellites are these? (Move the mouse to center the cursor on a satellite and its name will appear. If you cannot see all eight, try moving further from Saturn (increase the viewing location elevation with the elevation button showing the Up arrowhead). (**b**) The plane of Saturn's rings is the same as the plane of Saturn's equator. Which satellites appear to be the farthest from this plane?

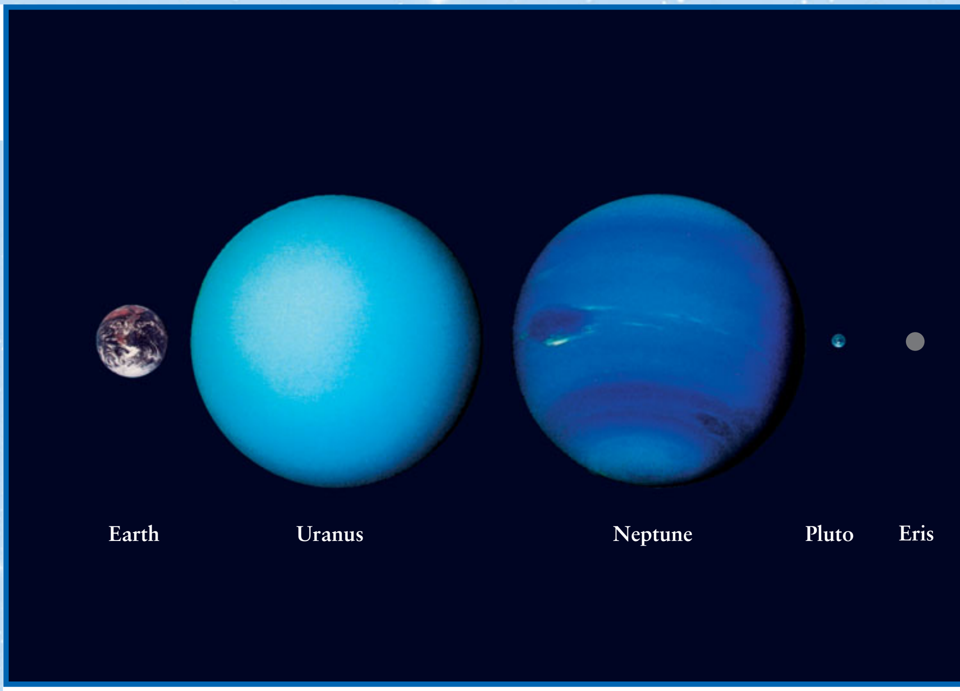
Collaborative Exercises

68. Imagine that scientists are proposing to send a robotic lander to Jupiter's satellite Callisto. Create a 100-word written proposal describing a robotic lander mission to another of the Galilean satellites, explaining why your group found it to be the most interesting and why the government should allocate the money for your alternative project. In your proposal, be sure to demonstrate your knowledge of Callisto and the other satellite.

69. From the data and the accompanying images for Table 13-1, "The Galilean Satellites Compared to the Moon, Mercury, and Mars," use someone's shoe to represent the 150,000 km diameter of Jupiter and determine about how many "shoes" away would each of the Galilean satellites be from Jupiter.

14

Uranus, Neptune, Pluto, and the Kuiper Belt: Remote Worlds



WEB LINK 14.1

R I V U X G

Earth, Uranus, Neptune, Pluto, and Eris to scale. (Alan Stern, Southwest Research Institute; Marc Buie, Lowell Observatory; NASA; and ESA)

Beyond Saturn, in the cold, dark recesses of the solar system, orbit three planets that have long been shrouded in mystery. These planets—Uranus, Neptune, and Pluto (all shown here to the same scale as the Earth)—are so distant, so dimly lit by the Sun, and so slow in their motion against the stars that they were unknown to ancient astronomers and were discovered only after the invention of the telescope. Even then, little was known about Uranus and Neptune until *Voyager 2* flew past these planets during the 1980s.

Surrounded by a system of small moons and thin, dark rings, Uranus is tipped on its side so that its axis of rotation lies nearly in its orbital plane. This remarkable orientation suggests that Uranus may have been the victim of a staggering impact by a massive planetesimal. Neptune is, at first glance, a denser, more massive version of Uranus, but it is a far more active world. It has an internal energy source that Uranus seemingly lacks, as well as atmospheric bands and storm activity resembling those on Jupiter.

Neptune also has dark rings, small, icy moons, and an intriguing large satellite, Triton, which is nearly devoid of impact craters and has geysers that squirt nitrogen-rich vapors. Triton's retrograde orbit suggests that this strange world may have been gravitationally captured by Neptune.

Close cousins of Triton are Pluto and thousands of other trans-Neptunian objects, among the most remote members of the solar system. These objects harbor many mysteries, in part because they have not yet been visited by a spacecraft.

Learning Goals

By reading the sections of this chapter, you will learn

14-1 How Uranus and Neptune were discovered

14-2 The unusual properties of the orbit and atmosphere of Uranus

14-3 What gives Neptune's clouds and atmosphere their distinctive appearance

14-4 The internal structures of Uranus and Neptune

14-5 The unique orientations of the magnetic fields of Uranus and Neptune

14-6 Why the rings of Uranus and Neptune are hard to see

14-7 What could have powered geologic activity on Uranus's moderate-sized moons

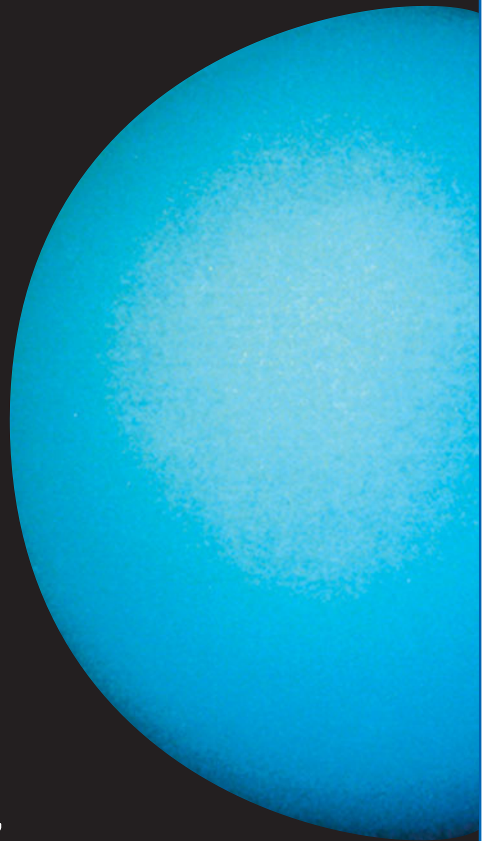
14-8 Why Neptune's satellite Triton is destined to be torn apart

14-9 How Pluto came to be discovered

14-10 What shapes the orbits of Pluto and the thousands of other objects that orbit beyond Neptune

Table 14-1 Uranus Data

Average distance from Sun:	19.194 AU = 2.871×10^9 km
Maximum distance from Sun:	20.017 AU = 2.995×10^9 km
Minimum distance from Sun:	18.371 AU = 2.748×10^9 km
Eccentricity of orbit:	0.0429
Average orbital speed:	6.83 km/s
Orbital period:	84.099 years
Rotation period (internal):	17.24 hours
Inclination of equator to orbit:	97.86°
Inclination of orbit to ecliptic:	0.77°
Diameter:	51,118 km = 4.007 Earth diameters (equatorial)
Mass:	8.682×10^{25} kg = 14.53 Earth masses
Average density:	1318 kg/m^3
Escape speed:	21.3 km/s
Surface gravity (Earth = 1):	0.90
Albedo:	0.56
Average temperature at cloudtops:	−218°C = −360°F = 55 K
Atmospheric composition (by number of molecules):	82.5% hydrogen (H_2), 15.2% helium (He), 2.3% methane (CH_4)
Average distance from Sun:	30.066 AU = 4.498×10^9 km
Maximum distance from Sun:	30.367 AU = 4.543×10^9 km



WEB LINK 14.2

R I V U X G

(NASA/JPL)

14-1 Uranus was discovered by chance, but Neptune's existence was predicted by applying Newtonian mechanics

Before the eighteenth century, only six planets were known to orbit the Sun: Mercury, Venus, Earth, Mars, Jupiter, and Saturn. That state of affairs changed thanks to William Herschel, a German-born musician who moved to England in 1757 and became fascinated by astronomy.

Discovering the "Georgian Star"

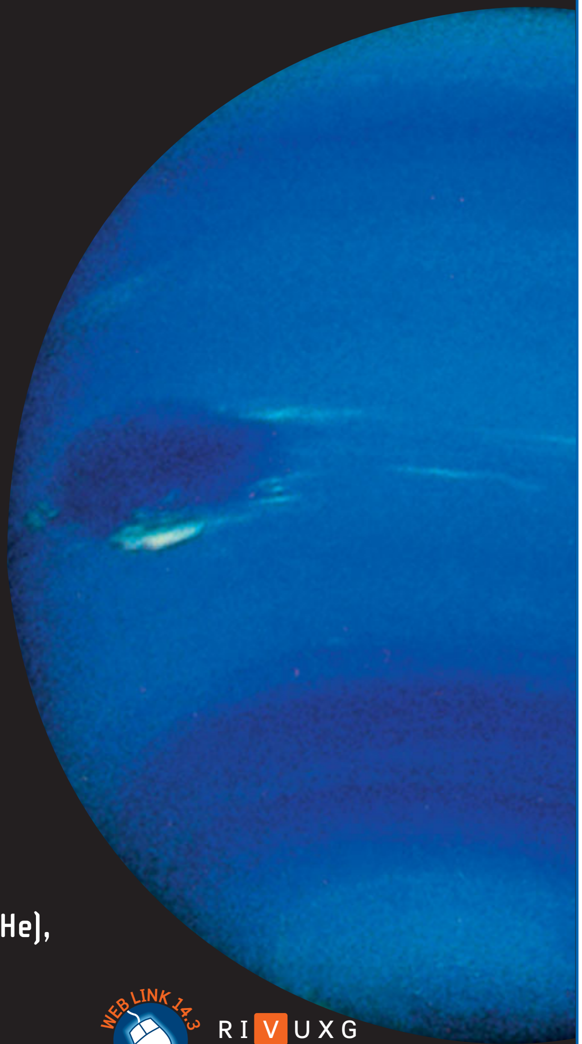
Using a telescope that he built himself, Herschel was systematically surveying the sky on March 13, 1781, when he noticed a faint, fuzzy object that he first thought to be a distant comet. By the end of 1781, however, his observations had revealed that the object's orbit was relatively circular and was larger than Saturn's orbit. Comets, by contrast, can normally be seen only when they follow elliptical orbits that bring them much closer to the Sun.

Herschel had discovered the seventh planet from the Sun. In doing so, he had doubled the radius of the known solar system from 9.5 AU (the semimajor axis of Saturn's orbit) to 19.2 AU (the distance from the Sun to Uranus). Herschel originally named his discovery Georgium Sidus (Latin for "Georgian star") in honor of the reigning monarch, George III. The name Uranus—in Greek mythology, the personification of Heaven—came into currency only some decades later.

Although Herschel received the credit for discovering Uranus, he was by no means the first person to have seen it. At opposition, Uranus is just barely bright enough to be seen with the naked eye under good observing conditions, so it was probably seen by the ancients. Many other astronomers with telescopes had

Table 14-2 Neptune Data

Average distance from Sun:	30.066 AU = 4.498×10^9 km
Maximum distance from Sun:	30.367 AU = 4.543×10^9 km
Minimum distance from Sun:	29.765 AU = 4.453×10^9 km
Eccentricity of orbit:	0.010
Average orbital speed:	5.5 km/s
Orbital period:	164.86 years
Rotation period (internal):	16.11 hours
Inclination of equator to orbit:	29.56°
Inclination of orbit to ecliptic:	1.77°
Diameter:	49,528 km = 3.883 Earth diameters (equatorial)
Mass:	1.024×10^{26} kg = 17.15 Earth masses
Average density:	1638 kg/m^3
Escape speed:	23.5 km/s
Surface gravity (Earth = 1):	1.1
Albedo:	0.51
Average temperature at cloudtops:	−218°C = −360°F = 55 K
Atmospheric composition (by number of molecules):	79% hydrogen (H_2), 18% helium (He), 3% methane (CH_4)
Average distance from Sun:	30.066 AU = 4.498×10^9 km
Maximum distance from Sun:	30.367 AU = 4.543×10^9 km



WEB LINK 14.3

RIVUXG

(NASA/JPL)

sighted this planet before Herschel; it is plotted on at least 20 star charts drawn between 1690 and 1781. But all these other observers mistook Uranus for a dim star. Herschel was the first to track its motion relative to the stars and recognize it as a planet. This was no small task, because Uranus moves very slowly on the celestial sphere, just over 4° in the space of a year (compared to about 12° for Saturn and about 35° for Jupiter).

The Discovery of Neptune

It was by carefully tracking Uranus's slow motions that astronomers were led to discover Neptune. By the beginning of the nineteenth century, it had become painfully clear to astronomers that they could not accurately predict the orbit of Uranus using Newtonian mechanics. By 1830 the discrepancy between the planet's predicted and observed positions had become large enough (2 arcmin) that some scientists suspected that Newton's law of gravitation might not be accurate at great distances from the Sun.

By the mid-1840s, two scientists working independently—the French astronomer Urbain Jean Joseph Le Verrier and the English mathematician John Couch Adams—were exploring an earlier and sounder suggestion. Perhaps the gravitational pull of an as yet undiscovered planet was causing Uranus to deviate slightly from its predicted orbit. Calculations by both scientists concluded that Uranus had indeed caught up with and had passed a more distant planet. Uranus had accelerated slightly as it approached the unknown planet, then decelerated slightly as it receded from the planet.

Inspired by Le Verrier's results, astronomers at Cambridge University Observatory undertook a six-week search for the proposed new planet in the summer of 1846. They were unsuccessful, in part because they lacked accurate star maps of the part of the sky being searched. The Cambridge astronomers also did not receive supporting information from Adams in a timely manner.

Meanwhile, Le Verrier wrote to Johann Gottfried Galle at the Berlin Observatory with detailed predictions of where to search for the new planet. Galle received the letter on September 23, 1846. That very night, aided by more complete star maps and after just a half-hour of searching, Galle and Heinrich d'Arrest located an uncharted star with the expected brightness in the predicted location. Subsequent observations confirmed that this new "star" showed a planetlike motion with respect to other stars.

Le Verrier proposed that the planet be called Neptune. After years of debate between English and French astronomers, the credit for its discovery came to be divided between Adams and Le Verrier. Neptune is thus the only planet whose existence was revealed by calculation rather than chance discovery.

At opposition, Neptune can be bright enough to be visible through small telescopes. The first person thus to have seen it was probably Galileo. His drawings from January 1613, when he was using his telescope to observe Jupiter and its four large satellites, show a "star" less than 1 arcmin from Neptune's location. Galileo even noted in his observation log that on one night this star seemed to have moved in relation to the other stars. But Galileo would have been hard pressed to identify Neptune as a planet, because its motion against the background stars is so slow (just over 2° per year).

Observing Uranus and Neptune

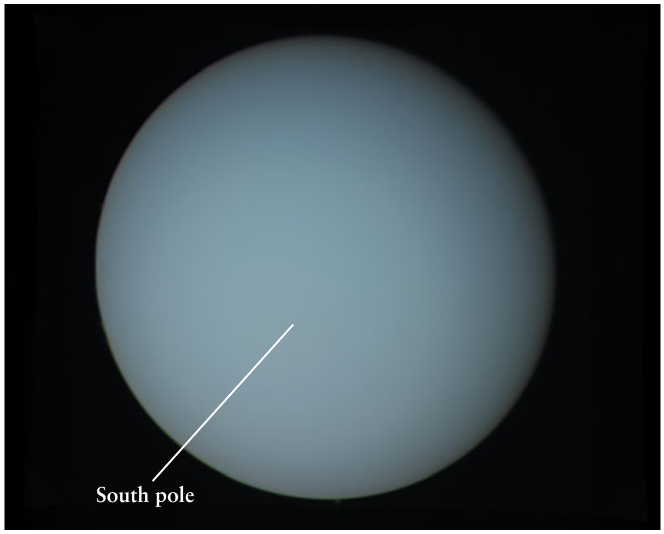
Through a large, modern telescope, both Uranus and Neptune are dim, uninspiring sights. Each planet appears as a hazy, featureless disk with a faint greenish-blue tinge. Although Uranus and Neptune are both about 4 times larger in diameter than Earth, they are so distant that their angular diameters as seen from the Earth are tiny—no more than 4 arcsec for Uranus and just over 2 arcsec for Neptune. To an Earth-based observer, Uranus is roughly the size of a golf ball seen at a distance of 1 kilometer.

Uranus and Neptune are in such large orbits that they move very slowly on the celestial sphere

From 2007 through 2011, Uranus and Neptune are less than 40° apart in the sky in the adjacent constellations of Pisces, Aquarius, and Capricornus (the Sea Goat). During these years, the two planets are at opposition in either August or September, which are thus the best months to view them with a telescope. Table 14-1 and Table 14-2 give basic data about Uranus and Neptune, respectively.

14-2 Uranus is nearly featureless and has an unusually tilted axis of rotation

Scientists had hoped that *Voyager 2* would reveal cloud patterns in Uranus's atmosphere when it flew past the planet in January 1986. But even images recorded at close range showed Uranus to be remarkably featureless (Figure 14-1). Faint cloud markings became visible in images of Uranus only after extreme computer enhancement (Figure 14-2).



WEB LINK 14.4

Figure 14-1 R I V U X G

Uranus from *Voyager 2* This image looks nearly straight down onto Uranus's south pole, which was pointed almost directly at the Sun when *Voyager 2* flew past in 1986. None of the *Voyager 2* images of Uranus show any pronounced cloud patterns. The color is due to methane in the planet's atmosphere, which absorbs red light but reflects green and blue. (JPL/NASA)

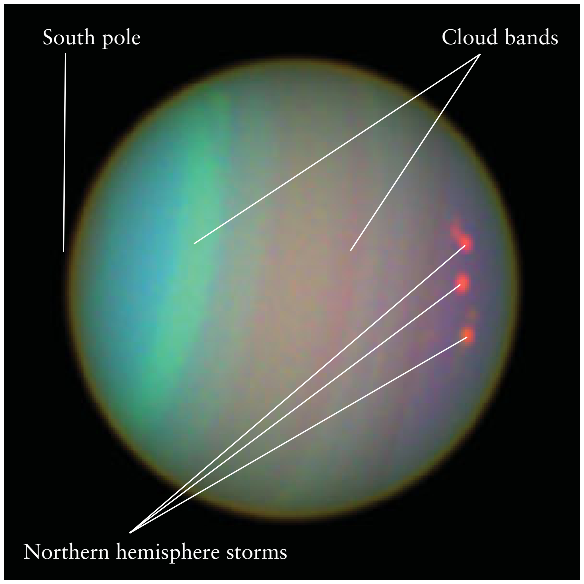
Uranus's Atmosphere

Voyager 2 data confirmed that the Uranian atmosphere is dominated by hydrogen (82.5%) and helium (15.2%), similar to the atmospheres of Jupiter and Saturn. Uranus differs, however, in that 2.3% of its atmosphere is methane (CH_4), which is 5 to 10 times the percentage found on Jupiter and Saturn. In fact, Uranus has a higher percentage of all heavy elements—including carbon atoms, which are found in molecules of methane—than Jupiter and Saturn. (In Section 14-4 we will investigate how this could have come about.)

Methane preferentially absorbs the longer wavelengths of visible light, so sunlight reflected from Uranus's upper atmosphere is depleted of its reds and yellows. This gives the planet its distinct greenish-blue appearance. As on Saturn's moon Titan (see Section 13-8), ultraviolet light from the Sun turns some of the methane gas into a hydrocarbon haze, making it difficult to see the lower levels of the atmosphere.

Ammonia (NH_3), which constitutes 0.01 to 0.03% of the atmospheres of Jupiter and Saturn, is almost completely absent from the Uranian atmosphere. The reason is that Uranus is colder than Jupiter or Saturn: The temperature in its upper atmosphere is only 55 K (−218°C = −360°F). Ammonia freezes at these very low temperatures, so any ammonia has long since precipitated out of the atmosphere and into the planet's interior. For the same reason, Uranus's atmosphere is also lacking in water. Hence, the substances that make up the clouds on Jupiter and Saturn—ammonia, ammonium hydrosulfide (NH_4SH), and water—are not available on Uranus. This helps to explain the bland, uniform appearance of the planet shown in Figure 14-1.

The few clouds found on Uranus are made primarily of methane, which condenses into droplets only if the pressure is suf-



WEB LINK 14.5

Figure 14-2 R I V U X G

Uranus from the Hubble Space Telescope Images made at ultraviolet, visible, and infrared wavelengths were combined and enhanced to give this false-color view of cloud features on Uranus. These images were captured in August 2004 (18 years after the image in Figure 14-1), during springtime in Uranus's northern hemisphere. (NASA and Erich Karkoschka, University of Arizona)

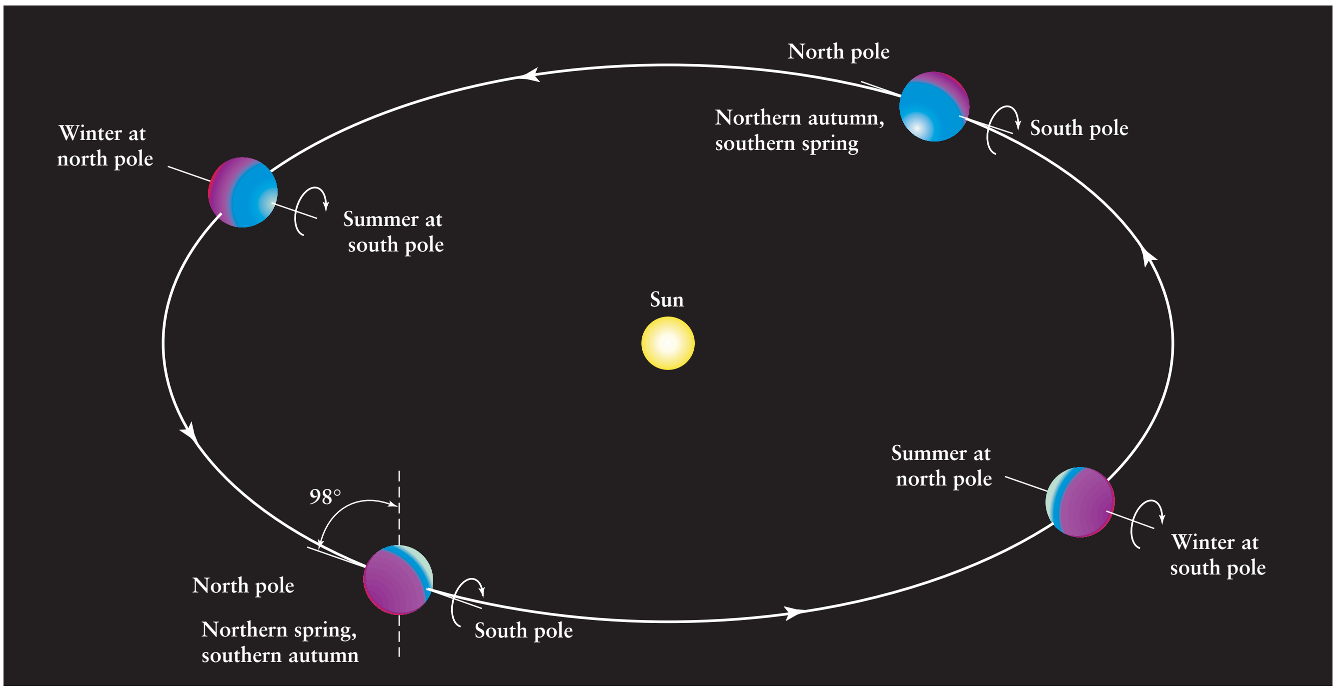
ficiently high. Hence, methane clouds form only at lower levels within the atmosphere, where they are difficult to see.

An Oddly Tilted Planet

The rotation period of Uranus's atmosphere is about 16 hours. Like Jupiter and Saturn, Uranus rotates differentially, so this period depends on the latitude. This can be measured by tracking the motions of clouds. To determine the rotation period for the underlying body of the planet, scientists looked to Uranus's magnetic field, which is presumably anchored in the planet's interior, or at least in the deeper and denser layers of its atmosphere. Data from *Voyager 2* indicate that Uranus's internal period of rotation is 17.24 hours.

Voyager 2 also confirmed that Uranus's rotation axis is tilted in a unique and bizarre way. Herschel found the first evidence of this in 1787, when he discovered two moons orbiting Uranus in a plane that is almost perpendicular to the plane of the planet's orbit around the Sun. Because the large moons of Jupiter and Saturn were known to orbit in the same plane as their planet's equator and in the same direction as their planet's rotation, it was thought that the same must be true for the moons of Uranus. Thus, Uranus's equator must be almost perpendicular to the plane of its orbit, and its rotation axis must lie very nearly in that plane (Figure 14-3).

Careful measurement shows that Uranus's axis of rotation is tilted by 98°, as compared to $23\frac{1}{2}$° for Earth (compare Figure 14-3 with Figure 2-12). A tilt angle greater than 90° means that



ANIMATION 14.1

Figure 14-3

Exaggerated Seasons on Uranus For most planets, the rotation axis is roughly perpendicular to the plane of the planet's orbit around the Sun. But for Uranus the rotation axis is tilted by 98° from the perpendicular. This causes severely exaggerated seasons. For example, during midsummer at Uranus's south pole, the Sun appears nearly overhead for many Earth years, while the planet's northern regions are in continuous darkness. Half an orbit later, the seasons are reversed.

Uranus exhibits retrograde (backward) rotation like that of Venus, shown in Figure 11-7*b*. Astronomers suspect that Uranus might have acquired its large tilt angle billions of years ago, when another massive body collided with Uranus while the planet was still forming.

The radical tilt of its axis means that as Uranus moves along its 84-year orbit, its north and south poles alternately point toward or away from the Sun. This produces highly exaggerated seasonal changes. For example, when *Voyager 2* flew by in 1986, Uranus's south pole was pointed toward the Sun. Most of the planet's southern hemisphere was bathed in continuous sunlight, while most of the northern hemisphere was subjected to a continuous, frigid winter night. But over the following quarter of a Uranian year later (21 of our years), sunlight has gradually been returning to the northern hemisphere, triggering immense storms there. Figure 14-2 shows some of these storms, which are much more visible at infrared wavelengths than with visible light.

Atmospheric Motions on Uranus

By following the motions of clouds and storm systems on Uranus, scientists find that the planet's winds flow to the east—that is, in the same direction as the planet's rotation—at northern and southern latitudes, but to the west near the equator. This is quite unlike the situation on Jupiter and Saturn, where the zonal winds alternate direction many times between the north and south poles (see Section 12-4). The fastest Uranian winds (about 700 km/h, or 440 mi/h) are found at the equator.

We saw in Section 12-4 that the internal heat of Jupiter and Saturn plays a major role in driving atmospheric activity on those worlds. Uranus is different: It appears to have little or no internal source of thermal energy. Measurements to date show that unlike Jupiter or Saturn, Uranus radiates into space only as much energy as it receives from the Sun. With only feeble sunlight to provide energy to its atmosphere, Uranus lacks the dramatic wind and cloud patterns found on Jupiter and Saturn. It is not known why Uranus should be different from Jupiter and Saturn.

Unlike the other Jovian planets, Uranus's interior heat has no effect on its atmosphere's motions

Although Uranus's equatorial region was receiving little sunlight at the time of the *Voyager 2* flyby, the atmospheric temperature there (about 55 K = −218°C = −359°F) was not too different from that at the sunlit pole. Heat must therefore be efficiently transported from the poles to the equator. This north-south heat transport, which is perpendicular to the wind flow, may have mixed and homogenized the atmosphere to make Uranus nearly featureless.

14-3 Neptune is a cold, bluish world with Jupiterlike atmospheric features

VIDEO 14.1

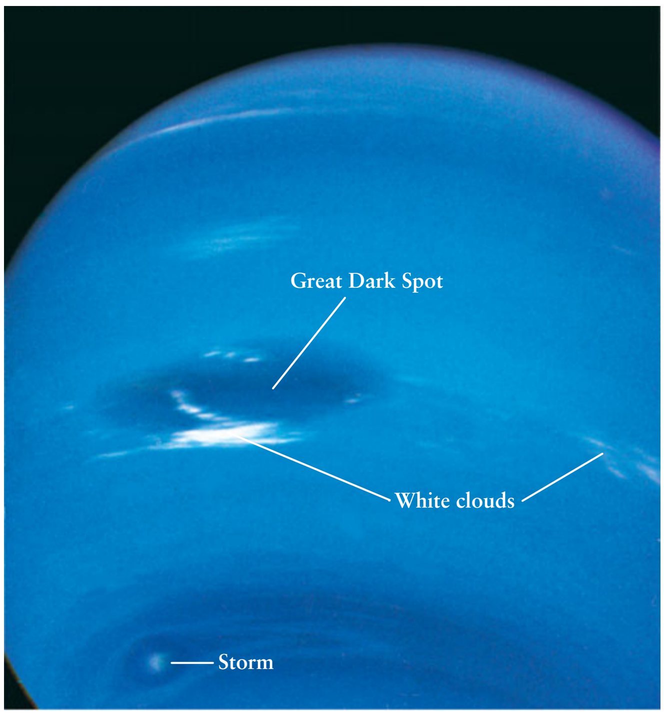
At first glance, Neptune appears to be the twin of Uranus. (See the image that opens this chapter, and compare Tables 14-1 and 14-2.) But these two planets are by no means identical. While Neptune and Uranus have almost the same diameter, Neptune is 18% more massive. Neptune's axis of rotation also has a more modest $29\frac{1}{2}$° tilt. When *Voyager 2* flew past Neptune in August 1989, it revealed that the planet has a more active and dynamic atmosphere than Uranus. This activity suggests that Neptune, unlike Uranus, has a powerful source of energy in its interior.

Neptune's Atmosphere

The *Voyager 2* data showed that Neptune has essentially the same atmospheric composition as Uranus: 79% hydrogen, 18% helium, 3% methane, and almost no ammonia or water vapor. As for Uranus, the presence of methane gives Neptune a characteristic bluish-green color. The temperature in the upper atmosphere is also the same as on Uranus, about 55 K. That this should be so, even though Neptune is much farther from the Sun, is further evidence that Neptune has a strong internal heat source.

Unlike Uranus, however, Neptune has clearly visible cloud patterns in its atmosphere. At the time that *Voyager 2* flew past Neptune, the most prominent feature in the planet's atmosphere was a giant storm called the Great Dark Spot (Figure 14-4). The Great Dark Spot had a number of similarities to Jupiter's Great Red Spot (see Sections 12-1 and 12-3). The storms

Unlike its close relative Uranus, Neptune has truly immense storms



WEB LINK 14.6 VIDEO 14.2

Figure 14-4 R I V U X G
Neptune from *Voyager 2* When this picture of Neptune's southern hemisphere was taken in 1989, the Great Dark Spot measured about 12,000 by 8000 km, comparable in size to the Earth. (A smaller storm appears at the lower left.) The white clouds are thought to be composed of crystals of methane ice. The color contrast in this image has been exaggerated to emphasize the differences between dark and light areas. (NASA/JPL)

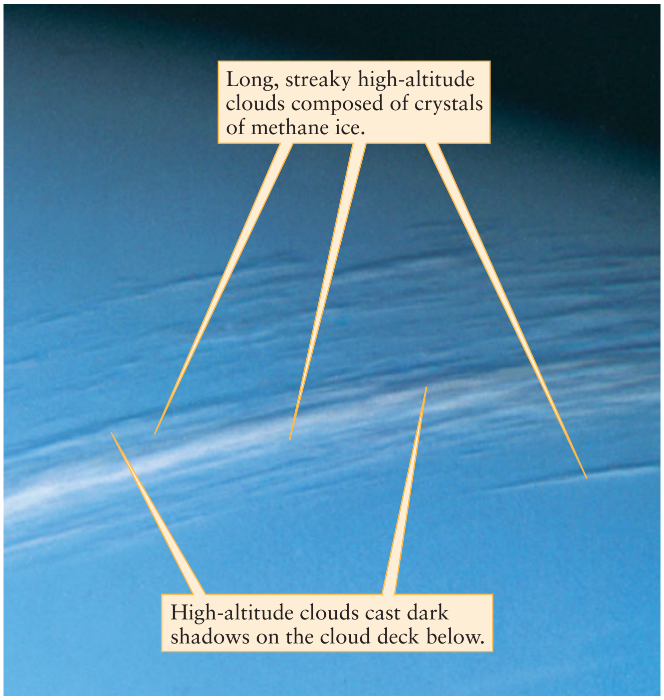


Figure 14-5 R I V U X G

Cirrus Clouds over Neptune *Voyager 2* recorded this image of clouds near Neptune's terminator (the border between day and night on the planet). Like wispy, high-altitude cirrus clouds in the Earth's atmosphere, these clouds are thought to be made of ice crystals. The difference is that Neptune's cirrus clouds are probably methane ice, not water ice as on Earth. (NASA/JPL)

on both planets were comparable in size to the Earth's diameter, both appeared at about the same latitude in the southern hemisphere, and the winds in both storms circulated in a counterclockwise direction (see Figure 12-5). But Neptune's Great Dark Spot appears not to have been as long-lived as the Great Red Spot on Jupiter. When the Hubble Space Telescope first viewed Neptune in 1994, the Great Dark Spot had disappeared. Another dark storm appeared in 1995 in Neptune's northern hemisphere.

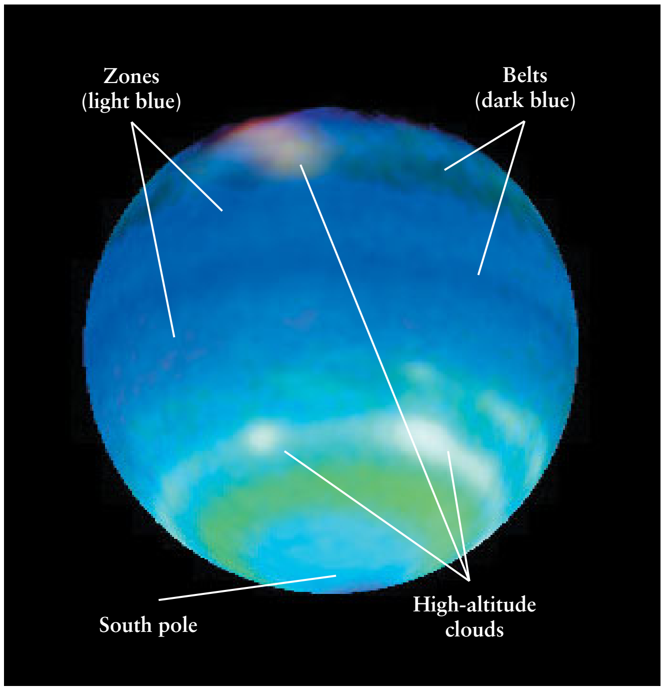
Voyager 2 also saw a few conspicuous whitish clouds on Neptune. These clouds are thought to be produced when winds carry methane gas into the cool, upper atmosphere, where it condenses into crystals of methane ice. *Voyager 2* images show these high-altitude clouds casting shadows onto lower levels of Neptune's atmosphere (Figure 14-5). Images from the Hubble Space Telescope also show the presence of high-altitude clouds (Figure 14-6).

Neptune's Internal Heat

Thanks to its greater distance from the Sun, Neptune receives less than half as much energy from the Sun as Uranus. But with less solar energy available to power atmospheric motions, why are there high-altitude clouds and huge, dark storms on Neptune but not on Uranus? At least part of the answer is probably that Neptune is still slowly contracting, thus converting gravitational energy into thermal energy that heats the planet's core. (The same is true for Jupiter and Saturn, as we saw in Section 12-4.) The evidence for this is that Neptune, like Jupiter and Saturn but unlike Uranus, emits more energy than it receives from the Sun. The combination of a warm interior and a cold outer atmosphere can cause convection in Neptune's atmosphere, producing the up-and-down motion of gases that generates clouds and storms. Neptune also resembles Jupiter in having faint belts and zones parallel to the planet's equator (see Figure 14-6).

Like those on Uranus, most of Neptune's clouds are probably made of droplets of liquid methane. Because these droplets form fairly deep within the atmosphere, the clouds are more difficult to see than the ones on Jupiter. Hence, Neptune's belts and zones are less pronounced than Jupiter's, although more so than those on Uranus (thanks to the extra cloud-building energy from Neptune's interior). As described above, Neptune's high-altitude clouds (see Figure 14-5) are probably made of *frozen* methane.

Although Neptune displays more evidence of up-and-down motion in its atmosphere than Uranus, the global pattern of east and west winds is almost identical on the two planets. This is rather strange. The two planets are heated very differently by the Sun, thanks to their different distances from the Sun and the different tilts of their axes of rotation, so we might have expected the wind patterns on Uranus and Neptune also to be different. Perhaps the explanation of these wind patterns will involve understanding how heat is transported not only within the atmospheres of Uranus and Neptune but within their interiors as well.



WEB LINK 14.7

Figure 14-6 R I V U X G

Neptune's Banded Structure This enhanced-color Hubble Space Telescope image shows Neptune's belts and zones. White areas denote high-altitude clouds; the very highest clouds (near the top of the image) are shown in yellow-red. The green belt near the south pole is a region where the atmosphere absorbs blue light, perhaps indicating a different chemical composition there. (Lawrence Sromovsky, University of Wisconsin-Madison; STScI/NASA)

14-4 Uranus and Neptune contain a higher proportion of heavy elements than Jupiter and Saturn

At first glance, Uranus and Neptune might seem to be simply smaller and less massive versions of Jupiter or Saturn (see Table 7-1). But like many first impressions, this one is misleading because it fails to take into account what lies beneath the surface. We saw in Section 14-2 that the interiors of both Jupiter and Saturn are composed primarily of hydrogen and helium, in nearly the same abundance as the Sun. But Uranus and Neptune must have a different composition, because their average densities are too high.

Uranus and Neptune: Curiously Dense

If Uranus and Neptune had solar abundances of the elements, the smaller masses of these planets would produce less gravitational compression, and Uranus and Neptune would both have lower average densities than Jupiter or Saturn. In fact, however, Uranus and Neptune have average densities (1320 kg/m^3 and 1640 kg/m^3, respectively) that are comparable to or greater than those of Jupiter (1330 kg/m^3) or Saturn (690 kg/m^3). Therefore, we can conclude that Uranus and Neptune contain greater proportions of the heavier elements than do Jupiter or Saturn.

This picture is not what we might expect. According to our discussion in Section 8-4 of how the solar system formed, hydrogen and helium should be relatively more abundant as distance from the vaporizing heat of the Sun increases. But Uranus and Neptune contain a greater percentage of heavy elements, and, therefore, a *smaller* percentage of hydrogen and helium, than Jupiter and Saturn.

A related problem is explaining the masses of Uranus and Neptune. At the locations of Uranus (19 AU from the Sun) and Neptune (30 AU from the Sun) the solar nebula was probably so sparse that these planets would have taken tens of hundreds of millions of years to grow to their present sizes around a core of icy planetesimals. But protoplanetary disks observed around other stars (see Section 8-3) do not appear to survive for that long before they dissipate. In other words, the problem is not how to explain why Uranus and Neptune are smaller than Jupiter and Saturn; it is how to explain why Uranus and Neptune should exist at all!

Models of How Uranus and Neptune Formed

In one model for the origin of Uranus and Neptune, the planets formed in denser regions of the solar nebula between 4 and 10 AU from the Sun. In this region they could have grown rapidly to their present sizes. Had the planets remained at these locations, they could eventually have accumulated enough hydrogen and helium to become as large as Jupiter or Saturn. But before that could happen, gravitational interactions with Jupiter and Saturn would have pushed Uranus and Neptune outward into their present large orbits. Because the solar nebula was very sparse at those greater distances, Uranus and Neptune stopped growing and remained at the sizes we see today.

Explaining the origins of Uranus and Neptune poses a challenge to planetary scientists

An alternative idea is that Uranus and Neptune formed in their present locations, not around a core of planetesimals but directly from the gases of the solar nebula. (We introduced this idea in Section 8-4.) Once a ball of gas formed, icy particles within the ball would have settled to its center, forming a solid core. This

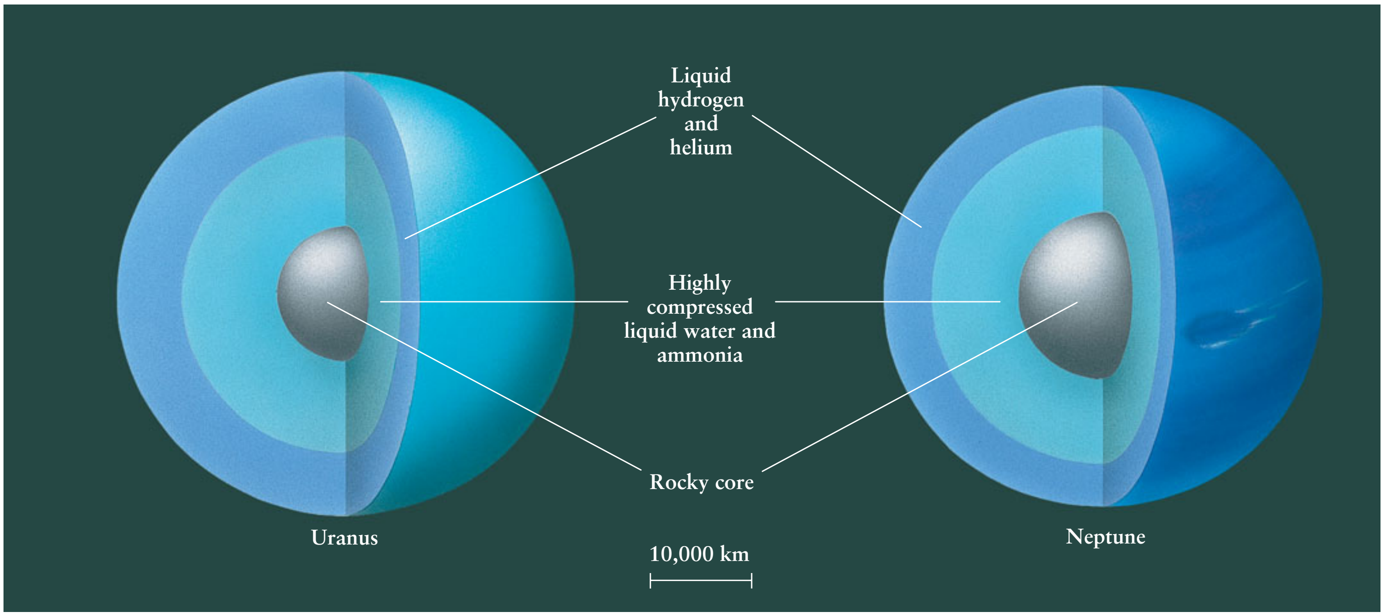


Figure 14-7

The Internal Structures of Uranus and Neptune In the model shown here, both Uranus and Neptune have a rocky core, resembling a terrestrial planet; a mantle of liquid water with ammonia dissolved in it; and an outer layer of liquid molecular hydrogen and liquid helium. The atmospheres are very thin shells on top of the hydrogen-helium layer. Uranus and Neptune have about the same diameter, but since Neptune is more massive it may have a somewhat larger core.

process could have formed Uranus and Neptune in only a few hundred years. In this model, Uranus and Neptune would actually have ended up larger than they are today, with substantially more hydrogen and helium. But if there happened to be a hot star relatively near the solar nebula, it would have emitted substantial amounts of ultraviolet radiation (see Section 5-3 and Figure 5-11). Ultraviolet photons have high energy (see Section 5-5), and a shower of them could have stripped Uranus and Neptune of many of their lightweight hydrogen and helium atoms, leaving shrunken planets with a larger percentage of heavy elements. (In this picture, Jupiter and Saturn were protected from the star's ultraviolet radiation because they were shrouded by the thicker clouds of material in their part of the solar nebula.) More research will be needed to better determine the origins of Uranus and Neptune.

Figure 14-7 shows one model for the present-day internal structures of Uranus and Neptune. (Compare this to the internal structures of Jupiter and Saturn, shown in Figure 12-12.) In this model each planet has a rocky core roughly the size of Earth, although more massive. Each planet's core is surrounded by a mantle of liquid water and ammonia. (This means that the mantle is chemically similar to household window cleaning fluid.) Around the mantle is a layer of liquid molecular hydrogen and liquid helium, with a small percentage of liquid methane. This layer is relatively shallow compared to those on Jupiter and Saturn, and the pressure is not high enough to convert the liquid hydrogen into liquid metallic hydrogen.

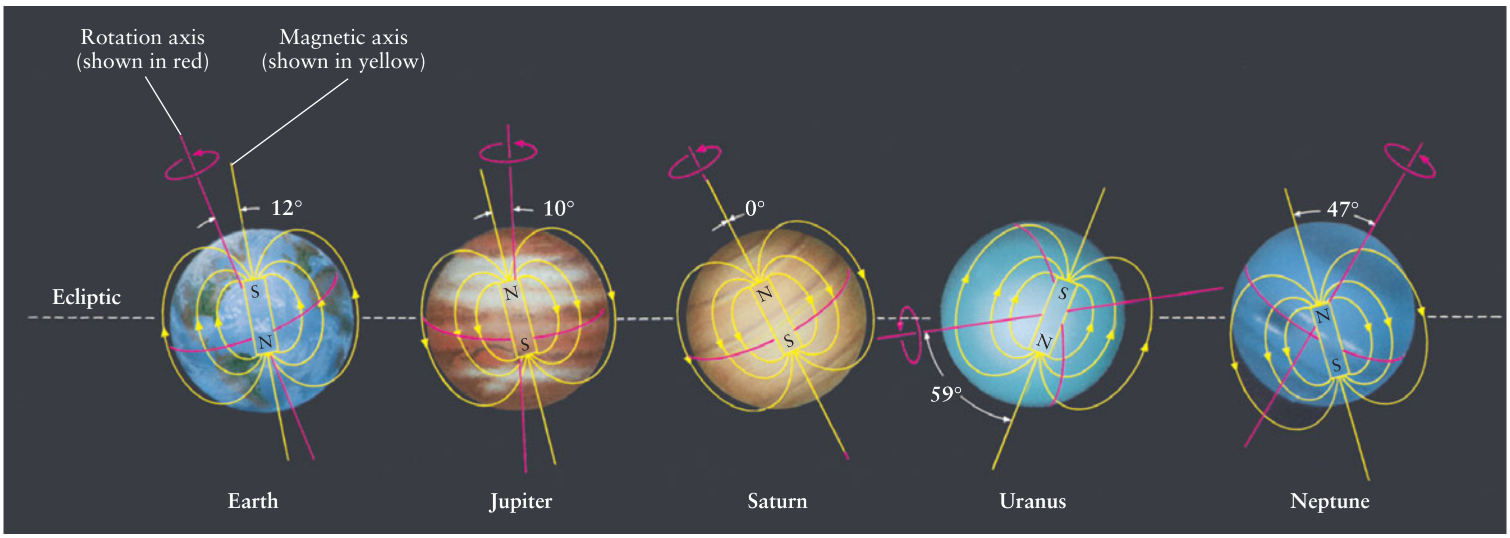
14-5 The magnetic fields of both Uranus and Neptune are oriented at unusual angles

Astronomers were quite surprised by the data sent back from *Voyager 2*'s magnetometer as the spacecraft sped past Uranus and Neptune. These data, as well as radio emissions from charged particles in their magnetospheres, showed that the magnetic fields of both Uranus and Neptune are tilted at steep angles to their axes of rotation. Uranus's **magnetic axis,** the line connecting its north and south magnetic poles, is inclined by 59° from its axis of rotation; Neptune's magnetic axis is tilted by 47°. By contrast, the magnetic axes of Earth, Jupiter, and Saturn are all nearly aligned with their rotation axes; the angle between their magnetic and rotation axes is 12° or less (Figure 14-8). Scientists were also surprised to find that the magnetic fields of Uranus and Neptune are offset from the centers of the planets.

CAUTION! The drawing of the Earth's magnetic field at the far left of Figure 14-8 may seem to be mislabeled, because it shows the *south* pole of a magnet at the Earth's *north* pole. But, in fact, this is correct, as you can understand by thinking about how magnets work. On a magnet that is free to swivel, like the magnetized needle in a compass, the "north pole" is called that because it tends to point north on Earth. Likewise, a compass needle's south pole points toward the south on Earth. Furthermore, opposite magnetic poles attract. If you take two magnets and put them next to each other, they try to align themselves so that one magnet's north pole is next to the other magnet's south pole. Now, if you think of a compass needle as one magnet and the entire Earth as the other magnet, it makes sense that the compass needle's north pole is being drawn toward a magnetic *south* pole—which happens to be located near the Earth's geographic north pole. The Earth's magnetic pole nearest its geographic North Pole is called the "magnetic north pole." Note that the magnets drawn inside Jupiter, Saturn, and Neptune are oriented opposite to the Earth's. On any of these planets, the north pole of a compass needle would point south, not north!

Explaining Misaligned Magnetism

Why are the magnetic axes and axes of rotation of Uranus and Neptune so badly misaligned? And why are the magnetic fields offset from the centers of the planets? One possibility is that their



INTERACTIVE EXERCISE 14.1

Figure 14-8

The Magnetic Fields of Five Planets This illustration shows how the magnetic fields of the Earth, Jupiter, Saturn, Uranus, and Neptune are tilted relative to their rotation axes (shown in red). For both Uranus and Neptune, the magnetic fields are offset from the planet's center and steeply inclined to the rotation axis.

magnetic fields might be undergoing a reversal; geological data show that the Earth's magnetic field has switched north to south and back again many times in the past. Another possibility is that the misalignments resulted from catastrophic collisions with planet-sized bodies. As we discussed in Section 14-2, the tilt of Uranus's rotation axis and its system of moons suggest that such collisions occurred long ago. As we will see in Section 14-7, Neptune may have gravitationally captured its largest moon, Triton, but no one knows if that incident was responsible for offsetting Neptune's magnetic axis.

Because neither Uranus nor Neptune is massive enough to compress hydrogen to a metallic state, their magnetic fields cannot be generated in the same way as those of Jupiter and Saturn. Instead, under the high pressures found in the watery mantles of Uranus and Neptune, dissolved molecules such as ammonia lose one or more electrons and become electrically charged (that is, they become *ionized*; see Section 5-8). Water is a good conductor of electricity when it has such electrically charged molecules dissolved in it, and electric currents in this fluid are probably the source of the magnetic fields of Uranus and Neptune.

Unlike any other planets, Uranus and Neptune have magnetic fields caused by ionized ammonia

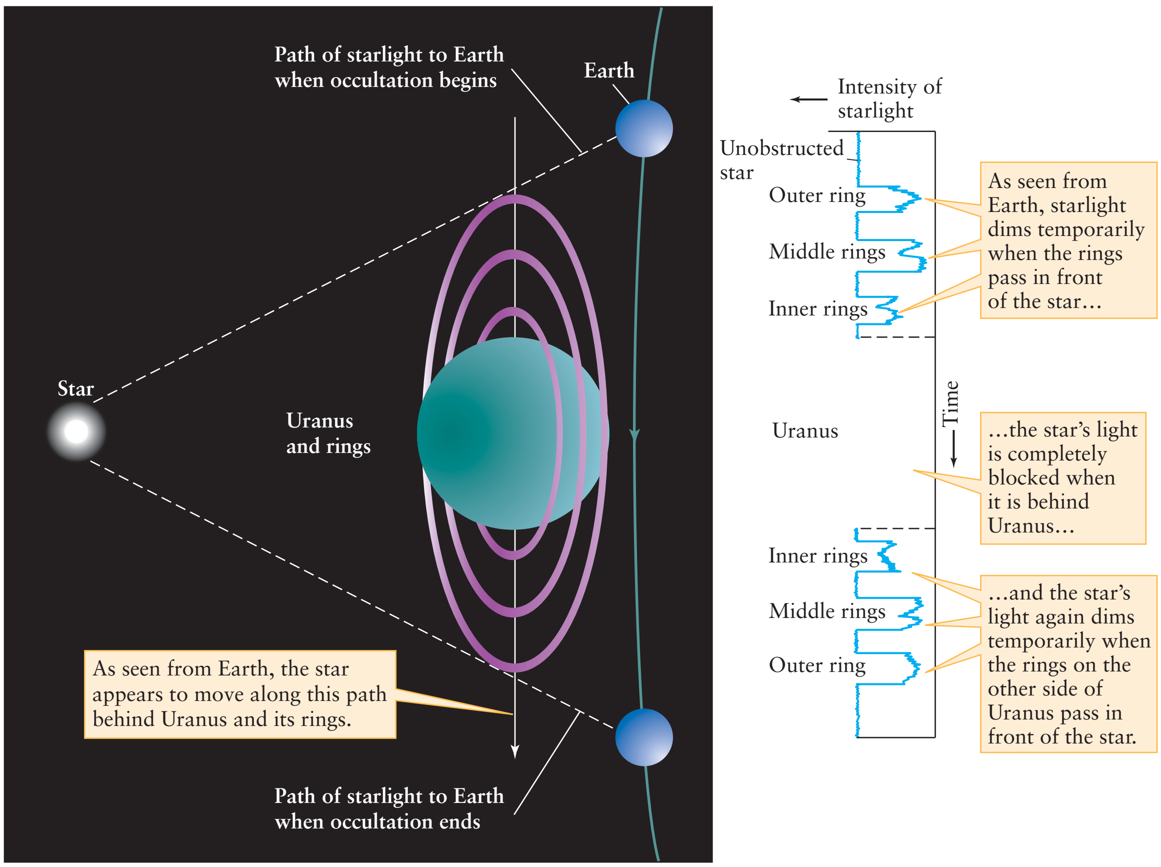
The *Cosmic Connections* figure summarizes the key properties of Uranus and Neptune and how they compare with those of Jupiter and Saturn.

14-6 Uranus and Neptune each has a system of thin, dark rings

On March 10, 1977, Uranus was scheduled to move in front of a faint star, as seen from the Indian Ocean. A team of astronomers headed by James L. Elliot of Cornell University observed this event, called an **occultation,** from a NASA airplane equipped with a telescope. They hoped that by measuring how long the star was hidden, they could accurately measure Uranus's size. In addition, by carefully measuring how the starlight faded when Uranus passed in front of the star, they planned to deduce important properties of Uranus's upper atmosphere.

The Surprising Rings of Uranus

To everyone's surprise, the background star briefly blinked on and off several times just before the star passed behind Uranus and again immediately after (Figure 14-9). The astronomers



ANIMATION 14.2

Figure 14-9

How the Rings of Uranus Were Discovered As seen from Earth, Uranus occasionally appears to move in front of a distant star. Such an event is called an occultation. The star's light is completely blocked when it is behind the planet. But before and after the occultation by Uranus, the starlight dims temporarily as the rings pass in front of the star.

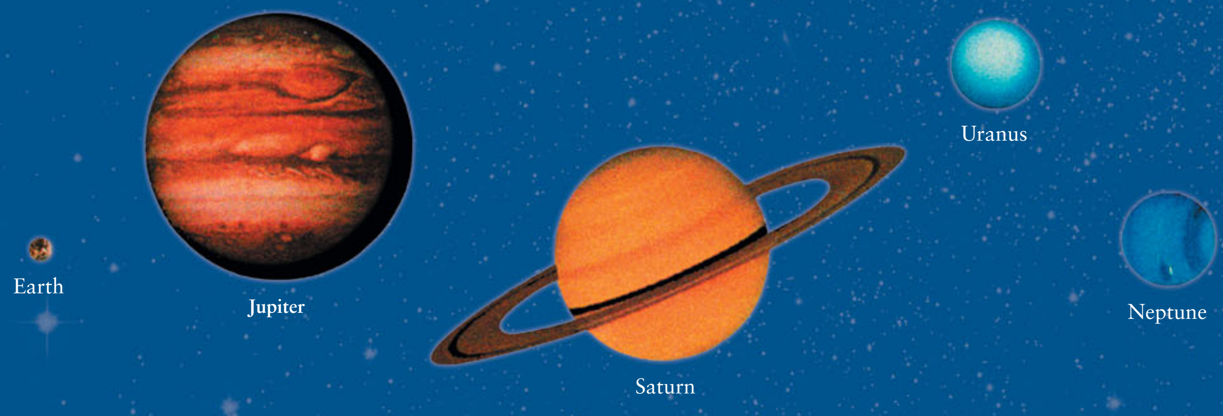
COSMIC CONNECTIONS

The Outer Planets: A Comparison

Uranus and Neptune are not simply smaller versions of Jupiter and Saturn. This table summarizes the key differences among the four Jovian planets.

	Interior	Surface	Rings	Atmosphere	Magnetic Field
Jupiter	Terrestrial core, liquid metallic hydrogen shell, liquid hydrogen mantle	No solid surface, atmosphere gradually thickens to liquid state, belt and zone structure, hurricanelike features	Yes	Primarily H, He	19,000× Earth's total field; at its cloud layer, 14× stronger than Earth's surface field
Saturn	Similar to Jupiter, with bigger terrestrial core and less metallic hydrogen	No solid surface, less distinct belt and zone structure than Jupiter	Yes	Primarily H, He	570× Earth's total field; at its cloud layer, $^2/_3$× Earth's surface field
Uranus	Terrestrial core, liquid water shell, liquid hydrogen and helium mantle	No solid surface, weak belt and zone system, hurricanelike features, color from methane absorption of red, orange, yellow	Yes	Primarily H, He, some CH_4	50× Earth's total field; at its cloud layer, 0.7× Earth's surface field
Neptune	Similar to Uranus	Like Uranus	Yes	Primarily H, He, some CH_4	35× Earth's total field; at its cloud layer, 0.4× Earth's surface field

For detailed comparisons between planets, see Appendices 1 and 2.
* To see the orientations of these magnetic fields relative to the rotation axes of the planets, see Figure 14-8.



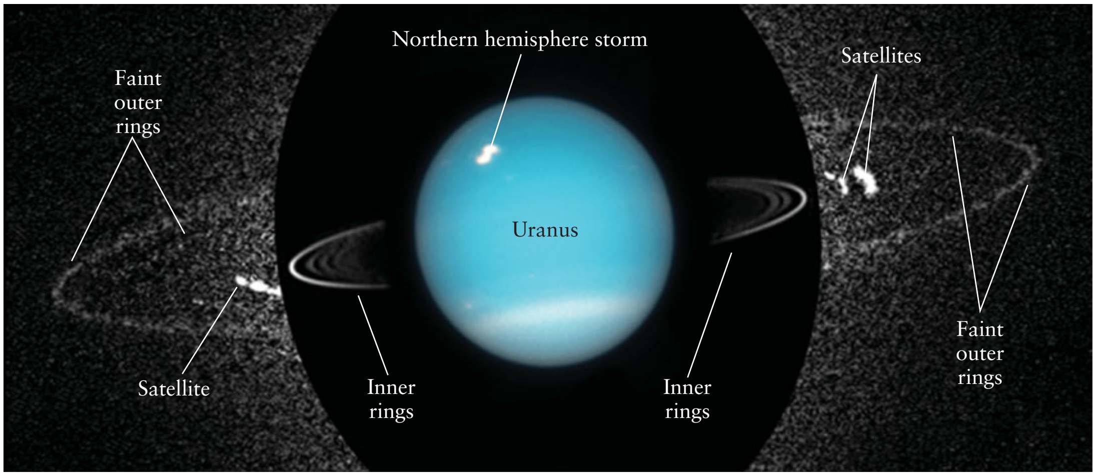


Figure 14-10 R I V U X G

Uranus's Rings from the Hubble Space Telescope Several Hubble Space Telescope images were assembled into this view of Uranus's rings. Relatively short exposure times reveal the planet and the inner rings, while much longer exposure times were needed to reveal the outer rings. During these long exposures several of Uranus's satellites moved noticeably, leaving bright trails on the image. (NASA; ESA; and M. Showalter, SETI Institute)

concluded that Uranus must be surrounded by a series of nine narrow rings. In addition to these nine rings, *Voyager 2* discovered two others that lie even closer to the cloudtops of Uranus. Two other extremely faint rings, much larger in diameter than any of the others, were found in 2005 using the Hubble Space Telescope (Figure 14-10).

Like the planet itself, the rings of Uranus were discovered by accident

Unlike Saturn's rings, the rings of Uranus are dark and narrow: Most are less than 10 km wide. Typical particles in Saturn's rings are chunks of reflective ice the size of snowballs (a few centimeters across), but typical particles in Uranus's rings are 0.1 to 10 meters in size and are no more reflective than lumps of coal. It is not surprising that these narrow, dark rings escaped detection for so long. Figure 14-11 is a *Voyager 2* image of the rings from the side of the planet away from the Sun, where light scattering from exceptionally small ring particles makes them more visible. (We discussed light scattering in Section 5-6 and Box 5-4.) Figure 14-10 shows the Sun-facing side of the rings.

The Uranian rings are so dark because sunlight at Uranus is only one-quarter as intense as at Saturn. As a result, Uranus's ring particles are so cold that they can retain methane ice. Scientists speculate that eons of impacts by electrons trapped in the magnetospheres of the planets have converted this methane ice into dark carbon compounds. This **radiation darkening** can account for the low reflectivity of the rings.

Uranus's major rings are located less than 2 Uranian radii from the planet's center, well within the planet's Roche limit (see Section 12-9). Some sort of mechanism, possibly one involving shepherd satellites, efficiently confines particles to their narrow orbits. (In Section 12-11 we discussed how shepherd satellites help keep planetary rings narrow.) *Voyager 2* searched for shepherd satellites but found only two. The others may be so small and dark that they have simply escaped detection.

The two faint rings discovered in 2005 (see Figure 14-10) are different: They lie well *outside* Uranus's Roche limit. The outer of these two rings owes its existence to a miniature satellite called Mab (named for a diminutive fairy in English folklore) that orbits within this ring. Meteorites colliding with Mab knock dust off the satellite's surface. Since Mab is so small (just 36 km in diameter) and hence has little gravity, the ejected dust escapes from

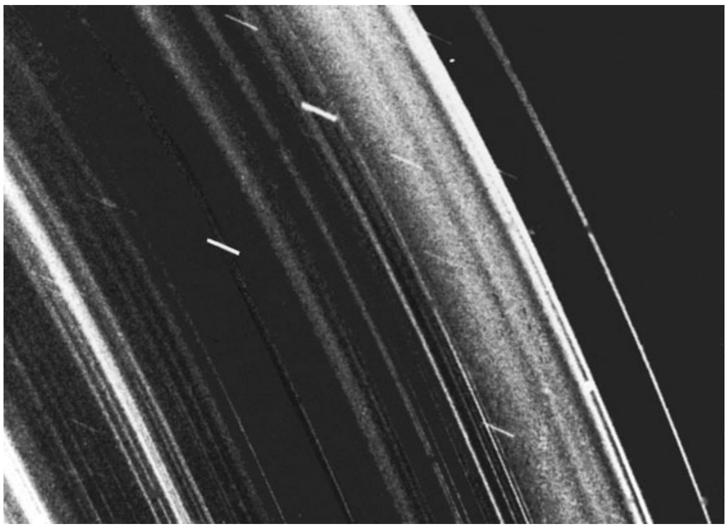
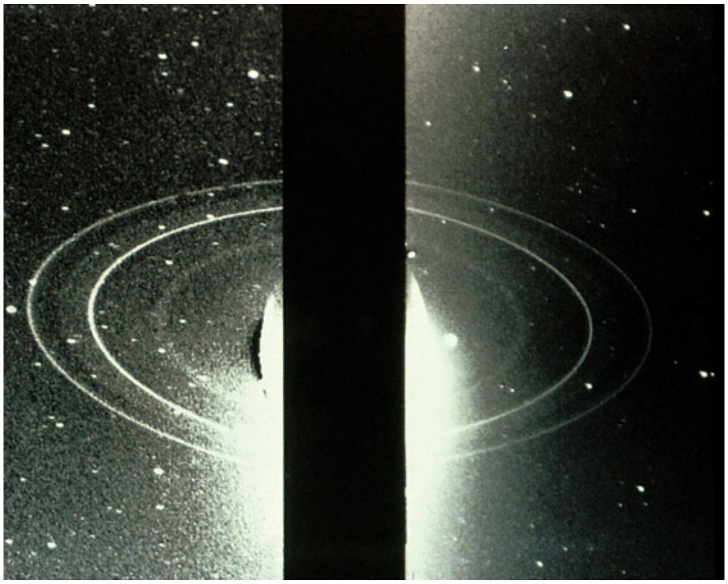


Figure 14-11 R I V U X G

The Shaded Side of Uranus's Rings This view from *Voyager 2*, taken when the spacecraft was in Uranus's shadow, looks back toward the Sun. Numerous fine-grained dust particles gleam in the spaces between the main rings. This dust is probably debris from collisions between larger particles in the main rings. The short horizontal streaks are star images blurred by the spacecraft's motion during the exposure. (NASA/JPL)



VIDEO 14.3

Figure 14-12 R I V U X G

Neptune's Rings The two main rings of Neptune and a faint, inner ring can easily be seen in this composite of two *Voyager 2* images. There is also a faint sheet of particles whose outer edge is located between the two main rings and extends inward toward the planet. The overexposed image of Neptune itself has been blocked out. (NASA)

the satellite and goes into orbit around Uranus, forming a ring. The inner of the two faint rings may be formed in the same way; however, no small satellite has yet been found within this ring.

The Decaying Rings of Neptune

Like Uranus, Neptune is surrounded by a system of thin, dark rings that were first detected in stellar occultations. Figure 14-12 is a *Voyager 2* image of Neptune's rings. The ring particles vary in size from a few micrometers (1 μm = 10^{-6} m) to about 10 meters. As for Uranus's rings, the particles that make up Neptune's rings reflect very little light because they have undergone radiation darkening.

In 2002 and 2003, a team led by Imke de Pater of the University of California, Berkeley, used the 10-meter Keck telescope in Hawaii to observe in detail the rings of Neptune. Remarkably, they found that all of the rings had become fainter since the *Voyager 2* flyby in 1989. Apparently, the rings are losing particles faster then new ones are being added. If the rings continue to decay at the same rate, one of them may vanish completely within a century. More research into the nature of Neptune's rings is needed to explain this curious and rapid decay.

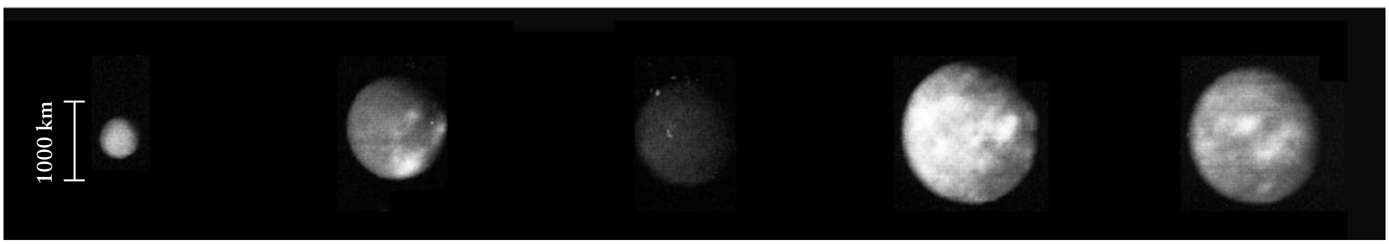
14-7 Some of Uranus's satellites show evidence of past tidal heating

Before the *Voyager 2* flyby of Uranus, five moderate-sized satellites—Titania, Oberon, Ariel, Umbriel, and Miranda—were known to orbit the planet (Figure 14-13). Most are named after sprites and spirits in Shakespeare's plays. They range in diameter from about 1600 km (1000 mi) for Titania and Oberon to less than 500 km (300 mi) for Miranda. All these moons have average densities around 1500 kg/m^3, which is consistent with a mixture of ice and rock.

Voyager 2 discovered 11 other small Uranian satellites, most of which have diameters of less than 100 km (60 mi); 11 more were found using ground-based telescopes and the Hubble Space Telescope between 1997 and 2003. Only Jupiter and Saturn have more known satellites than Uranus. Figure 14-14 shows several of the smaller satellites. Appendix 3 summarizes information about all 27 known moons of Uranus.

Uranus's Unusual Moderate-Sized Satellites

Umbriel and Oberon both appear to be geologically dead worlds, with surfaces dominated by impact craters. By contrast, Ariel's surface appears to have been cracked at some time in the past, allowing some sort of ice lava to flood low-lying areas. A similar process appears to have taken place on Titania. This geologic activity may be due to a combination of the satellites' internal heat and tidal heating like that which powers volcanism on Jupiter's satellite Io (see Section 13-4). Io's tidal heating is only possible because of the 1:2:4 ratio of the orbital periods of Io, Europa, and Ganymede. While there are no such simple ratios between the



Miranda Ariel Umbriel Titania Oberon

INTERACTIVE EXERCISE 14.2

Figure 14-13 R I V U X G

Uranus's Principal Satellites This "family portrait" (a montage of five *Voyager 2* images) shows Uranus's five largest moons to the same scale and correctly displays their respective reflectivities. (The darkest satellite, Umbriel, is actually more reflective than the Earth's Moon.) All five satellites have grayish surfaces, with only slight variations in color. (NASA/JPL)

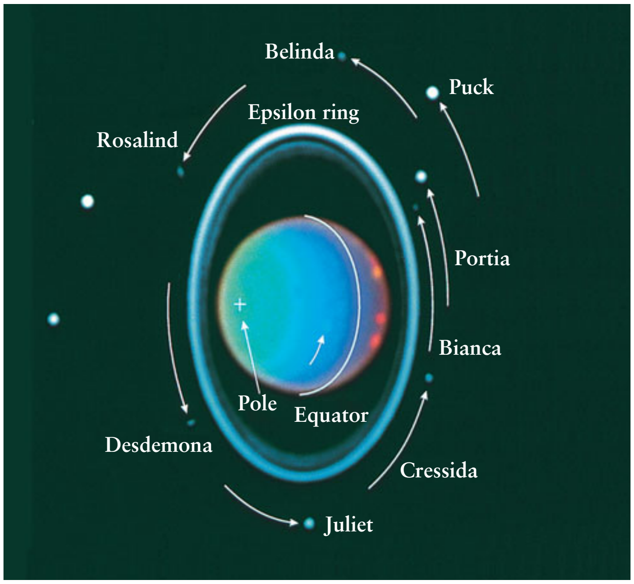
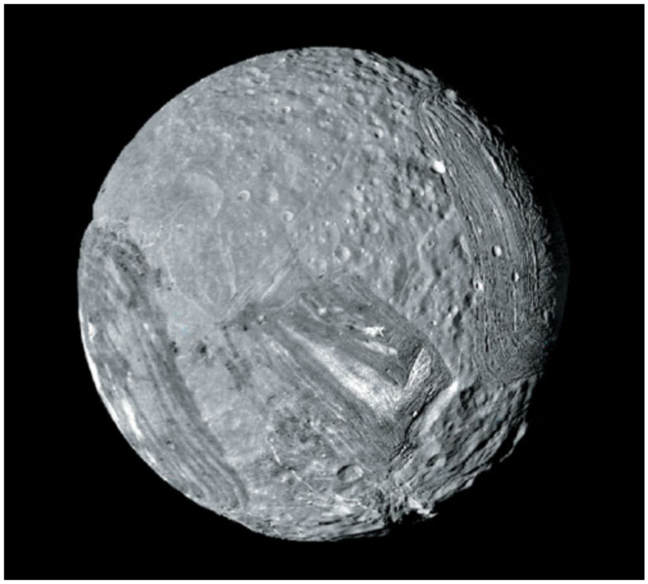


Figure 14-14 R I V U X G

Uranus's Rings and Small Satellites This false-color infrared image from the Hubble Space Telescope shows eight of Uranus's satellites, all of which were discovered by *Voyager 2* when it flew past Uranus in 1986. They all lie within 86,000 km of the planet's center (only about one-fifth of the distance from the Earth to our Moon) and all are less than 160 km (100 mi) in diameter. The arcs show how far each satellite moves around its orbit in 90 minutes. (Erich Karkoschka, University of Arizona; and NASA)



VIDEO 14.4

Figure 14-15 R I V U X G

Miranda This composite of *Voyager 2* images shows that part of Miranda's surface is ancient and heavily cratered, while other parts are dominated by parallel networks of valleys and ridges. At the very bottom of the image—where a "bite" seems to have been taken out of Miranda—is a range of enormous cliffs that jut upward to an elevation of 20 km, twice as high as Mount Everest. (NASA/JPL)

present-day orbital periods of Uranus's satellites, there may have been in the past. If so, tidal heating could have helped reshape the surfaces of some of the satellites.

Unique among Uranus's satellites is Miranda, which has a landscape unlike that of any other world in the solar system. Much of the surface is heavily cratered, as we would expect for a satellite only 470 km in diameter, but several regions have unusual and dramatic topography (Figure 14-15). Detailed analysis of Miranda's geology suggests that this satellite's orbital period was once in a whole-number ratio with that of more massive Umbriel or Ariel. The resulting tidal heating melted Miranda's interior, causing dense rocks in some locations on the surface to settle toward the satellite's center as blocks of less dense ice were forced upward toward the surface, thus creating Miranda's resurfaced terrain. Tidal heating must have ceased before this process could run its full course, which explains why some ancient, heavily cratered regions remain on Miranda's surface.

Uranus's satellite Miranda has cliffs twice as high as Mount Everest

Uranus's Perplexing Small Satellites

While Uranus's small satellites are unlikely to show the kind of geology found in Miranda, some of them move in curious orbits. Eight of the nine small satellites in large orbits beyond Oberon are in retrograde orbits: They go around Uranus oppositely to the direction in which Uranus rotates. Many of the small outer satellites of Jupiter and Saturn have retrograde orbits and are probably captured asteroids (see Section 13-9 and Section 13-10); the same is probably true of the outer satellites of Uranus.

The 13 satellites that orbit closest to Uranus (inside the orbit of Miranda) are all in prograde orbits, so they travel around Uranus in the same direction as the planet rotates. However, when Mark Showalter and Jack Lissauer compared 11 of these satellites' orbits in 2005 with their orbits in 1994, they found surprisingly large differences. (The other two inner satellites were only discovered in 2003, so it is not known whether their orbits underwent similar changes.) These satellites can pass within a few thousand to a few hundred kilometers from each other, so they can exert strong gravitational forces on each other. Over time these forces can modify their orbits. The magnitude of the orbital changes is such that the inner satellites will begin colliding with each other within a few million years. If these inner satellites are in such unstable orbits, is it possible that they have been in orbit around Uranus since it formed more than 4.5 billion years ago? Or did Uranus somehow acquire these satellites in the relatively recent past? More research will be needed to answer these questions.

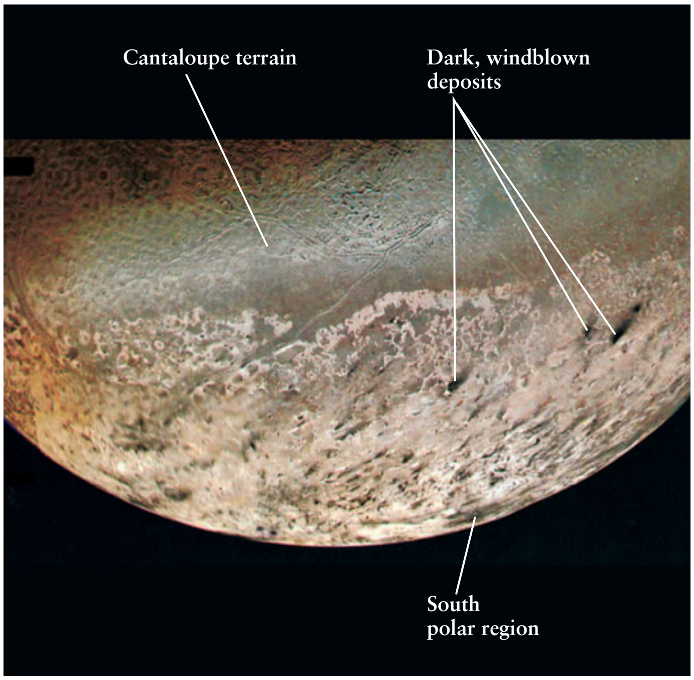
14-8 Neptune's satellite Triton is an icy world with a young surface and a tenuous atmosphere

Neptune has 13 known satellites, listed in Appendix 3. They are named for mythological beings related to bodies of water. (Neptune itself is named for the Roman god of the sea.) Most of these worlds are small, icy bodies, probably similar to the smaller satellites of Uranus. The one striking exception is Triton, Neptune's largest satellite (see Table 7-2). In many ways Triton is quite unlike any other world in the solar system.

Triton's Backward Orbit and Young Surface

Like many of the small satellites of Jupiter, Saturn, Uranus, and Neptune, Triton is in a retrograde orbit, so that it moves around Neptune oppositely to the direction of Neptune's rotation. Furthermore, the plane of Triton's orbit is inclined by 23° from the plane of Neptune's equator. It is difficult to imagine how a satellite might form out of the same cloud of material as a planet and end up orbiting in a direction opposite the planet's rotation and in such a tilted plane. Hence, Triton probably formed elsewhere in the solar system, collided long ago with a now-vanished satellite of Neptune, and was captured by Neptune's gravity. With a diameter of 2706 km, a bit smaller than our Moon but much larger than any other satellite in a retrograde orbit, Triton is certainly the largest captured satellite.

Figure 14-16 shows the icy, reflective surface of Triton as imaged by *Voyager 2*. There is a conspicuous absence of large craters, which immediately tells us that Triton has a young surface on which the scars of ancient impacts have largely been erased by tectonic activity. There are areas that resemble frozen lakes and may be the calderas of extinct ice volcanoes. Some of Triton's surface features resemble the long cracks seen on Europa (Section 13-6) and Ganymede (Section 13-7). Still other features are unique to Triton. For example, in the upper portion of Figure 14-16 you can see dimpled, wrinkled terrain that resembles the skin of a cantaloupe.



VIDEO 14.5 VIDEO 14.6 **Figure 14-16** R I V U X G

Triton Several high-resolution *Voyager 2* images were combined to create this mosaic. The pinkish material surrounding Triton's south polar region is probably nitrogen frost. Some of this presumably evaporates when summer comes to the south pole; the northward flow of the evaporated gas may cause the dark surface markings. Farther north is a brown area of "cantaloupe terrain" that resembles the skin of a melon. (NASA/JPL)

Triton's tectonically active history is probably related to its having been captured into orbit around Neptune. After its capture, Triton most likely started off in a highly elliptical orbit, but today the satellite's orbit is quite circular. The satellite's original elliptical orbit would have been made circular by tidal forces exerted on Triton by Neptune's gravity. These tidal forces would also have stretched and flexed Triton, causing enough tidal heating to melt much of the interior. The resulting volcanic activity (with lavas made of ice rather than molten rock) would have obliterated Triton's original surface features, including craters.

There may still be warmth in Triton's interior today. *Voyager 2* observed plumes of dark material being ejected from the surface to a height of 8 km (5 mi). These plumes may have been generated from a hot spot far below Triton's surface, similar to geysers on the Earth. Alternatively, the energy source for the plumes may be sunlight that warms the surface, producing subsurface pockets of gas and creating fissures in the icy surface through which the gas can escape.

Triton's surface temperature is only 38 K ($-235°C = -391°F$), the lowest of any world yet visited by spacecraft. This temperature is low enough to solidify nitrogen, and indeed the spectrum of sunlight reflected from Triton's surface shows absorption lines due to nitrogen ice as well as methane ice. But Triton is also warm enough to allow some nitrogen to evaporate from the surface, like the steam that rises from ice cubes when you first take them out of the freezer. *Voyager 2* confirmed that Triton has a very thin nitrogen atmosphere with a surface pressure of only 1.6×10^{-5} atmosphere, about the same as at an altitude of 100 km above the Earth's surface. Despite its thinness, Triton's atmosphere has noticeable effects. *Voyager 2* saw areas on Triton's surface where dark material has been blown downwind by a steady breeze (see Figure 14-16). It also observed dark material ejected from the geyserlike plumes being carried as far as 150 km by high-altitude winds.

Tidal Forces and Triton's Doom

Just as tidal forces presumably played a large role in Triton's history, they also determine its future. Triton raises a tidal bulge on Neptune, just as our Moon distorts the Earth (recall Figure 10-17). In the case of the Earth-Moon system, the gravitational pull of the Earth's tidal bulge causes the Moon to spiral away from the Earth. But because Triton's orbit is retrograde, the tidal bulge on Neptune exerts a force on Triton that makes the satellite slow down rather than speed up. (In Figure 10-17, imagine that the moon is orbiting toward the bottom of the figure rather than toward the top.) This is causing Triton to spiral gradually in toward Neptune. In approximately 100 million years, Triton will be inside Neptune's Roche limit, and the satellite will eventually be torn to pieces by tidal forces. When this happens, the planet will develop

Within 100 million years Triton will be torn apart by Neptune's tidal forces

a spectacular ring system—overshadowing by far its present-day set of narrow rings—as rock fragments gradually spread out along Triton's former orbit.

Prior to *Voyager 2,* only one other satellite was known to orbit Neptune. Nereid, which was first sighted in 1949, is in a prograde orbit. Hence, it orbits Neptune in the direction opposite to Triton. Nereid also has the most eccentric orbit of any satellite in the solar system; its distance from Neptune varies from 1.4 million to 9.7 million kilometers. One possible explanation is that when Triton was captured by Neptune's gravity, the interplay of gravitational forces exerted on Nereid by both Neptune and Triton moved Nereid from a relatively circular orbit (like those of Neptune's other, smaller moons) into its present elliptical one.

14-9 Pluto is smaller than any planet

Speculations about worlds beyond Neptune date back to the late 1800s, when a few astronomers suggested that Neptune's orbit was being perturbed by an unknown object. Encouraged by the fame of Adams and Le Verrier, several people set out to discover "Planet X." Two prosperous Boston gentlemen, William Pickering and Percival Lowell, were prominent in this effort. (Lowell also enthusiastically promoted the idea of canals on Mars; see Section 11-4.) Modern calculations show that there are, in fact, no unaccounted perturbations of Neptune's orbit. It is thus not surprising that no orbiting body was found at the positions predicted by Pickering, Lowell, and others. Yet the search continued.

The Discovery of Pluto

Before he died in 1916, Lowell urged that a special camera with a wide field of view be constructed to help search for Planet X. After many delays, the camera was finished in 1929 and installed at the Lowell Observatory in Flagstaff, Arizona, where a young astronomer, Clyde W. Tombaugh, had joined the staff to carry on the project. On February 18, 1930, Tombaugh finally discovered the long-sought object. It was disappointingly faint—a thousand times dimmer than the dimmest object visible with the naked eye and 250 times dimmer than Neptune at opposition—and presented no discernible disk. The new world was named for Pluto, the mythological god of the underworld; fittingly, this name has Percival Lowell's initials as its first two letters. The discovery was publicly announced on March 13, 1930, the 149th anniversary of the discovery of Uranus. The two photographs in Figure 14-17 show one day's motion of Pluto.

Pluto's orbit about the Sun is more elliptical (eccentricity 0.2501) and more steeply inclined to the plane of the ecliptic (17.15°) than the orbit of any of the planets. In fact, Pluto's orbit is so eccentric that it is sometimes closer to the Sun than Neptune. This was the case from 1979 until 1999; indeed, when Pluto was at perihelion in 1989, it was more than 10^8 km closer to the Sun than was Neptune. Happily, the orbits of Neptune and Pluto are such that the two worlds will never collide.

Pluto is so far away that it subtends only a very small angle of 0.15 arcsec. Hence, it is extraordinarily difficult to resolve any of its surface features. But by observing Pluto with the Hubble Space Telescope over the course of a solar day on Pluto (6.3872 Earth days) and using computer image processing, Alan Stern and Marc Buie generated the maps of Pluto's surface shown in Figure 14-18. These maps show bright polar ice caps as well as regions of different reflectivity near the planet's equator. Observations of Pluto's rotation confirm that the planet's rotation axis is tipped by more than 90°, so that Pluto has retrograde rotation like Uranus.

Pluto's Satellites

In 1978, while examining some photographs of Pluto, James W. Christy of the U.S. Naval Observatory noticed that the image of the planet on a photographic plate appeared slightly elongated, as though Pluto had a lump on one side. He promptly examined a number of other photographs of Pluto, and found a series that showed the lump moving clockwise around Pluto with a period of about 6 days.

Christy concluded that the lump was actually a satellite of Pluto. He proposed that the newly discovered moon be christened Charon (pronounced KAR-en), after the mythical boatman who ferried souls across the River Styx to Hades, the domain ruled by Pluto. (Christy also chose the name because of its similarity to Charlene, his wife's name.) The average distance between Charon

Figure 14-17 R I V U X G

Pluto's Motion across the Sky Pluto was discovered in 1930 by searching for a dim, starlike object that moves slowly in relation to the background stars. These two photographs were taken one day apart. Even when its apparent motion on the celestial sphere is fastest, Pluto moves only about 1 arcmin per day relative to the stars. (Lick Observatory)

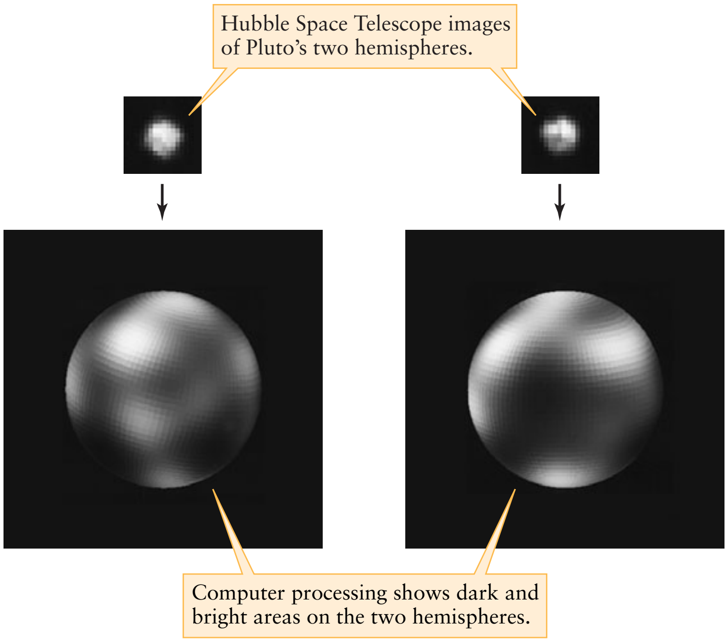


Figure 14-18 R I V U X G

Pluto's Surface The bright regions at the top and bottom of each of Pluto's hemispheres may be polar ice caps. The bright regions nearer the equator may be impact basins where more reflective subsurface ice has been exposed. (Alan Stern, Southwest Research Institute; Marc Buie, Lowell Observatory; NASA; and ESA)

and Pluto is a scant 19,640 km, less than 5% of the distance between the Earth and our Moon. The best pictures of Pluto and Charon have been obtained using the Hubble Space Telescope (HST). In 2005 a team of astronomers using HST discovered two additional, smaller satellites that orbit Pluto at two to three times the distance of Charon (Figure 14-19).

Observations show that Charon's orbital period of 6.3872 days is the same as the rotational period of Pluto *and* the rotational period of Charon. In other words, both Pluto and Charon rotate synchronously with their orbital motion, and so both always keep the same face toward each other. As seen from the Charon-facing side of Pluto, the satellite remains perpetually suspended above the horizon. Likewise, Pluto neither rises nor sets as seen from Charon.

Pluto and its satellite Charon both rotate in lockstep, always keeping the same faces toward each other

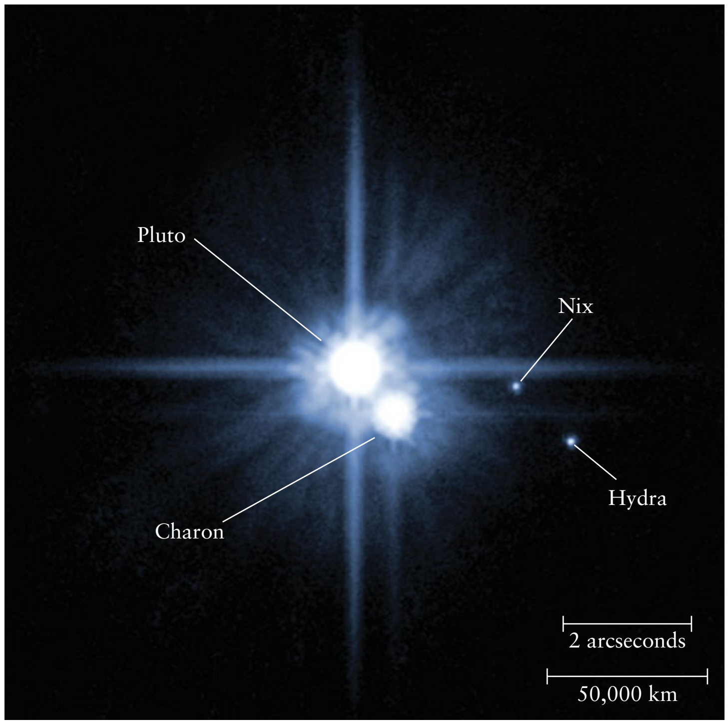
Soon after the discovery of Charon, astronomers witnessed an alignment that occurs only once every 124 years. From 1985 through 1990, Charon's orbital plane appeared nearly edge-on as seen from the Earth, allowing astronomers to view mutual eclipses of Pluto and its moon. As the bodies passed in front of each other, their combined brightness diminished in ways that revealed their sizes and surface characteristics. Data from the eclipses combined with subsequent measurements give Pluto's diameter as about 2274 km and Charon's as about 1190 km. For comparison, our Moon's diameter (3476 km) nearly equals the diameters of Pluto and Charon added together. The brightness variations during eclipses suggest that Charon has a bright southern polar cap.

The average densities of Pluto and Charon, at about 2000 kg/m^3, are essentially the same as that of Triton (2070 kg/m^3). All three worlds are therefore probably composed of a mixture of rock and ice, as we might expect for small bodies that formed in the cold outer reaches of the solar nebula.

Pluto's spectrum shows absorption lines of various solid ices that cover the planet's surface, including nitrogen (N_2), methane (CH_4), and carbon monoxide (CO). Stellar occultation measurements have shown that Pluto, like Triton, has a very thin atmosphere. Exposed to a daytime temperature of around 40 K, N_2 and CO ices turn to gas more easily than frozen CH_4. For this reason, most of Pluto's tenuous atmosphere probably consists of N_2 and CO. Charon's weaker gravity has apparently allowed most of the N_2, CH_4, and CO to escape into space; only water ice is found on its surface.

The Origin of Pluto's Satellites

Pluto and Charon are remarkably like each other in mass, size, and density. Throughout the rest of the solar system, planets are many times larger and more massive than any of their satellites.



VIDEO 14.7 VIDEO 14.8

Figure 14-19 R I V U X G

Pluto and Its Satellites This false-color Hubble Space Telescope shows Pluto (discovered 1930), its large satellite Charon (discovered 1978), and the two small satellites Nix and Hydra discovered in 2005. Nix is named for the mythological goddess of the night who was the mother of Charon; Hydra in mythology was a nine-headed, poisonous serpent who guarded the entrance to Hades. (NASA; ESA; H. Weaver, JHU/APL; A. Stern, SwRI; and the HST Pluto Companion Search Team)

The exceptional similarities between Pluto and Charon suggest to some astronomers that this binary system probably formed when Pluto collided with a similar body. Perhaps chunks of matter were stripped from the second body, leaving behind a mass, now called Charon, that was captured into orbit by Pluto's gravity. This same collision probably also left behind the small satellites Nix and Hydra shown in Figure 14-19.

In support of this model is the observation that all three satellites have almost the same color and reflection spectrum, suggesting a common origin. Pluto, by contrast, has a redder color than its satellites, which is consistent with the idea that Pluto and the satellites formed in different ways.

For this scenario to be feasible, there must have been many Plutolike objects in the outer regions of the solar system. Astronomers estimate that in order for a collision or close encounter between two of them to have occurred at least once since the solar system formed some 4.56 billion years ago, there must have been at least a thousand Plutos. As we will see, astronomers have begun to discover this population of Plutolike objects in the dark recesses of the solar system beyond Neptune.

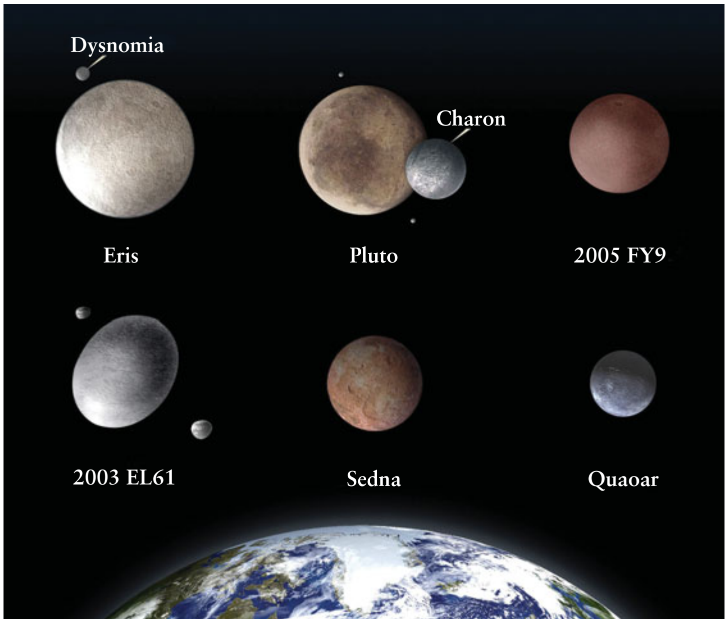


Figure 14-20

The Largest Trans-Neptunian Objects This artist's impression depicts the Earth and the largest objects known beyond Neptune (as of 2006) to the same scale. The largest of these, Eris, has a diameter of about 2900 km versus 2274 km for Pluto. Note the differences in color among these objects and that three of them (Eris, Pluto, and 2003 EL61) have satellites of their own. Table 7-4 in Section 7-5 lists the properties of these objects. (NASA; ESA; and A. Field, STScI)

14-10 Pluto is just one of thousands of icy objects that orbit the Sun beyond Neptune

For many years astronomers attempted to find other worlds beyond Neptune using the same technique used to discover Pluto (see Figure 14-17). The first to succeed were David Jewitt and Jane Luu, who in 1992 found an object with a semimajor axis of 42 AU and a diameter estimated to be only 240 km. This object, named 1992 QB1, has a reddish color like Pluto, possibly because frozen methane has been degraded by eons of radiation exposure. As of this writing (2006), more than 1100 more **trans-Neptunian objects**—icy worlds whose orbits have semimajor axes larger than that of Neptune—have been discovered. Most of these are relatively small like 1992 QB1, but a number are larger than Charon and one is larger than Pluto itself (Figure 14-20). Thus, Pluto is by no means unique: We can best regard it as a particularly large (but not the largest) trans-Neptunian object rather than a planet.

With improvements in the sensitivity of telescopes, new objects beyond Neptune are being discovered at a rapid pace; 67 new trans-Neptunian objects were found in 2005 alone. Based on these observations, it is thought that there could be 35,000 or more objects beyond Neptune with diameters greater than 100 km. One of the larger objects could have collided with Pluto in the past, giving rise to Pluto's retinue of satellites (see Section 14-9). At least 29 trans-Neptunian objects have satellites, which suggests that such collisions have taken place many times.

The Kuiper Belt and Beyond

Most of the trans-Neptunian objects lie within the **Kuiper belt,** which extends from about 30 to 50 AU from the Sun and is relatively close to the ecliptic. When the solar system was young, a large number of icy planetesimals formed in the region beyond Jupiter. Over time, the gravitational forces of the massive Jovian planets pushed most of these planetesimals beyond Neptune's orbit, concentrating them into a belt centered on the plane of the ecliptic. Most of the trans-Neptunian objects within the belt are in orbits that are only slightly inclined to the ecliptic; these are thought to have formed beyond Neptune and to be in roughly their original orbits. Other objects such as 2005 FY9 and 2003 EL61 (see Figure 14-20 and Table 7-4) are in orbits that are inclined by about 30° from the ecliptic. These are thought to have been pushed into their steeply inclined orbits by gravitational interactions with Neptune.

The processes that gave rise to the Kuiper belt in our solar system also appear to have taken place around other stars. Figure 14-21 shows a disk of material surrounding the young star HD 139664. This disk has dimensions comparable to our Kuiper belt. A number of other young stars have been found with disks of this same type.

There are relatively few members of the Kuiper belt between the orbits of Neptune and Pluto. Remarkably, there are about 100 objects that have nearly the same semimajor axis as Pluto. These so-called **plutinos,** which include Pluto itself, have the property that they complete nearly two orbits around the Sun in the same time that Neptune completes three orbits. The plutinos thus experience rhythmic gravitational pulls from Neptune, and these pulls keep them in their orbits. (In Section 13-1 we saw a similar relationship among the orbital periods of Jupiter's satellites Io, Europa, and Ganymede, though the ratio of their orbital periods is 1:2:4 rather

Neptune's gravity shapes the orbits of many icy worlds, including Pluto

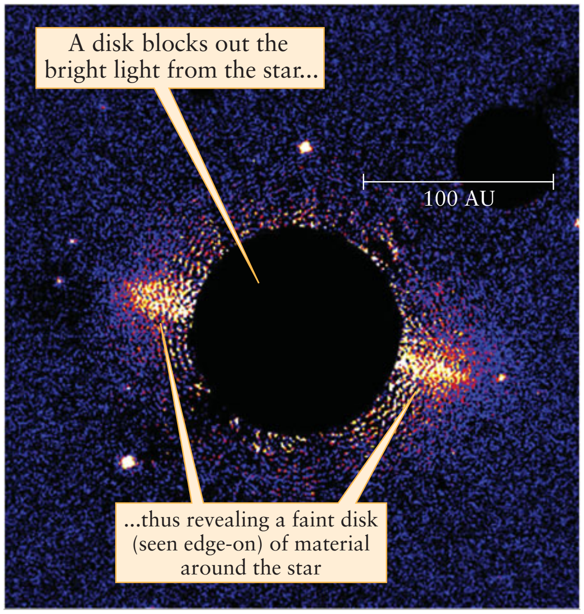


Figure 14-21 R I V U X G

A "Kuiper Belt" Around Another Star The star HD 139664 is only slightly more massive than the Sun but is thought to be just 300 million years old. This Hubble Space Telescope image shows a ring of material around HD 139664 that is similar in size to the Kuiper belt in our solar system. HD 139664 is 57 light-years (17 parsecs) from Earth in the constellation Lupus (the Wolf). (NASA; ESA; and P. Kalas/University of California, Berkeley)

than the 2:3 ratio for Neptune and the plutinos.) Most Kuiper belt objects orbit at distances beyond the plutinos but within about 50 AU from the Sun, at which distance an object completes one orbit for every two orbits of Neptune. At this distance the rhythmic gravitational forces of Neptune appear to pull small objects inward, giving the Kuiper belt a relatively sharp outer edge.

Two examples of trans-Neptunian objects that are not members of the Kuiper belt are shown in Figure 14-20. Eris, the largest known trans-Neptunian object, has a semimajor axis of more than 67 AU and an orbital eccentricity of 0.442. This orbit takes Eris from inside the orbit of Pluto to far beyond the Kuiper belt. An even more extreme case is Sedna: Its orbital semimajor axis is 489 AU, giving it an orbital period of more than 10,000 years, and the orbital eccentricity has the remarkably high value of 0.844. It is not well understood how Sedna could have been moved into such an immense and elongated orbit.

Excitement about the worlds beyond Neptune has motivated the development of a spacecraft called *New Horizons*. Launched in 2006, *New Horizons* swung by Jupiter in 2007 and will make the first-ever flyby of Pluto and Charon in 2015. It will then be aimed toward one or two other targets in the Kuiper belt. The choice of targets will not be made until a year before the Pluto-Charon flyby, giving astronomers extra time to discover new members of the Kuiper belt that might be suitable destinations. The high-resolution images and other data to be returned by *New Horizons* promise to revolutionize our understanding of these most remote members of the solar system.

Key Words

Kuiper belt, p. 370	plutino, p. 370
magnetic axis, p. 361	radiation darkening, p. 364
occultation, p. 362	trans-Neptunian object, p. 370

Key Ideas

Discovery of the Outer Planets: Uranus was discovered by chance, while Neptune was discovered at a location predicted by applying Newtonian mechanics. Pluto was discovered after a long search.

Atmospheres of Uranus and Neptune: Both Uranus and Neptune have atmospheres composed primarily of hydrogen, helium, and a few percent methane.

• Methane absorbs red light, giving Uranus and Neptune their greenish-blue color.

• No white ammonia clouds are seen on Uranus or Neptune. Presumably the low temperatures have caused almost all the ammonia to precipitate into the interiors of the planets. All of these planets' clouds are composed of methane.

• Much more cloud activity is seen on Neptune than on Uranus. This is because Uranus lacks a substantial internal heat source.

Interiors and Magnetic Fields of Uranus and Neptune: Both Uranus and Neptune may have a rocky core surrounded by a mantle of water and ammonia. Electric currents in these mantles may generate the magnetic fields of the planets.

• The magnetic axes of both Uranus and Neptune are steeply inclined from their axes of rotation. The magnetic and rotational axes of all the other planets are more nearly parallel. The magnetic fields of Uranus and Neptune are also offset from the centers of the planets.

Uranus's Unusual Rotation: Uranus's axis of rotation lies nearly in the plane of its orbit, producing greatly exaggerated seasonal changes on the planet.

• This unusual orientation may be the result of a collision with a planetlike object early in the history of our solar system. Such a collision could have knocked Uranus on its side.

Ring Systems of Uranus and Neptune: Uranus and Neptune are both surrounded by systems of thin, dark rings. The low reflectivity of the ring particles may be due to radiation-darkened methane ice.

Satellites of Uranus and Neptune: Uranus has five satellites similar to the moderate-sized moons of Saturn, plus at least 22 more small satellites. Neptune has 13 satellites, one of which (Triton) is comparable in size to our Moon or the Galilean satellites of Jupiter.

• Triton has a young, icy surface indicative of tectonic activity. The energy for this activity may have been provided by tidal heating that occurred when Triton was captured by Neptune's gravity into a retrograde orbit.

• Triton has a tenuous nitrogen atmosphere.

Worlds Beyond Neptune: Pluto and its moon, Charon, move together in a highly elliptical orbit steeply inclined to the plane of the ecliptic.

• More than a thousand icy worlds have been discovered beyond Neptune. Pluto and Charon are part of this population.

• Most trans-Neptunian objects lie in a band called the Kuiper belt that extends from 30 to 50 AU from the Sun. Neptune's gravity shapes the orbits of objects within the Kuiper belt.

Questions

Review Questions

1. Could astronomers in antiquity have seen Uranus? If so, why was it not recognized as a planet?
2. Why do you suppose that the discovery of Neptune is rated as one of the great triumphs of science, whereas the discoveries of Uranus and Pluto are not?
3. Why is it so difficult to see features in the atmosphere of Uranus?
4. (**a**) Draw a figure like Figure 14-3, and indicate on it where Uranus was in 1986 and 2004. Explain your reasoning. (*Hint:* See Figure 14-1 and Figure 14-2.) (**b**) In approximately what year will the Sun next be highest in the sky as seen from Uranus's south pole? Explain your reasoning.
5. Why do you suppose the tilt of Uranus's rotation axis was deduced from the orbits of its satellites and not by observing the rotation of the planet itself?
6. Describe the seasons on Uranus. In what ways are the Uranian seasons different from those on Earth?
7. Explain the statement "Methane is to Uranus's atmosphere as water is to Earth's atmosphere."
8. A number of storms in the Uranian atmosphere can be seen in Figure 14-2, but none are visible in Figure 14-1. How can you account for the difference?
9. Why are Uranus and Neptune distinctly blue-green in color, while Jupiter or Saturn are not?
10. How does the energy source for Uranus's atmospheric motions differ from those from Jupiter, Saturn, and Neptune?
11. Why are fewer white clouds seen on Uranus and Neptune than on Jupiter and Saturn?
12. Why do Uranus and Neptune have higher densities than Jupiter and Saturn?
13. Discuss some competing explanations of why Uranus and Neptune are substantially smaller than Jupiter and Saturn.
14. How do the orientations of Uranus's and Neptune's magnetic axes differ from those of other planets?
15. Briefly describe the evidence supporting the idea that Uranus was struck by a large planetlike object several billion years ago.
16. Compare the rings that surround Jupiter, Saturn, Uranus, and Neptune. Briefly discuss their similarities and differences.
17. The 1977 occultation that led to the discovery of Uranus's rings was visible from the Indian Ocean. Explain why it could not be seen from other parts of the Earth's night side.
18. As *Voyager 2* flew past Uranus, it produced images only of the southern hemispheres of the planet's satellites. Why do you suppose this was?
19. Why do astronomers think that the energy needed to resurface parts of Miranda came from tidal heating rather than the satellite's own internal heat?
20. Using the data in Appendix 3, explain why Uranus's satellites Caliban and Sycorax (both discovered in 1997) were probably captured from space rather than having formed at the same time as the planet itself.
21. Briefly describe the evidence supporting the idea that Triton was captured by Neptune.
22. If you were floating in a balloon in Neptune's upper atmosphere, in what part of the sky would you see Triton rise? Explain your reasoning.
23. Why is it reasonable to suppose that Neptune will someday be surrounded by a broad system of rings, perhaps similar to those that surround Saturn?
24. How can astronomers distinguish a faint solar system object like Pluto from background stars within the same field of view?
25. What is the evidence that the three moons of Pluto have a common origin?
26. How do the presence of Pluto's moons suggest that there must be other worlds beyond Neptune?
27. Why are there a large number of objects with the same semimajor axis as Pluto?
28. Are there any trans-Neptunian objects that are not members of the Kuiper belt? Are there any members of the Kuiper belt that are not trans-Neptunian objects? Explain.

Advanced Questions

Questions preceded by an asterisk (*) involve topics discussed in Box 1-1.

Problem-solving tips and tools

See Box 1-1 for the small-angle formula. Recall that the volume of a sphere of radius r is $4\pi r^3/3$. Section 4-7 discusses the gravitational force between two objects and Newton's form of Kepler's third law. You will find discussions of the original form of Kepler's third law in Section 4-4, tidal forces in Section 4-8, Wien's law for blackbody radiation in Section 5-4, and the transparency of the Earth's atmosphere to various wavelengths of light in Section 6-7.

29. For which configuration of the Sun, Uranus, and Neptune is the gravitational force of Neptune on Uranus at a maximum? For this configuration, calculate the gravitational force exerted by the Sun on Uranus and by Neptune on Uranus. Then

calculate the fraction by which the sunward gravitational pull on Uranus is reduced by Neptune at that configuration. Based on your calculations, do you expect that Neptune has a relatively large or relatively small effect on Uranus's orbit?

30. At certain points in its orbit, a stellar occultation by Uranus would *not* reveal the existence of the rings. What points are those? How often does this circumstance arise? Explain using a diagram.

31. According to one model for the internal structure of Uranus, the rocky core and the surrounding shell of water and methane ices together make up 80% of the planet's mass. This interior region extends from the center of Uranus to about 70% of the planet's radius. (**a**) Find the average density of this interior region. (**b**) How does your answer to (a) compare with the average density of Uranus as a whole? Is this what you would expect? Why?

32. Uranus's epsilon (ε) ring has a radius of 51,150 km. (**a**) How long does it take a particle in the ε ring to make one complete orbit of Uranus? (**b**) If you were riding on one of the particles in the ε ring and watching a cloud near Uranus's equator, would the cloud appear to move eastward or westward as Uranus rotates? Explain your answer.

33. Suppose you were standing on Pluto. Describe the motions of Charon relative to the Sun, the stars, and your own horizon. Would you ever be able to see a total eclipse of the Sun? (*Hint:* You will need to calculate the angles subtended by Charon and by the Sun as seen by an observer on Pluto.) In what circumstances would you *never* see Charon?

34. It is thought that Pluto's tenuous atmosphere may become even thinner as the planet moves toward aphelion (which it will reach in 2113), then regain its present density as it again moves toward perihelion. Why should this be?

35. The brightness of sunlight is inversely proportional to the square of the distance from the Sun. For example, at a distance of 4 AU from the Sun, sunlight is only $(^1/_4)^2 = {}^1/_{16} = 0.0625$ as bright as at 1 AU. Compared with the brightness of sunlight on the Earth, what is its brightness (**a**) on Pluto at perihelion (29.649 AU from the Sun) and (**b**) on Pluto at aphelion (49.425 AU from the Sun)? (**c**) How much brighter is it on Pluto at perihelion compared with aphelion? (Even this brightness is quite low. Noon on Pluto is about as dim as it is on the Earth a half hour after sunset on a moonless night.)

***36.** If Earth-based telescopes can resolve angles down to 0.25 arcsec, how large could a trans-Neptunian object be at Pluto's average distance from the Sun and still not present a resolvable disk?

37. The observations of Pluto shown in Figure 14-19 were made using blue and ultraviolet light. What advantages does this have over observations made with red or infrared light?

***38.** Calculate the maximum angular separation between Pluto and Charon as seen from Earth. Assume that Pluto is at perihelion (29.649 AU from the Sun) and that Pluto is at opposition as seen from Earth.

39. Presumably Pluto and Charon raise tidal bulges on each other. Explain why the average distance between Pluto and Charon is probably constant, rather than increasing like the Earth-Moon distance or decreasing like the Neptune-Triton distance. Include a diagram like Figure 10-17 as part of your answer.

40. (**a**) Find the semimajor axis of the orbit of an object whose period is 3/2 of the orbital period of Neptune. How does your result compare to the semimajor axis of Pluto's orbit? (**b**) A number of Kuiper belt objects called plutinos have been discovered with the same orbital period and hence the same semimajor axis as Pluto. Explain how these objects can avoid colliding with Pluto.

41. Find the semimajor axis of the orbit of an object whose period is twice the orbital period of Neptune. How does your result compare to the outer limit of the Kuiper belt?

42. Suppose you wanted to search for trans-Neptunian objects. Why might it be advantageous to do your observations at infrared rather than visible wavelengths? (*Hint:* At visible wavelengths, the light we see from planets is reflected sunlight. At what wavelengths would you expect distant planets to *emit* their own light most strongly? Use Wien's law to calculate the wavelength range best suited for your search.) Could such observations be done at an observatory on the Earth's surface? Explain.

43. The *New Horizons* spacecraft will swing by Jupiter to get a boost from that planet's gravity, enabling it to reach Pluto relatively quickly. To see what would happen if this technique were not used, consider a spacecraft trajectory that is an elliptical orbit around the Sun. The perihelion of this orbit is at 1 AU from the Sun (at the Earth) and the aphelion is at 30 AU (at Pluto's position). Calculate how long it would take a spacecraft in this orbit to make the one-way trip from Earth to Pluto. Based on the information in Section 14-10, how much time is saved by making a swing by Jupiter instead?

Discussion Questions

44. Discuss the evidence presented by the outer planets that suggests that catastrophic impacts of planetlike objects occurred during the early history of our solar system.

45. Some scientists are discussing the possibility of placing spacecraft in orbit about Uranus and Neptune. What kinds of data should be collected, and what questions would you like to see answered by these missions?

46. If Triton had been formed along with Neptune rather than having been captured, would you expect it to be in a prograde or retrograde orbit? Would you expect the satellite to show signs of tectonic activity? Explain your answers.

47. Would you expect the surfaces of Pluto and Charon to be heavily cratered? Explain why or why not.

48. Imagine that you are in charge of planning the *New Horizons* flyby of Pluto and Charon. In your opinion, what data should be collected and what kinds of questions should the mission attempt to answer?

49. In 2006 the International Astronomical Union changed Pluto's designation from *planet* to *dwarf planet*. One criterion that Pluto failed to meet was that a planet must have "cleared the neighborhood" around its orbit. In what sense has Pluto not done so? In what sense have the eight planets (Mercury through Uranus) cleared their neighborhoods? Do you agree with this criterion?

Web/eBook Questions

50. **Miranda.** Access and view the video "Uranus's Moon Miranda" in Chapter 14 of the *Universe* Web site or eBook. Discuss some of the challenges that would be involved in launching a spacecraft from Earth to land on the surface of Miranda.
51. Charon was discovered by an astronomer at the U.S. Naval Observatory. Why do you suppose the U.S. Navy carries out work in astronomy? Search the World Wide Web for the answer.
52. Search the World Wide Web for a list of trans-Neptunian objects. What are the largest and smallest objects of this sort that have so far been found, and how large are they? Have any objects larger than Eris been found?
53. **Separation of Pluto and Charon.** Pluto is located about 4.5 billion km from Earth and has a maximum observable separation from Charon of about 0.9 arcseconds. Access the AIMM (Active Integrated Media Module) called "Small-Angle Toolbox" in Chapter 1 of the *Universe* Web site or eBook. Use this AIMM and the above data to determine the distance between Pluto and Charon. How does your answer compare with the value given in the text?

Activities

Observing Projects

> **Observing tips and tools**
>
> During the period 2007–2011, Uranus will be at opposition in September, Neptune in August, and Pluto in June. You can find Uranus and Neptune with binoculars if you know where to look (a good star chart is essential), but Pluto is so dim that it can be a challenge to spot even with a 25-cm (10-inch) telescope. Each year, star charts that enable you to find these planets are printed in the issue of *Sky & Telescope* for the month in which each planet is first visible in the nighttime sky. You can also locate the outer planets using the *Starry Night Enthusiast™* program on the CD-ROM that accompanies certain printed copies of this textbook.

54. Make arrangements to view Uranus through a telescope. The planet is best seen at or near opposition. Use a star chart at the telescope to find the planet. Are you certain that you have found Uranus? Can you see a disk? What is its color?
55. If you have access to a large telescope, make arrangements to view Neptune. Like Uranus, Neptune is best seen at or near opposition and can most easily be found using a star chart. Can you see a disk? What is its color?
56. If you have access to a large telescope (at least 25 cm in diameter), make arrangements to view Pluto. Using the star chart from *Sky & Telescope* referred to above, view the part of the sky where Pluto is expected to be seen and make a careful sketch of all of the stars that you see. Repeat this process on a later night. Can you identify the "star" that has moved?
57. Use the *Starry Night Enthusiast™* program to observe the five large satellites of Uranus. Open the **Favourites pane** and click on **Guides > Atlas** to display the entire celestial sphere. Open the **Find** pane and double-click the entry for Uranus to center this planet in the view. (Clicking once on the **Space** bar will speed up this centering). You can reduce the confusion in this view by removing the background stars by clicking on **View > Stars > Stars** and by ensuring that the celestial grid is removed by clicking on **View > Celestial Grid.** Using the controls at the right-hand end of the toolbar, zoom in to a field of view of about $2' \times 1'$. In the toolbar, set the year to 1986 and the **Time Flow Rate** to 1 hour. Then click on the **Run Time Forward** button, the right-pointing triangle on the toolbar. You can scroll on and off the labels for the moons by clicking on **Labels > Planets-Moons.** (**a**) Describe how the satellites move, and relate your observations to Kepler's third law (see Sections 4-4 and 4-7). (**b**) Set the year to 2007 and again click on the **Run Time Forward** button. How do the orbits look different than in (a)? Explain any differences.
58. Use the *Starry Night Enthusiast™* program to examine the satellites of Uranus. (**a**) Select **Solar System > Uranus** from the **Favourites** menu. Remove the image of the astronaut's spacesuit by clicking on **View > Feet** in the menu and remove the background stars by selecting **View > Stars > Stars** from the menu. Use the Elevation buttons in the **Viewing Location** section of the toolbar to change the distance from the planet to about 0.004 AU. You should now be able to see at least five satellites of the planet Uranus. Which satellites are these? Select **Label > Planets-Moons** from the menu to confirm your identification of these satellites. (**b**) You can rotate the image of the planet and its moons by holding down the **Shift** key while clicking the mouse button and moving the mouse. Use this technique to rotate Uranus until you are looking at the plane of the satellites' orbits edge-on. Do all of the satellites appear to lie in the same plane? (To display the orbits of each of the moons, open the **Find** pane, expand the layer for Uranus, and click in the right-hand box next to each moon.) How do you imagine that this plane relates to the plane of Uranus's equator? Why do you suspect that this is so?
59. Use the *Starry Night Enthusiast™* program to observe Neptune and its satellites Triton and Nereid. Display the entire celestial sphere (select **Guides > Atlas** in the **Favourites** menu). Open the **Find** pane and double-click on the entry for Neptune to center this planet in the view. In the Find Pane, expand the layer for Neptune and then click the box to the left of the entries for Triton and Nereid to label them in the view. To reduce confusion, remove the background stars and the coordinate grid by selecting **Stars > Stars** and **> Celestial Grid** under the View menu. Use the controls at the right-hand end of the toolbar to zoom in to a field of view of about $9' \times 6'$. In the toolbar, set the Time Flow Rate to 1 hour and click on the **Run**

Time Forward button (a triangle that points to the right). (**a**) The satellite you see orbiting close to Neptune is Triton. In which direction (clockwise or counterclockwise) does it orbit Neptune? (**b**) Nereid is in a rather elongated orbit that takes it far from Neptune. You can use another viewpoint to show the relative positions and orientations of the orbits of these moons more clearly. In the **Options** menu, open the **Viewing Location** pane and, in the **View from** boxes, select **position hovering over** and **Neptune.** On the map that appears, you will note that you will be hovering over Neptune's equator. To choose this location, click on **Set Location.** Increase your elevation to **0.05 a.u.** above Neptune by clicking and holding the **up** arrow in the **Viewing Location.** Again, you can remove the background stars by clicking on **View > Stars > Stars.** Open the **Find** pane and expand the list of moons under Neptune. You can now label **Triton** and **Nereid** and display their orbits by clicking in the boxes to the left and the right of their names respectively. Set the Time Flow Rate to 1 day and click on the **Run Time Forward** button. In which direction (clockwise or counterclockwise) does Nereid orbit Neptune? When it is closest to Neptune, does Nereid approach closer to Uranus than does Triton? You can view these orbits from different perspectives by holding down the Shift key while moving the mouse over Neptune to tilt this planetary system to different angles. (Note: You are at a point directly above a moon's orbit when its orbital track appears uniformly bright.) You can reduce the elevation above Neptune and set the Time Flow Rate to 10 minutes to show the planet's rotation and see the different clouds and dark spots on its surface. What is the planet's rotation period?

STARRY NIGHT

60. Use the *Starry Night Enthusiast™* program to observe Pluto and Charon. First select **Options > Viewing Location** from the menu. At the top of the Viewing Location dialog box, select **position hovering over** and **Pluto** in the drop boxes. Then click on the center of the map of Pluto that appears in the dialog window and click the **Set Location** button. Use the elevation buttons in the toolbar to increase the distance from the surface of Pluto to about **35,000 km.** Use the **Location scroller** (hold down the Shift key while holding down the mouse button and moving the mouse, to rotate the view around Pluto. In the toolbar, set the **Time Flow Rate** to 1 hour, then click on the **Run Time Forward** button (a triangle that points to the right). (**a**) Estimate Charon's orbital period. (**b**) By following a spot on Pluto's surface, estimate Pluto's rotation period. How does it compare to your answer in part (a)? (**c**) Select **Options > Viewing Location** from the menu and set the dropdown boxes at the top of the Viewing Location dialog window to read **position moving with** and **Pluto.** Then select the **Above orbital plane** option and click the **Set Location** button. Open the **Find** pane and double-click the entry for the **Sun** to center the Sun in the view. Set the Time Flow Rate to 1 year, and click on the **Run Time Forward** button to see the apparent motion of the Sun as seen from Pluto. Observe the motion for several centuries of simulated time. Does the Sun always appear to move at the same speed? Use the properties of Pluto's orbit to explain why or why not.

Collaborative Exercise

61. Sir William Herschel, a British astronomer, discovered Uranus in 1781 and named it Georgium Sidus, after the reigning monarch, George III. What name might Uranus have been given in 1781 if an astronomer in your country had discovered it? Why? What if it had been discovered in your country in 1881? In 1991?

This page intentionally left blank

15 Vagabonds of the Solar System

R I V U X G

Comet Hale-Bopp, the great comet of 1997. (Courtesy of Johnny Horne)

On the Tuesday night after Easter in the year 1066, a strange new star appeared in the European sky. Seemingly trailing fire, this "star" hung in the night sky for weeks. Some, including the English king Harold, may have taken it as an ill omen; others, such as Harold's rival William, Duke of Normandy, may have regarded it as a sign of good fortune. Perhaps they were both right, because by the end of the year Harold was dead and William was installed on the English throne.

Neither man ever knew that he was actually seeing the trail of a city-sized ball of ice and dust, part of which evaporated as it rounded the Sun to produce a long tail. They were seeing a comet—remarkably, the same Comet Halley that last appeared in 1986 and will next be seen in 2061. Although we now understand that comets are natural phenomena rather than supernatural omens, they still have the power to awe and inspire us, as did Comet Hale-Bopp (shown here) in 1997.

While comets are made of ice and dust, asteroids are rocky bits of the solar nebula that never formed into a full-sized planet. Some asteroids break into fragments, as do some comets, and some of these fragments fall to the Earth as meteorites. On extraordinarily rare occasions, an entire asteroid strikes Earth. Such an asteroid impact may have led to the extinction of the dinosaurs some 65 million years ago. Thus, these minor members of our solar system can have major consequences for our planet.

15-1 A search for a planet between Mars and Jupiter led to the discovery of asteroids

After William Herschel's discovery of Uranus in 1781 (which we described in Section 14-1), many astronomers began to wonder if there were other, as yet undiscovered, planets. If these planets were too dim to be seen by the naked eye, they might still be visible through telescopes. These planets would presumably be found close to the plane of the ecliptic, because all the other planets orbit the Sun in or near that plane (see Section 7-1). But how far from the Sun might these additional planets be found?

The Hunt for the "Missing Planet"

LOOKING DEEPER 15.1

Astronomers in the late eighteenth century had a simple rule of thumb, called the *Titius-Bode Law*, relating the sizes of planetary orbits: From one planet to

Learning Goals

By reading the sections of this chapter, you will learn

15-1 What led to the discovery of the asteroids

15-2 Why the asteroids never formed into a planet

15-3 What asteroids look like

15-4 How an asteroid led to the demise of the dinosaurs

15-5 What meteorites tell us about the nature of asteroids

15-6 What meteorites may reveal about the origin of the solar system

15-7 What comets are and why they have tails

15-8 How comets originate from the outer solar system

15-9 The connection between comets and meteor showers the next, the semimajor axis of the orbit increases by a factor between approximately 1.4 and 2. For example, the semimajor axis of Mercury's orbit is 0.39 AU. Venus's orbit has a semimajor axis of 0.72 AU, which is larger by a factor of (0.72)/(0.39) = 1.85. (Unlike Newton's laws, the Titius-Bode "Law" is not a fundamental law of nature. It is probably just a reflection of how our solar system happened to form.)

The glaring exception is Jupiter (semimajor axis 5.20 AU), which is more than *3 times* farther from the Sun than Mars (semimajor axis 1.52 AU). Table 7-1 depicts this large space between the orbits of Mars and Jupiter. If the rule of thumb is correct, astronomers reasoned, there should be a "missing planet" with an orbit about 1.4 to 2 times larger than that of Mars. This planet should therefore have a semimajor axis between 2 and 3 AU.

Six German astronomers, who jokingly called themselves the "Celestial Police," organized an international group to begin a careful search for this missing planet. Before their search got under way, however, surprising news reached them from Giuseppe Piazzi, a Sicilian astronomer. Piazzi had been carefully mapping faint stars in the constellation of Taurus when on January 1, 1801, the first night of the nineteenth century, he noticed a dim, previously uncharted star. This star's position shifted slightly over the next several nights. Suspecting that he might have found the "missing planet," Piazzi excitedly wrote to Johann Bode, the director of the Berlin Observatory and a member of the Celestial Police.

Unfortunately, Piazzi's letter did not reach Bode until late March. By that time, the Earth had moved around its orbit so that Piazzi's object appeared too near the Sun to be visible in the night sky. With no way of knowing where to look after it emerged from the Sun's glare, astronomers feared that Piazzi's object might have been lost.

Upon hearing of this dilemma, the brilliant young German mathematician Karl Friedrich Gauss took up the challenge. He developed a general method of computing an object's orbit from only three separate measurements of its position on the celestial sphere. (With slight modifications, this same method is used by astronomers today.) In November 1801, Gauss predicted that Piazzi's object would be found in the predawn sky in the constellation of Virgo. And indeed Piazzi's object was sighted again on December 31, 1801, only a short distance from the position Gauss had calculated. Piazzi named the object Ceres (pronounced SEE-reez), after the patron goddess of Sicily in Roman mythology.

WEB LINK 15.2

Ceres orbits the Sun once every 4.6 years at an average distance of 2.77 AU, just where astronomers had expected to find the missing planet. But Ceres is very small; its equatorial diameter is a mere 974.6 km (Figure 15-1). (The pole-to-pole diameter is even smaller, just 909.4 km.) Hence, Ceres reflects only a little sunlight, which is why it cannot be seen with the naked eye even at opposition; it can be seen with binoculars, but it looks like just another faint star. Because of its small size, it is known today as a **minor planet**, or **asteroid**.

On March 28, 1802, Heinrich Olbers discovered another faint, starlike object that moved against the background stars. He called it Pallas, after the Greek goddess of wisdom. Like Ceres, Pallas orbits the Sun every 4.6 years at an average distance of 2.77 AU, but its orbit is more steeply inclined from the plane of the ecliptic and is somewhat more eccentric. Pallas is even smaller and dimmer than Ceres, with an estimated diameter of only 522 km. Obviously, Pallas was also not the missing planet.

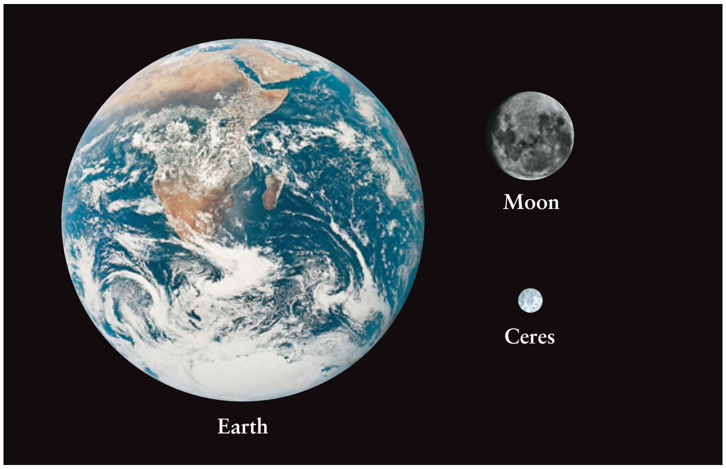


Figure 15-1 R I V U X G

Ceres Compared with the Earth and Moon A drawing of Ceres, the largest of the asteroids and the first one discovered, is shown here to the same scale as the Earth and the Moon. Ceres is too small to be considered a planet, and it cannot be regarded as a moon because it does not orbit any other body. To denote their status, asteroids are also called minor planets. (NASA)

Did the missing planet even exist? Some astronomers speculated that perhaps there had once been such a full-size planet, but it had somehow broken apart or exploded to produce a population of asteroids orbiting between Mars and Jupiter. The search was on to discover this population. Two more such asteroids were discovered in the next few years, Juno in 1804 and Vesta in 1807. Several hundred more were discovered beginning in the mid-1800s, by which time telescopic equipment and techniques had improved.

An Abundance of Asteroids

The next major breakthrough came in 1891, when the German astronomer Max Wolf began using photographic techniques to search for asteroids. Before this, asteroids had to be painstakingly discovered by scrutinizing the skies for faint, uncharted, starlike objects whose positions move slightly from one night to the next. With photography, an astronomer simply aims a camera-equipped telescope at the stars and takes a long exposure. If an asteroid happens to be in the field of view, its orbital motion during the long exposure leaves a distinctive trail on the photographic plate (Figure 15-2). Using this technique, Wolf alone discovered 228 asteroids.

WEB LINK 15.3

Today, more than 300,000 asteroids have been discovered, of which 100,000 have been studied well enough so that their orbits are known. About 5000 more are discovered every month, some by amateur astronomers. After the discoverer reports a new find to the Minor Planet Center of the Smithsonian Astrophysical Observatory, the new asteroid is given

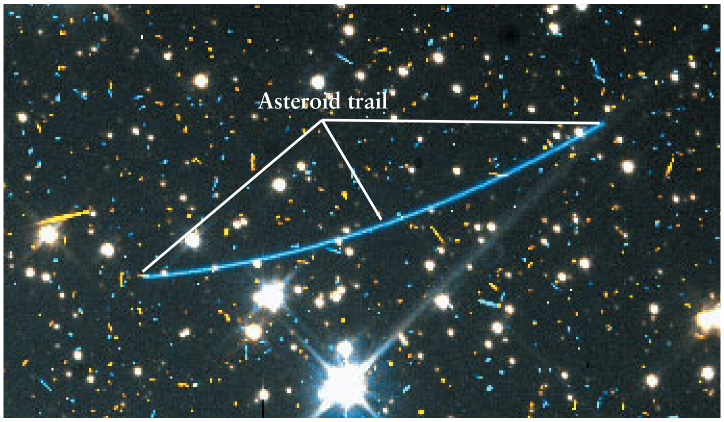


Figure 15-2 R I V U X G

The Trail of an Asteroid Telescopes used to photograph the stars are motorized to follow the apparent motion of the celestial sphere. Because asteroids orbit the Sun, their positions change with respect to the stars, and they leave blurred trails on time exposure images of the stars. This bluish asteroid trail was recorded quite by accident by the Hubble Space Telescope. (R. Evans and K. Stapelfeldt, Jet Propulsion Laboratory; and NASA)

a provisional designation. For example, asteroid 1980 JE was the fifth ("E") to be discovered during the second half of May (the tenth half-month of the year, "J") 1980. The same scheme for provisional designations is also used for small objects found in the outer solar system beyond Neptune, or trans-Neptunian objects (see Section 14-10).

If the asteroid is located again on at least four succeeding oppositions—a process that can take decades—the asteroid is assigned an official sequential number (Ceres is 1, Pallas is 2, and so forth), and the discoverer is given the privilege of suggesting a name for the asteroid. The suggested name must then be approved by the International Astronomical Union. For example, the asteroid 1980 JE was officially named 3834 Zappafrank (after the American musician Frank Zappa) in 1994.

Ceres—or, in modern nomenclature, 1 Ceres—is unquestionably the largest asteroid; it has about 25% of the mass of all the asteroids combined. In recognition of its size, 1 Ceres is also called a **dwarf planet,** a designation that it shares with the large trans-Neptunian objects Eris and Pluto (see Section 14-10).

Besides 1 Ceres, only two asteroids—2 Pallas, and 4 Vesta—have diameters greater than 300 km. Thirty other asteroids have diameters between 200 and 300 km, and 200 more are bigger than 100 km across. The vast majority of asteroids are less than 1 km across.

CAUTION! Be careful not to confuse asteroids with the trans-Neptunian objects that we discussed in Section 14-10. Asteroids are found in the inner solar system, are made primarily of rocky materials, and are generally quite small. By contrast, trans-Neptunian objects are found primarily in the Kuiper belt beyond Neptune, are a mixture of ices and rock, and range in size from a few kilometers across to 2900 km in diameter—more than 3 times the diameter of 1 Ceres. The combined mass of all trans-Neptunian objects is not known, but is probably thousands of times greater than the combined mass of all asteroids. Asteroids make up only a tiny fraction of the total mass of the solar system!

The Asteroid Belt

Like Ceres, Pallas, Vesta, and Juno, most asteroids orbit the Sun at distances between 2 and 3.5 AU. This region of our solar system between the orbits of Mars and Jupiter is called the **asteroid belt** (Figure 15-3).

CAUTION! You may wonder how spacecraft such as *Voyager 1*, *Voyager 2*, *Galileo*, and *Cassini* were able to cross the asteroid belt to reach Jupiter or Saturn. Aren't asteroids hazards to space navigation, as they are sometimes depicted in science-fiction movies? Wouldn't these spacecraft have been likely to run into an asteroid? Happily, the answer to both questions is no! While it is true that there are more than 10^5 asteroids, they are spread over a belt with a total area (as seen from above the plane of the ecliptic) of about 10^{17} square kilometers—about 10^8 times greater than the Earth's entire surface area. Hence, the average distance between asteroids in the ecliptic plane is about 10^6 kilometers, about twice the distance between the Earth and Moon. Furthermore, many asteroids have orbits that are tilted out of the ecliptic. The *Galileo* spacecraft did pass within a few thousand kilometers of two asteroids, but these passes were intentional and required that the spacecraft be carefully aimed.

Science-fiction movies to the contrary, asteroids are not hazards to space navigation

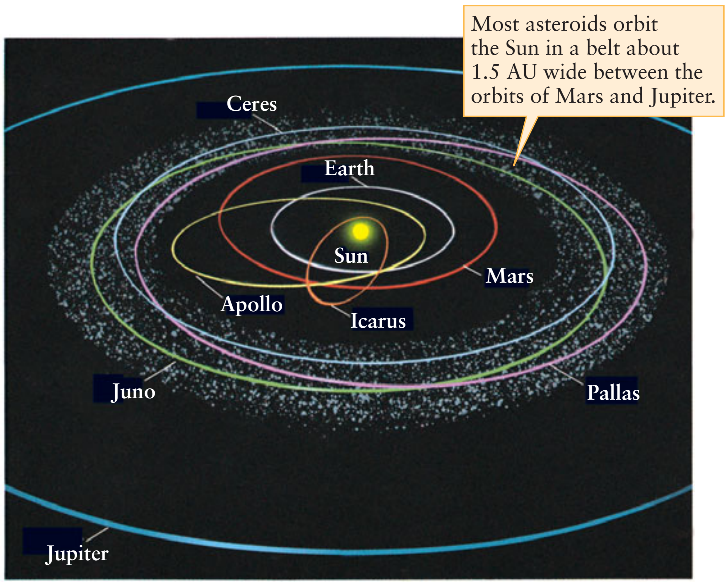


Figure 15-3

The Asteroid Belt Most asteroids, including the large asteroids 1 Ceres, 2 Pallas, and 3 Juno, have roughly circular orbits that lie within the asteroid belt. By contrast, some asteroids, such as 1862 Apollo and 1566 Icarus, have eccentric orbits that carry them inside the orbit of Earth.

What of the nineteenth-century notion of a missing planet? This idea has long since been discarded, because the combined matter of all the asteroids (including an estimate for those not yet officially known) would produce an object barely 1500 km in diameter. This is considerably smaller than anything that could be considered a missing planet. We now understand that asteroids are actually debris left over from the formation of the solar system.

15-2 Jupiter's gravity helped shape the asteroid belt

Where did the asteroid belt come from? Why did asteroids, rather than a planet, form in the solar nebula (see Section 8-5) in the region between Mars and Jupiter? Important insights into these questions come from supercomputer simulations of the formation of the terrestrial planets, like the one depicted in Figure 8-12 but more elaborate.

Jupiter and the Formation of the Asteroid Belt

A typical simulation starts off with a billion (10^9) or so planetesimals, each with a mass of 10^{15} kg or more, so that their combined mass equals that of the terrestrial planets. The computer then follows these planetesimals as they collide and accrete to form the planets. By selecting different speeds and positions for the planetesimals at the start of the simulation, a scientist can study alternative scenarios of the development of the inner solar system.

If the effects of Jupiter's gravity are not included in a simulation, an Earth-sized planet usually forms in the asteroid belt, giving us five terrestrial planets instead of four. But when Jupiter's gravity is added, this fifth terrestrial planet is less likely to form. Jupiter's strong pull "clears out" the asteroid belt by disrupting the orbits of planetesimals in the belt, ejecting most of them from the solar system altogether. As a result, the asteroid belt becomes depleted of planetesimals before a planet has a chance to form. The few planetesimals that remain in the simulation—only a small fraction of those that originally orbited between Mars and Jupiter—become the asteroids that we see today.

The asteroids were never part of a larger planet

Jupiter's gravitational influence, however, probably cannot explain why asteroid orbits have the wide variety of semimajor axes, eccentricities, and inclinations to the ecliptic shown in Figure 15-3. Even with the influence of Jupiter, collisions between planetesimals in the solar nebula would have left the asteroids in roughly circular orbits close to the ecliptic plane, like the orbits of the planets. Therefore, something else must have "stirred up" the asteroids to put them into their current state. Recent analyses suggest that one or more Mars-sized objects—that is, terrestrial planets intermediate in size between the Earth and the Moon—did in fact form within the asteroid belt. Planetesimals that passed within close range of these Mars-sized objects would have experienced strong gravitational forces, and these forces could have deflected the planetesimals into the eccentric or inclined orbits that many asteroids follow today.

Mars-sized objects would also have helped clear out the asteroid belt by deflecting planetesimals into orbits that crossed Jupiter's orbit. Once a planetesimal wandered close to Jupiter, it could be ejected from the solar system by the giant planet's gravitational force.

In this scheme, the Mars-sized objects disappeared when gravitational forces from Jupiter either accelerated them out of the solar system entirely or caused them to slow down and fall into the Sun. A Mars-sized object colliding with the Earth is thought to have produced the Moon (see Figure 10-18). Perhaps this object was an "escapee" from the asteroid belt.

Kirkwood Gaps

Jupiter's gravity continues to influence the asteroid belt down to the present day. The American astronomer Daniel Kirkwood found the first evidence for this in 1867, when he discovered gaps in the asteroid belt. These features, today called **Kirkwood gaps**, can best be seen in a histogram like Figure 15-4, which shows asteroid orbital periods. The gaps in this histogram show regions where there are relatively few asteroids. Curiously, these gaps occur for asteroid orbits whose periods are simple fractions (such as 1/3, 2/5, 3/7, and 1/2) of Jupiter's orbital period.

To understand why the Kirkwood gaps exist, imagine an asteroid within the belt that circles the Sun once every 5.93 years, exactly half of Jupiter's orbital period. On every second trip around the Sun, the asteroid finds itself lined up between Jupiter

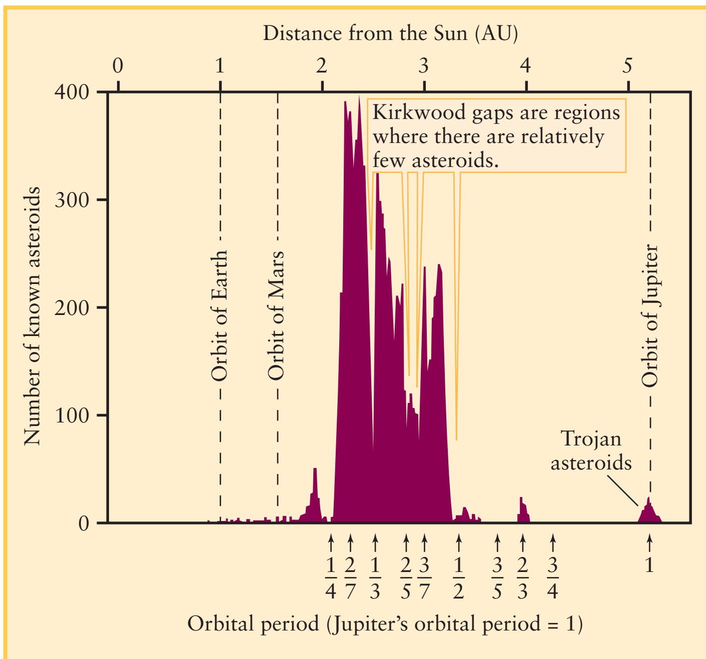


Figure 15-4

The Kirkwood Gaps This graph displays the numbers of asteroids at various distances from the Sun. Very few asteroids are found in orbits whose orbital periods are equal to the simple fractions 1/3, 2/5, 3/7, and 1/2 of Jupiter's orbital period. The Trojan asteroids (at far right) follow the same orbit as Jupiter and thus have the same orbital period. We describe these in Section 15-4.

and the Sun, always at the same location and with the same orientation. Because of these repeated alignments, called a *2-to-1 resonance,* Jupiter's gravity deflects the asteroid from its original 5.93-year orbit, ultimately ejecting it from the asteroid belt. Similar resonances help clear out the Cassini division in Saturn's rings (see Section 12-11), drive the tidal heating of Jupiter's moon Io (see Section 13-4), and give an outer edge to the Kuiper belt beyond Neptune (see Section 14-10). According to Kepler's third law, a period of 5.93 years corresponds to a semimajor axis of 3.28 AU. Because of Jupiter's gravity, almost no asteroids orbit the Sun at this average distance.

Another Kirkwood gap corresponds to an orbital period of one-third Jupiter's period, or 3.95 years (a *3-to-1 resonance*). Additional gaps exist for other simple ratios between the periods of asteroids and Jupiter. Any Mars-sized bodies that existed in the asteroid belt were probably ejected after wandering into one of the Kirkwood gaps.

15-3 Astronomers use a variety of techniques to study asteroids

Although the asteroid belt is mostly empty space, collisions between asteroids should occur from time to time. Asteroids move in orbits with a variety of different eccentricities and inclinations (see Figure 15-3), and some of these orbits intersect each other. A collision between asteroids must be an awe-inspiring event. Typical collision speeds are estimated to be from 1 to 5 km/s (3600 to 18,000 km/h, or 2000 to 11,000 mi/h), which is more than sufficient to shatter rock. Recent observations show that these titanic impacts play an important role in determining the nature of asteroids.

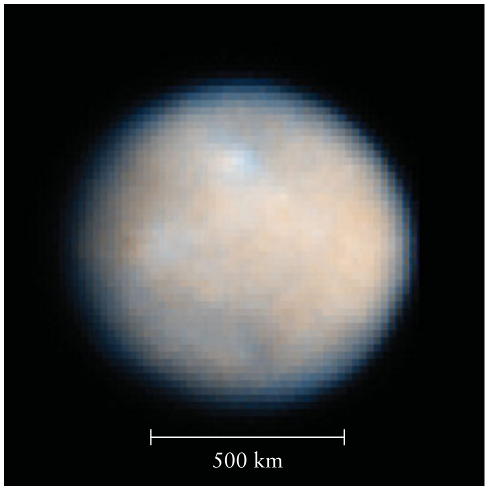
Many asteroids have been battered by collisions with other asteroids

Observing Asteroids with Telescopes, Radar, and Spacecraft

State-of-the-art optical telescopes have good enough resolution to reveal surface details on large asteroids such as 1 Ceres, 2 Pallas, and 4 Vesta (Figure 15-5). These asteroids resemble miniature terrestrial planets. They have enough mass, and hence enough gravity, to pull themselves into roughly spherical shapes. These large asteroids are likely to have extensively cratered surfaces due to collisions with other asteroids in the asteroid belt.

Optical telescopes reveal much less detail about the smaller asteroids. But by studying how the brightness of an asteroid changes as it rotates, astronomers can infer the asteroid's shape. An elongated asteroid appears dimmest when its long axis points toward the Sun so that it presents less of its surface to reflect sunlight. By contrast, a rotating spherical asteroid will have a more constant brightness. Such observations show that most smaller asteroids are not spherical at all. Instead, they retain the odd shapes produced by previous collisions with other asteroids.

An important recent innovation in studying asteroids has been the use of radar. An intense radio beam is sent toward an asteroid, and the reflected signal is detected and analyzed to determine the asteroid's shape (Figure 15-6). The amount of the radio signal that is reflected also gives clues about the texture of the surface.



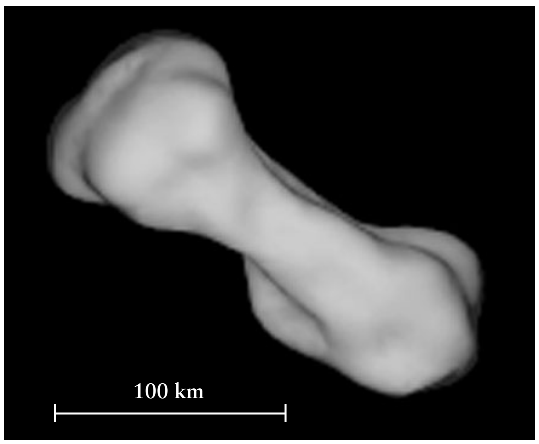
WEB LINK 15.4

Figure 15-5 R I V U X G

A Hubble Space Telescope View of a Large Asteroid With an equatorial diameter of 974.6 km (605.6 mi), 1 Ceres is the largest of all the asteroids. Its very nearly spherical shape suggests that it was once molten throughout its volume, in which case it probably underwent chemical differentiation like the terrestrial planets (see Box 7-1 and Section 8-5). (NASA; ESA; J. Parker, SwRI; P. Thomas, Cornell U., L. McFadden, U. of Maryland; and M. Mutchler and Z. Levay, STScI)

VIDEO 15.1 WEB LINK 15.6

The most powerful technique for studying asteroids is to send a spacecraft to make observations at close range. The Jupiter-bound *Galileo* spacecraft gave us our first close-up view of an asteroid



WEB LINK 15.5

Figure 15-6 R I V U X G

A Radar View of a Medium-Sized Asteroid This radar image reveals the curious dog-bone shape of the asteroid 216 Kleopatra. The 300-m Arecibo radio telescope in Puerto Rico was used both to send radio waves toward Kleopatra and to detect the waves that were reflected back. Scientists then created this image by computer analysis of the reflected waves. (Arecibo Observatory, JPL/NASA)

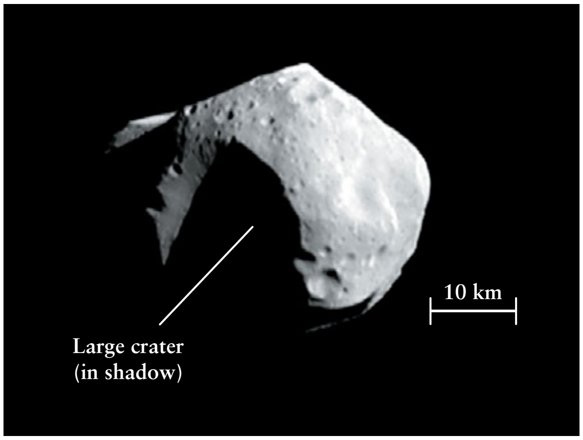


Figure 15-7 R I V U X G

A Spacecraft View of a Small Asteroid The asteroid 253 Mathilde has an albedo of only 0.04, making it half as reflective as a charcoal briquette. It has an irregular shape, with dimensions 66 × 48 × 46 km. The partially shadowed crater indicated by the label is about 20 km across and 10 km deep. The *NEAR Shoemaker* spacecraft imaged Mathilde at close range during its close flyby in June 1997. (Johns Hopkins University, Applied Physics Laboratory)

WEB LINK 15.7

in 1991, when it flew past 951 Gaspra. Two years later *Galileo* made a close pass by 243 Ida. The *NEAR Shoemaker* spacecraft made a close approach to 253 Mathilde in 1997 (Figure 15-7). It later went into orbit around 433 Eros (see Figure 7-7) and touched down on Eros in 2001. It was the first spacecraft to orbit and land on an asteroid. A number of other spacecraft have since made close flybys of asteroids.

One of the most ambitious asteroid missions to date is *Hayabusa* ("Falcon"), a project of the Japan Aerospace Exploration Agency. In 2005 it touched down briefly on the surface of 25143 Itokawa, where it was to have gathered small samples of the asteroid's surface. Scientists will not know whether this attempt was successful until *Hayabusa* returns to Earth in 2010. In 2007 the *Dawn* spacecraft was scheduled to be launched on a mission that would place it in orbit around 4 Vesta in 2011 and in orbit around 1 Ceres in 2015.

Collisions and the Nature of Asteroids

What have spacecraft and Earth-based radar told us about the asteroids? To appreciate the answer, we should first consider what scientists expected to find with these new tools. A few years ago, the consensus among planetary scientists was that the smaller asteroids were solid, rocky objects. The surfaces were presumed to be cratered by impacts, but only relatively small craters were expected. The reason is that an impact forceful enough to make a very large crater would probably fracture a rocky asteroid into two or more smaller asteroids.

The truth proves to be rather different. It is now suspected that most asteroids larger than about a kilometer are not entirely solid, but are actually composed of several pieces. An extreme example is Mathilde (see Figure 15-7). This asteroid has such a low density (1300 kg/m^3) that it cannot be made of solid rock, which typically has densities of 2500 kg/m^3 or higher. Instead, it is probably a "rubble pile" of small fragments that fit together loosely. Billions of years of impacts have totally shattered this asteroid. But the impacts were sufficiently gentle that the resulting fragments drifted away relatively slowly, only to be pulled back onto the asteroid by gravitational attraction.

Another surprising discovery is that some asteroids have extremely large craters, comparable in size to the asteroid itself. An example is the immense crater on Mathilde, shown in Figure 15-7. The object that formed this crater probably collided with Mathilde at a speed of 1 to 5 km/s (3600 to 18,000 km/h, or 2000 to 11,000 mi/h), which is more than sufficient to shatter rock. Why didn't this collision break the asteroid completely apart, sending the pieces flying in all directions? The explanation may be that the asteroid had already been broken into a loose collection of fragments before this major impact took place. Such a fragmented asteroid can more easily absorb the energy of a collision than can a solid, rocky one.

ANALOGY If you fire a rifle bullet at a wine glass, the glass will shatter into tiny pieces. But if you fire the same bullet into a bag full of sand (which has much the same chemical composition as glass), the main damage will be a hole in the bag. In a similar way, an asteroid that is a collection of small pieces can survive a major impact more easily than can a solid asteroid.

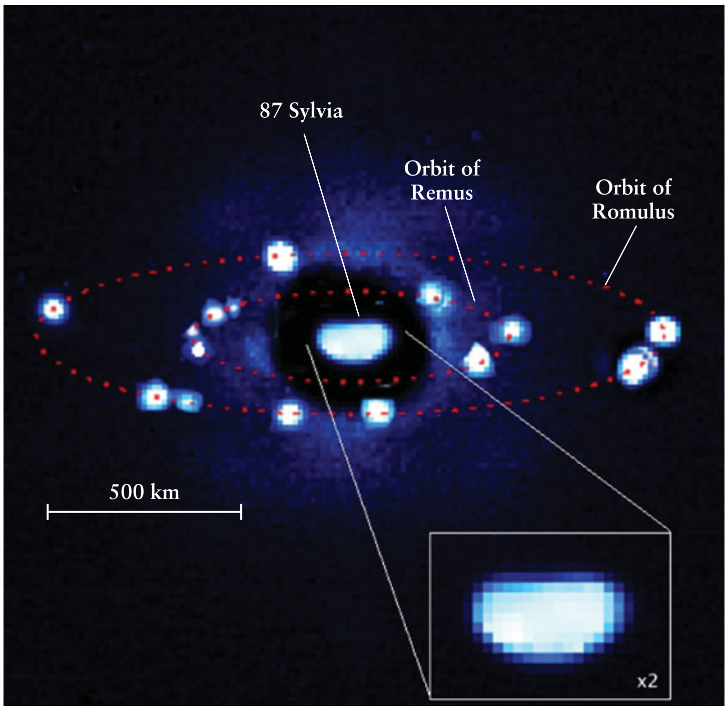


Figure 15-8 R I V U X G

An Asteroid with Two Satellites Although it is only about 280 km across, the potato-shaped asteroid 87 Sylvia has two satellites of its own. The outer satellite, Romulus, is 18 km across; the inner satellite, Remus, is just 7 km across. This image is actually a composite of 9 observations made with the adaptive optics system at the Yepun telescope of the European Southern Observatory (see Figure 6-17). (F. Marchis, U. of California; P. Descamps, D. Hestroffer, and J. Berthier, Observatoire de Paris; and ESO)

In some cases a collision that produces a rubble-pile asteroid can eject a piece that goes into orbit around the asteroid. This may explain why a handful of asteroids are known to have satellites. There is even one known case of an asteroid with two satellites, like a miniature planetary system (Figure 15-8).

If a collision is sufficiently energetic, it may shatter an asteroid permanently. Evidence for this was first pointed out in 1918 by the Japanese astronomer Kiyotsugu Hirayama. He drew attention to families of asteroids that share nearly identical orbits. Each of these groupings, now called **Hirayama families,** presumably resulted from a parent asteroid that was broken into fragments by a high-speed, high-energy collision with another asteroid.

15-4 Asteroids are found outside the asteroid belt—and have struck the Earth

While the majority of asteroids lie within the asteroid belt, there are some notable exceptions. A few actually share the same orbit as Jupiter. Other asteroids venture well inside the orbit of Mars, and can pose a serious threat to life on Earth.

The Lagrange Points and Trojan Asteroids

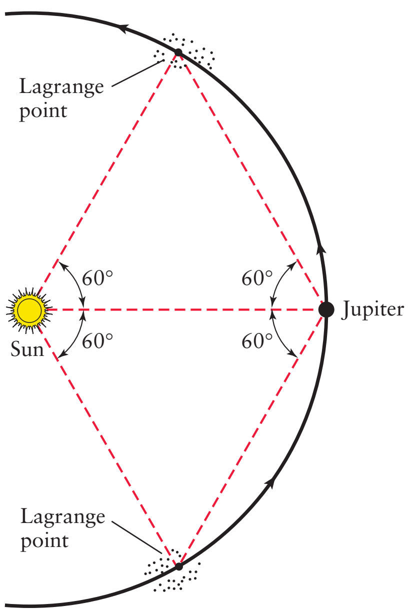
We saw in Section 15-3 that Jupiter's gravitational pull by itself depletes certain orbits in the asteroid belt, forming the Kirkwood gaps. But the gravitational forces of the Sun and Jupiter work together to *capture* asteroids at two locations outside the asteroid belt called the **stable Lagrange points.** One of these points is one-sixth of the way around Jupiter's orbit ahead of the planet, while the other point is the same distance behind the planet (Figure 15-9*a*). The French mathematician Joseph Louis Lagrange predicted the existence of these points in 1772; asteroids were first discovered there in 1906.

The asteroids trapped at Jupiter's Lagrange points are called **Trojan asteroids,** named individually after heroes of the Trojan War. More than 1600 Trojan asteroids have been catalogued so far (see Figure 15-9*b*). (Four objects have also been found at the Lagrange points of Neptune. These are icy trans-Neptunian objects rather than rocky asteroids.)

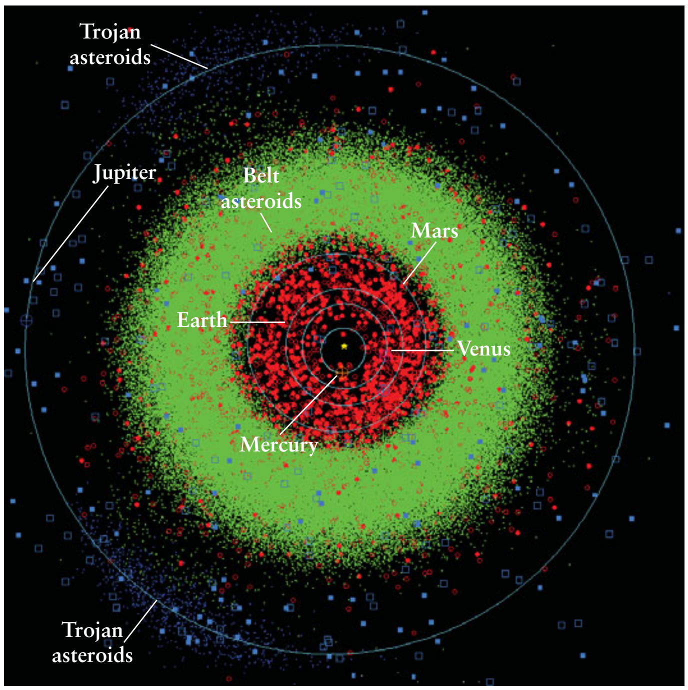
Near-Earth Objects and Impacts from Space

Some asteroids have highly elliptical orbits that bring them into the inner regions of the solar system. Asteroids that cross Mars's orbit, or whose orbits lie completely within that of Mars, are called **near-Earth objects** or **NEOs.** These are shown in red in Figure 15-9*b*.

Occasionally, a near-Earth object may pass relatively close to the Earth. On October 30, 1937, Hermes passed the Earth at a distance of 900,000 km (560,000 mi), only a little more than twice the distance to the Moon. One of the closest calls to date occurred on December 9, 1994, when an asteroid called 1994 XM1 passed within a mere 105,000 km (60,000 mi) of Earth.



(a) A clump of asteroids is found at each of Jupiter's Lagrange points



(b) A map of all asteroids within Jupiter's orbit

Figure 15-9

Asteroids Outside the Belt (a) The combined gravitational forces of Jupiter and the Sun tend to trap asteroids near the two stable Lagrange points along Jupiter's orbit. Note that the Sun, Jupiter, and either one of these points lie at the vertices of an equilateral triangle. **(b)** This plot shows the actual positions of all known asteroids at Jupiter's orbit or closer. Green dots denote belt asteroids, deep blue squares denote Trojan asteroids, and red circles denote asteroids that come within 1.3 AU of the Sun. Comets are filled and unfilled light-blue squares. (b: Minor Planet Center)

Figure 15-10 R I V U X G

The Barringer Crater This 1.2-km-(3⁄4-mi-) wide crater was formed by a meteoroid impact about 50,000 years ago. Fewer than 200 other impact craters between 3 and 150 km across have been identified on the Earth. None of these is more than 500 million years old, because the reshaping of the Earth's surface by erosion and by tectonic processes tends to erase craters over time. (Meteor Crater Enterprises)

WEB LINK 15.9

Because 1994 XM1 is only about 10 meters across (about the size of a house), even if it had been aimed directly at the Earth it would probably have broken up in the atmosphere before reaching the surface.

More than 4100 near-Earth objects are known, and there may be thousands more that we have not yet detected. Fortunately, space is a big place and the Earth is a rather small target, so asteroids strike the Earth only very rarely.

However, collisions between asteroids produce numerous smaller chunks of rock, and many of these do eventually rain down on the terrestrial planets. Fortunately for us, the majority of these fragments, usually called **meteoroids,** are quite small. But on rare occasions a large fragment collides with our planet. When such a collision takes place, the result is an impact crater whose diameter depends on the mass and speed of the impinging object.

A relatively young, pristine terrestrial impact crater is the Barringer Crater near Winslow, Arizona (Figure 15-10). This crater was formed 50,000 years ago when an iron-rich object approximately 50 m across struck the ground at a speed of more than 11 km/s (40,000 km/h, or 25,000 mi/h). The resulting blast was equivalent to the detonation of a 20-megaton hydrogen bomb and left a crater 1.2 km (3⁄4 mi) wide and 200 m (650 ft) deep at its center. (Figure 7-10*b* shows a much larger but much older crater in Canada.)

Iridium and the Demise of the Dinosaurs

Iron is an important constituent of asteroids and meteoroids. Another element, iridium, is common in iron-rich minerals but rare in other types of rocks. Measurements of iridium in the Earth's crust can thus tell us the rate at which meteoritic material has been deposited on the Earth over the ages. The geologist Walter Alvarez and his physicist father, Luis Alvarez, from the University of California, Berkeley, made such measurements in the late 1970s.

An asteroid impact probably ended the age of dinosaurs 65 million years ago

Working at a site of exposed marine limestone in the Apennine Mountains in Italy, the Alvarez team discovered an exceptionally high abundance of iridium in a dark-colored layer of clay between limestone strata (Figure 15-11). Since this discovery in 1979, a comparable layer of iridium-rich material has been uncovered at numerous sites around the world. In every case, geological dating reveals that this apparently worldwide layer of iridium-rich clay was deposited about 65 million years ago.

Paleontologists were quick to realize the significance of this date, because it was 65 million years ago that all the dinosaurs rather suddenly became extinct. In fact, more than 75% of all the species on the Earth disappeared within a relatively brief span of time.

The Alvarez discovery suggests a startling explanation for this dramatic mass extinction: Perhaps an asteroid hit the Earth at that time. An asteroid 10 km in diameter slamming into our planet would have produced a fiery plume that ignited tremendous wildfires, killing untold numbers of animals and consuming much of the Earth's vegetation. The dust thrown up from the impact and soot from wildfires hung in the atmosphere, blackening the sky for months after the impact. The dust eventually drifted to the ground, forming a worldwide layer of iridium-rich clay. (Evidence for this picture is that a layer of soot 65 million years old is also found at various sites around the world. From the thickness of this layer, scientists calculate that the wildfires produced nearly 70 *billion* tons of soot.) Temperatures would have plummeted due to the blocking out of sunlight by dust and soot, only to increase later due to global warming brought on by greenhouse gases (see Section 9-1 and Section 9-7) released by the wildfires. The *Cosmic Connections* figure depicts the sequence of events that may have followed the asteroid impact.

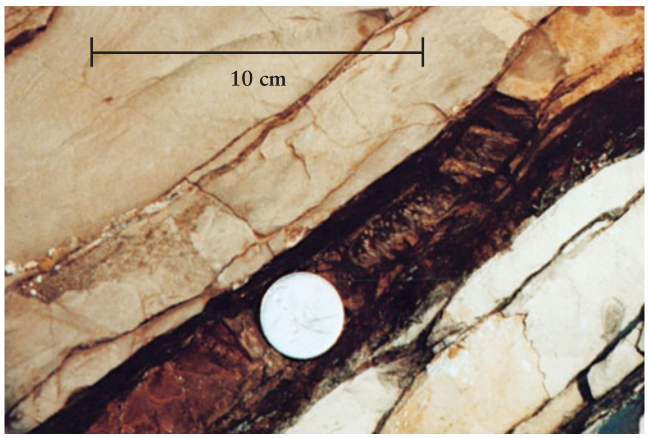


Figure 15-11 R I V U X G

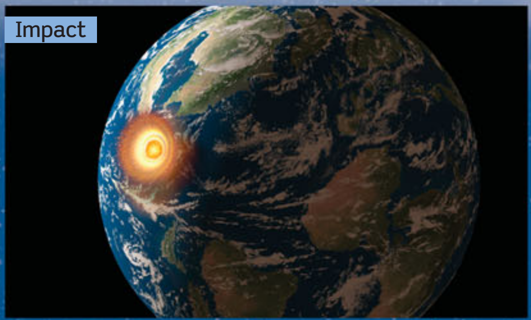
Iridium-Rich Clay This photograph of strata in Italy's Apennine Mountains shows a dark-colored layer of iridium-rich clay. This is sandwiched between older layers of white limestone (lower right) and younger layers of grayish limestone (upper left). The iridium-rich layer is thought to be the result of an asteroid impact 65 million years ago. The coin is the size of a U.S. quarter. (Courtesy of W. Alvarez)

COSMIC CONNECTIONS

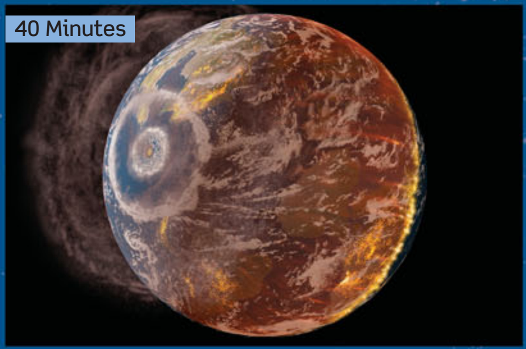
A Killer Asteroid

There is compelling evidence that 65 million years ago the Earth was struck by an asteroid about 10 km in diameter. This impact triggered a biological cataclysm that wiped out 75 percent of all plant and animal species on our planet. (From "The Day the World Burned" by D. A. Kring and D. D. Durda, *Scientific American*, December 2003; art by Chris Butler)

Late Cretaceous swamps and rivers in North America had a mix of coniferous, broad-leaved evergreen, and deciduous trees. They formed canopied forests and open woodlands with understories of ferns, aquatic plants and flowering shrubs.



The Chicxulub Impact occurred in a shallow sea and immediately lofted rocky, molten and vaporous debris into the atmosphere. The bulk of the debris rained down on nearby continental regions, but much of it rose all the way into space.



The vapor-rich plume of material expanded to envelop Earth. As material in that plume fell back to the ground, it streaked through the atmosphere like trillions of meteors, heating it in some places by hundreds of degrees.

After fires had ravaged the landscape, only a few stark trunks and skeletons remained. Soot from the fires and dust from the impact slowly settled to the ground. Sunlight was dramatically, if not totally, attenuated for months.

The postimpact environment was less diverse. Ferns and algae were the first to recover. Plant species in swamps and swamp margins generally survived better than species in other types of ecosystems. Conifers fared particularly badly.

Shrubs took advantage of the vacant landscape and began to cover it. Species pollinated by the wind did better than those that relied on insects. Trees began to grow, but it took years for forest canopies to rebuild. The recovery time is uncertain.

Tiny rodentlike creatures capable of ferreting out seeds and nuts were among the animals that managed to survive this catastrophe, setting the stage for the rise of mammals in the eons that followed. As the dominant mammals on the planet, we may owe our very existence to an ancient asteroid.

In 1992 a team of geologists suggested that this asteroid crashed into a site in Mexico. They based this conclusion on glassy debris and violently shocked grains of rock ejected from the 180-km-diameter Chicxulub Crater on the Yucatán Peninsula. Using radioactive dating techniques (see Section 8-3 and Box 8-1), the scientists found that the asteroid struck 64.98 million years ago, in remarkable agreement with the age of the iridium-rich layer of clay.

Some geologists and paleontologists are not convinced that an asteroid impact led to the extinction of the dinosaurs. But many scientists agree that this hypothesis fits the available evidence better than any other explanation that has been offered so far.

Tunguska and the Threat of Asteroid Impact

On June 30, 1908, a spectacular explosion occurred over the Tunguska region of Siberia that released about 10^{15} joules of energy, equivalent to a nuclear detonation of several hundred kilotons. The blast knocked a man off his porch some 60 km away and could be heard more than 1000 km away. Millions of tons of dust were injected into the atmosphere, darkening the sky as far away as California.

Preoccupied with political upheaval and World War I, Russia did not send a scientific expedition to the site until 1927. Researchers found that trees had been seared and felled radially outward in an area about 50 km in diameter (Figure 15-12). There was no clear evidence of a crater. In fact, the trees at "ground zero" were left standing upright, although they were completely stripped of branches and leaves.

Several teams of modern astronomers have argued that the Tunguska explosion was caused by an asteroid traveling at supersonic speed. They arrived at this conclusion after assessing the effects of various impactor sizes, speeds, and compositions. Even data from above-ground nuclear detonations of the 1940s and 1950s were worked into the calculations. The Tunguska event is well-matched by an asteroid about 80 m (260 ft) in diameter entering the Earth's atmosphere at 22 km/s (80,000 km/h = 50,000 mph).

Large objects, such as that which formed the Chicxulub Crater and may have led to the demise of the dinosaurs, occasionally strike the Earth with the destructive force of as much as a million megatons of TNT. An asteroid 1 or 2 kilometers in diameter, striking the Earth at 30 km/s, would destroy an area the size of California. Such an impact would throw more than 10^{13} kg of microscopic dust particles into the atmosphere, which would decrease the amount of sunlight reaching the Earth's surface by enough to threaten the health of the world's agriculture. The loss of a year's crops could lead to the demise of a quarter of the world's population and would place our civilization—although perhaps not the survival of our species—in grave jeopardy.

Figure 15-12 R I V U X G

Aftermath of the Tunguska Event In 1908 a stony asteroid with a mass of about 10^8 kilograms entered the Earth's atmosphere over the Tunguska region of Siberia. Its passage through the atmosphere made a blazing trail in the sky some 800 km long. The asteroid apparently exploded before reaching the surface, blowing down trees for hundreds of kilometers around "ground zero." (Courtesy of Sovfoto)

The good news is that such large objects strike the Earth *very* infrequently. An asteroid 1 km in diameter might be expected to hit the Earth only once every 300,000 years. Statistically, the probability that an average person will die because of such an event is about 1 in 20,000, or only about 1% as great as the probability of dying in an automobile accident. While the threat of asteroid or comet impact exists, a catastrophic impact is very unlikely to occur in our lifetimes.

15-5 Meteorites are classified as stones, stony irons, or irons, depending on their composition

Asteroids can only be studied at close range using spacecraft. But from time to time pieces of asteroids fall to Earth where they can be examined by scientists. These asteroid fragments are called *meteorites*.

Meteoroids, Meteors, and Meteorites

A *meteoroid,* like an asteroid, is a chunk of rock in space. There is no official dividing line between meteoroids and asteroids, but the term *asteroid* is generally applied only to objects larger than 50 m across.

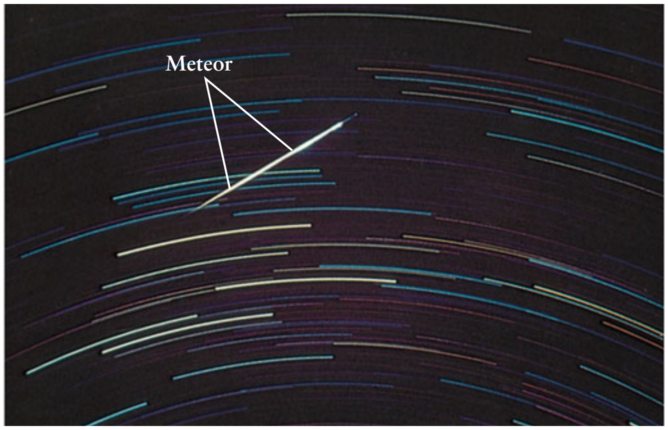


Figure 15-13 R I V U X G

A Meteor A meteor is produced when a meteoroid—a piece of interplanetary rock or dust—strikes the Earth's atmosphere at high speed (typically 5 to 30 km/s). The heat of air friction makes the meteoroid glow as it disintegrates. Most meteoroids burn up completely at altitudes of 100 km (60 mi) or so. This long exposure (notice the star trails) shows an exceptionally bright meteor called a fireball. (Courtesy of R. A. Oriti)

A **meteor** is the brief flash of light (sometimes called a *shooting star*) that is visible at night when a meteoroid enters the Earth's atmosphere (Figure 15-13). Frictional heat is generated as the meteoroid plunges through the atmosphere, leaving a fiery trail across the night sky.

WEB LINK 15.12

If a piece of rock survives its fiery descent through the atmosphere and reaches the ground, it is called a **meteorite.** As incredible as it may seem, an estimated total of 300 tons of extraterrestrial matter falls on the Earth each day. Most of this is dust rather than meteorites, however, and is hardly noticeable. It is fortunate that large meteorite falls are rather rare events, because these fast-moving rock fragments can cause substantial damage when they strike the ground (Figure 15-14). Because they are so rare, meteorites are prized by scientists and collectors alike. Almost all meteorites are asteroid fragments and thus provide information about the chemical composition of asteroids. A very few meteorites have been identified as pieces of the Moon or Mars that were blasted off the surfaces of those worlds by asteroid impacts and eventually landed on the Earth.

Studying meteorites reveals the properties of the asteroids from which they came

People have been finding specimens of meteorites for thousands of years, and descriptions of them appear in ancient Chinese, Greek, and Roman literature. Many civilizations have regarded meteorites as objects of veneration. The sacred Black Stone, enshrined at the Ka'aba in the Great Mosque at Mecca, may be a relic of a meteorite impact.

The extraterrestrial origin of meteorites was hotly debated until as late as the eighteenth century. Upon hearing a lecture by two Yale professors, President Thomas Jefferson is said to have remarked, "I could more easily believe that two Yankee professors could lie than that stones could fall from Heaven." Although several falling meteorites had been widely witnessed and specimens from them had been collected, many scientists were reluctant to accept the idea that rocks could fall to the Earth from outer space. Conclusive evidence for the extraterrestrial origin of meteorites came on April 26, 1803, when fragments pelted the French town of L'Aigle. This event was analyzed by the noted physicist Jean-Baptiste Biot, and his conclusions helped finally to convince scientists that meteorites were indeed extraterrestrial.

Types of Meteorites

Meteorites are classified into three broad categories: stones, stony irons, and irons. As their name suggests, **stony meteorites,** or **stones,** look like ordinary rocks at first glance, but they are sometimes covered with a **fusion crust** (Figure 15-15*a*). This crust is produced by the melting of the meteorite's outer layers during its fiery descent through the atmosphere. When a stony meteorite is cut in two and polished, tiny flecks of iron can sometimes be found in the rock (Figure 15-15*b*).

Although stony meteorites account for about 95% of all the meteoritic material that falls to the Earth, they are the most difficult meteorite specimens to find. If they lie undiscovered and become exposed to the weather for a few years, they look almost indistinguishable from common terrestrial rocks. (The asteroid that caused the Tunguska explosion, described in Section 15-4, was probably of stony composition—which would account for the lack of obvious meteoritic debris on the ground.) Meteorites with high iron content can be found much more easily, because they can be located with a metal detector. Consequently, iron and stony iron meteorites dominate most museum collections.

Figure 15-14 R I V U X G

A Meteorite "Fender-Bender" On the evening of October 9, 1992, Michelle Knapp of Peekskill, New York, heard a noise from outside that sounded like a car crash. She discovered that the trunk of her car had been smashed by a 12-kilogram (27-pound) meteorite, which was lying beside the car and was still warm to the touch. Meteorite damage is not covered by automobile insurance, but Ms. Knapp was offered several tens of thousands of dollars for the meteorite (and the car). (R. A. Langheinrich Meteorites, Ilion, N.Y.)

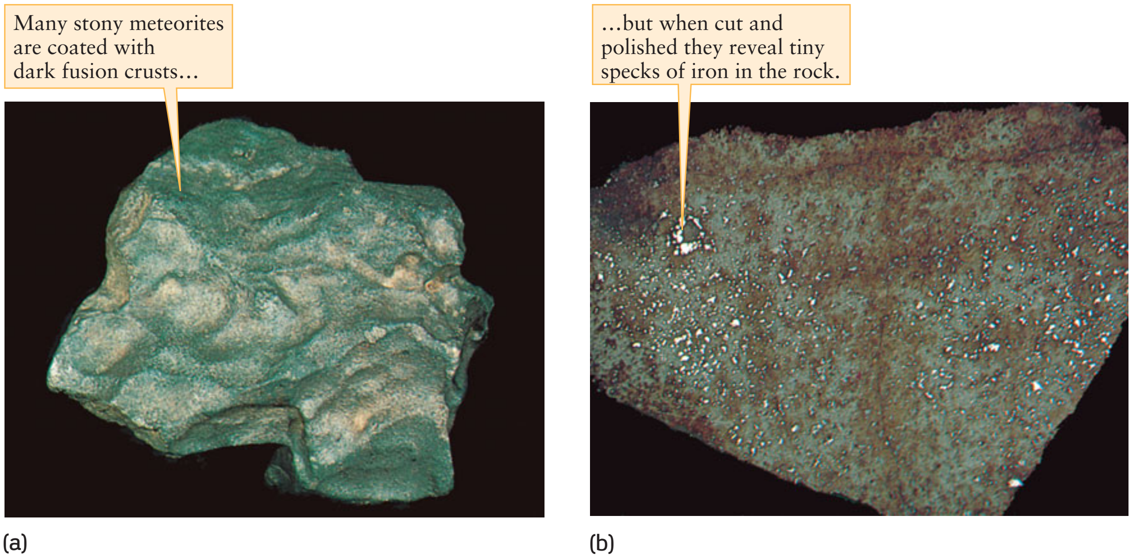


Figure 15-15 RIVUXG

Stony Meteorites Of all meteorites that fall on the Earth, 95% are stones. **(a)** Many freshly discovered specimens, like the one shown here, are coated with dark fusion crusts. This particular stone fell in Texas. **(b)** Some stony meteorites are found to contain tiny specks of iron mixed in the rock. This specimen was discovered in California. (From the collection of R. A. Oriti)

Stony iron meteorites consist of roughly equal amounts of rock and iron. Figure 15-16, for example, shows the mineral olivine suspended in a matrix of iron. Only about 1% of the meteorites that fall to the Earth are stony irons.

Iron meteorites (Figure 15-17), or **irons,** account for about 4% of the material that falls on the Earth. Iron meteorites usually contain no stone, but many contain from 10% to 20% nickel. Before humans began extracting iron from ores around 2000 B.C., the only source of the metal was from iron meteorites. Because these are so rare, iron was regarded as a precious metal like gold or silver.

Figure 15-16 RIVUXG

A Stony Iron Meteorite Stony irons account for about 1% of all the meteorites that fall to the Earth. This particular specimen, a variety of stony iron called a pallasite, fell in Chile. (Chip Clark)

In 1808, Count Alois von Widmanstätten, director of the Imperial Porcelain Works in Vienna, discovered a conclusive test for the most common type of iron meteorite. About 75% of all iron meteorites have a unique crystalline structure. These **Widmanstätten patterns** (pronounced VIT-mahn-shtetten) become visible when the meteorite is cut, polished, and briefly dipped into a dilute solution of acid (Figure 15-17*b*). Nickel-iron crystals of this type can form only if the molten metal cools slowly over many millions of years. Therefore, Widmanstätten patterns are never found in counterfeit meteorites, or "meteorwrongs."

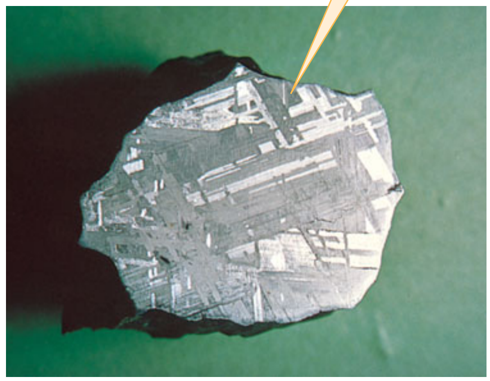
Meteorites and the Early History of Asteroids

Widmanstätten patterns suggest that some asteroids were originally at least partly molten inside and remained partly molten long after they were formed. As soon as an asteroid accreted from planetesimals some 4.56 billion years ago, rapid decay of short-lived radioactive isotopes could have heated the asteroid's interior to temperatures above the melting point of rock. If the asteroid was large enough—at least 200 to 400 km in diameter—to insulate the interior and prevent cooling, the interior would have remained molten over the next few million years and chemical differentiation would have occurred (see Box 7-1 and Section 8-5). Iron and nickel would have sunk toward the asteroid's center, forcing less dense rock upward toward the asteroid's surface.

Like a terrestrial planet, such a **differentiated asteroid** would have a distinct core and crust. After a differentiated asteroid cooled and its core solidified, collisions with other asteroids would have fragmented the parent body into meteoroids. Iron meteorites are therefore specimens from a differentiated asteroid's core, while stony irons come from regions between a differentiated asteroid's core and its crust.

Unlike irons and stony irons, stony meteorites may or may not come from differentiated asteroids. Smaller asteroids would

(a)



(b)

Figure 15-17 R I V U X G

Iron Meteorites **(a)** The surface of a typical iron is covered with small depressions caused by ablation (removal by melting) during its high-speed descent through the atmosphere. Note the one-cent coin for scale. **(b)** When cut, polished, and etched with a weak acid solution, most iron meteorites exhibit interlocking crystals in designs called Widmanstätten patterns. Both of these iron meteorites were found in Australia. (From the collection of R. A. Oriti)

not have retained their internal heat long enough for chemical differentiation to occur. Some stony meteorites are fragments of such **undifferentiated asteroids** (also called *primitive* asteroids), while others are relics of the crust of differentiated asteroids.

15-6 Some meteorites retain traces of the early solar system

A class of rare stony meteorites called **carbonaceous chondrites** shows no evidence of ever having been melted as part of asteroids. As suggested by their name, carbonaceous chondrites contain substantial amounts of carbon and carbon compounds, including complex organic molecules and as much as 20% water bound into the minerals. These compounds would have been broken down and the water driven out if these meteorites had been subjected to heating and melting. Carbonaceous chondrites may therefore be samples of the original material from which our solar system was created. The asteroid 253 Mathilde, shown in Figure 15-7, has a very dark gray color and the same sort of spectrum as a carbonaceous chondrite. It, too, is likely composed of material that predates the formation of the solar system.

Amino acids, the building blocks of proteins upon which terrestrial life is based, are among the organic compounds occasionally found inside carbonaceous chondrites. Interstellar organic material has thus probably been falling on our planet since its formation. Some scientists suspect that carbonaceous chondrites may have played a role in the origin of life on Earth.

Shortly after midnight on February 8, 1969, the night sky around Chihuahua, Mexico, was illuminated by a brilliant blue-white light moving across the heavens. As the light crossed the sky, it disintegrated in a spectacular, noisy explosion that dropped thousands of rocks and pebbles over terrified onlookers. Within hours, teams of scientists were on their way to collect specimens of a carbonaceous chondrite, collectively named the Allende meteorite after the locality (Figure 15-18).

Specimens were scattered in an elongated ellipse approximately 50 km long by 10 km wide. Most fragments were coated with a fusion crust, but minerals immediately beneath the crust showed no signs of damage. Surface material is peeled away as it becomes heated during flight through the atmosphere; this process forms the fusion crust. Because the heat has little time to penetrate the meteorite's interior, compounds there are left intact.

One of the most striking discoveries to come from the Allende meteorite was evidence suggesting that a nearby star exploded into a **supernova** about 4.6 billion years ago. In a supernova, a massive star reaches the end of its life cycle and blows itself apart in a cataclysm that hurls matter outward at tremendous speeds (see Figure 1-8). During this detonation, violent collisions between nuclei produce a host of radioactive elements, including ^{26}Al, a radioactive isotope of aluminum.

Figure 15-18 R I V U X G

A Piece of the Allende Meteorite This carbonaceous chondrite fell near Chihuahua, Mexico, in February 1969. The meteorite's dark color is due to its high abundance of carbon. Radioactive age-dating indicates that this meteorite is 4.56 billion years old, suggesting that this meteorite is a specimen of primitive planetary material that predates the formation of the planets. The ruler is 15 cm (6 in.) long. (Courtesy of J. A. Wood)

Researchers found clear evidence for the former presence of ^{26}Al in the Allende meteorite. Chemical analyses revealed a high abundance of a stable isotope of magnesium (^{26}Mg), which is produced by the radioactive decay of ^{26}Al. Some astronomers interpret this as evidence for a supernova in our vicinity at about the time the Sun was born. Indeed, by compressing interstellar gas and dust, the supernova's shock wave may have triggered the birth of our solar system.

15-7 A comet is a chunk of ice and dust that partially vaporizes as it passes near the sun

Just as heat from the protosun produced two classes of planets, the terrestrial and the Jovian, two main types of small bodies formed in the solar system. Near the Sun, interplanetary debris consists of the rocky objects called asteroids or meteoroids. Far from the Sun, where temperatures in the early solar system were low enough to permit ices of water, methane, ammonia, and carbon dioxide to form, interplanetary debris took the form of loose collections of ices and small rocky particles. These objects are like icy versions of "rubble pile" asteroids such as 253 Mathilde, which we described in Section 15-3. The Harvard astronomer Fred Whipple, who devised this model in 1950, called them "icy conglomerates" or "dirty snowballs." Some of these "snowballs" are trans-Neptunian objects that orbit forever in the outer regions of the solar system. Some of them, however, are in noncircular orbits that take them inward toward the Sun, where the ice begins to evaporate. These objects are known as **comets.**

The Structure of a Comet

WEB LINK 15.13

Kuiper belt objects and asteroids travel around the Sun along roughly circular orbits that lie close to the plane of the ecliptic. In sharp contrast, comets travel around the Sun along highly elliptical orbits inclined at random angles to the ecliptic. Comets are just a few kilometers across, so when they are far from the Sun they are difficult to spot with even a large telescope. As a comet approaches the Sun, however, solar heat begins to vaporize the comet's ices, liberating gases as well as dust particles. The liberated gases begin to glow, producing a fuzzy, luminous ball called a **coma** that is typically 1 million km in diameter. Some of these luminous gases stream outward into a long, flowing **tail.** This tail, which can be more than 100 million kilometers in length—comparable to the distance from the Earth to the Sun—is one of the most awesome sights that can be seen in the night sky (see Figure 15-19 and the image that opens this chapter).

LOOKING DEEPER 15.2

Many comets are discovered when they are still far from the Sun, long before their tails grow to full size. A number of these discoveries are made by amateur astronomers, many of whom use only binoculars or small telescopes to aid in their search.

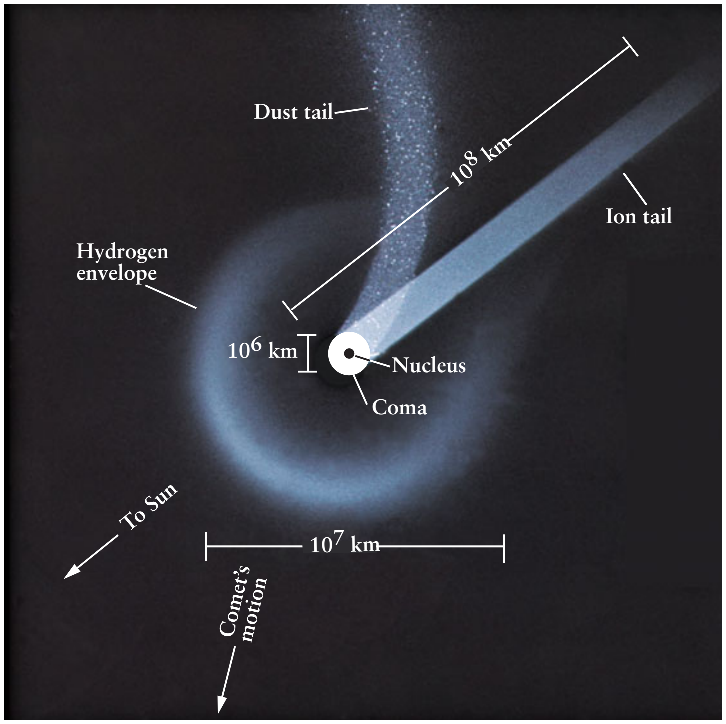
WEB LINK 15.15

Figure 15-20 depicts the structure of a comet. The solid part of the comet, from which the coma and tail emanate, is called the **nucleus.** It is a mixture of ice

WEB LINK 15.14

Figure 15-19 R I **V** U X G

Comet Hyakutake Using binoculars, Japanese amateur astronomer Yuji Hyakutake first noticed this comet on the morning of January 30, 1996. Two months later, Comet Hyakutake came within 0.1 AU (15 million km, or 9 million mi) of Earth and became one of the brightest comets of the twentieth century. This image was made in April 1996, when the comet's tail extended more than 30° across the sky (more than 60 times the diameter of the full moon). The stars appear as short streaks because the camera followed the comet's motion during the exposure. (Courtesy of Gary Goodman)



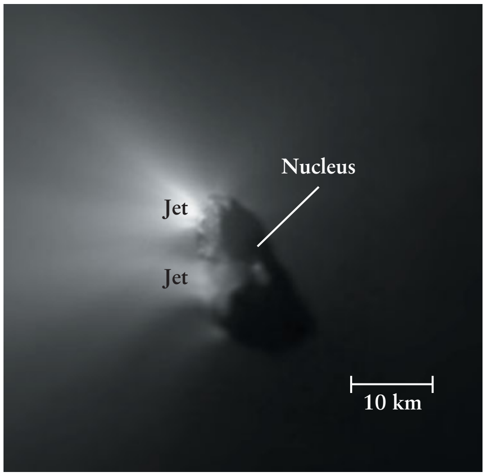
INTERACTIVE EXERCISE 15.2

Figure 15-20

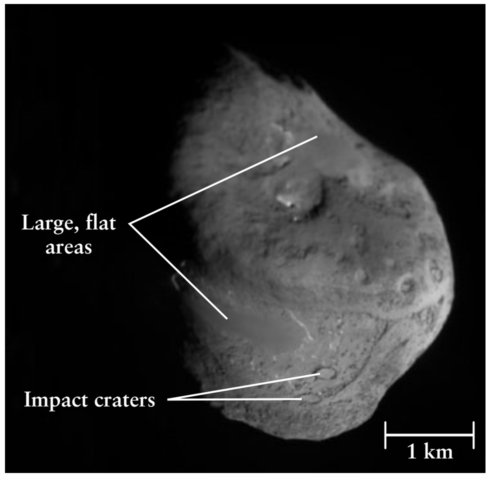
The Structure of a Comet The solid part of a comet is the nucleus. It is typically about 10 km in diameter. The coma typically measures 1 million kilometers in diameter, and the hydrogen envelope usually spans 10 million kilometers. A comet's tail can be as long as 1 AU.

Figure 15-21 R I V U X G

Comet Nuclei (a) This image from the European Space Agency spacecraft *Giotto* shows the dark, potato-shaped nucleus of Comet Halley in March 1986. Sunlight from the left illuminates half of the nucleus, which is about 15 km long and 8 km wide. Two bright jets of dust extend from the nucleus toward the Sun. **(b)** The *Deep Impact* spacecraft recorded this image of the nucleus of Comet Tempel 1 in July 2005. This nucleus, which measures 7.6 by 4.9 km, has a variety of different terrain on its surface. (a: Max-Planck-Institut für Aeronomie; b: NASA/JPL/U. of Maryland)



(a) Comet Halley



(b) Comet Tempel 1

and dust that typically measures a few kilometers across. Before 1986, no one had seen the nucleus of a comet. It is small, dim, and buried in the glare of the coma. In that year, the first close-up pictures of a comet's nucleus were obtained by a spacecraft that flew past Comet Halley (Figure 15-21*a*). Comet Halley's potato-shaped nucleus is darker than coal, reflecting only about 4% of the light that strikes it. This dark color is probably caused by a layer of carbon-rich compounds and dust that is left behind as the comet's ice evaporates.

Several 15-km-long jets of dust particles were seen emanating from bright areas on Halley's nucleus. These seem to be active only when exposed to the Sun. The bright areas, which probably cover about 10% of the surface of the nucleus, are presumably places where the dark layer covering the nucleus is particularly thin. When exposed to the Sun, these areas evaporate rapidly, producing the jets. When the nucleus's rotation brings the bright areas into darkness, away from the Sun, the jets shut off.

VIDEO 15.2

A few other comet nuclei have been imaged at close range at spacecraft. The best images to date have come from the *Deep Impact* spacecraft, which flew to within 500 km of Comet Tempel 1 in 2005 (Figure 15-21*b*). As it approached, *Deep Impact* launched a 372-kg copper projectile that slammed into the surface of the nucleus at 37,000 km/h (23,000 mi/h). The collision excavated about 11,000 tons of material from the nucleus. The character of the ejected material confirmed that the nucleus of Tempel 1 is a "rubble pile" held together not by chemical forces between its component parts (like a chocolate chip cookie) but by the gravitational forces between those parts.

The *Deep Impact* mission also measured the mass and hence the density of the comet nucleus. The very low density of 600 kg/m^3 (60% of the density of water) shows that the nucleus cannot be solid. It is best regarded as a porous jumble of rock, fine dust, and ice.

An even more ambitious mission to explore a comet is the European Space Agency's *Rosetta*. This spacecraft was lunched in 2004 and is scheduled to fly past at least one asteroid on its way to intercept Comet 67P/Churyumov-Gerasimenko in 2014. If all goes well, *Rosetta* will go into orbit around the comet nucleus and release a small probe that will actually land on the nucleus.

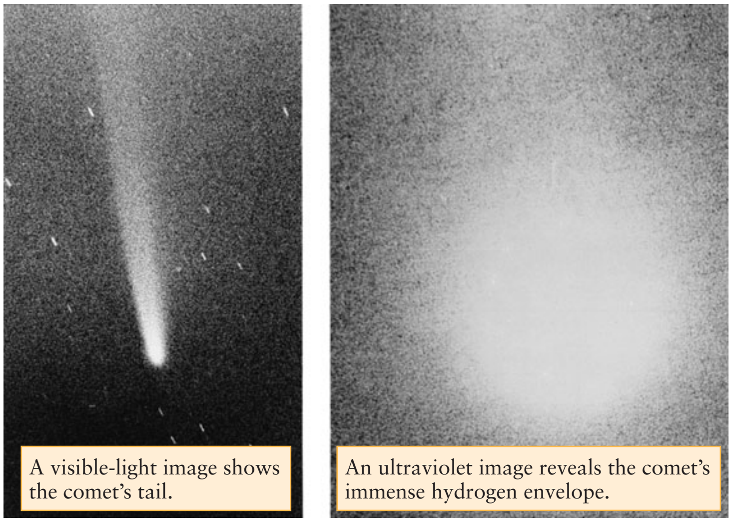
Comet Envelopes and Comet Tails

A part of a comet that is not visible to the human eye is the **hydrogen envelope,** a huge sphere of tenuous gas surrounding the nucleus. This hydrogen comes from water molecules (H_2O) that escape from the comet's evaporating ice and then break apart when they absorb ultraviolet photons from the Sun. The hydrogen atoms also absorb solar ultraviolet photons, which excite the atoms from the ground state into an excited state (see Section 5-8). When the atoms return to the ground state, they emit ultraviolet photons.

Unfortunately for astronomers, these emissions from a comet's hydrogen envelope cannot be detected by Earth-based telescopes because our atmosphere is largely opaque to ultraviolet light (see Section 6-7). Instead, cameras above the Earth's atmosphere must be used. Figure 15-22 shows two views of the same comet, one as it appeared in visible light to Earth-based observers and one as photographed by an ultraviolet camera aboard a rocket. The ultraviolet image shows the enormous extent of the hydrogen envelope, which can span 10 million kilometers.

As the diagram in Figure 15-20 suggests, comet tails always point away from the Sun. This is true regardless of the direction of the comet's motion (Figure 15-23). The implication that something from the Sun was "blowing" the comet's gases radially outward led Ludwig Biermann to predict the existence of a *solar wind,* a stream of particles rushing away from the Sun (see Section 8-5). A decade later, in 1962, Biermann's prediction was confirmed when the solar wind was detected for the first time by instruments on the *Mariner 2* spacecraft (see Figure 11-15).

A comet's tail does not stream behind it; instead, it points generally away from the Sun



R I V U X G R I V U X G

Figure 15-22

A Comet and Its Hydrogen Envelope These visible-light and ultraviolet images of Comet Kohoutek, which made an appearance in 1974, are reproduced to the same scale. The hydrogen envelope is only visible in the ultraviolet image. (Johns Hopkins University, Naval Research Laboratory)

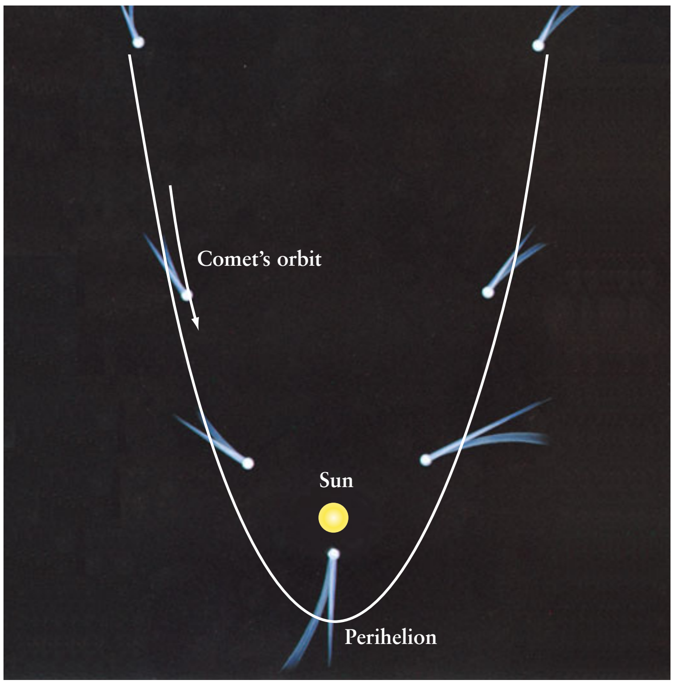
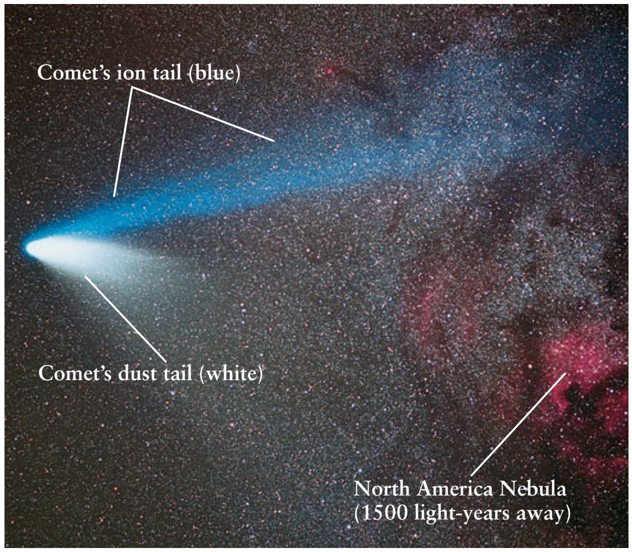


Figure 15-23

The Orbit and Tail of a Comet The solar wind and the pressure of sunlight blow a comet's dust particles and ionized atoms away from the Sun. Consequently, a comet's tail points generally away from the Sun. In particular, the tail does *not* always stream behind the nucleus. At the upper right of this figure, the comet is moving upward along its orbit and is literally chasing its own tail.



WEB LINK 15.16

Figure 15-24 R I V U X G

The Two Tails of Comet Hale-Bopp A comet's dust tail is white because it reflects sunlight, while the molecules in the ion tail emit their own light with a characteristic blue color. When this picture was taken on March 8, 1997, the ion tail extended more than 10° across the sky. The red object to the right is the North America Nebula, a star-forming region some 1500 light-years beyond the solar system. (Courtesy of Tony and Daphne Hallas Astrophotos)

In fact, the Sun usually produces *two* comet tails—an **ion tail** and a **dust tail** (Figure 15-24). Ionized atoms and molecules—that is, atoms and molecules missing one or more electrons—are swept directly away from the Sun by the solar wind to form the relatively straight ion tail. The distinct blue color of the ion tail is caused by emissions from carbon-bearing molecules such as CN and C_2. The dust tail is formed when photons of light strike dust particles freed from the evaporating nucleus. (Figure 8-6 shows a dust particle of this type.) Light exerts a pressure on any object that absorbs or reflects it. This pressure, called **radiation pressure,** is quite weak but is strong enough to make fine-grained dust particles in a comet's coma drift away from the comet, thus producing a dust tail. The solar wind has less of an effect on dust particles than on (much smaller) ions, so the dust tail ends up being curved rather than straight. On rare occasions a comet is oriented in such a way that its dust tail appears to stick out in front of the comet (Figure 15-25).

After the comet passes perihelion, it recedes from the Sun back into the cold regions of the outer solar system. The ices stop vaporizing, the coma and tail disappear, and the comet goes back into an inert state—until the next time its orbit takes it toward the inner solar system. An object in an elliptical orbit spends most of its time far from the Sun, so it is only during a relatively brief period before and after perihelion that a comet can have a prominent tail. An example is Comet Halley, which orbits the Sun along a highly elliptical path that stretches from just inside the Earth's orbit to slightly beyond the orbit of Neptune (Figure 15-26). The comet's tail is visible to the naked eye or with binoc-

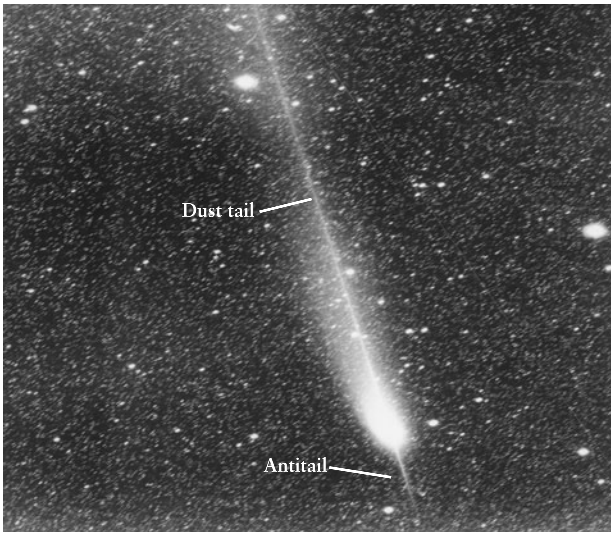
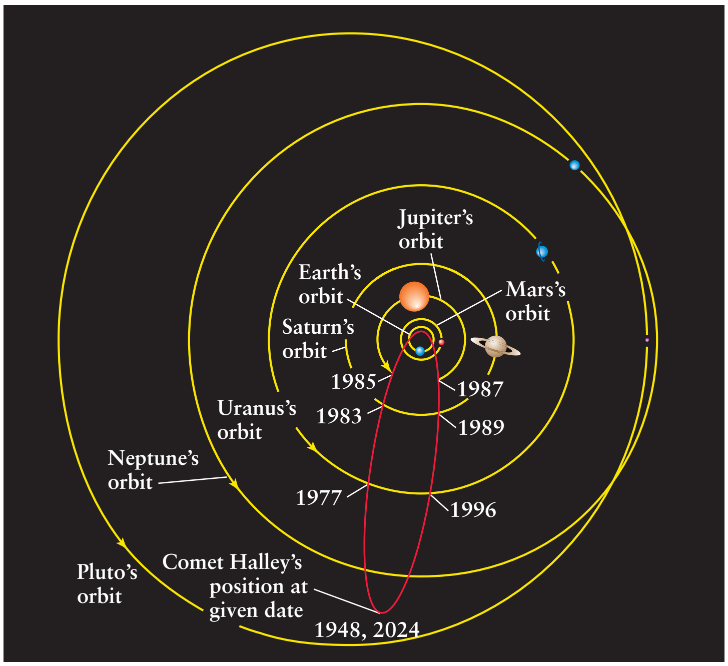


Figure 15-25 R I V U X G

The Antitail of Comet Hale-Bopp In January 1998, nine months after passing perihelion, Comet Hale-Bopp was oriented in such a way that the end of its arched dust tail looked like a spike sticking out of the comet's head. This spike is called an "antitail" because it appears to protrude in front of the comet. (European Southern Observatory)



ANIMATION 15.1

Figure 15-26

Comet Halley's Eccentric Orbit Like most comets, Comet Halley has an elongated orbit and spends most of its time far from the Sun. Figure 4-23 shows the comet's tail, which appears only when the comet is close to perihelion. Comet Halley has been observed at intervals of about 76 years—the period of its orbit—since 88 B.C. (see Section 4-7).

ulars only during a few months around perihelion, which last occurred in 1986 and will happen again in 2061.

Other comets have orbits that are larger and more elongated, with even longer periods. The orbit of Comet Hyakutake (see Figure 15-19) takes it to a distance of about 2000 AU from the Sun, 50 times the size of Pluto's orbit, with an orbital period of around 30,000 years!

15-8 Comets originate either from the Kuiper belt or from a vast cloud in near-interstellar space

Both comets and Kuiper belt objects are mixtures of ice and rock that must have formed in the outer solar system. (Had they formed in the warm inner solar system, they would not have included ices.) But most Kuiper belt objects have relatively circular orbits that lie close to the plane of the ecliptic, while comets have very elongated orbits that are often steeply inclined to the ecliptic. What gives comets their distinctive orbits? As we will see, the answer to this question depends on whether the comet's orbital period is relatively short or relatively long.

Jupiter-Family Comets and the Kuiper Belt

As we saw in Section 14-10, the gravitational effects of Neptune shape the orbits of objects in the **Kuiper belt.** Most of these objects are found between the orbit of Pluto (where an object makes two orbits for every three orbits of Neptune) and about 50 AU from the Sun (where an object makes one orbit for every two orbits of Neptune).

However, collisions between Kuiper belt objects can break off relatively small chunks. Gravitational perturbations from Neptune can occasionally launch one of these chunks into a highly elliptical orbit that takes it close to the Sun. When one of these chunks approaches the Sun, it develops a visible tail and appears as a comet. These are called **Jupiter-family comets** because their orbits tend to be influenced strongly by Jupiter's gravitational pull. Jupiter-family comets orbit the Sun in fewer than 20 years. One example is Comet Tempel 1 (see Figure 15-21*b*), which has an orbital period of 5.5 years.

There may be tens of thousands of objects in the Kuiper belt that could eventually be perturbed into an elongated orbit and become comets. It is thought that a few of these refugees from the Kuiper belt have been captured by the gravitational pull of Saturn, becoming that planet's small outer satellites (see Section 13-10).

Intermediate-Period Comets, Long-Period Comets, and the Oort Cloud

The majority of comets are **intermediate-period comets,** with orbital periods between 20 and 200 years, and **long-period comets,** which take more than 200 years to complete one orbit of the Sun. Comet Halley, shown in Figure 15-21*a,* has an orbital period of 76 years and so is an intermediate-period comet; Comet Hyakutake (see Figure 15-19) is a long-period comet that takes more than 70,000 years to complete one orbit.

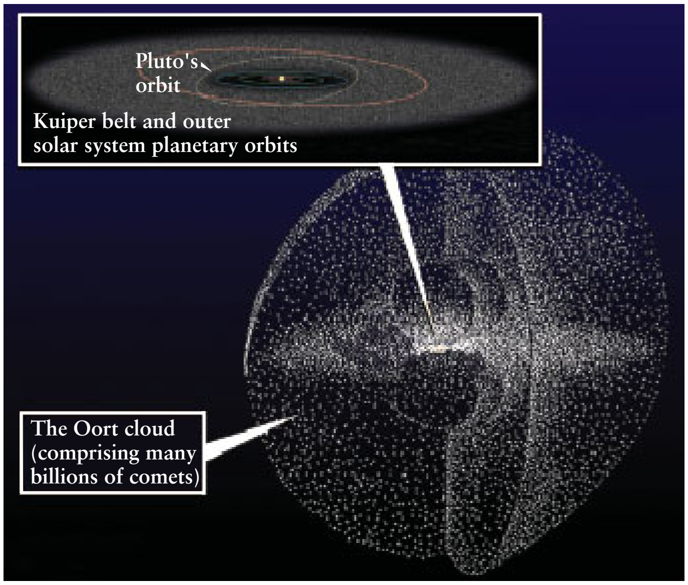


Figure 15-27

The Kuiper Belt and the Oort Cloud Objects in the Kuiper belt orbit close to the plane of the ecliptic. The Oort cloud, by contrast, is a spherical distribution of icy objects whose orbits are randomly oriented relative to the ecliptic. (NASA and A. Field/Space Telescope Science Institute)

Long-period comets travel along extremely elongated orbits and consequently spend most of their time at distances of roughly 10^4 to 10^5 AU from the Sun, or about one-fifth of the way to the nearest star. These orbits extend far beyond the Kuiper belt. Furthermore, intermediate-period and long-period comets have orbital planes that are often steeply inclined to the plane of the ecliptic. (Comet Halley's orbital plane is inclined by 162° to the ecliptic, which means that it orbits the Sun in the opposite direction to the planets.)

The best explanation for these observations is that intermediate-period and long-period comets come from a reservoir that extends from the Kuiper belt to some 50,000 AU from the Sun. This reservoir, first hypothesized by Dutch astronomer Jan Oort in 1950 and now called the **Oort cloud,** does not lie in or near the ecliptic; rather, it is a spherical distribution centered on the Sun (Figure 15-27). This explains why many intermediate-period and long-period comets have steeply inclined orbits.

Because astronomers discover long-period comets at the rate of about one per month, it is reasonable to suppose that there is an enormous population of comets in the Oort cloud. Estimates of the number of "dirty snowballs" in the Oort cloud range as high as 5 trillion (5×10^{12}). Only such a large reservoir of comet nuclei would explain why we see so many long-period comets, even though each one takes several million years to travel once around its orbit. Because the Oort cloud is so distant, it has not yet been possible to detect objects in the Oort cloud directly.

The Oort cloud was probably created 4.56 billion years ago from numerous icy planetesimals that orbited the Sun in the vicinity of the newly formed Jovian planets. During near-collisions with the giant planets, many of these chunks of ice and dust were catapulted by gravity into highly elliptical orbits, in much the same way that *Pioneer* and *Voyager* spacecraft were flung far from the Sun during their flybys of the Jovian planets. Gravitational perturbations from nearby stars could have tilted the planes of the orbits in all directions, giving the Oort cloud its spherical shape.

Changing a Comet's Orbit

The distinction between long-period, intermediate-period, and Jupiter-family comets can be blurred by the effects of gravitational perturbations. During a return trip toward the Sun, an encounter with a Jovian planet may force a comet into a much larger orbit. Alternatively, such an encounter can move a long-period comet into a smaller orbit (Figure 15-28). Several comets have been perturbed in this way into orbits that always remain within the inner solar system. One such is Comet Tempel 1 (see Figure 15-21*b*), whose orbit extends from the orbit of Mars out to the orbit of Jupiter.

Our description of cometary origins suggests that different comets are of different ages. Computer simulations suggest that many Jupiter-family comets were produced no more than a few hundred million years ago by collisions between larger icy objects in the Kuiper belt. Comets emanating from the Oort cloud may

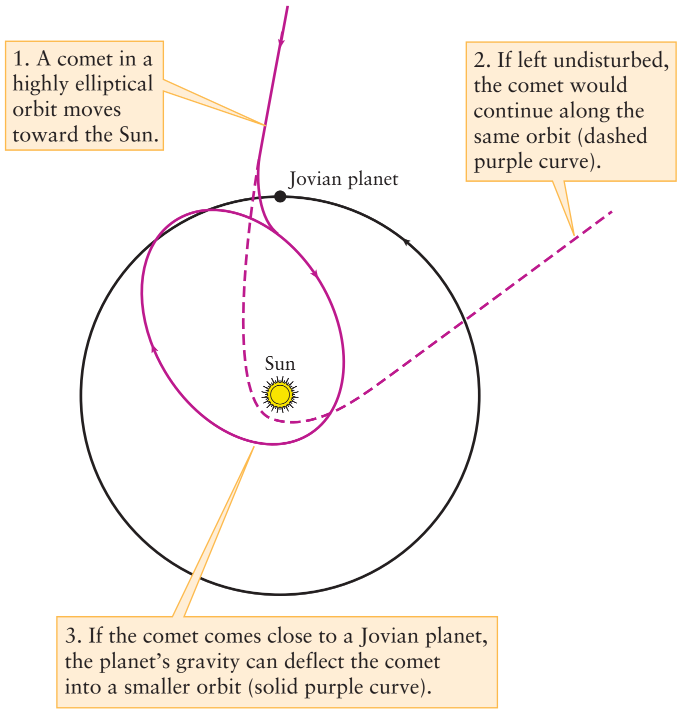


Figure 15-28

Transforming a Comet's Orbit The gravitational force of a planet can deflect a comet from a highly elliptical orbit into a less elliptical one. In some cases a comet ends up with an orbit that keeps it within the inner solar system.

Figure 15-29 R I V U X G

The Fragmentation of Comet LINEAR The Lincoln Near Earth Asteroid Research (LINEAR) project discovered this comet in 1999. As it passed the perihelion of its orbit in July 2000, Comet LINEAR broke into more than a dozen small fragments. These continue to orbit the Sun along the same trajectory that the comet followed prior to its breakup. This image was made using the Very Large Telescope (see Figure 6-18). (European Southern Observatory)

date from the very early solar system, but may have originated at somewhat different times at different distances from the center of the solar nebula before being catapulted into the Oort cloud. This variety of ages is one reason why comets are of particular interest to planetary scientists: Comets carry information about the history of the cold, outer regions of the solar system, but are brought into the inner solar system where we can examine them at closer range using telescopes or spacecraft.

Comet Breakup and Meteor Showers

Because they evaporate away part of their mass each time they pass near the Sun, comets cannot last forever. A typical comet may lose about 0.5% to 1% of its ice each time it passes near the Sun. Hence, the ice completely vaporizes after about 100 or 200 perihelion passages, leaving only a swarm of dust and pebbles. Astronomers have observed some comet nuclei in the process of fragmenting (Figure 15-29).

Comets eventually break apart, and their fragments give rise to meteor showers

WEB LINK 15.17

One remarkable comet, called Shoemaker-Levy 9, broke apart for a different reason. As the comet swung by Jupiter in July 1992, the giant planet's tidal forces tore the nucleus into more than 20 fragments. Jupiter's gravity then deflected the trajectories of these fragments so that they plummeted into the planet's atmosphere.

As a comet's nucleus evaporates, residual dust and rock fragments form a **meteoritic swarm,** a loose collection of debris that continues to circle the Sun along the comet's orbit (Figure 15-30). If the Earth's orbit happens to pass through this swarm, a **meteor shower** is seen as the dust particles strike the Earth's upper atmosphere.

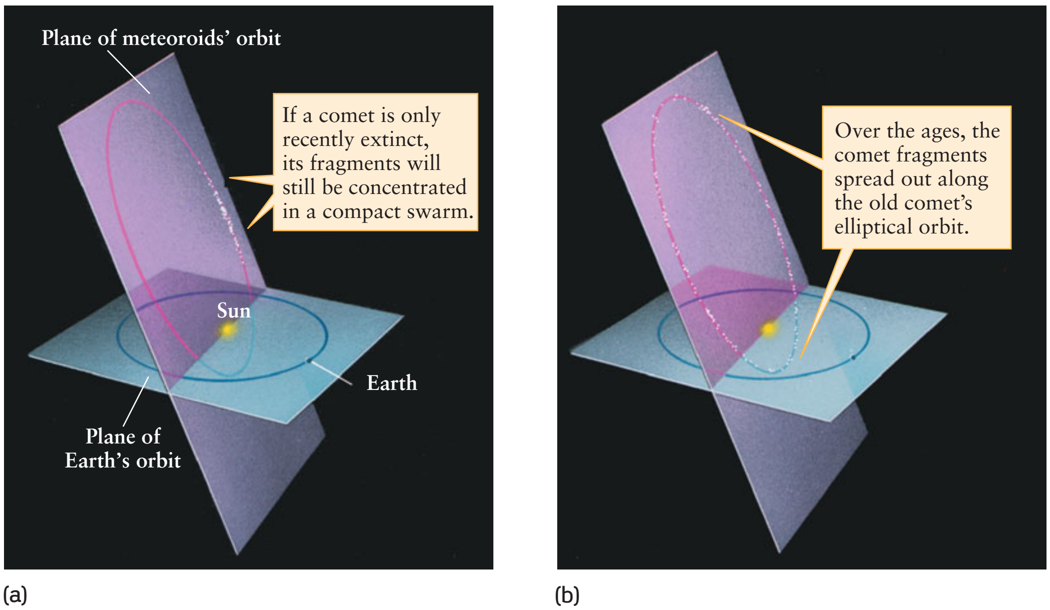


Figure 15-30

Meteoritic Swarms Rock fragments and dust from "burned out" comets continue to circle the Sun. **(a)** The most spectacular meteor showers occur when the Earth passes through compact swarms of debris from recently extinct comets. **(b)** Meteor showers associated with older comets are more predictable, because the Earth passes through the evenly distributed swarm on each trip around the Sun.

LOOKING DEEPER 15.3

Table 15-1 Prominent Yearly Meteor Showers

Shower name	Date of maximum intensity*	Typical hourly rate	Average speed (km/s)	Radiant constellation
Quadrantids	January 3	40	40	Boötes
Lyrids	April 22	15	50	Lyra
Eta Aquarids	May 4	20	64	Aquarius
Delta Aquarids	July 30	20	40	Aquarius
Perseids	August 12	50	60	Perseus
Orionids	October 21	20	66	Orion
Taurids	November 4	15	30	Taurus
Leonids	November 16	15	70	Leo
Geminids	December 13	50	35	Gemini
Ursids	December 22	15	35	Ursa Minor

**The date of maximum intensity is the best time to observe a particular shower, although good displays can often be seen a day or two before or after the maximum. The typical hourly rate is given for an observer under optimum viewing conditions. The average speed refers to how fast the meteoroids are moving when they strike the atmosphere.*

Nearly a dozen meteor showers can be seen each year (Table 15-1). Note that the **radiants** for these showers—that is, the places among the stars from which the meteors appear to come—are not confined to the constellations of the zodiac. This means that the swarms do not all orbit in the plane of the ecliptic. The reason is that the orbits of their parent comets were inclined at random angles to the plane of the ecliptic (see Figure 15-30).

Key Words

amino acids, p. 389
asteroid, p. 378
asteroid belt, p. 379
carbonaceous chondrite, p. 389
coma (of a comet), p. 390
comet, p. 390
differentiated asteroid, p. 388
dust tail, p. 392
dwarf planet, p. 379
fusion crust, p. 387
Hirayama family, p. 383
hydrogen envelope, p. 391
intermediate-period comet, p. 393
iron meteorite (iron), p. 388
ion tail, p. 392
Jupiter-family comet, p. 393
Kirkwood gaps, p. 380
Kuiper belt, p. 393
long-period comet, p. 393
meteor, p. 387
meteor shower, p. 395
meteorite, p. 387
meteoritic swarm, p. 395
meteoroid, p. 384
minor planet, p. 378
near-Earth object (NEO), p. 383
nucleus (of a comet), p. 390
Oort cloud, p. 394
radiant (of a meteor shower), p. 396
radiation pressure, p. 392
stable Lagrange points, p. 383
stony iron meteorite, p. 388
stony meteorite (stone), p. 387
supernova, p. 389
tail (of a comet), p. 390
Trojan asteroid, p. 383
undifferentiated asteroid, p. 389
Widmanstätten patterns, p. 388

Key Ideas

Discovery of the Asteroids: Astronomers first discovered the asteroids while searching for a "missing planet."

• Thousands of asteroids with diameters ranging from a few kilometers up to 1000 kilometers orbit within the asteroid belt between the orbits of Mars and Jupiter.

Origin of the Asteroids: The asteroids are the relics of planetesimals that failed to accrete into a full-sized planet, thanks to the effects of Jupiter and other Mars-sized objects.

• Even today, gravitational perturbations by Jupiter deplete certain orbits within the asteroid belt. The resulting gaps, called Kirkwood gaps, occur at simple fractions of Jupiter's orbital period.

• Jupiter's gravity also captures asteroids in two locations, called Lagrangian points, along Jupiter's orbit.

Asteroid Collisions: Asteroids undergo collisions with each other, causing them to break up into smaller fragments.

• Some asteroids, called near-Earth objects, move in elliptical orbits that cross the orbits of Mars and Earth. If such an asteroid strikes the Earth, it forms an impact crater whose diameter depends on both the mass and the speed of the asteroid.

• An asteroid struck the Earth 65 million years ago, possibly causing the extinction of the dinosaurs and many other species.

Meteoroids, Meteors, and Meteorites: Small rocks in space are called meteoroids. If a meteoroid enters the Earth's atmosphere, it produces a fiery trail called a meteor. If part of the object survives the fall, the fragment that reaches the Earth's surface is called a meteorite.

• Meteorites are grouped into three major classes, according to composition: iron, stony iron, and stony meteorites. Irons and stony irons are fragments of the core of an asteroid that was large enough and hot enough to have undergone chemical differentiation, just like a terrestrial planet. Some stony meteorites come from the crust of such differentiated meteorites, while others are fragments of small asteroids that never underwent differentiation.

• Rare stony meteorites called carbonaceous chondrites may be relatively unmodified material from the solar nebula. These meteorites often contain organic material and may have played a role in the origin of life on Earth.

• Analysis of isotopes in certain meteorites suggests that a nearby supernova may have triggered the formation of the solar system 4.56 billion years ago.

Comets: A comet is a chunk of ice with imbedded rock fragments that generally moves in a highly elliptical orbit about the Sun.

• As a comet approaches the Sun, its icy nucleus develops a luminous coma, surrounded by a vast hydrogen envelope. An ion tail and a dust tail extend from the comet, pushed away from the Sun by the solar wind and radiation pressure, respectively.

Origin and Fate of Comets: Comets are thought to originate from two regions, the Kuiper belt and the Oort cloud.

• The Kuiper belt lies in the plane of the ecliptic at distances between 30 and 50 AU from the Sun. It is thought to contain many tens of thousands of comet nuclei. Many Jupiter-family comets probably come from the Kuiper belt, and hundreds of larger objects have been observed in the Kuiper belt.

• The Oort cloud contains billions of comet nuclei in a spherical distribution that extends out to 50,000 AU from the Sun. Intermediate-period and long-period comets are thought to originate in the Oort cloud. As yet no objects in the Oort cloud have been detected directly.

• Fragments of "burned out" comets produce meteoritic swarms. A meteor shower is seen when the Earth passes through a meteoritic swarm.

Questions

Review Questions

1. What led astronomers to suspect that there were members of the solar system that orbit between Mars and Jupiter?
2. How did the first asteroid come to be discovered? What role did theoretical calculations play in confirming the discovery? How did this discovery differ from what astronomers had expected to find?
3. How do modern astronomers discover new asteroids?
4. What are the differences between asteroids and trans-Neptunian objects?
5. Describe the asteroid belt. Does it lie completely within the plane of the ecliptic? What are its inner and outer radii?
6. If Jupiter was not present in our solar system, would the asteroid belt exist? Why or why not?
7. What are Kirkwood gaps? What causes them?
8. Compare the explanation of the Kirkwood gaps in the asteroid belt to the way in which Saturn's moons help produce divisions in that planet's rings (see Section 12-11).
9. How is it possible to tell that some asteroids are nonspherical even though we do not have images of those asteroids?
10. What is the evidence that some asteroids are made of a loose conglomeration of smaller pieces?
11. The asteroid 243 Ida, which was viewed by the *Galileo* spacecraft, is a member of a Hirayama family. Discuss what this tells us about the history of this asteroid. Where might you look to find other members of the same family?
12. What are the Trojan asteroids, and where are they located? What holds them in this location?
13. What are near-Earth objects? What is the evidence that the Earth has been struck by these objects?
14. What is the evidence that an asteroid impact could have contributed to the demise of the dinosaurs?
15. What is the difference between a meteoroid, a meteor, and a meteorite?
16. Is there anywhere on the Earth where you might find large numbers of stony meteorites that are not significantly weathered? If so, where? If not, why not?
17. Scientists can tell that certain meteorites came from the interior of an asteroid rather than from its outer layers. Explain how this is done.
18. Why are some asteroids differentiated while others are not?
19. Suppose you found a rock you suspect might be a meteorite. Describe some of the things you could do to determine whether it was a meteorite or a "meteorwrong."
20. What is the evidence that carbonaceous chondrites are essentially unaltered relics of the early solar system? What do they suggest about how the solar system may have formed?
21. With the aid of a drawing, describe the structure of a comet.
22. Why is the phrase "dirty snowball" an appropriate characterization of a comet's nucleus?
23. What did scientists learn about the structure of comet nuclei from the *Deep Impact* mission?
24. Why do the ion tail and dust tail of a comet point in different directions?
25. Why do comets have prominent tails for only a short time during each orbit?
26. Why is it that Jupiter and Saturn can be seen in the night sky every year, while seeing specific comets such as Halley and Hyakutake is a once-in-a-lifetime event?
27. What is the relationship between the Kuiper belt and comets?
28. What is the Oort cloud? How might it be related to planetesimals left over from the formation of the solar system?
29. Why are comets more likely to break apart at perihelion than at aphelion?
30. Why do astronomers think that meteorites come from asteroids, while meteor showers are related to comets?
31. Why are asteroids, meteorites, and comets all of special interest to astronomers who want to understand the early history and subsequent evolution of the solar system?

Advanced Questions

Questions preceded by an asterisk (*) involve topics discussed in Box 1-1 or 7-2.

Problem-solving tips and tools

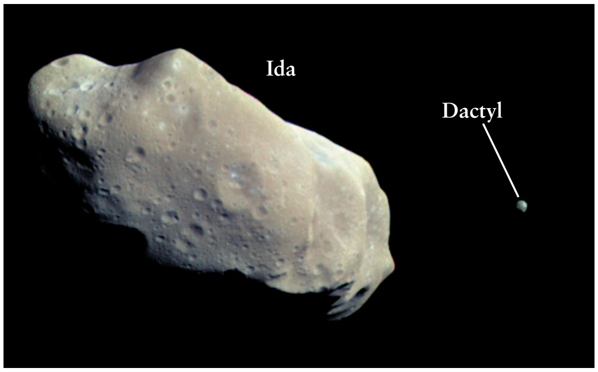
The small-angle formula is described in Box 1-1. We discussed retrograde motion in Section 4-1 and described its causes in Section 4-2. You will need to use Kepler's third law, described in Section 4-4 and Box 4-2, in some of the problems below. Box 7-2 discusses the concept of escape speed. A spherical object of radius r intercepts an amount of sunlight proportional to its cross-sectional area, equal to πr^2. The volume of a sphere of radius r is $4\pi r^3/3$.

32. When Olbers discovered Pallas in March 1802, the asteroid was moving from east to west relative to the stars. At what time of night was Pallas highest in the sky over Olbers's observatory? Explain your reasoning.

33. Consider the Kirkwood gap whose orbital period is two-fifths of Jupiter's 11.86-year period. Calculate the distance from the Sun to this gap. Does your answer agree with Figure 15-4?

***34.** When the image in Figure 15-5 was made, the asteroid 1 Ceres was 1.63 AU, or 2.44×10^8 km, from the Hubble Space Telescope. (**a**) What was the angular size of the asteroid as a whole? (**b**) You can see individual pixels in the image shown in Figure 15-5. Using a ruler and the scale bar in the figure, determine how many kilometers on the surface of 1 Ceres are contained in the width of one pixel. (**c**) What is the angular width of each pixel in arcseconds?

35. Suppose that a binary asteroid (two asteroids orbiting each other) is observed in which one member is 16 times brighter than the other. Suppose that both members have the same albedo and that the larger of the two is 120 km in diameter. What is the diameter of the other member?



R I V U X G

(JPL/NASA)

36. The accompanying image from the *Galileo* spacecraft shows the asteroid 243 Ida, which has dimensions $56 \times 24 \times 21$ km. *Galileo* discovered a tiny moon called Dactyl, just $1.6 \times 1.4 \times 1.2$ km in size, which orbits Ida at a distance of about 100 km. (In Greek mythology, the Dactyli were beings who lived on the slopes of Mount Ida.) Describe a scenario that could explain how Ida came to have a moon.

***37.** Assume that Ida's tiny moon Dactyl (see Question 36) has a density of 2500 kg/m^3. (**a**) Calculate the mass of Dactyl in kilograms. For simplicity, assume that Dactyl is a sphere 1.4 km in diameter. (**b**) Calculate the escape speed from the surface of Dactyl. If you were an astronaut standing on Dactyl's surface, could you throw a baseball straight up so that it would never come down? Professional baseball pitchers can throw at speeds around 40 m/s (140 km/h, or 90 mi/h); your throwing speed is probably a bit less.

38. Imagine that you are an astronaut standing on the surface of a Trojan asteroid. How will you see the phase of Jupiter change with the passage of time? How will you see Jupiter move relative to the distant stars? Explain your answers.

39. Use the percentages of stones, irons, and stony iron meteorites that fall to Earth to estimate what fraction of their parent asteroids' interior volume consisted of an iron core. Assume that the percentages of stones and irons that fall to the Earth indicate the fractions of a parent asteroid's interior volume occupied by rock and iron, respectively. How valid do you think this assumption is?

40. On March 8, 1997, Comet Hale-Bopp was 1.39 AU from the Earth and 1.00 AU from the Sun. Use this information and that given in the caption to Figure 15-24 to estimate the length of the comet's ion tail on that date. Give your answer in kilometers and astronomical units.

41. *Sun-grazing comets* come so close to the Sun that their perihelion distances are essentially zero. Find the orbital periods of Sun-grazing comets whose aphelion distances are (**a**) 100 AU, (**b**) 1000 AU, (**c**) 10,000 AU, and (**d**) 100,000 AU. Assuming that these comets can survive only a hundred perihelion passages, calculate their lifetimes. (*Hint:* Remember that the semimajor axis of an orbit is one-half the length of the orbit's long axis.)

42. Comets are generally brighter a few weeks after passing perihelion than a few weeks before passing perihelion. Explain why might this be. (*Hint:* Water, including water ice, does an excellent job of retaining heat.)

43. The hydrogen envelopes of comets are especially bright at an ultraviolet wavelength of 122 nm. Use Figure 5-24 to explain why.

44. A *very* crude model of a typical comet nucleus is a cube of ice (density 1000 kg/m^3) 10 km on a side. (**a**) What is the mass of this nucleus? (**b**) Suppose 1% of the mass of the nucleus evaporates away to form the comet's tail. Suppose further that the tail is 100 million (10^8) km long and 1 million (10^6) km wide. Estimate the average density of the tail (in kg/m^3). For comparison, the density of the air you breathe is about 1.2 kg/m^3. (**c**) In 1910 the Earth actually passed through the tail of Comet Halley. At the time there was some concern among the general public that this could have deleterious effects on human health. Was this concern justified? Why or why not?

45. For many years it was thought that the Tunguska event was caused by a comet striking the Earth. This idea was rejected because a small comet would have broken up too high in the atmosphere to cause significant damage on the ground. Explain why, using your knowledge of a comet's structure.

Discussion Questions

46. Discuss the idea that 1 Ceres should be regarded as the smallest dwarf planet rather than the largest asteroid. What are the advantages of this scheme? What are the disadvantages?

47. From the abundance of craters on the Moon and Mercury, we know that numerous asteroids and meteoroids struck the inner planets early in the history of our solar system. Is it reasonable to suppose that numerous comets also pelted the planets 3.5 to 4.5 billion years ago? Speculate about the effects of such a cometary bombardment, especially with regard to the evolution of the primordial atmospheres on the terrestrial planets.

48. In the 1998 movie *Armageddon,* an asteroid "the size of Texas" is on a collision course with Earth. The asteroid is first discovered by astronomers just 18 days prior to impact. To avert disaster, a team of astronauts blasts the asteroid into two pieces just 4 hours before impact. Discuss the plausibility of this scenario. (*Hint:* On average, the state of Texas extends for about 750 km from north to south and from east to west. How does this compare with the size of the largest known asteroids?)

49. Suppose astronomers discover that a near-Earth object the size of 1994 XM1 is on a collision course with Earth. Describe what humanity could do within the framework of present technology to counter such a catastrophe.

Web/eBook Questions

50. A NASA spacecraft called *Dawn* is intended to go into orbit around two asteroids, 1 Ceres and 4 Vesta. Search the World Wide Web for information about this mission. Why were these two particular asteroids selected for study? What types of observations will the spacecraft make?

51. Search the World Wide Web to find out why some scientists disagree with the idea that a tremendous impact led to the demise of the dinosaurs. (They do not dispute that the impact took place, only what its consequences were.) What are their arguments? From what you learn, what is your opinion?

52. Several scientific research programs are dedicated to the search for near-Earth objects (NEOs), especially those that might someday strike our planet. Search the World Wide Web for information about at least one of these programs. How does this program search for NEOs? How many NEOs has this program discovered? Will any of these pose a threat in the near future?

53. Estimating the Speed of a Comet. Access and view the video "Two Comets and an Active Sun" in Chapter 15 of the *Universe* Web site or eBook. (**a**) Why don't the comets reappear after passing the Sun? (**b**) The white circle shows the size of the Sun, which has the diameter 1.39×10^6 km. Using this to set the scale, step through the video and measure how the position of one of the comets changes. Use your measurements and the times displayed in the video to estimate the comet's speed in km/h and km/s. (Assume that the comet moved in the same plane as that shown in the video.) As part of your answer, explain the technique and calculations you used. (**c**) How does your answer in (b) compare with the orbital speed of Mercury, the innermost and fastest-moving planet (see Table 11-1)? Why is there a difference? (**d**) If the comet's motion was not in the same plane as that shown in the video, was its actual speed more or less than your estimate in (b)? Explain.

Activities

Observing Projects

Observing tips and tools

Meteors: Informative details concerning upcoming meteor showers appear on the Web sites for *Sky & Telescope* and *Astronomy* magazines.

Comets: Because astronomers discover dozens of comets each year, there is usually a comet visible somewhere in the sky. Unfortunately, they are often quite dim, and you will need to have access to a moderately large telescope (at least 35 cm, or 14 in.). You can get up-to-date information from the Minor Planet Center Web site. Also, if there is an especially bright comet in the sky, useful information about it might be found at the Web sites for *Sky & Telescope* and *Astronomy* magazines.

Asteroids: At opposition, some of the largest asteroids are bright enough to be seen through a modest telescope. Check the Minor Planet Center Web site to see if any bright asteroids are near opposition. If so, check the current issue as well as the most recent January issue of *Sky & Telescope* magazine for a star chart showing the asteroid's path among the constellations. You will need such a chart to distinguish the asteroid from background stars. Also, you can locate Ceres, Pallas, Juno, and Vesta using the ***Starry Night Enthusiast*™** program on the CD-ROM that accompanies certain printed copies of this textbook.

54. Make arrangements to view a meteor shower. Table 15-1 lists the dates of major meteor showers. Choose a shower that will occur near the time of a new moon. Set your alarm clock for the early morning hours (1 to 3 A.M.). Get comfortable in a reclining chair or lie on your back so that you can view a large portion of the sky. Record how long you observe, how many meteors you see, and what location in the sky they seem to come from. How well does your observed hourly rate agree with published estimates, such as those in Table 15-1? Is the radiant of the meteor shower apparent from your observations?

55. If a comet is visible with a telescope at your disposal, make arrangements to view it. Can you distinguish the comet from background stars? Can you see its coma? Can you see a tail?

56. Make arrangements to view an asteroid. Observe the asteroid on at least two occasions, separated by a few days. On each night, draw a star chart of the objects in your telescope's field of view. Has the position of one starlike object shifted between observing sessions? Does the position of the moving object agree with the path plotted on published star charts? Do you feel confident that you have in fact seen the asteroid?

57. Use the *Starry Night Enthusiast™* program to observe two asteroids. Display the entire celestial sphere by selecting **Guides > Atlas** in the **Favourites** menu. Center on **Ceres** by typing its name in the **Search All Databases** box in the **Find** pane and label it by clicking in the box to the left of its name as it appears in the Find list. In the toolbar, set the Time Flow Rate to 1 day. Then click the Run Time Forward button. (**a**) Watch Ceres for at least two years of simulated time. Describe how Ceres moves. How can you tell that Ceres orbits the Sun in the same direction as the planets? Does Ceres remain as close to the ecliptic (shown as a solid green line) as the planets do? (If you do not see a green line representing the ecliptic, make sure that the item **The Ecliptic** is checked in the View menu.) (**b**) Click on the **Stop** button and use the **Find** pane to locate and center the view on Pallas. Again click on the Run Time Forward button, and watch how Pallas moves for at least two years of simulated time. How does the motion of Pallas differ from that of Ceres? How are the two motions similar? Which asteroid's orbit is more steeply inclined to the plane of the Earth's orbit? How can you tell?

58. Use the *Starry Night Enthusiast™* program to study the motion of a comet. First set up the field of view so that you are observing the inner solar system from a distance (select **Solar System > Inner Solar system** in the **Favourites** menu). In the toolbar, click on the **Stop** button to halt the animation, and then set the date to January 1, 1995, and the time step to 1 day. Select **View > Solar System > Asteroids** in the menu to remove the asteroids from the view. Open the **Find** pane and center on Comet Hyakutake by typing "*Hyakutake*" in the **Search All Databases** box and then pressing the **Enter** key. Use the Zoom controls to decrease the field of view to about 25° × 17°. Then click on the **Run Time Forward** button. (**a**) Watch the motion of Comet Hyakutake for at least two years of simulated time. Describe what you see. Is the comet's orbit in about the same plane as the orbits of the inner planets, or is it steeply inclined to that plane? (You can tilt the plane of the solar system by holding down the **Shift** key while clicking on and moving the mouse to investigate this off-ecliptic motion.) How does the comet's speed vary as it moves along its orbit? During which part of the orbit is the tail visible? In what direction does the tail point? (**b**) Click on the **Stop** button to halt the animation, and set up the field of view so that you are observing from the center of a transparent Earth by selecting **Guides > Atlas** in the **Favourites** menu. Set the date to **January 1, 1995**, and the **Time Flow Rate** to **1 day**, and again center on Comet Hyakutake. Use the controls at the right-hand end of the toolbar to zoom out as far as possible. Then click on the **Run Time Forward** button and watch the comet's motion for at least two years of simulated time. Describe the motion, and explain why it is more complicated than the motion you observed in part (a). (**c**) Stop the animation, set the date to today's date, set the **Time Flow Rate** to **1 month** ("lunar m."), and restart the animation. Comet Hyakutake is currently moving almost directly away from the Sun and so, as seen from the Sun, its position on the celestial sphere should not change. Is this what you see in *Starry Night Enthusiast™*? Explain any differences. (*Hint:* You are observing from the Earth, not the Sun.)

59. Use the *Starry Night Enthusiast™* program to examine Comet Halley as it would have been seen during its last visit to the Sun and the inner solar system. Display the entire celestial sphere as seen from the center of a transparent Earth by clicking on **Guides > Atlas** in the **Favourites** pane. **Stop** time flow and set the date to **February 28, 1986,** and the time to midnight, **12:00:00 A.M.** Center on Halley's Comet by using **Edit > Find** and entering the name ***Halle***y in the **Search** box. Set the **Time Flow Rate** to **6 hours**. From this view at the Earth's center with daylight turned off, you will note that the Sun is to the left of the view, on the ecliptic plane, as expected. (**a**) Based on the direction of the comet's tail, can you tell in which direction the comet is moving at this instant in time? Step time forward by a few steps to check your prediction. (**b**) Did Halley's Comet move in the direction that you predicted in (a) above? Run time backward to about January 7, 1986, and then run time forward and backward again several times to observe Halley's comet and particularly the direction of its tail during this encounter with the Sun. (**c**) In what direction does the tail of the comet point? Explain why.

Pluto and the Kuiper Belt

by Scott Sheppard

The "what is a planet" controversy began with the discovery of Ceres and several other objects between Mars and Jupiter in the early 1800s. At first, Ceres and then Pallas, Juno, and Vesta were considered planets, even though they were orders of magnitude less massive than the other known planets and all had similar orbits. The definition of a planet at this time was simple. Any object in orbit about the Sun that did not show cometary effects was a planet.

By the mid 1800s, the discovery of more and more objects between Mars and Jupiter saw the words "asteroid" and "minor planet" start to be used in order to signify these objects as an ensemble of bodies. That is, astronomers felt a major planet should be a rare and unique thing. The major planet club thus was left with eight members: Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus, and Neptune.

The planet controversy arose once again in 1930 with the discovery of Pluto, where it was quickly regarded as the ninth planet. During the ensuing years, new observations downgraded Pluto's size from being near that of Earth to being smaller than our Moon. Its orbit also proved to be peculiar because it was highly inclined and eccentric, having it even crossing Neptune's orbit. Its uniqueness kept Pluto in the planet club with relatively few detractors until 1992, when the first Kuiper belt object was discovered. As it became apparent that hundreds of thousands of Kuiper belt objects existed on stable orbits just beyond Neptune, Pluto's planet status was becoming less and less certain. With the discovery of Eris (2003 UB313) in 2003, an object larger than Pluto, Pluto's last unique claim of being the largest Kuiper belt object was lost. Thus, the planet controversy had finally reached a point where a major planet's minimum size must be precisely defined.

No matter how you count them, there are no longer nine major planets in our solar system. The same situation of the main belt asteroids in the mid 1800's is now apparent with Pluto, Eris, and the rest of the Kuiper belt objects. Although the largest Kuiper belt objects are significantly bigger than the largest main belt asteroids, they are still an ensemble of bodies with a continuous size distribution and similar formation and evolution histories. The reason Pluto's status as a planet is a controversy now is simply because it took 62 years between the discovery of Pluto and the next Kuiper belt object in 1992. Within that time span Pluto became known as the ninth planet through several generations and was ingrained in society as such.

Classification is not that important to scientists but it is the understanding of how the object formed and its history. Even so, any classification should have a scientific base and not be based on tradition, as seems to be the case for Pluto in the recent past. Eris as well as Pluto and the rest of the Kuiper belt objects had very similar histories and formation scenarios and thus they should be considered as an ensemble of objects just like the main belt asteroids between Mars and Jupiter. As technology advances, our knowledge and thus understanding advances and so should our classification.

Several minimum-size planet definitions were debated recently. The first is that anything larger than a certain minimum size is a planet. The chosen size would likely be arbitrary and may be a nice round number like 1000 km or the size of Pluto. This stems from the traditional belief that Pluto has been considered a planet in the past and should continue as such. This would give us ten current planets including Eris. A second possible definition is that anything that is spherical should be a planet. This definition is based on physics in which the gravitational force of an object is large enough to overcome any other forces in determining the overall shape of the object. Determining if an object meets this criteria or not is complicated. This definition would also significantly change the status quo by increasing the number of known planets by several factors overnight. A third definition is that any object with a unique orbit and gravitationally dominates its local orbital environment should be a planet. This definition is the most scientifically based, in that objects with similar formation and evolutionary histories would be grouped together. Using this definition would give us eight known major planets, dropping Pluto from the list.

A final proposal as to what is a planet takes into account that the current known major planets themselves are remarkably different and can be split into individual categories themselves.

(continued on the next page)

Scott S. Sheppard is an astronomer at the Carnegie Institution of Washington in the Department of Terrestrial Magnetism located in Washington, D.C. His main research currently involves discovering small bodies in our solar system to understand their dynamical and physical properties that help constrain solar system and planet formation. He has discovered several small moons around all the giant planets, been a pioneer in the discovery of Neptune Trojans, as well as found tens of Kuiper Belt objects.

Pluto and the Kuiper Belt *(continued)*

Mercury, Venus, Earth and Mars are terrestrial planets whose compositions are dominated by rock. Jupiter and Saturn are gas giant planets dominated by their hydrogen and helium envelopes. Uranus and Neptune are ice giant planets dominated by gases other than hydrogen and helium. The trans-Neptunian objects or "ice dwarf planets," as some are calling them, are probably composed of large amounts of volatiles such as methane ice and water ice.

The International Astronomical Union (IAU) recently choose to use a combination of the second and third definitions above to define a planet. That is, a planet is an object that is both spherical and has a unique orbit in which it is gravitationally dominant. An object that only satisfies definition two above and not three is now being called a dwarf planet, of which Ceres, Pluto and Eris qualify, as well as several more objects in the Kuiper belt. In reality, it's the public that must accept this definition, and we will only know if that is the case a few generations from now.

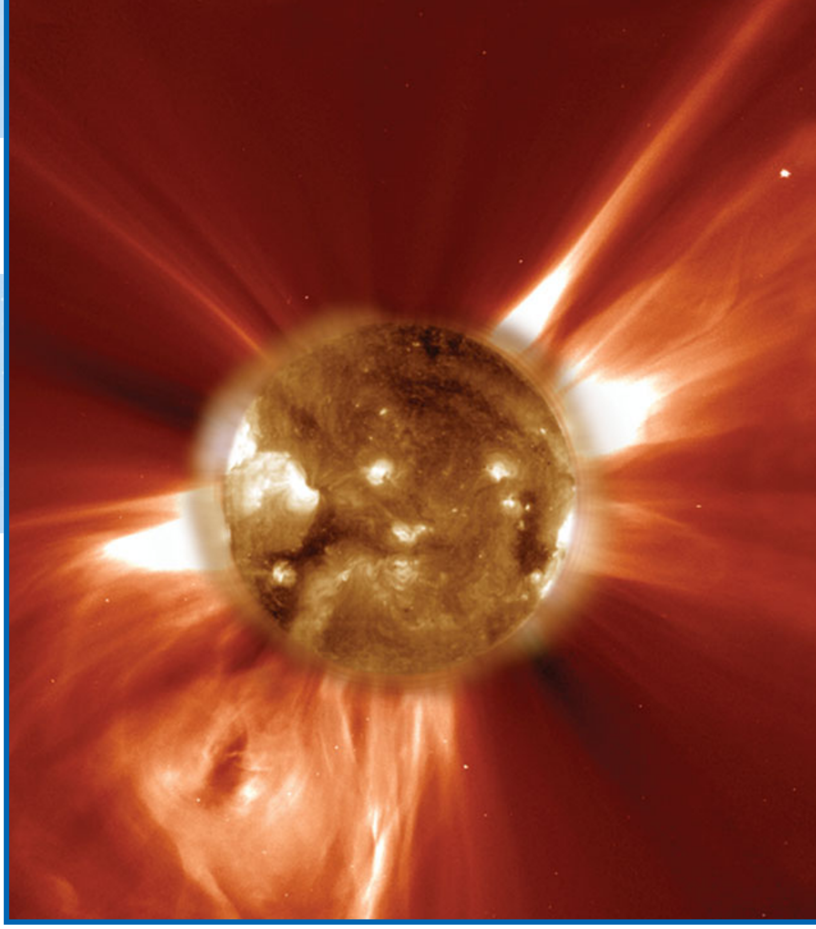
Defining the word "planet" is like defining what an ocean is. At first it appears to be a simple term, but is very hard to precisely define. Further planet discoveries will be made and there are sure to be borderline cases. This is just the way nature is; it does not have little bins to nicely classify objects, but there is usually a continuous array of objects. The important thing to take away from all of this is to understand this array of objects by using our rapidly expanding scientific knowledge and understanding our place in the solar system.

16

Our Star, the Sun

The Sun is by far the brightest object in the sky. By earthly standards, the temperature of its glowing surface is remarkably high, about 5800 K. Yet there are regions of the Sun that reach far higher temperatures of tens of thousands or even millions of kelvins. Gases at such temperatures emit ultraviolet light, which makes them prominent in the accompanying image from an ultraviolet telescope in space. Some of the hottest and most energetic regions on the Sun spawn immense disturbances. These can propel solar material across space to reach the Earth and other planets.

In recent decades, we have learned that the Sun shines because at its core hundreds of millions of tons of hydrogen are converted to helium every second. We have confirmed this picture by detecting the by-products of this transmutation—strange, ethereal particles called neutrinos—streaming outward from the Sun into space. We have discovered that the Sun has a surprisingly violent atmosphere, with a host of features such as sunspots whose numbers rise and fall on a predictable 11-year cycle. By studying the Sun's vibrations, we have begun to probe beneath its surface into hitherto unexplored realms. And we have just begun to investigate how changes in the Sun's activity can affect the Earth's environment as well as our technological society.



R I V U X G

A composite view of the Sun (at ultraviolet wavelengths) and an upheaval in the Sun's outer atmosphere, or corona (at visible wavelengths). (*SOHO*/LASCO/EIT/ESA/NASA)

16-1 The Sun's energy is generated by thermonuclear reactions in its core

The Sun is the largest member of the solar system. It has almost a thousand times more mass than all the planets, moons, asteroids, comets, and meteoroids put together. But the Sun is also a star. In fact, it is a remarkably typical star, with a mass, size, surface temperature, and chemical composition that are roughly midway between the extremes exhibited by the myriad other stars in the heavens. Table 16-1 lists essential data about the Sun.

Learning Goals

By reading the sections of this chapter, you will learn

16-1 The source of the Sun's heat and light

16-2 How scientists model the Sun's internal structure

16-3 How the Sun's vibrations reveal what lies beneath its glowing surface

16-4 How scientists are able to probe the Sun's energy-generating core

16-5 Why the gaseous Sun appears to have a sharp outer edge

16-6 Why the upper regions of the solar atmosphere have an emission spectrum

16-7 The relationship between the Sun's corona and the solar wind

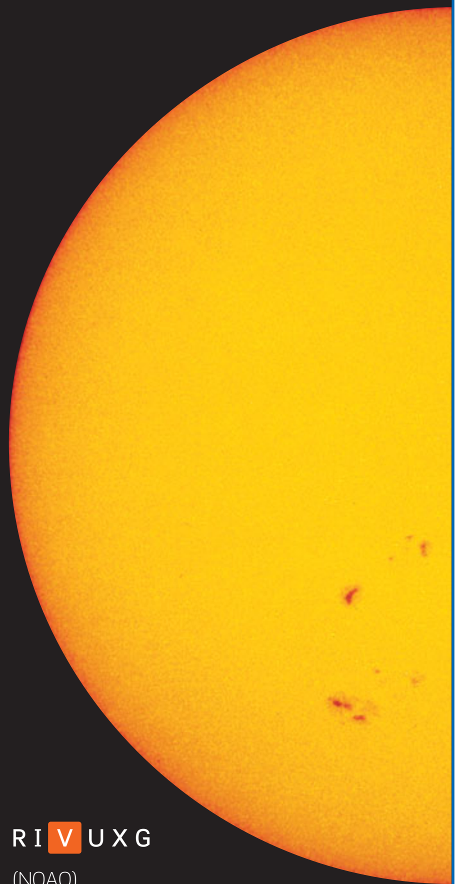
16-8 The nature of sunspots

16-9 The connection between sunspots and the Sun's magnetic field

16-10 How magnetic reconnection can power immense solar eruptions

Table 16-1 Sun Data

Distance from the Earth:	Mean: 1 AU = 149,598,000 km Maximum: 152,000,000 km Minimum: 147,000,000 km
Light travel time to the Earth:	8.32 min
Mean angular diameter:	32 arcmin
Radius:	696,000 km = 109 Earth radii
Mass:	1.9891×10^{30} kg = 3.33×10^{5} Earth masses
Composition (by mass):	74% hydrogen, 25% helium, 1% other elements
Composition (by number of atoms):	92.1% hydrogen, 7.8% helium, 0.1% other elements
Mean density:	1410 kg/m^3
Mean temperatures:	Surface: 5800 K; Center: 1.55×10^{7} K
Luminosity:	3.90×10^{26} W
Distance from center of Galaxy:	8000 pc = 26,000 ly
Orbital period around center of Galaxy:	220 million years
Orbital speed around center of Galaxy:	220 km/s



WEB LINK 16.1

R I V U X G

(NOAO)

Solar Energy

For most people, what matters most about the Sun is the energy that it radiates into space. Without the Sun's warming rays, our atmosphere and oceans would freeze into an icy layer coating a desperately cold planet, and life on Earth would be impossible. To understand why we are here, we must understand the nature of the Sun.

Why is the Sun such an important source of energy? An important part of the answer is that the Sun has a far higher surface temperature than any of the planets or moons. The Sun's spectrum is close to that of an idealized blackbody with a temperature of 5800 K (see Sections 5-3 and 5-4, especially Figure 5-12). Thanks to this high temperature, each square meter of the Sun's surface emits a tremendous amount of radiation, principally at visible wavelengths. Indeed, the Sun is the only object in the solar system that emits substantial amounts of visible light. The light that we see from the Moon and planets is actually sunlight that struck those worlds and was reflected toward Earth.

The Sun's size also helps us explain its tremendous energy output. Because the Sun is so large, the total number of square meters of radiating surface—that is, its surface area—is immense. Hence, the total amount of energy emitted by the Sun each second, called its **luminosity,** is very large indeed: about 3.9×10^{26} watts, or 3.9×10^{26} joules of energy emitted per second. (We discussed the relation among the Sun's surface temperature, radius, and luminosity in Box 5-2.) Astronomers denote the Sun's luminosity by the symbol $L_\odot$. (A circle with a dot in the center is the astronomical symbol for the Sun and was also used by ancient astrologers.)

The Source of the Sun's Energy: Early Ideas

These ideas lead us to a more fundamental question: What keeps the Sun's visible surface so hot? Or, put another way, what is the fundamental source of the tremendous energies that the Sun radiates into space? For centuries, this was one of the greatest mysteries in science. The mystery deepened in the nineteenth century, when geologists and biologists found convincing evidence that life had existed on Earth for at least several hundred million years. (We now know that the Earth is 4.56 billion years old and that life has existed on it for most of its history.) Since life as we know

Figure 16-1 R I V U X G

The Sun The Sun's visible surface has a temperature of about 5800 K. At this temperature, all solids and liquids vaporize to form gases. It was only in the twentieth century that scientists discovered what has kept the Sun so hot for billions of years: the thermonuclear fusion of hydrogen nuclei in the Sun's core. (Jeremy Woodhouse/PhotoDisc)

it depends crucially on sunlight, the Sun must be as old. This posed a severe problem for physicists. What source of energy could have kept the Sun shining for so long (Figure 16-1)?

One attempt to explain solar energy was made in the mid-1800s by the English physicist Lord Kelvin (for whom the temperature scale is named) and the German scientist Hermann von Helmholtz. They argued that the tremendous weight of the Sun's outer layers should cause the Sun to contract gradually, compressing its interior gases. Whenever a gas is compressed, its temperature rises. (You can demonstrate this with a bicycle pump: As you pump air into a tire, the temperature of the air increases and the pump becomes warm to the touch.) Kelvin and Helmholtz thus suggested that gravitational contraction could cause the Sun's gases to become hot enough to radiate energy out into space.

This process, called *Kelvin-Helmholtz contraction,* actually does occur during the earliest stages of the birth of a star like the Sun (see Section 8-4). But Kelvin-Helmholtz contraction cannot be the major source of the Sun's energy today. If it were, the Sun would have had to be much larger in the relatively recent past. Helmholtz's own calculations showed that the Sun could have started its initial collapse from the solar nebula no more than about 25 million years ago. But the geological and fossil record shows that the Earth is far older than that, and so the Sun must be as well. Hence, this model of a Sun that shines because it shrinks cannot be correct.

On Earth, a common way to produce heat and light is by burning fuel, such as a log in a fireplace or coal in a power plant. Is it possible that a similar process explains the energy released by the Sun? The answer is no, because this process could not continue for a long enough time to explain the age of the Earth. The chemical reactions involved in burning release roughly 10^{-19} joule of energy per atom. Therefore, the number of atoms that would have to undergo chemical reactions each second to generate the Sun's luminosity of 3.9×10^{26} joules per second is approximately

$$\frac{3.9 \times 10^{26} \text{ joules per second}}{10^{-19} \text{ joule per atom}} = 3.9 \times 10^{45} \text{ atoms per second}$$

From its mass and chemical composition, we know that the Sun contains about 10^{57} atoms. Thus, the length of time that would be required to consume the entire Sun by burning is

$$\frac{10^{57} \text{ atoms}}{3.9 \times 10^{45} \text{ atoms per second}} = 3 \times 10^{11} \text{ seconds}$$

There are about 3×10^{7} seconds in a year. Hence, in this model, the Sun would burn itself out in a mere 10,000 (10^4) years! This is far shorter than the known age of the Earth, so chemical reactions also cannot explain how the Sun shines.

The Source of the Sun's Energy: Discovering Thermonuclear Fusion

The source of the Sun's luminosity could be explained if there were a process that was like burning but released much more energy per atom. Then the rate at which atoms would have to be consumed would be far less, and the lifetime of the Sun could be long enough to be consistent with the known age of the Earth. Albert Einstein discovered the key to such a process in 1905. According to his *special theory of relativity,* a quantity m of mass can in principle be converted into an amount of energy E according to a now-famous equation:

Einstein's mass-energy equation

$$E = mc^2$$

m = quantity of mass, in kg

c = speed of light = 3×10^8 m/s

E = amount of energy into which the mass can be converted, in joules

The speed of light c is a large number, so c^2 is huge. Therefore, a small amount of matter can release an awesome amount of energy.

Inspired by Einstein's ideas, astronomers began to wonder if the Sun's energy output might come from the conversion of matter into energy. The Sun's low density of 1410 kg/m^3 indicates that it must be made of the very lightest atoms, primarily hydrogen and helium. In the 1920s, the British astronomer Arthur Eddington showed that temperatures near the center of the Sun must be so high that atoms become completely ionized. Hence, at the Sun's center we expect to find hydrogen nuclei and electrons flying around independent of each other.

Ideas from relativity and nuclear physics led to an understanding of how the Sun shines

Another British astronomer, Robert Atkinson, suggested that under these conditions hydrogen nuclei could fuse together to produce helium nuclei in a *nuclear reaction* that transforms a tiny amount of mass

BOX 16-1 Tools of the Astronomer's Trade

Converting Mass into Energy

The *Cosmic Connections* figure shows the steps involved in the thermonuclear fusion of hydrogen at the Sun's center. In these steps, four protons are converted into a single nucleus of ^{4}He, an isotope of helium with two protons and two neutrons. (As we saw in Box 5-5, different isotopes of the same element have the same number of protons but different numbers of neutrons.) The reaction depicted in Figure 16-2*a* also produces a neutral, nearly massless particle called the *neutrino*. Neutrinos respond hardly at all to ordinary matter, so they travel almost unimpeded through the Sun's massive bulk. Hence, the energy that neutrinos carry is quickly lost into space. This loss is not great, however, because the neutrinos carry relatively little energy. (See Section 16-4 for more about these curious particles.)

Most of the energy released by thermonuclear fusion appears in the form of gamma-ray photons. The energy of these photons remains trapped within the Sun for a long time, thus maintaining the Sun's intense internal heat. Some gamma-ray photons are produced by the reaction shown as Step 2 in the *Cosmic Connections* figure. Others appear when an electron in the Sun's interior annihilates a positively charged electron, or **positron,** which is a by-product of the reaction shown in Step 1 in the *Cosmic Connections* figure. An electron and a positron are respectively matter and antimatter, and they convert entirely into energy when they meet. (You may have thought that "antimatter" was pure science fiction. In fact, tremendous amounts of antimatter are being created and annihilated in the Sun as you read these words.)

We can summarize the thermonuclear fusion of hydrogen as follows:

$$4\ {}^1\text{H} \rightarrow {}^4\text{He} + \text{neutrinos} + \text{gamma-ray photons}$$

To calculate how much energy is released in this process, we use Einstein's mass-energy formula: The energy released is equal to the amount of mass consumed multiplied by c^2, where c is the speed of light. To see how much mass is consumed, we compare the combined mass of four hydrogen atoms (the ingredients) to the mass of one helium atom (the product):

$$\begin{array}{r} 4 \text{ hydrogen atoms} = 6.693 \times 10^{-27} \text{ kg} \\ -1 \text{ helium atom} = 6.645 \times 10^{-27} \text{ kg} \\ \hline \text{Mass lost} = 0.048 \times 10^{-27} \text{ kg} \end{array}$$

Thus, a small fraction (0.7%) of the mass of the hydrogen going into the nuclear reaction does not show up in the mass of the helium. This lost mass is converted into an amount of energy $E = mc^2$:

$$\begin{aligned} E = mc^2 &= (0.048 \times 10^{-27} \text{ kg})(3 \times 10^8 \text{ m/s})^2 \\ &= 4.3 \times 10^{-12} \text{ joule} \end{aligned}$$

This is the amount of energy released by the formation of a single helium atom. It would light a 10-watt lightbulb for almost one-half of a trillionth of a second.

EXAMPLE: How much energy is released when 1 kg of hydrogen is converted to helium?

Situation: We are given the initial mass of hydrogen. We know that a fraction of the mass is lost when the hydrogen undergoes fusion to make helium; our goal is to find the quantity of energy into which this lost mass is transformed.

Tools: We use the equation $E = mc^2$ and the result that 0.7% of the mass is lost when hydrogen is converted into hydrogen.

Answer: When 1 kilogram of hydrogen is converted to helium, the amount of mass lost is 0.7% of 1 kg, or 0.007 kg. (This means that 0.993 kg of helium is produced.) Using Einstein's equation, we find that this missing 0.007 kg of matter is transformed into an amount of energy equal to

$$E = mc^2 = (0.007 \text{ kg})(3 \times 10^8 \text{ m/s})^2 = 6.3 \times 10^{14} \text{ joules}$$

Review: The energy released by the fusion of 1 kilogram of hydrogen is the same as that released by burning *20,000 metric tons* (2×10^7 kg) of coal! Hydrogen fusion is a *much* more efficient energy source than ordinary burning.

The Sun's luminosity is 3.9×10^{26} joules per second. To generate this much power, hydrogen must be consumed at a rate of

$$\frac{3.9 \times 10^{26} \text{ joules per second}}{6.3 \times 10^{14} \text{ joules per kilogram}} = 6 \times 10^{11} \text{ kilograms per second}$$

That is, the Sun converts 600 million metric tons of hydrogen into helium every second.

into a large amount of energy. Experiments in the laboratory using individual nuclei show that such reactions can indeed take place. The process of converting hydrogen into helium is called **hydrogen fusion.** (It is also sometimes called *hydrogen burning,* even though nothing is actually burned in the conventional sense. Ordinary burning involves chemical reactions that rearrange the outer electrons of atoms but have no effect on the atoms' nuclei.) Hydrogen fusion provides the devastating energy released in a hydrogen bomb (see Figure 1-4).

The fusing together of nuclei is also called **thermonuclear fusion,** because it can take place only at extremely high temperatures. The reason is that all nuclei have a positive electric charge

COSMIC CONNECTIONS

The Proton-Proton Chain

The most common form of hydrogen fusion in the Sun involves three steps, each of which releases energy.

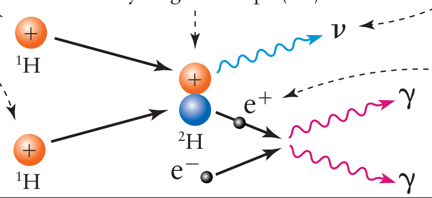
Hydrogen fusion in the Sun usually takes place in a sequence of steps called the proton-proton chain. Each of these steps releases energy that heats the Sun and gives it its luminosity.

STEP 1

(a) Two protons (hydrogen nuclei, ^{1}H) collide.

(b) One of the protons changes into a neutron (shown in blue). The proton and neutron form a hydrogen isotope (^{2}H).

(c) One byproduct of converting a proton to a neutron is a neutral, nearly massless neutrino (ν). This escapes from the Sun.

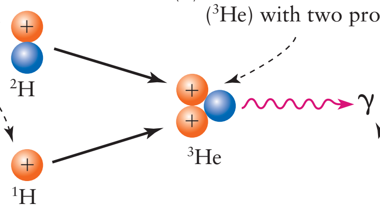


(d) The other byproduct of converting a proton to a neutron is a positively charged electron, or positron (e^+). This encounters an ordinary electron (e^-), annihilating both particles and converting them into gamma-ray photons (γ). The energy of these photons goes into sustaining the Sun's internal heat.

STEP 2

(a) The ^{2}H nucleus produced in Step 1 collides with a third proton (^{1}H).

(b) The result of the collision is a helium isotope (^{3}He) with two protons and one neutron.



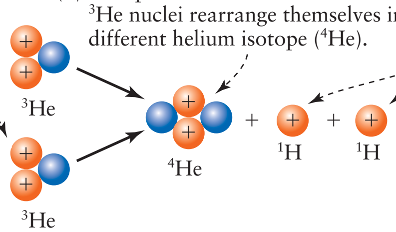
(c) This nuclear reaction releases another gamma-ray photon (γ). Its energy also goes into sustaining the internal heat of the Sun.

STEP 3

(a) The ^{3}He nucleus produced in Step 2 collides with another ^{3}He nucleus produced from three other protons.

(b) Two protons and two neutrons from the two ^{3}He nuclei rearrange themselves into a different helium isotope (^{4}He).

(c) The two remaining protons are released. The energy of their motion contributes to the Sun's internal heat.



(d) Six ^{1}H nuclei went into producing the two ^{3}He nuclei, which combine to make one ^{4}He nucleus. Since two of the original ^{1}H nuclei are returned to their original state, we can summarize the three steps as:

$$\boxed{4\ ^1\mathrm{H} \longrightarrow\ ^4\mathrm{He} + \text{energy}}$$

ANIMATION 16.1

Hydrogen fusion also takes place in all of the stars visible to the naked eye. (Fusion follows a different sequence of steps in the most massive stars, but the net result is the same.)

(Courtesy of Wally Pacholka)

and so tend to repel one another. But in the extreme heat and pressure at the Sun's center, hydrogen nuclei (protons) are moving so fast that they can overcome their electric repulsion and actually touch one another. When that happens, thermonuclear fusion can take place.

ANALOGY You can think of protons as tiny electrically charged spheres that are coated with a very powerful glue. If the spheres are not touching, the repulsion between their charges pushes them apart. But if the spheres are forced into contact, the strength of the glue "fuses" them together.

CAUTION! Be careful not to confuse thermonuclear fusion with the similar-sounding process of *nuclear fission*. In nuclear fusion, energy is released by joining together nuclei of lightweight atoms such as hydrogen. In nuclear fission, by contrast, the nuclei of very massive atoms such as uranium or plutonium release energy by fragmenting into smaller nuclei. Nuclear power plants produce energy using fission, not fusion. (Generating power using fusion has been a goal of researchers for decades, but no one has yet devised a commercially viable way to do this.)

Converting Hydrogen to Helium

We learned in Section 5-8 that the nucleus of a hydrogen atom (H) consists of a single proton. The nucleus of a helium atom (He) consists of two protons and two neutrons. In the nuclear process that Atkinson described, four hydrogen nuclei combine to form one helium nucleus, with a concurrent release of energy:

$$4\ \mathrm{H} \rightarrow \mathrm{He} + \text{energy}$$

In several separate reactions, two of the four protons are changed into neutrons, and eventually combine with the remaining protons to produce a helium nucleus. This sequence of reactions is called the **proton-proton chain.** The *Cosmic Connections* figure depicts the proton-proton chain in detail.

Each time this process takes place, a small fraction (0.7%) of the combined mass of the hydrogen nuclei does not show up in the mass of the helium nucleus. This "lost" mass is converted into energy. Box 16-1 describes how to use Einstein's mass-energy equation to calculate the amount of energy released.

CAUTION! You may have heard the idea that mass is always conserved (that is, it is neither created nor destroyed), or that energy is always conserved in a reaction. Einstein's ideas show that neither of these statements is quite correct, because mass can be converted into energy and vice versa. A more accurate statement is that the total amount of mass *plus* energy is conserved. Hence, the destruction of mass in the Sun does not violate any laws of nature.

For every four hydrogen nuclei converted into a helium nucleus, 4.3×10^{-12} joule of energy is released. This may seem like only a tiny amount of energy, but it is about 10^7 times larger than the amount of energy released in a typical chemical reaction, such as occurs in ordinary burning. Thus, thermonuclear fusion can explain how the Sun could have been shining for billions of years.

To produce the Sun's luminosity of 3.9×10^{26} joules per second, 6×10^{11} kg (600 million metric tons) of hydrogen must be converted into helium each second. This rate is prodigious, but there is a literally astronomical amount of hydrogen in the Sun. In particular, the Sun's core contains enough hydrogen to have been giving off energy at the present rate for as long as the solar system has existed, about 4.56 billion years, and to continue doing so for more than 6 billion years into the future.

LOOKING DEEPER 16.1

The proton-proton chain is also the energy source for many of the stars in the sky. In stars with central temperatures that are much hotter than that of the Sun, however, hydrogen fusion proceeds according to a different set of nuclear reactions, called the **CNO cycle,** in which carbon, nitrogen, and oxygen nuclei absorb protons to produce helium nuclei. Still other thermonuclear reactions, such as helium fusion, carbon fusion, and oxygen fusion, occur late in the lives of many stars.

16-2 A theoretical model of the Sun shows how energy gets from its center to its surface

While thermonuclear fusion is the source of the Sun's energy, this process cannot take place everywhere within the Sun. As we have seen, extremely high temperatures—in excess of 10^7 K—are required for atomic nuclei to fuse together to form larger nuclei. The temperature of the Sun's visible surface, about 5800 K, is too low for these reactions to occur there. Hence, thermonuclear fusion can be taking place only within the Sun's interior. But precisely where does it take place? And how does the energy produced by fusion make its way to the surface, where it is emitted into space in the form of photons?

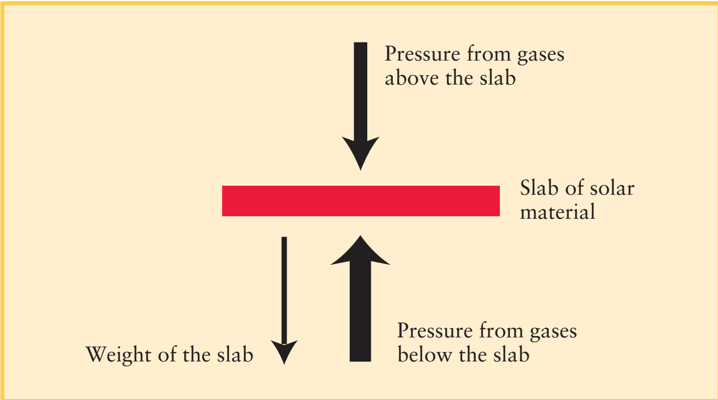
To answer these questions, we must understand conditions in the Sun's interior. Ideally, we would send an exploratory spacecraft to probe deep into the Sun; in practice, the Sun's intense heat would vaporize even the sturdiest spacecraft. Instead, astronomers use the laws of physics to construct a theoretical model of the Sun. (We discussed the use of models in science in Section 1-1.) Let's see what ingredients go into building a model of this kind.

Hydrostatic Equilibrium

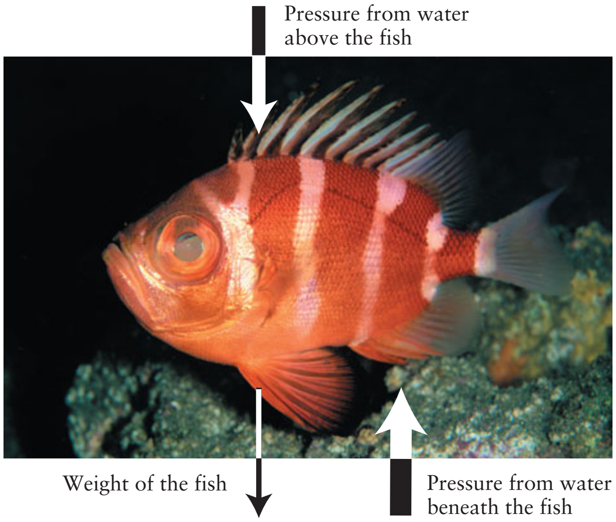
Note first that the Sun is not undergoing any dramatic changes. The Sun is not exploding or collapsing, nor is it significantly heating or cooling. The Sun is thus said to be in both *hydrostatic equilibrium* and *thermal equilibrium*.

To understand what is meant by **hydrostatic equilibrium,** imagine a slab of material in the solar interior (Figure 16-2*a*). In equilibrium, the slab on average will move neither up nor down. (In fact, there are upward and downward motions of material inside the Sun, but these motions average out.) Equilibrium is maintained by a balance among three forces that act on this slab:

1. The downward pressure of the layers of solar material above the slab.
2. The upward pressure of the hot gases beneath the slab.
3. The slab's weight—that is, the downward gravitational pull it feels from the rest of the Sun.



(a) Material inside the sun is in hydrostatic equilibrium, so forces balance



(b) A fish floating in water is in hydrostatic equilibrium, so forces balance

Figure 16-2

Hydrostatic Equilibrium (a) Material in the Sun's interior tends to move neither up nor down. The upward forces on a slab of solar material (due to pressure of gases below the slab) must balance the downward forces (due to the slab's weight and the pressure of gases above the slab). Hence, the pressure must increase with increasing depth. **(b)** The same principle applies to a fish floating in water. In equilibrium, the forces balance and the fish neither rises nor sinks. (Ken Usami/PhotoDisc)

The pressure from below must balance both the slab's weight and the pressure from above. Hence, the pressure below the slab must be greater than that above the slab. In other words, pressure has to increase with increasing depth. For the same reason, pressure increases as you dive deeper into the ocean (Figure 16-2*b*) or as you move toward lower altitudes in our atmosphere.

Hydrostatic equilibrium also tells us about the density of the slab. If the slab is too dense, its weight will be too great and it will sink; if the density is too low, the slab will rise. To prevent this, the density of solar material must have a certain value at each depth within the solar interior. (The same principle applies to objects that float beneath the surface of the ocean. Scuba divers wear weight belts to increase their average density so that they will neither rise nor sink but will stay submerged at the same level.)

Thermal Equilibrium

Another consideration is that the Sun's interior is so hot that it is completely gaseous. Gases compress and become more dense when you apply greater pressure to them, so density must increase along with pressure as you go to greater depths within the Sun. Furthermore, when you compress a gas, its temperature tends to rise, so the temperature must also increase as you move toward the Sun's center.

While the temperature in the solar interior is different at different depths, the temperature at each depth remains constant in time. This principle is called **thermal equilibrium.** For the Sun to be in thermal equilibrium, all the energy generated by thermonuclear reactions in the Sun's core must be transported to the Sun's glowing surface, where it can be radiated into space. If too much energy flowed from the core to the surface to be radiated away, the Sun's interior would cool down; alternatively, the Sun's interior would heat up if too little energy flowed to the surface.

Transporting Energy Outward from the Sun's Core

But exactly how is energy transported from the Sun's center to its surface? There are three methods of energy transport: *conduction, convection,* and *radiative diffusion.* Only the last two are important inside the Sun.

If you heat one end of a metal bar with a blowtorch, energy flows to the other end of the bar so that it too becomes warm. The efficiency of this method of energy transport, called **conduction,** varies significantly from one substance to another. For example, copper is a good conductor of heat, but wood is not (which is why copper pots often have wooden handles). Conduction is not an efficient means of energy transport in substances with low average densities, including the gases inside stars like the Sun.

Inside stars like our Sun, energy moves from center to surface by two other means: convection and radiative diffusion. **Convection** is the circulation of fluids—gases or liquids—between hot and cool regions. Hot gases rise toward a star's surface, while cool gases sink back down toward the star's center. This physical movement of gases transports heat energy outward in a star, just as the physical movement of water boiling in a pot transports energy from the bottom of the pot (where the heat is applied) to the cooler water at the surface.

In **radiative diffusion,** photons created in the thermonuclear inferno at a star's center diffuse outward toward the star's surface. Individual photons are absorbed and reemitted by atoms and electrons inside the star. The overall result is an outward migration from the hot core, where photons are constantly created, toward the cooler surface, where they escape into space.

Modeling the Sun

To construct a model of a star like the Sun, astrophysicists express the ideas of hydrostatic equilibrium, thermal equilibrium, and energy transport as a set of equations. To ensure that the model applies to the particular star under study, they also make

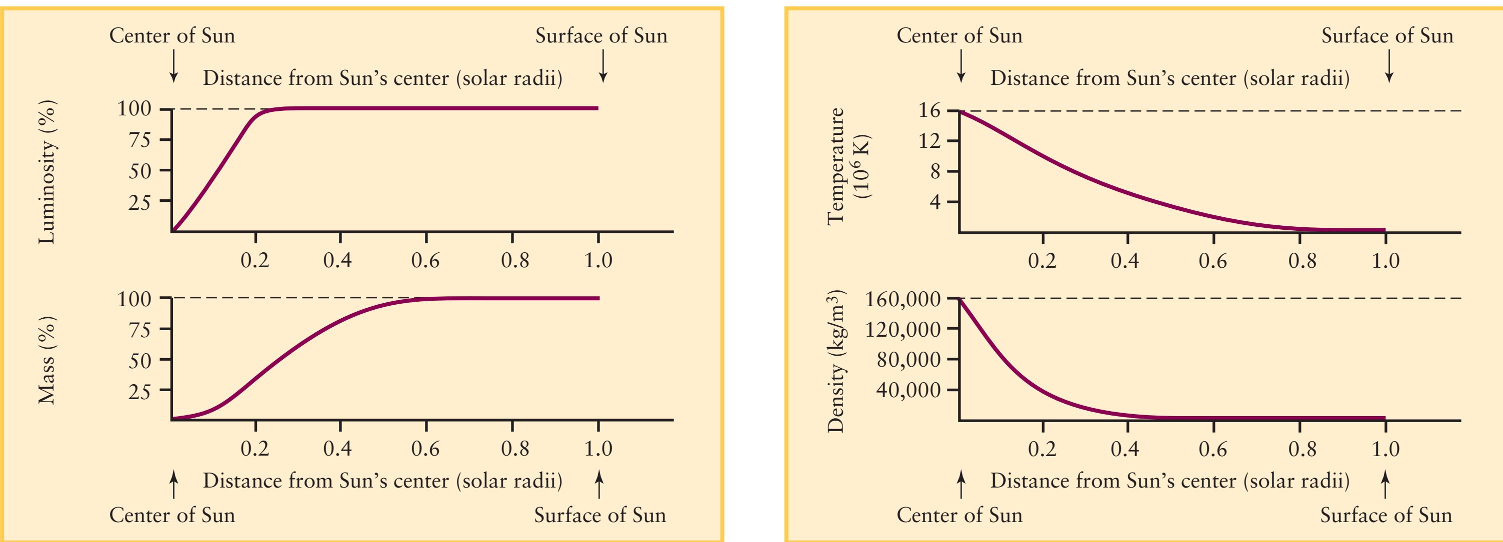
Table 16-2 A Theoretical Model of the Sun

Distance from the Sun's center (solar radii)	Fraction of luminosity	Fraction of mass	Temperature ($\times 10^6$ K)	Density (kg/m^3)	Pressure relative to pressure at center
0.0	0.00	0.00	15.5	160,000	1.00
0.1	0.42	0.07	13.0	90,000	0.46
0.2	0.94	0.35	9.5	40,000	0.15
0.3	1.00	0.64	6.7	13,000	0.04
0.4	1.00	0.85	4.8	4,000	0.007
0.5	1.00	0.94	3.4	1,000	0.001
0.6	1.00	0.98	2.2	400	0.0003
0.7	1.00	0.99	1.2	80	4×10^{-5}
0.8	1.00	1.00	0.7	20	5×10^{-6}
0.9	1.00	1.00	0.3	2	3×10^{-7}
1.0	1.00	1.00	0.006	0.00030	4×10^{-13}

Note: The distance from the Sun's center is expressed as a fraction of the Sun's radius ($R_\odot$). Thus, 0.0 is at the center of the Sun and 1.0 is at the surface. The fraction of luminosity is that portion of the Sun's total luminosity produced within each distance from the center; this is equal to 1.00 for distances of 0.25 $R_\odot$ or more, which means that all of the Sun's nuclear reactions occur within 0.25 solar radius from the Sun's center. The fraction of mass is that portion of the Sun's total mass lying within each distance from the Sun's center. The pressure is expressed as a fraction of the pressure at the center of the Sun.

use of astronomical observations of the star's surface. (For example, to construct a model of the Sun, they use the data that the Sun's surface temperature is 5800 K, its luminosity is 3.9×10^{26} W, and the gas pressure and density at the surface are almost zero.) The astrophysicists then use a computer to solve their set of equations and calculate conditions layer by layer in toward the star's center. The result is a model of how temperature, pressure, and density increase with increasing depth below the star's surface.

Table 16-2 and Figure 16-3 show a theoretical model of the Sun that was calculated in just this way. Different models of the Sun use slightly different assumptions, but all models give essen-



INTERACTIVE EXERCISE 16.1

Figure 16-3

A Theoretical Model of the Sun's Interior These graphs depict what percentage of the Sun's total luminosity is produced within each distance from the center (upper left), what percentage of the total mass lies within each distance from the center (lower left), the temperature at each distance (upper right), and the density at each distance (lower right). (See Table 16-2 for a numerical version of this model.)

tially the same results as those shown here. From such computer models we have learned that at the Sun's center the density is 160,000 kg/m^3 (14 times the density of lead!), the temperature is 1.55×10^7 K, and the pressure is 3.4×10^{11} atm. (One atmosphere, or 1 atm, is the average atmospheric pressure at sea level on Earth.)

Table 16-2 and Figure 16-3 show that the solar luminosity rises to 100% at about one-quarter of the way from the Sun's center to its surface. In other words, the Sun's energy production occurs within a volume that extends out only to 0.25 $R_\odot$. (The symbol $R_\odot$ denotes the solar radius, or radius of the Sun as a whole, equal to 696,000 km.) Outside 0.25 $R_\odot$, the density and temperature are too low for thermonuclear reactions to take place. Also note that 94% of the total mass of the Sun is found within the inner 0.5 $R_\odot$. Hence, the outer 0.5 $R_\odot$ contains only a relatively small amount of material.

How energy flows from the Sun's center toward its surface depends on how easily photons move through the gas. If the solar gases are comparatively transparent, photons can travel moderate distances before being scattered or absorbed, and energy is thus transported by radiative diffusion. If the gases are comparatively opaque, photons cannot get through the gas easily and heat builds up. Convection then becomes the most efficient means of energy transport. The gases start to churn, with hot gas moving upward and cooler gas sinking downward.

From the center of the Sun out to about 0.71 $R_\odot$, energy is transported by radiative diffusion. Hence, this region is called the **radiative zone.** Beyond about 0.71 $R_\odot$, the temperature is low enough (a mere 2×10^6 K or so) for electrons and hydrogen nuclei to join into hydrogen atoms. These atoms are very effective at absorbing photons, much more so than free electrons or nuclei, and this absorption chokes off the outward flow of photons. Therefore, beyond about 0.71 $R_\odot$, radiative diffusion is not an effective way to transport energy. Instead, convection dominates the energy flow in this outer region, which is why it is called the **convective zone.** Figure 16-4 shows these aspects of the Sun's internal structure.

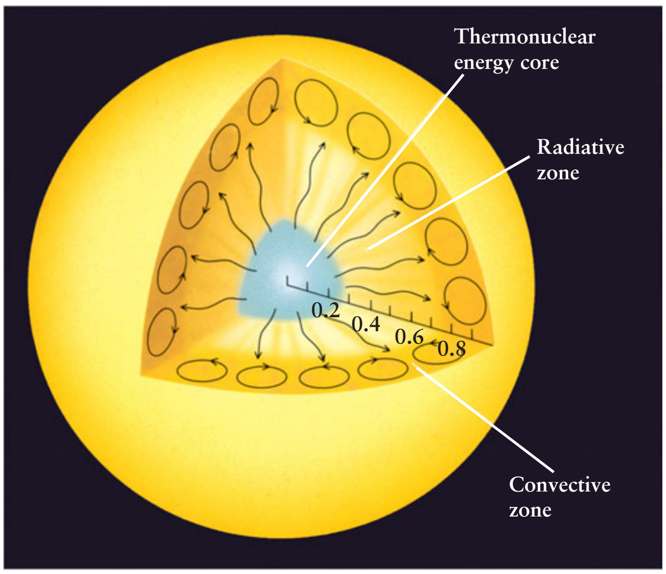


Figure 16-4

The Sun's Internal Structure Thermonuclear reactions occur in the Sun's core, which extends out to a distance of 0.25 $R_\odot$ from the center. Energy is transported outward, via radiative diffusion, to a distance of about 0.71 $R_\odot$. In the outer layers between 0.71 $R_\odot$ and 1.00 $R_\odot$, energy flows outward by convection.

Although energy travels through the radiative zone in the form of photons, the photons have a difficult time of it. Table 16-2 shows that the material in this zone is extremely dense, so photons from the Sun's core take a long time to diffuse through the radiative zone. As a result, it takes approximately 170,000 years for energy created at the Sun's center to travel 696,000 km to the solar surface and finally escape as sunlight. The energy flows outward at an average rate of 50 centimeters per hour, or about 20 times slower than a snail's pace.

Once the energy escapes from the Sun, it travels much faster—at the speed of light. Thus, solar energy that reaches you today took only 8 minutes to travel the 150 million kilometers from the Sun's surface to the Earth. But this energy was actually produced by thermonuclear reactions that took place in the Sun's core hundreds of thousands of years ago.

The sunlight that reaches Earth today results from thermonuclear reactions that took place about 170,000 years ago

16-3 Astronomers probe the solar interior using the Sun's own vibrations

WEB LINK 16.2

We have described how astrophysicists construct models of the Sun. But since we cannot see into the Sun's opaque interior, how can we check these models to see if they are accurate? What is needed is a technique for probing the Sun's interior. A very powerful technique of just this kind involves measuring vibrations of the Sun as a whole. This field of solar research is called **helioseismology.**

Vibrations are a useful tool for examining the hidden interiors of all kinds of objects. Food shoppers test whether melons are ripe by tapping on them and listening to the vibrations. Geologists can determine the structure of the Earth's interior by using seismographs to record vibrations during earthquakes.

Although there are no true "sunquakes," the Sun does vibrate at a variety of frequencies, somewhat like a ringing bell. These vibrations were first noticed in 1960 by Robert Leighton of the California Institute of Technology, who made high-precision Doppler shift observations of the solar surface. These measurements revealed that parts of the Sun's surface move up and down about 10 meters every 5 minutes. Since the mid-1970s, several astronomers have reported slower vibrations, having periods ranging from 20 to 160 minutes. The detection of extremely slow vibrations has inspired astronomers to organize networks of telescopes around and in orbit above the Earth to monitor the Sun's vibrations on a continuous basis.

WEB LINK 16.3

The vibrations of the Sun's surface can be compared with sound waves. If you could somehow survive within the Sun's outermost layers, you would first notice a deafening roar, somewhat like a jet engine, produced by turbulence in the Sun's gases. But superimposed on this noise would be a variety of nearly pure tones. You would need greatly enhanced hearing to detect these tones, however; the strongest has a frequency of just 0.003 hertz, 13 octaves below the lowest frequency audible to humans. (Recall from Section 5-2 that one hertz is one oscillation per second.)

In 1970, Roger Ulrich, at UCLA, pointed out that sound waves moving upward from the solar interior would be reflected back inward after reaching the solar surface. However, as a reflected sound wave descends back into the Sun, the increasing density and pressure bend the wave severely, turning it around and aiming it back toward the solar surface. In other words, sound waves bounce back and forth between the solar surface and layers deep within the Sun. These sound waves can reinforce each other if their wavelength is the right size, just as sound waves of a particular wavelength resonate inside an organ pipe.

The Sun oscillates in millions of ways as a result of waves resonating in its interior. Figure 16-5 is a computer-generated illustration of one such mode of vibration. Helioseismologists can deduce information about the solar interior from measurements of these oscillations. For example, they have been able to set limits on the amount of helium in the Sun's core and convective zone and to determine the thickness of the transition region between the radiative zone and convective zone. They have also found that the convective zone is thicker than previously thought.

The Sun's oscillations resemble those inside an organ pipe, but are much more complex

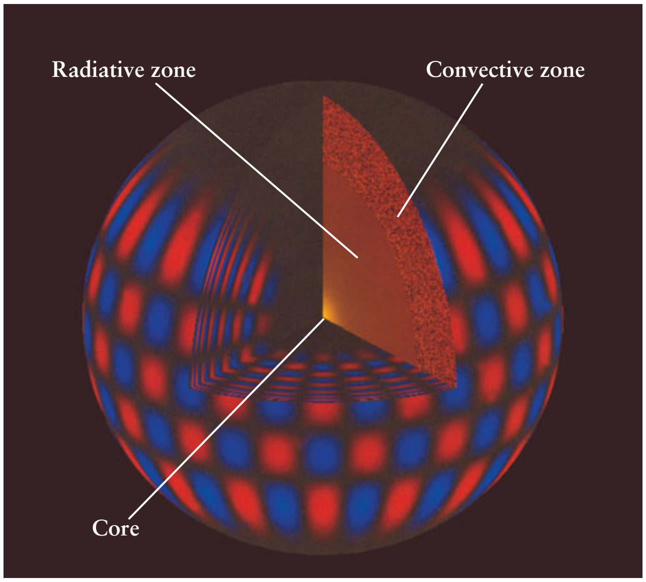


Figure 16-5

A Sound Wave Resonating in the Sun This computer-generated image shows one of the millions of ways in which the Sun's interior vibrates. The regions that are moving outward are colored blue, those moving inward, red. As the cutaway shows, these oscillations are thought to extend into the Sun's radiative zone (compare Figure 16-4). (National Solar Observatory)

16-4 Neutrinos reveal information about the Sun's core—and have surprises of their own

We have seen circumstantial evidence that thermonuclear fusion is the source of the Sun's power. To be certain, however, we need more definitive evidence. How can we show that thermonuclear fusion really is taking place in the Sun's core?

What Sunlight Cannot Tell Us

Although the light energy that we receive from the Sun originates in the core, it provides few clues about conditions there. The problem is that this energy has changed form repeatedly during its passage from the core: It appeared first as photons diffusing through the radiative zone, then as heat transported through the outer layers by convection, and then again as photons emitted from the Sun's glowing surface. As a result of these transformations, much of the information that the Sun's radiated energy once carried about conditions in the core has been lost.

ANALOGY If you make a photocopy of a photocopy of a photocopy of an original document, the final result may be so blurred as to be unreadable. In an analogous way, because solar energy is transformed many times while en route to Earth, the story it could tell us about the Sun's core is hopelessly blurred.

Helioseismology also cannot reveal what is happening in the core. The vibrations that astronomers see on the Sun's surface do not penetrate that far into the interior.

Solar Neutrinos

Happily, there is a way for scientists to learn about conditions in the Sun's core and to get direct evidence that thermonuclear fusion really does happen there. The trick is to detect the subatomic by-products of thermonuclear fusion reactions.

Scientists use the most ethereal of subatomic particles to learn about the Sun, and vice versa

As part of the process of hydrogen fusion, protons change into neutrons and release **neutrinos** (see the *Cosmic Connections* figure in Section 16-1 as well as Box 16-1). Like photons, neutrinos are particles that have no electric charge. Unlike photons, however, neutrinos interact only very weakly with matter. Even the vast bulk of the Sun offers little impediment to their passage, so neutrinos must be streaming out of the core and into space. Indeed, the conversion of hydrogen into helium at the Sun's center produces 10^{38} neutrinos each second. Every second, about 10^{14} neutrinos from the Sun—that is, **solar neutrinos**—must pass through each square meter of the Earth.

If it were possible to detect these solar neutrinos, we would have direct evidence that thermonuclear reactions really do take place in the Sun's core. Beginning in the 1960s, scientists began to build neutrino detectors for precisely this purpose.

The challenge is that neutrinos are exceedingly difficult to detect. Just as neutrinos pass unimpeded through the Sun, they also pass through the Earth almost as if it were not there. We stress the word "almost," because neutrinos can and do interact with matter, albeit infrequently.

On rare occasions a neutrino will strike a neutron and convert it into a proton. This effect was the basis of the original solar neutrino detector, designed and built by Raymond Davis of the Brookhaven National Laboratory in the 1960s. This device used 100,000 gallons of perchloroethylene (C_2Cl_4), a fluid used in dry cleaning, in a huge tank buried deep underground. Most of the solar neutrinos that entered Davis's tank passed right through it with no effect whatsoever. But occasionally a neutrino struck the nucleus of one of the chlorine atoms (^{37}Cl) in the cleaning fluid and converted one of its neutrons into a proton, creating a radioactive atom of argon (^{37}Ar).

The rate at which argon is produced is related to the neutrino flux—that is, the number of neutrinos from the Sun arriving at the Earth per square meter per second. By counting the number of newly created argon atoms, Davis was able to determine the neutrino flux from the Sun. (Other subatomic particles besides neutrinos can also induce reactions that create radioactive atoms. By placing the experiment deep underground, however, the body of the Earth absorbs essentially all such particles—with the exception of neutrinos.)

The Solar Neutrino Problem

Davis and his collaborators found that solar neutrinos created one radioactive argon atom in the tank every three days. But this rate corresponded to only one-third of the neutrino flux predicted from standard models of the Sun. This troubling discrepancy between theory and observation, called the **solar neutrino problem,** motivated scientists around the world to conduct further experiments to measure solar neutrinos.

One key question was whether the neutrinos that Davis had detected had really come from the Sun. (The Davis experiment had no way to determine the direction from which neutrinos had entered the tank of cleaning fluid.) This was resolved by an experiment in Japan called Kamiokande, which was designed by the physicist Masatoshi Koshiba. A large underground tank containing 3000 tons of water was surrounded by 1100 light detectors. From time to time, a high-energy solar neutrino struck an electron in one of the water molecules, dislodging it and sending it flying like a pin hit by a bowling ball. The recoiling electron produced a flash of light, which was sensed by the detectors. By analyzing the flashes, scientists could tell the direction from which the neutrinos were coming and confirmed that they emanated from the Sun. These results in the late 1980s gave direct evidence that thermonuclear fusion is indeed occurring in the Sun's core. (Davis and Koshiba both received the 2002 Nobel Prize in Physics for their pioneering research on solar neutrinos.)

Like Davis's experiment, however, Koshiba and his colleagues at Kamiokande detected only a fraction of the expected flux of neutrinos. Where, then, were the missing solar neutrinos?

One proposed solution had to do with the energy of the detected neutrinos. The vast majority of neutrinos from the Sun are created during the first step in the proton-proton chain, in which two protons combine to form a heavy isotope of hydrogen (see the *Cosmic Connections* figure in Section 16-1). But these neutrinos have too little energy to convert chlorine into argon. Both Davis's and Koshiba's experiments responded only to high-energy neutrinos produced by reactions that occur only part of the time near the end of the proton-proton chain. (The *Cosmic Connections* figure in Section 16-1 does not show these reactions.) Could it be that the discrepancy between theory and observation would go away if the flux of low-energy neutrinos could be measured?

To test this idea, two teams of physicists constructed neutrino detectors that used several tons of gallium (a liquid metal) rather than cleaning fluid. Low-energy neutrinos convert gallium (^{71}Ga) into a radioactive isotope of germanium (^{71}Ge). By chemically separating the germanium from the gallium and counting the radioactive atoms, the physicists were able to measure the flux of low-energy solar neutrinos. These experiments—GALLEX in Italy and SAGE (Soviet-American Gallium Experiment) in Russia—detected only 50% to 60% of the expected neutrino flux. Hence, the solar neutrino problem was a discrepancy between theory and observation for neutrinos of all energies.

Another proposed solution to the neutrino problem was that the Sun's core is cooler than predicted by solar models. If the Sun's central temperature were only 10% less than the current estimate, fewer neutrinos would be produced and the neutrino flux would agree with experiments. However, a lower central temperature would cause other obvious features, such as the Sun's size and surface temperature, to be different from what we observe.

Finding the Missing Neutrinos

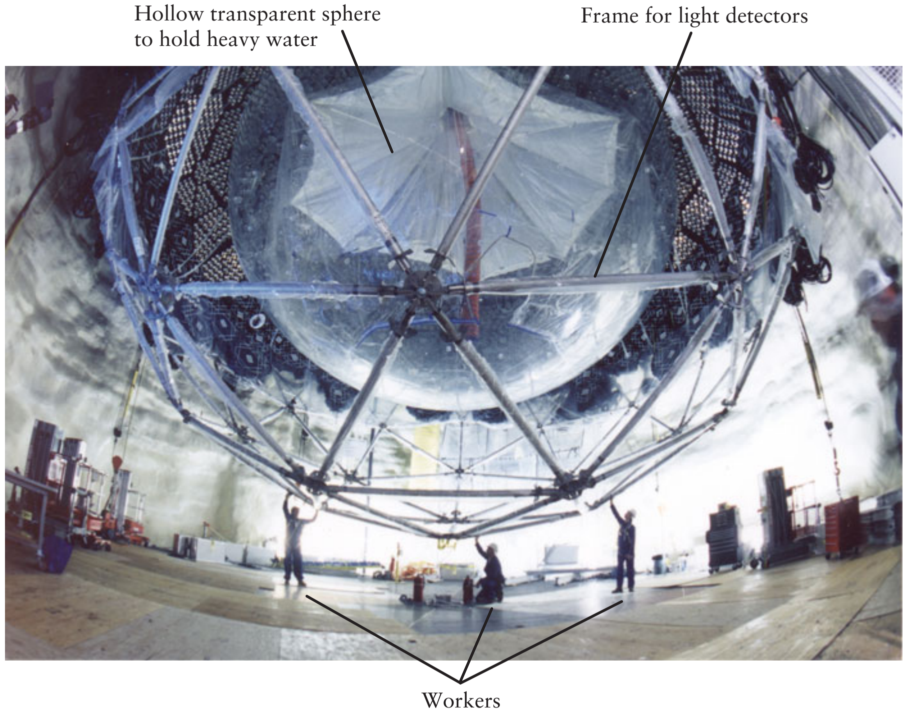
Only very recently has the solution to the neutrino problem been found. The answer lies not in how neutrinos are produced, but rather in what happens to them between the Sun's core and detectors on the Earth. Physicists have found that there are actually three types of neutrinos. Only one of these types is produced in the Sun, and it is only this type that can be detected by the experiments we have described. But if some of the solar neutrinos change in flight into a different type of neutrino, the detectors in these experiments would record only a fraction of the total neutrino flux. This effect is called **neutrino oscillation.** In June 1998, scientists at the Super-Kamiokande neutrino observatory (a larger and more sensitive device than Kamiokande) revealed evidence that neutrino oscillation does indeed take place.

The best confirmation of this idea has come from the Sudbury Neutrino Observatory (SNO) in Canada. Like Kamiokande and Super-Kamiokande, SNO uses a large tank of water placed deep underground (Figure 16-6). But unlike those earlier experiments, SNO can detect all of the three types of neutrinos. It does this by using *heavy water*. In ordinary, or "light," water, each hydrogen atom in the H_2O molecule has a solitary proton as its nucleus. In heavy water, by contrast, each of the hydrogen nuclei has a nucleus made up of a proton and a neutron. (This is the isotope ^2H shown in the *Cosmic Connections* figure in Section 16-1.) If a high-energy solar neutrino of any type passes through SNO's tank of heavy water, it can knock the neutron out of one of the ^2H nuclei. The ejected neutron can then be captured by

WEB LINK 16.7

Figure 16-6 R I V U X G

The Sudbury Neutrino Observatory Under Construction The transparent acrylic sphere holds 1000 tons of heavy water. Any of the three types of solar neutrino produces a flash of light when it interacts with the heavy water. The flash is sensed by 9600 light detectors surrounding the tank. (The detectors were not all installed when this photograph was taken.) (Photo courtesy of SNO)



another nucleus, and this capture releases energy that manifests itself as a tiny burst of light. As in Kamiokande, detectors around SNO's water tank record these light flashes.

Scientists using SNO have found that the combined flux of all three types of neutrinos coming from the Sun is *equal* to the theoretical prediction. Together with the results from earlier neutrino experiments, this strongly suggests that the Sun is indeed producing neutrinos at the predicted rate as a by-product of thermonuclear reactions. But before these neutrinos can reach the Earth, about two-thirds of them undergo an oscillation and change their type. Thus, there is really no solar neutrino problem—scientists merely needed the right kind of detectors to observe all the neutrinos, including the ones that transformed in flight.

The story of the solar neutrino problem illustrates how two different branches of science—in this case, studies of the solar interior and investigations of subatomic particles—can sometimes interact, to the mutual benefit of both. While there is still much we do not understand about the Sun and about neutrinos, a new generation of neutrino detectors in Japan, Canada, and elsewhere promises to further our knowledge of these exotic realms of astronomy and physics.

16-5 The photosphere is the lowest of three main layers in the Sun's atmosphere

Although the Sun's core is hidden from our direct view, we can easily see sunlight coming from the high-temperature gases that make up the Sun's atmosphere. These outermost layers of the Sun prove to be the sites of truly dramatic activity, much of which has a direct impact on our planet. By studying these layers, we gain further insight into the character of the Sun as a whole.

Observing the Photosphere

A visible-light photograph like Figure 16-7 makes it appear that the Sun has a definite surface. This is actually an illusion; the Sun is gaseous throughout its volume because of its high internal temperature, and the gases simply become less and less dense as you move farther away from the Sun's center.

Figure 16-7 R I V U X G

The Photosphere The photosphere is the layer in the solar atmosphere from which the Sun's visible light is emitted. Note that the Sun appears darker around its limb, or edge; here we are seeing the upper photosphere, which is relatively cool and thus glows less brightly. (The dark sunspots, which we discuss in Section 16-8, are also relatively cool regions.) (Celestron International)

Why, then, does the Sun appear to have a sharp, well-defined surface? The reason is that essentially all of the Sun's visible light emanates from a single, thin layer of gas called the **photosphere** ("sphere of light"). Just as you can see only a certain distance through the Earth's atmosphere before objects vanish in the haze, we can see only about 400 km into the photosphere. This distance is so small compared with the Sun's radius of 696,000 km that the photosphere appears to be a definite surface.

The photosphere is actually the lowest of the three layers that together constitute the solar atmosphere. Above it are the *chromosphere* and the *corona,* both of which are transparent to visible light. We can see them only using special techniques, which we discuss later in this chapter. Everything below the photosphere is called the *solar interior*.

The Lesson of Limb Darkening

The photosphere is heated from below by energy streaming outward from the solar interior. Hence, temperature should decrease as you go upward in the photosphere, just as in the solar interior (see Table 16-2 and Figure 16-3). We know this is the case because the photosphere appears darker around the edge, or *limb,* of the Sun than it does toward the center of the solar disk, an effect called **limb darkening** (examine Figure 16-7). This happens because when we look near the Sun's limb, we do not see as deeply into the photosphere as we do when we look near the center of the disk (Figure 16-8). The high-altitude gas we observe at the limb is not as hot and thus does not glow as brightly as the deeper, hotter gas seen near the disk center.

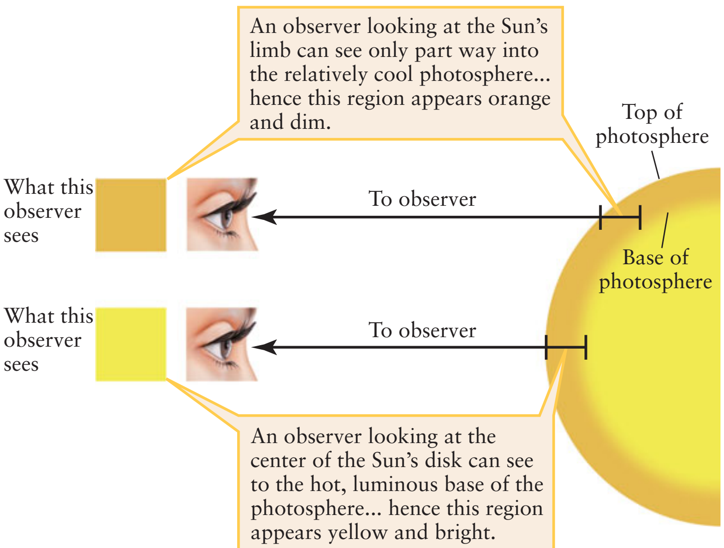


Figure 16-8

The Origin of Limb Darkening Light from the Sun's limb and light from the center of its disk both travel about the same straight-line distance through the photosphere to reach us. Because of the Sun's curved shape, light from the limb comes from a greater height within the photosphere, where the temperature is lower and the gases glow less brightly. Hence, the limb appears darker and more orange.

The spectrum of the Sun's photosphere confirms how its temperature varies with altitude. As we saw in Section 16-1, the photosphere shines like a nearly perfect blackbody with an average temperature of about 5800 K. However, superimposed on this spectrum are many dark absorption lines (see Figure 5-12). As discussed in Section 5-6, we see an *absorption line spectrum* of this sort whenever we view a hot, glowing object through a relatively cool gas. In this case, the hot object is the lower part of the photosphere; the cooler gas is in the upper part of the photosphere, where the temperature declines to about 4400 K. All the absorption lines in the Sun's spectrum are produced in this relatively cool layer, as atoms selectively absorb photons of various wavelengths streaming outward from the hotter layers below.

CAUTION! You may find it hard to think of 4400 K as "cool." But keep in mind that the ratio of 4400 K to 5800 K, the temperature in the lower photosphere, is the same as the ratio of the temperature on a Siberian winter night to that of a typical day in Hawaii.

Granules and Supergranules in the Photosphere

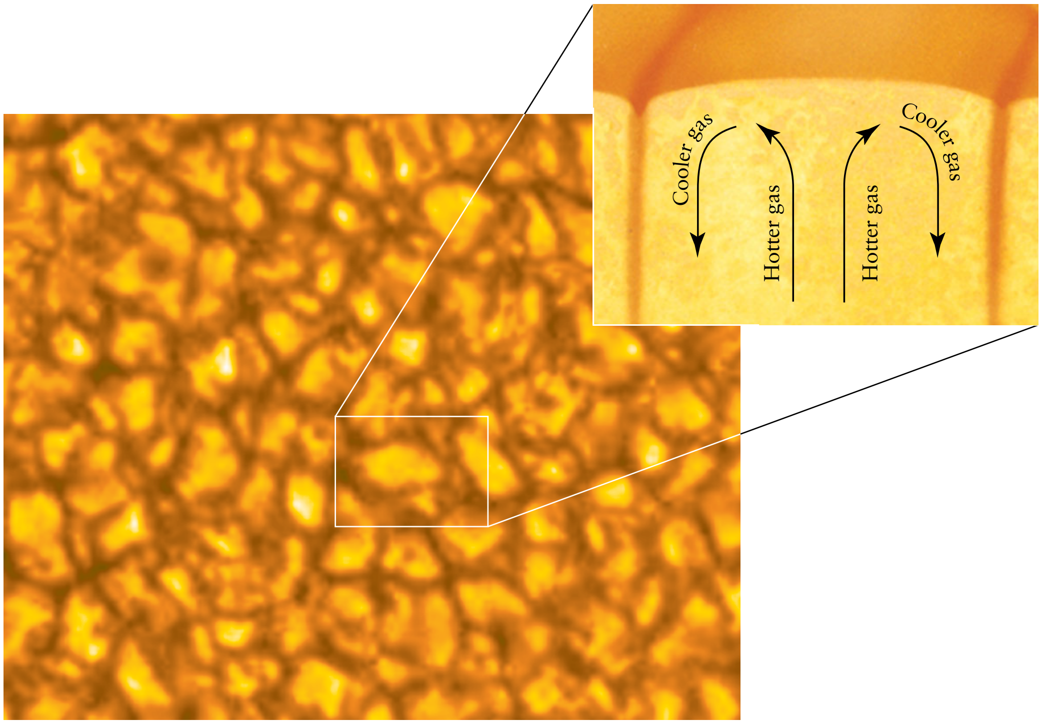
We can learn still more about the photosphere by examining it with a telescope—but only when using special dark filters to prevent eye damage. ***Looking directly at the Sun without the correct filter, whether with the naked eye or with a telescope, can cause permanent blindness!*** Under good observing conditions, astronomers using such filter-equipped telescopes can often see a blotchy pattern in the photosphere, called **granulation** (Figure 16-9). Each light-colored **granule** measures about 1000 km (600 mi) across—equal in size to the areas of Texas and Oklahoma combined—and is surrounded by a darkish boundary. The difference in brightness between the center and the edge of a granule corresponds to a temperature drop of about 300 K.

Like water boiling on a stove, the photosphere bubbles with convection cells

Granulation is caused by convection of the gas in the photosphere. The inset in Figure 16-9 shows how gas from lower levels rises upward in granules, cools off, spills over the edges of the granules, and then plunges back down into the Sun. This can occur only if the gas is heated from below, like a pot of water being heated on a stove (see Section 16-2). Along with limb darkening and the Sun's absorption line spectrum, granulation shows that the upper part of the photosphere must be cooler than the lower part.

Time-lapse photography reveals more of the photosphere's dynamic activity. Granules form, disappear, and reform in cycles lasting only a few minutes. At any one time, about 4 million granules cover the solar surface.

Superimposed on the pattern of granulation are even larger convection cells called **supergranules** (Figure 16-10). As in granules, gases rise upward in the middle of a supergranule, move horizontally outward toward its edge, and descend back into the Sun. The difference is that a typical supergranule is about 35,000 km in diameter, large enough to enclose several hundred granules. This large-scale convection moves at only about 0.4 km/s (1400 km/h, or 900 mi/h), about one-tenth the speed of gases churning in a granule. A given supergranule lasts about a day.



VIDEO 16.1 VIDEO 16.2

Figure 16-9 R I V U X G

Solar Granulation High-resolution photographs of the Sun's surface reveal a blotchy pattern called granulation. Granules are convection cells about 1000 km (600 mi) wide in the Sun's photosphere. **Inset:** Rising hot gas produces bright granules. Cooler gas sinks downward along the boundaries between granules; this gas glows less brightly, giving the boundaries their dark appearance. This convective motion transports heat from the Sun's interior outward to the solar atmosphere. (MSFC/NASA; inset: Goran Scharmer, Lund Observatory)

ANALOGY Similar patterns of large-scale and small-scale convection can be found in the Earth's atmosphere. On the large scale, air rises gradually at a low-pressure area, then sinks gradually at a high-pressure area, which might be hundreds of kilometers away. This is analogous to the flow in a supergranule. Thunderstorms in our atmosphere are small but intense convection cells within which air moves rapidly up and down. Like granules, they last only a relatively short time before they dissipate.

The Photosphere: Hot, Thin, and Opaque

Although the photosphere is a very active place, it actually contains relatively little material. Careful examination of the spectrum shows that it has a density of only about 10^{-4} kg/m^3,

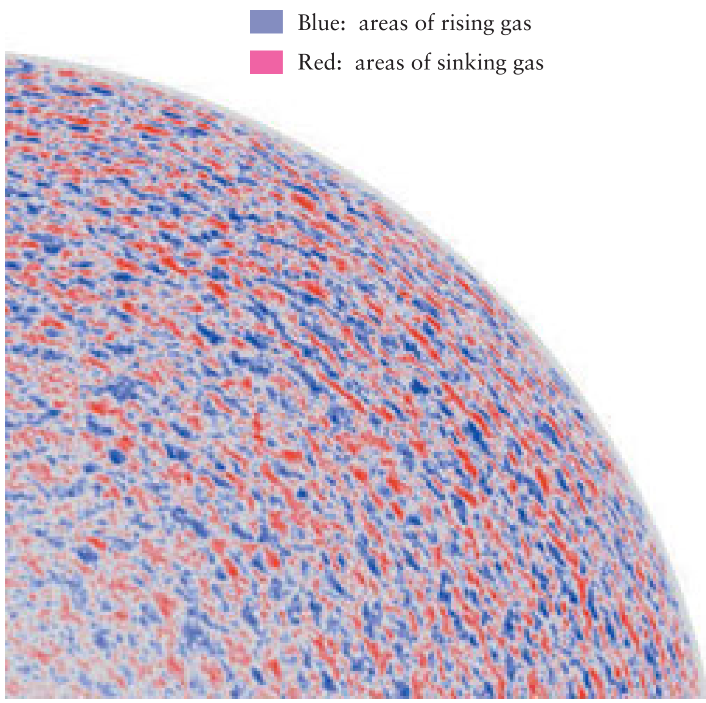


Figure 16-10 R I V U X G

Supergranules and Large-Scale Convection Supergranules display relatively little contrast between their center and edges, so they are hard to observe in ordinary images. But they can be seen in a false-color Doppler image like this one. Light from gas that is approaching us (that is, rising) is shifted toward shorter wavelengths, while light from receding gas (that is, descending) is shifted toward longer wavelengths (see Section 5-9). (David Hathaway, MSFC/NASA) roughly 0.01% the density of the Earth's atmosphere at sea level. The photosphere is made primarily of hydrogen and helium, the most abundant elements in the solar system (see Figure 8-3).

Despite being such a thin gas, the photosphere is surprisingly opaque to visible light. If it were not so opaque, we could see into the Sun's interior to a depth of hundreds of thousands of kilometers, instead of the mere 400 km that we can see down into the photosphere. What makes the photosphere so opaque is that its hydrogen atoms sometimes acquire an extra electron, becoming **negative hydrogen ions.** The extra electron is only loosely attached and can be dislodged if it absorbs a photon of any visible wavelength. Hence, negative hydrogen ions are very efficient light absorbers, and there are enough of these light-absorbing ions in the photosphere to make it quite opaque. Because it is so opaque, the photosphere's spectrum is close to that of an ideal blackbody.

16-6 Spikes of rising gas extend through the Sun's chromosphere

An ordinary visible-light image such as Figure 16-7 gives the impression that the Sun ends at the top of the photosphere. But during a total solar eclipse, the Moon blocks the photosphere from our view, revealing a glowing, pinkish layer of gas above the photosphere (Figure 16-11). This is the tenuous **chromosphere** ("sphere of color"), the second of the three major levels in the Sun's atmosphere. The chromosphere is only about one ten-thousandth (10^{-4}) as dense as the photosphere, or about 10^{-8} as dense as our own atmosphere. No wonder it is normally invisible!

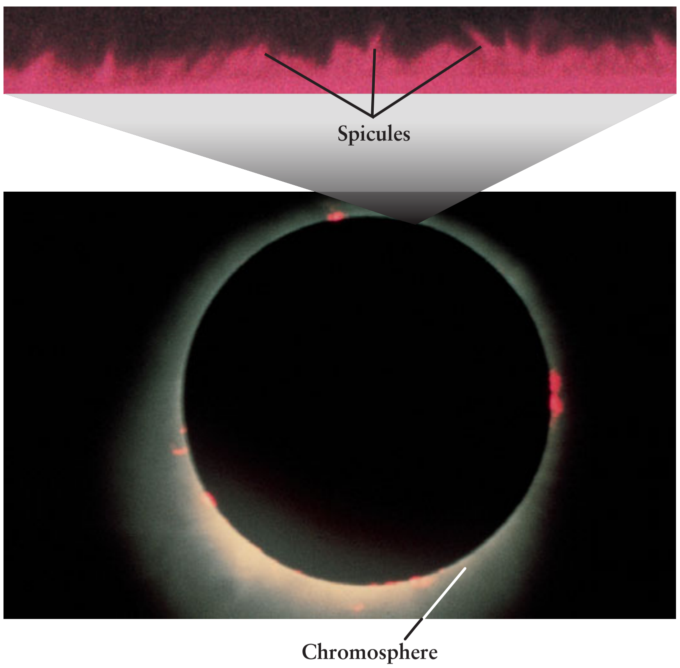


Figure 16-11 R I V U X G

The Chromosphere During a total solar eclipse, the Sun's glowing chromosphere can be seen around the edge of the Moon. It appears pinkish because its hot gases emit light at only certain discrete wavelengths, principally the H_α emission of hydrogen at a red wavelength of 656.3 nm. The expanded area above shows spicules, jets of chromospheric gas that surge upward into the Sun's outer atmosphere. (NOAO)

Comparing the Chromosphere and Photosphere

Unlike the photosphere, which has an absorption line spectrum, the chromosphere has a spectrum dominated by emission lines. An emission line spectrum is produced by the atoms of a hot, thin gas (see Section 5-6 and Section 5-8). As their electrons fall from higher to lower energy levels, the atoms emit photons.

One of the strongest emission lines in the chromosphere's spectrum is the H_α line at 656.3 nm, which is emitted by a hydrogen atom when its single electron falls from the $n = 3$ level to the $n = 2$ level (recall Figure 5-21*b*). This wavelength is in the red part of the spectrum, which gives the chromosphere its characteristic pinkish color. The spectrum also contains emission lines of singly ionized calcium, as well as lines due to ionized helium and ionized metals. In fact, helium was originally discovered in the chromospheric spectrum in 1868, almost 30 years before helium gas was first isolated on Earth.

Analysis of the chromospheric spectrum shows that temperature *increases* with increasing height in the chromosphere. This is just the opposite of the situation in the photosphere, where temperature decreases with increasing height. The temperature is about 4400 K at the top of the photosphere; 2000 km higher, at the top of the chromosphere, the temperature is nearly 25,000 K. This is very surprising, since temperatures should decrease as you move away from the Sun's interior. In Section 16-10 we will see how solar scientists explain this seeming paradox.

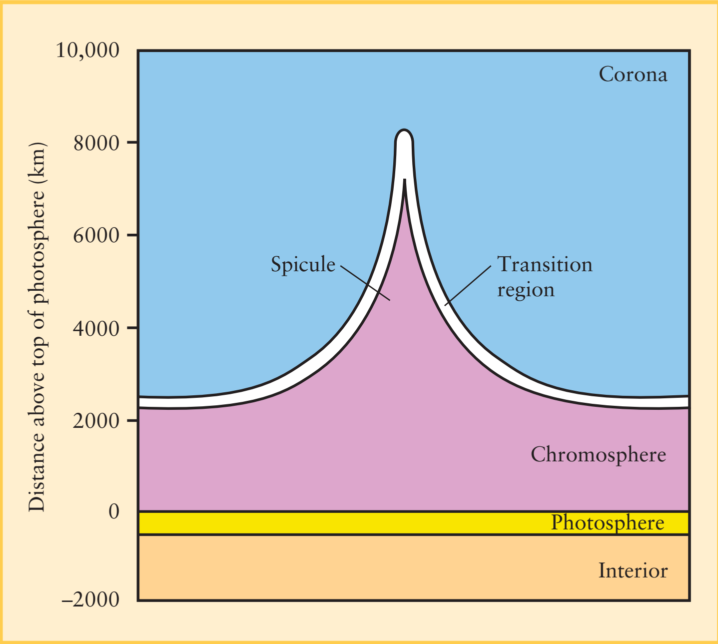
WEB LINK 16.8

The photospheric spectrum is dominated by absorption lines at certain wavelengths, while the spectrum of the chromosphere has emission lines at these same wavelengths. In other words, the photosphere appears dark at the wavelengths at which the chromosphere emits most strongly, such as the H_α wavelength of 656.3 nm. By viewing the Sun through a special filter that is transparent to light only at the wavelength of H_α, astronomers can screen out light from the photosphere and make the chromosphere visible. (The same technique can be used with other wavelengths at which the chromosphere emits strongly, including nonvisible wavelengths.) This makes it possible to see the chromosphere at any time, not just during a solar eclipse.

Spicules

The top photograph in Figure 16-11 is a high-resolution image of the Sun's chromosphere taken through an H_α filter. This image shows numerous vertical spikes, which are actually jets of rising gas called **spicules.** A typical spicule lasts just 15 minutes or so: It rises at the rate of about 20 km/s (72,000 km/h, or 45,000 mi/h), can reach a height of several thousand kilometers, and then collapses and fades away (Figure 16-12). Approximately 300,000 spicules exist at any one time, covering about 1% of the Sun's surface.

Jets of gas thousands of kilometers in height rise through the chromosphere



INTERACTIVE EXERCISE 16.2

Figure 16-12

The Solar Atmosphere This schematic diagram shows the three layers of the solar atmosphere. The lowest, the photosphere, is about 400 km thick. The chromosphere extends about 2000 km higher, with spicules jutting up to nearly 10,000 km above the photosphere. Above a transition region is the Sun's outermost layer, the corona, which we discuss in Section 16-7. It extends many millions of kilometers out into space. (Adapted from J. A. Eddy)

Spicules are generally located directly above the edges of supergranules (see Figure 16-10). This is a surprising result, because chromospheric gases are rising in a spicule while photospheric gases are *descending* at the edge of a supergranule. What, then, is pulling gases upward to form spicules? As we will see in Section 16-10, the answer proves to be the Sun's intense magnetic field. But before we delve into how this happens, let us complete our tour of the solar atmosphere by exploring its outermost, least dense, most dynamic, and most bizarre layer—the region called the corona.

16-7 The corona ejects mass into space to form the solar wind

The **corona,** or outermost region of the Sun's atmosphere, begins at the top of the chromosphere. It extends out to a distance of several million kilometers. Despite its tremendous extent, the corona is only about one-millionth (10^{-6}) as bright as the photosphere—no brighter than the full moon. Hence, the corona can be viewed only when the light from the photosphere is blocked out, either by use of a specially designed telescope or during a total eclipse.

Figure 16-13 is an exceptionally detailed photograph of the Sun's corona taken during a solar eclipse. It shows that the corona is not merely a spherical shell of gas surrounding the Sun. Rather, numerous streamers extend in different directions far above the solar surface. The shapes of these streamers vary on time scales of days or weeks. (For another view of the corona during a solar eclipse, see Figure 3-10*b*.)

Comparing the Corona, Chromosphere, and Photosphere

Like the chromosphere that lies below it, the corona has an emission line spectrum characteristic of a hot, thin gas. When the spectrum of the corona was first measured in the nineteenth century, astronomers found a number of emission lines at wavelengths that had never been seen in the laboratory. Their explanation was that the corona contained elements that had not yet been detected on the Earth. However, laboratory experiments in the 1930s revealed that these unusual emission lines were in fact caused by the same atoms found elsewhere in the universe—but in highly ionized states. For example, a prominent green line at 530.3 nm is caused by highly ionized iron atoms, each of which has been stripped of 13 of its 26 electrons. In order to strip that many electrons from atoms, temperatures in the corona must reach 2 million kelvins (2×10^6 K) or even higher—far greater than the temperatures in the chromosphere. Figure 16-14 shows how temperature in both the chromosphere and corona varies with altitude.

CAUTION! The corona is actually not very "hot"—that is, it contains very little thermal energy. The reason is that the corona is nearly a vacuum. In the corona there are only about 10^{11} atoms

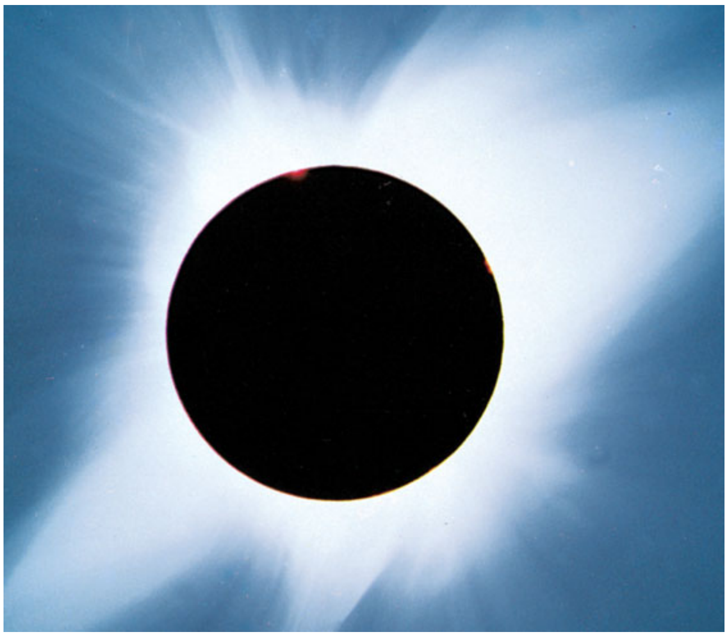


Figure 16-13 R I V U X G

The Solar Corona This striking photograph of the corona was taken during the total solar eclipse of July 11, 1991. Numerous streamers extend for millions of kilometers above the solar surface. The unearthly light of the corona is one of the most extraordinary aspects of experiencing a solar eclipse. (Courtesy of R. Christen and M. Christen, Astro-Physics, Inc.)

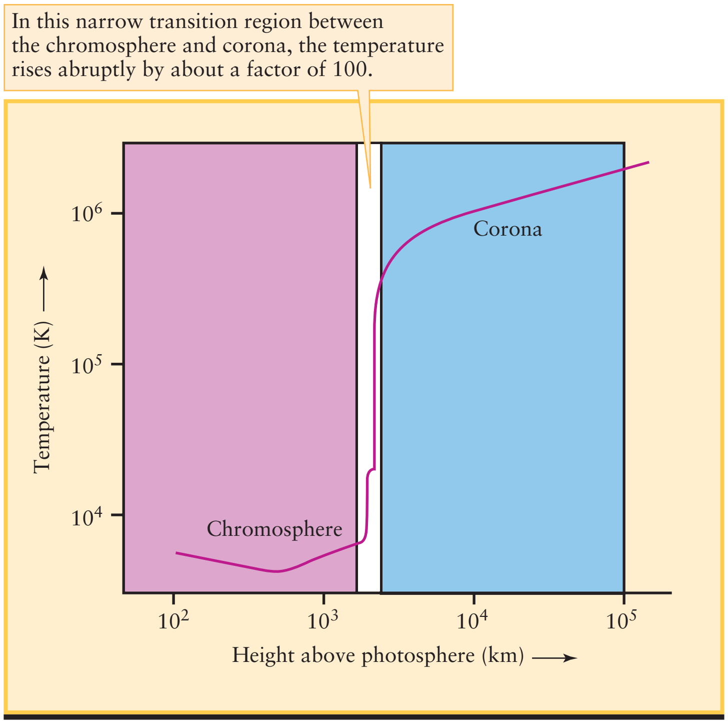


Figure 16-14

Temperatures in the Sun's Upper Atmosphere This graph shows how temperature varies with altitude in the Sun's chromosphere and corona and in the narrow transition region between them. In order to show a large range of values, both the vertical and horizontal scales are nonlinear. (Adapted from A. Gabriel)

per cubic meter, compared with about 10^{23} atoms per cubic meter in the Sun's photosphere and about 10^{25} atoms per cubic meter in the air that we breathe. Because of the corona's high temperature, the atoms there are moving at very high speeds. But because there are so few atoms in the corona, the total amount of energy in these moving atoms (a measure of how "hot" the gas is) is rather low. If you flew a spaceship into the corona, you would have to worry about becoming overheated by the intense light coming from the photosphere, but you would notice hardly any heating from the corona's ultra-thin gas.

ANALOGY The situation in the corona is similar to that inside a conventional oven that is being used for baking. Both the walls of the oven and the air inside the oven are at the same high temperature, but the air contains very few atoms and thus carries little energy. If you put your hand in the oven momentarily, the lion's share of the heat you feel is radiation from the oven walls.

The low density of the corona explains why it is so dim compared with the photosphere. In general, the higher the temperature of a gas, the brighter it glows. But because there are so few atoms in the corona, the net amount of light that it emits is very feeble compared with the light from the much cooler, but also much denser, photosphere.

The Solar Wind and Coronal Holes

The Earth's gravity keeps our atmosphere from escaping into space. In the same way, the Sun's powerful gravitational attraction keeps most of the gases of the photosphere, chromosphere, and corona from escaping. But the corona's high temperature means that its atoms and ions are moving at very high speeds, around a million kilometers per hour. As a result, some of the coronal gas can and does escape. This outflow of gas, which we first encountered in Section 8-5, is called the **solar wind.**

Each second the Sun ejects about a million tons (10^9 kg) of material into the solar wind. But the Sun is so massive that, even over its entire lifetime, it will eject only a few tenths of a percent of its total mass. The solar wind is composed almost entirely of electrons and nuclei of hydrogen and helium. About 0.1% of the solar wind is made up of ions of more massive atoms, such as silicon, sulfur, calcium, chromium, nickel, iron, and argon. The aurorae seen at far northern or southern latitudes on Earth are produced when electrons and ions from the solar wind enter our upper atmosphere.

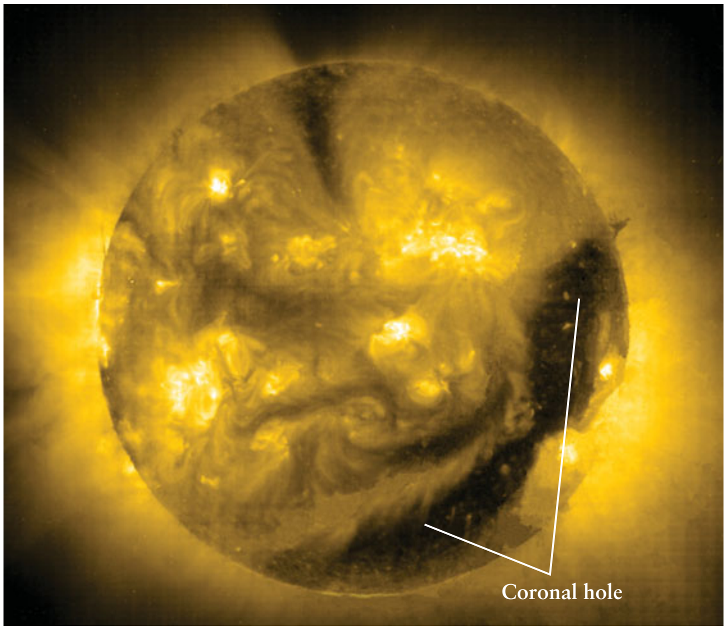
Special telescopes enable astronomers to see the origin of the solar wind. To appreciate what sort of telescopes are needed, note that because the temperature of the coronal gas is so high, ions in the corona are moving very fast (see Box 7-2). When ions collide, the energy of the impact is so great that the ion's electrons are boosted to very high energy levels. As the electrons fall back to lower levels, they emit high-energy photons in the ultraviolet and X-ray portions of the spectrum—wavelengths at which the photosphere and chromosphere are relatively dim. Hence, telescopes sensitive to these short wavelengths are ideal for studying the corona and the flow of the solar wind.

WEB LINK 16.9 The Earth's atmosphere is opaque to ultraviolet light and X rays, so telescopes for these wavelengths must be placed on board spacecraft (see Section 6-7, especially Figure 6-25). Figure 16-15 shows an ultraviolet view of the corona from the *SOHO* spacecraft (*Solar and Heliospheric Observatory*), a joint project of the European Space Agency (ESA) and NASA.

Figure 16-15 reveals that the corona is not uniform in temperature or density. The densest, highest-temperature regions appear bright, while the thinner, lower-temperature regions are dark. Note the large dark area, called a **coronal hole** because it is almost devoid of luminous gas. Particles streaming away from the Sun can most easily flow outward through these particularly thin regions. Therefore, it is thought that coronal holes are the main corridors through which particles of the solar wind escape from the Sun.

Unlike the lower levels of the Sun's atmosphere, the corona has immense holes that shift and reshape

WEB LINK 16.10 Evidence in favor of this picture has come from the *Ulysses* spacecraft, another joint ESA/NASA mission. In 1994 and 1995, *Ulysses* became the first spacecraft to fly over the Sun's north and south poles, where there are apparently permanent coronal holes. The spacecraft indeed measured a stronger solar wind emanating from these holes.



VIDEO 16.3

Figure 16-15 R I V U X G

The Ultraviolet Corona The *SOHO* spacecraft recorded this false-color ultraviolet view of the solar corona. The dark feature running across the Sun's disk from the top is a coronal hole, a region where the coronal gases are thinner than elsewhere. Such holes are often the source of strong gusts in the solar wind. (*SOHO*/EIT/ESA/NASA)

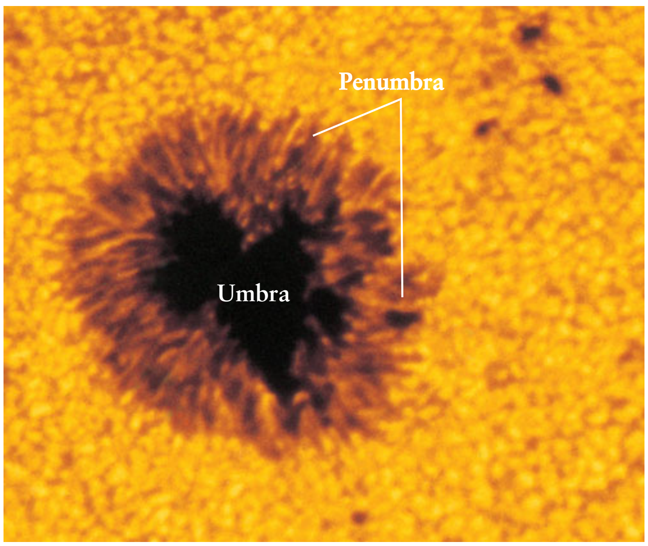
The temperatures in the corona and the chromosphere are not at all what we would expect. Just as you feel warm if you stand close to a campfire but cold if you move away, we would expect that the temperature in the corona and chromosphere would *decrease* with increasing altitude and, hence, increasing distance from the warmth of the Sun's photosphere. Why, then, does the temperature in these regions *increase* with increasing altitude? This has been one of the major unsolved mysteries in astronomy for the past half-century. As astronomers have tried to resolve this dilemma, they have found important clues in one of the Sun's most familiar features—sunspots.

16-8 Sunspots are low-temperature regions in the photosphere

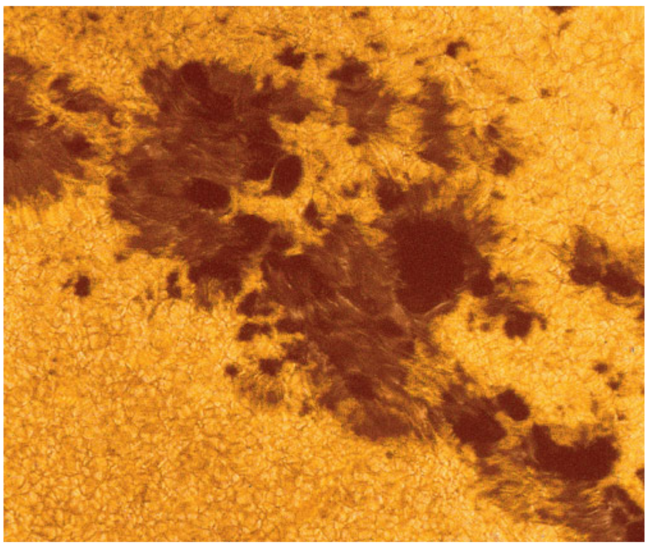
Granules, supergranules, spicules, and the solar wind occur continuously. These features are said to be aspects of the *quiet* Sun. But other, more dramatic features appear periodically, including massive eruptions and regions of concentrated magnetic fields. When these are present, astronomers refer to the *active* Sun. The features of the active Sun that can most easily be seen with even a small telescope (although only with an appropriate filter attached) are sunspots.

Observing Sunspots

Sunspots are irregularly shaped dark regions in the photosphere. Sometimes sunspots appear in isolation (Figure 16-16*a*), but frequently they are found in sunspot groups (Figure 16-16*b*; see also Figure 16-7). Although sunspots vary greatly in size, typical ones



(a)



(b)

VIDEO 16.4

Figure 16-16 R I V U X G

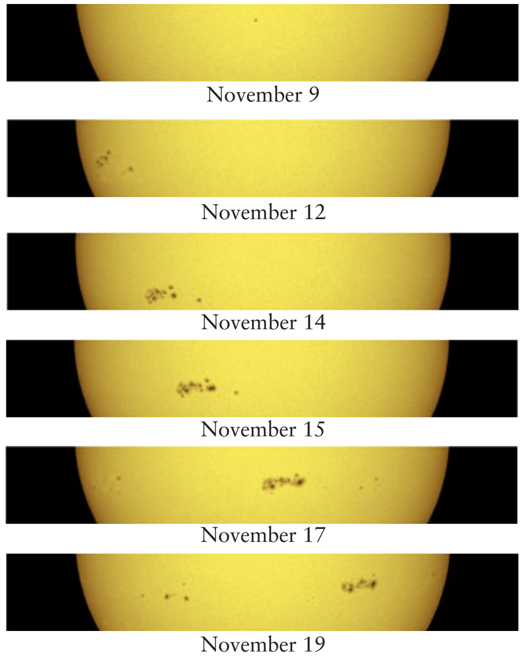
Sunspots **(a)** This high-resolution photograph of the photosphere shows a mature sunspot. The dark center of the spot is called the umbra. It is bordered by the penumbra, which is less dark and has a featherlike appearance. **(b)** In this view of a typical sunspot group, several sunspots are close enough to overlap. In both images you can see granulation in the surrounding, undisturbed photosphere. (NOAO) measure a few tens of thousands of kilometers across—comparable to the diameter of the Earth. Sunspots are not permanent features of the photosphere but last between a few hours and a few months.

Each sunspot has a dark central core, called the *umbra*, and a brighter border called the *penumbra*. We used these same terms in Section 3-4 to refer to different parts of the Earth's or the Moon's shadow. But a sunspot is not a shadow; it is a region in the photosphere where the temperature is relatively low, which makes it appear darker than its surroundings. If the surrounding photosphere is blocked from view, a sunspot's umbra appears red and the penumbra appears orange. As we saw in Section 5-4, Wien's law relates the color of a blackbody (which depends on the wavelength at which it emits the most light) to the blackbody's temperature. The colors of a sunspot indicate that the temperature of the umbra is typically 4300 K and that of the penumbra is typically 5000 K. While high by earthly standards, these temperatures are quite a bit lower than the average photospheric temperature of 5800 K.

The Stefan-Boltzmann law (see Section 5-4) tells us that the energy flux from a blackbody is proportional to the fourth power of its temperature. This law lets us compare the amounts of light energy emitted by a square meter of a sunspot's umbra and by a square meter of undisturbed photosphere. The ratio is:

$$\frac{\text{flux from umbra}}{\text{flux from photosphere}} = \left(\frac{4300\ \text{K}}{5800\ \text{K}}\right)^4 = 0.30$$

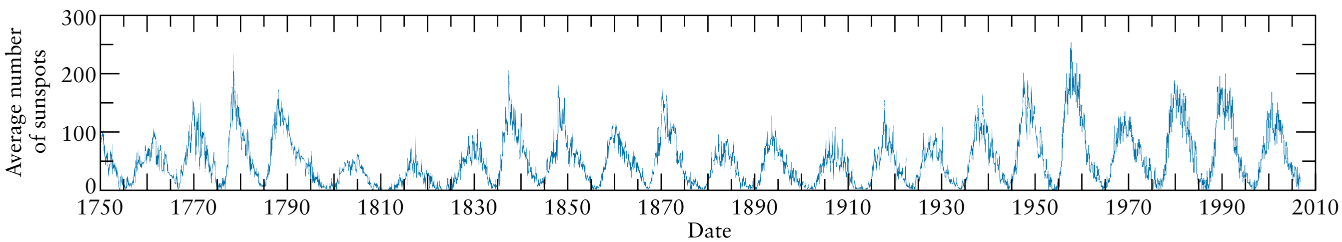
That is, the umbra emits only 30% as much light as an equally large patch of undisturbed photosphere. This is why sunspots appear so dark.



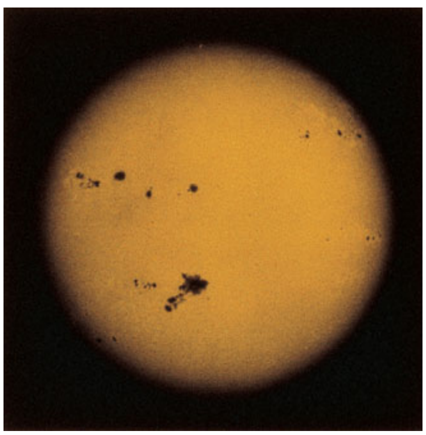
VIDEO 16.5

Figure 16-17 R I V U X G

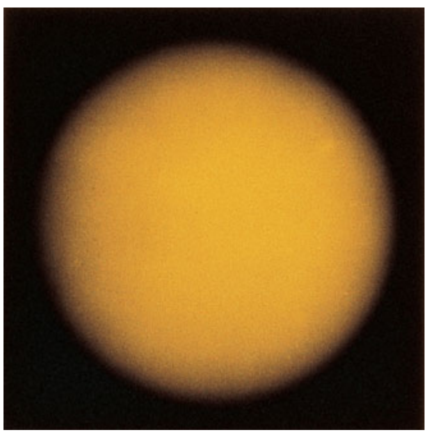
Tracking the Sun's Rotation with Sunspots This series of photographs taken in 1999 shows the rotation of the Sun. By observing the same group of sunspots from one day to the next, Galileo found that the Sun rotates once in about four weeks. (The equatorial regions of the Sun actually rotate somewhat faster than the polar regions.) Notice how the sunspot group shown here changed its shape. (The Carnegie Observatories)



(a)



(b) Near sunspot maximum



(c) Near sunspot minimum

WEB LINK 16.11

Figure 16-18 R I V U X G

The Sunspot Cycle **(a)** The number of sunspots on the Sun varies with a period of about 11 years. The most recent sunspot maximum occurred in 2000. **(b)** This photograph, taken near sunspot maximum in 1989, shows a number of sunspots and large sunspot groups. The sunspot group visible near the bottom of the Sun's disk has about the same diameter as the planet Jupiter. **(c)** Near sunspot minimum, as in this 1986 photograph, essentially no sunspots are visible. (NOAO)

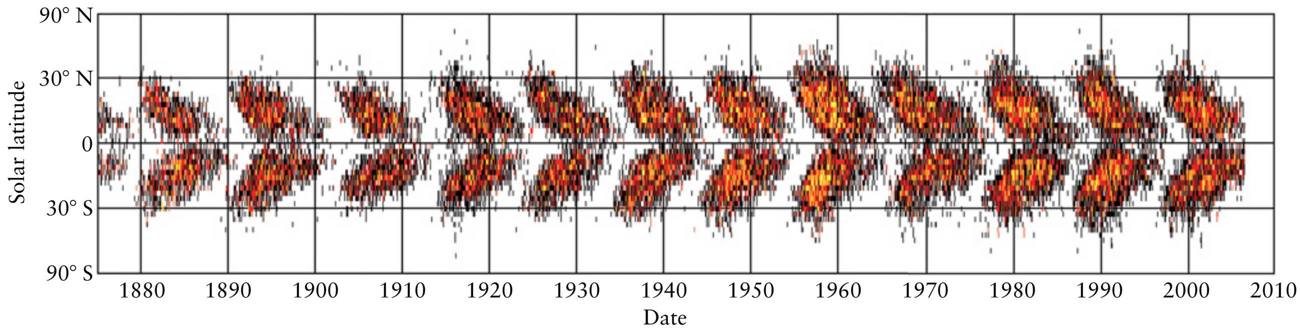


Figure 16-19

Variations in the Average Latitude of Sunspots The dots in this graph (sometimes called a "butterfly diagram") record how far north or south of the Sun's equator sunspots were observed. At the beginning of each sunspot cycle, most sunspots are found near latitudes 30° north or south. As the cycle goes on, sunspots typically form closer to the equator. (NASA Marshall Space Flight Center)

Sunspots and the Sun's Rotation

Occasionally, a sunspot group is large enough to be seen without a telescope. Chinese astronomers recorded such sightings 2000 years ago, and huge sunspot groups visible to the naked eye (with an appropriate filter) were seen in 1989 and 2003. But it was not until Galileo introduced the telescope into astronomy (see Section 4-5) that anyone was able to examine sunspots in detail.

Galileo discovered that he could determine the Sun's rotation rate by tracking sunspots as they moved across the solar disk (Figure 16-17). He found that the Sun rotates once in about four weeks. A typical sunspot group lasts about two months, so a specific one can be followed for two solar rotations.

Further observations by the British astronomer Richard Carrington in 1859 demonstrated that the Sun does not rotate as a rigid body. Instead, the equatorial regions rotate more rapidly than the polar regions. This phenomenon is known as **differential rotation.** Thus, while a sunspot near the solar equator takes only 25 days to go once around the Sun, a sunspot at 30° north or south of the equator takes $27\frac{1}{2}$ days. The rotation period at 75° north or south is about 33 days, while near the poles it may be as long as 35 days.

The Sunspot Cycle

The average number of sunspots on the Sun is not constant, but varies in a predictable **sunspot cycle** (Figure 16-18*a*). This phenomenon was first reported by the German astronomer Heinrich Schwabe in 1843 after many years of observing. As Figure 16-18*a* shows, the average number of sunspots varies with a period of about 11 years. A period of exceptionally many sunspots is a **sunspot maximum** (Figure 16-18*b*), as occurred in 1979, 1989, and 2000. Conversely, the Sun is almost devoid of sunspots at a **sunspot minimum** (Figure 16-18*c*), as occurred in 1976, 1986, and 1996 and is projected to occur in 2007.

The number of sunspots increases and decreases on an 11-year cycle

The locations of sunspots also vary with the same 11-year sunspot cycle. At the beginning of a cycle, just after a sunspot minimum, sunspots first appear at latitudes around 30° north and south of the solar equator (Figure 16-19). Over the succeeding years, the sunspots occur closer and closer to the equator.

16-9 Sunspots are produced by a 22-year cycle in the Sun's magnetic field

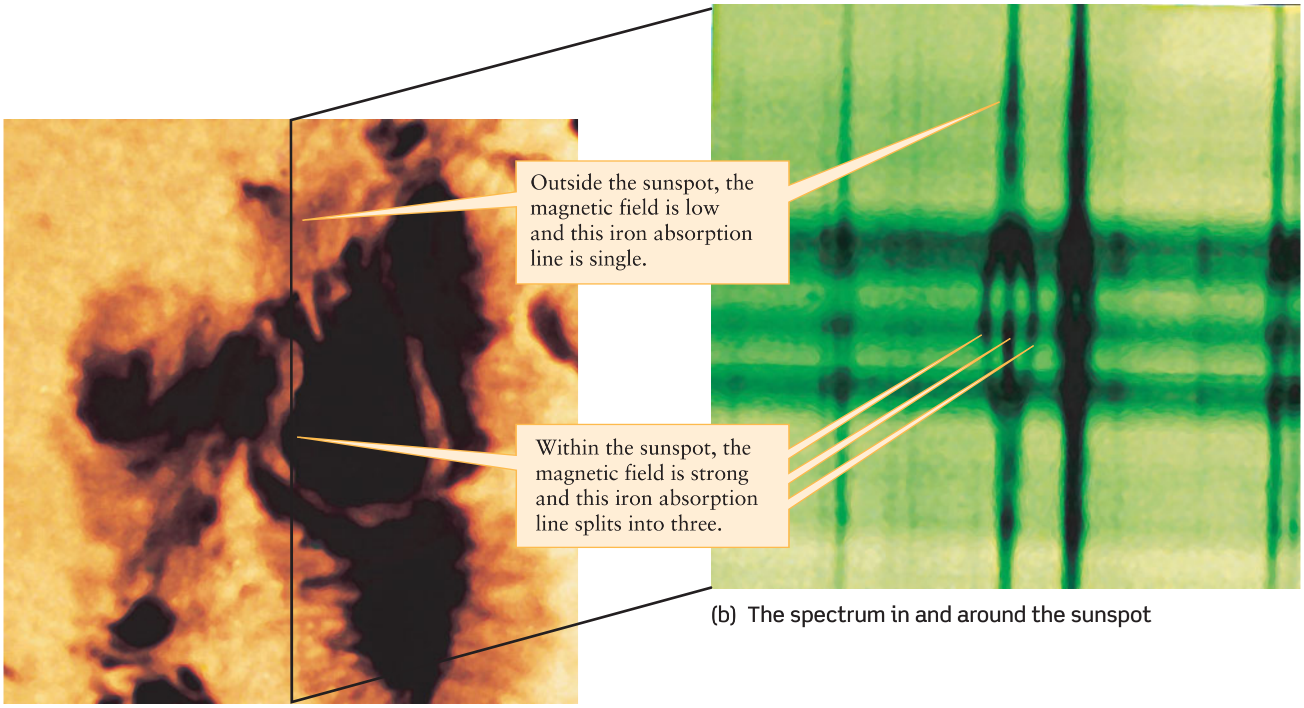
Why should the number of sunspots vary with an 11-year cycle? Why should their average latitude vary over the course of a cycle? And why should sunspots exist at all? The first step toward answering these questions came in 1908, when the American astronomer George Ellery Hale discovered that sunspots are associated with intense magnetic fields on the Sun.

Probing Solar Magnetism

LOOKING DEEPER 16.2

When Hale focused a spectroscope on sunlight coming from a sunspot, he found that many spectral lines appear to be split into several closely spaced lines (Figure 16-20). This "splitting" of spectral lines is called the **Zeeman effect,** after the Dutch physicist Pieter Zeeman, who first observed it in his laboratory in 1896. Zeeman showed that a spectral line splits when the atoms are subjected to an intense magnetic field. The more intense the magnetic field, the wider the separation of the split lines.

Hale's discovery showed that sunspots are places where the hot gases of the photosphere are bathed in a concentrated magnetic field. Many of the atoms of the Sun's atmosphere are ionized due to the high temperature. The solar atmosphere is thus a special type of gas called a **plasma,** in which electrically charged ions and electrons can move freely. Like any moving, electrically charged objects, they can be deflected by magnetic fields. Figure 16-21 shows how a magnetic field in the laboratory bends a beam of fast-moving electrons into a curved trajectory. Similarly, the paths of moving ions and electrons in the photosphere are deflected by the Sun's magnetic field. In particular, magnetic forces act on the hot plasma that rises from the Sun's interior due to convection. Where the magnetic field is particularly strong, these

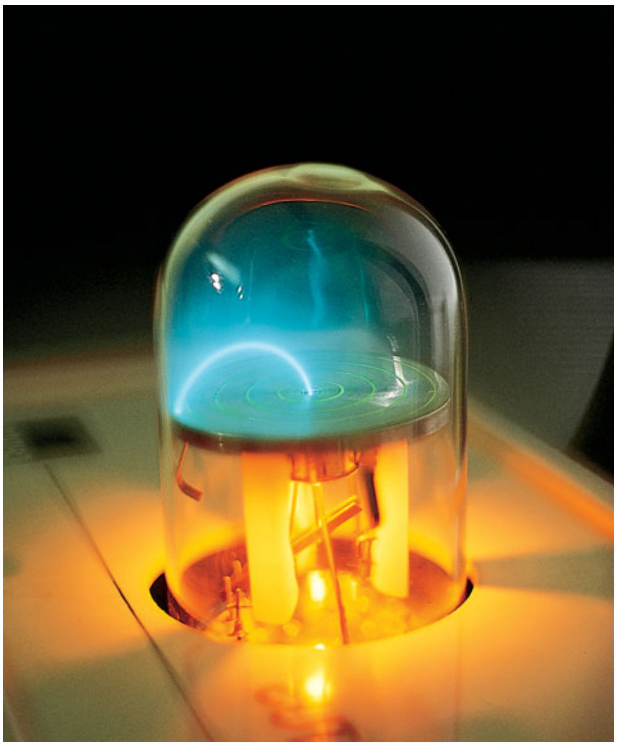


(b) The spectrum in and around the sunspot

(a) A sunspot

Figure 16-20 R I V U X G

Sunspots Have Strong Magnetic Fields **(a)** A black line in this image of a sunspot shows where the slit of a spectrograph was aimed. **(b)** This is a portion of the resulting spectrum, including a dark absorption line caused by iron atoms in the photosphere. The splitting of this line by the sunspot's magnetic field can be used to calculate the field strength. Typical sunspot magnetic fields are over 5000 times stronger than the Earth's field at its north and south poles. (NOAO)

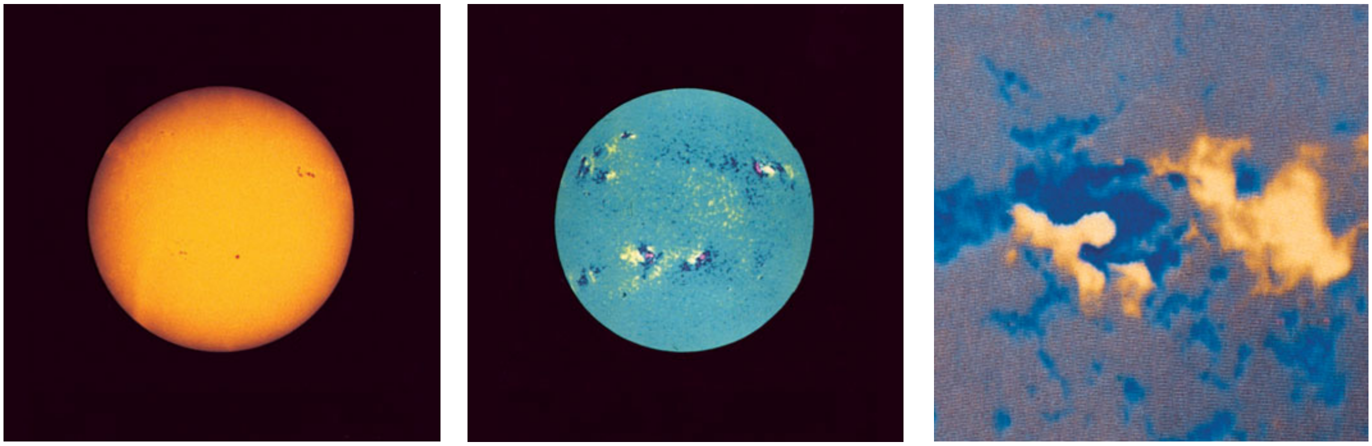


forces push the hot plasma away. The result is a localized region where the gas is relatively cool and thus glows less brightly—in other words, a sunspot.

To get a fuller picture of the Sun's magnetic fields, astronomers take images of the Sun at two wavelengths, one just less than and one just greater than the wavelength of a magnetically split spectral line. From the difference between these two images, they can construct a picture called a **magnetogram,** which displays the magnetic fields in the solar atmosphere. Figure 16-22*a* is an ordinary white-light photograph of the Sun taken at the same time as the magnetogram in Figure 16-22*b*. In the magnetogram, dark blue indicates areas of the photosphere with one magnetic polarity (north), and yellow indicates areas with the opposite (south) magnetic polarity. This image shows that many sunspot groups have roughly comparable areas covered by north

Figure 16-21 R I V U X G

Magnetic Fields Deflect Moving, Electrically Charged Objects In this laboratory experiment, a beam of negatively charged electrons (shown by a blue arc) is aimed straight upward from the center of the apparatus. The entire apparatus is inside a large magnet, and the magnetic field deflects the beam into a curved path. (Courtesy of Central Scientific Company)



(a) Visible-light image (b) Magnetogram (c) Magnetogram of a sunspot group

Figure 16-22 R I V U X G

Mapping the Sun's Magnetic Field **(a)** This visible-light image and **(b)** this false-color magnetogram were recorded at the same time. Dark blue and yellow areas in the magnetogram have north and south magnetic polarity, respectively; blue-green regions have weak magnetic fields. The highly magnetized regions in (b) correlate with the sunspots in (a). **(c)** The two ends of this large sunspot group have opposite magnetic polarities (colored blue and yellow), like the ends of a giant bar magnet. (NOAO)

and south magnetic polarities (see also Figure 16-22*c*). Thus, a sunspot group resembles a giant bar magnet, with a north magnetic pole at one end and a south magnetic pole at the other.

If different sunspot groups were unrelated to one another, their magnetic poles would be randomly oriented, like a bunch of compass needles all pointing in random directions. As Hale discovered, however, there is a striking regularity in the magnetization of sunspot groups. As a given sunspot group moves with the Sun's rotation, the sunspots in front are called the "preceding members" of the group. The spots that follow behind are referred to as the "following members." Hale compared the sunspot groups in the two solar hemispheres, north or south of the Sun's equator. He found that the preceding members in one solar hemisphere all have the same magnetic polarity, while the preceding members in the other hemisphere have the opposite polarity. Furthermore, in the hemisphere where the Sun has its north magnetic pole, the preceding members of all sunspot groups have north magnetic polarity. In the opposite hemisphere, where the Sun has its south magnetic pole, the preceding members all have south magnetic polarity.

Along with his colleague Seth B. Nicholson, Hale also discovered that the Sun's polarity pattern completely reverses itself every 11 years—the same interval as the time from one solar maximum to the next. The hemisphere that has preceding north magnetic poles during one 11-year sunspot cycle will have preceding south magnetic poles during the next 11-year cycle, and vice versa. The north and south magnetic poles of the Sun itself also reverse every 11 years. Thus, the Sun's magnetic pattern repeats itself only after two sunspot cycles, which is why astronomers speak of a **22-year solar cycle.**

The Magnetic-Dynamo Model

In 1960, the American astronomer Horace Babcock proposed a description that seems to account for many features of this 22-year solar cycle. Babcock's scenario, called a **magnetic-dynamo model,** makes use of two basic properties of the Sun's photosphere—differential rotation and convection. Differential rotation causes the magnetic field in the photosphere to become wrapped around the Sun (Figure 16-23). As a result, the magnetic field becomes concentrated at certain latitudes on either side of the solar equator. Convection in the photosphere creates tangles in the concentrated magnetic field, and "kinks" erupt through the solar surface. Sunspots appear where the magnetic field protrudes through the photosphere. The theory suggests that sunspots should appear first at northern and southern latitudes and later form nearer to the equator. This is just what is observed (see Figure 16-19). Note also that as shown on the far right in Figure 16-23, the preceding member of a sunspot group has the same polarity (N or S) as the Sun's magnetic pole in that hemisphere. This is just as Hale observed.

Differential rotation eventually undoes the twisted magnetic field. The preceding members of sunspot groups move toward the Sun's equator, while the following members migrate toward the poles. Because the preceding members from the two hemispheres have opposite magnetic polarities, their magnetic fields cancel each other out when they meet at the equator. The following members in each hemisphere have the opposite polarity to the Sun's pole in that hemisphere; hence, when they converge on the pole, the following members first cancel out and then reverse the Sun's overall magnetic field. The fields are now completely relaxed. Once again, differential rotation begins to twist the Sun's magnetic field, but now with all magnetic polarities reversed. In this way, Babcock's model helps to explain the change in field direction every 11 years.

The Sun's differential rotation makes the magnetic field twist like a rubber band

Recent discoveries in helioseismology (Section 16-3) offer new insights into the Sun's magnetic field. By comparing the

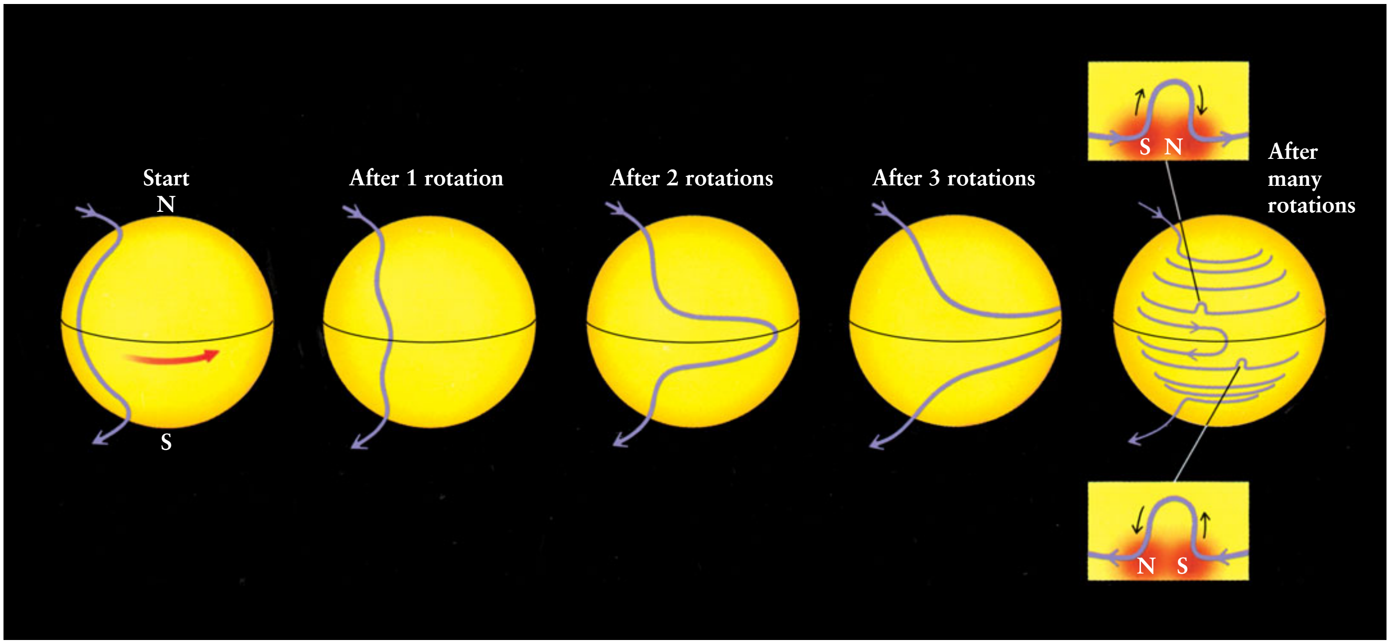


Figure 16-23

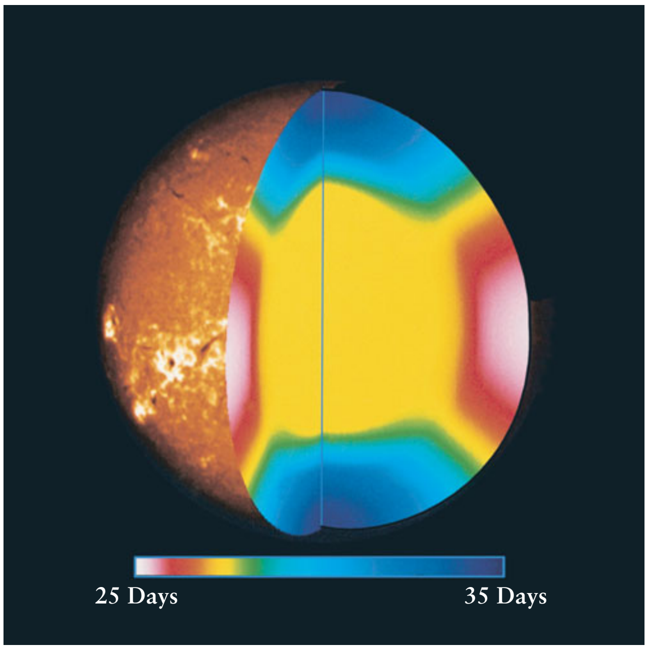
Babcock's Magnetic Dynamo Model Magnetic field lines tend to move along with the plasma in the Sun's outer layers. Because the Sun rotates faster at the equator than near the poles, a field line that starts off running from the Sun's north magnetic pole (N) to its south magnetic pole (S) ends up wrapped around the Sun like twine wrapped around a ball. The insets on the far right show how sunspot groups appear where the concentrated magnetic field rises through the photosphere.

speeds of sound waves that travel with and against the Sun's rotation, helioseismologists have been able to determine the Sun's rotation rate at different depths and latitudes. As shown in Figure 16-24, the Sun's surface pattern of differential rotation persists through the convective zone. Farther in, within the radiative zone, the Sun seems to rotate like a rigid object with a period of 27 days at all latitudes. Astronomers suspect that the Sun's magnetic field originates in a relatively thin layer where the radiative and convective zones meet and slide past each other due to their different rotation rates.

One dilemma about sunspots is that compressed magnetic fields tend to push themselves apart, which means that sunspots should dissipate rather quickly. Yet observations show that sunspots can persist for many weeks. The resolution of this paradox may have been found using helioseismology (see Section 16-3). Analysis of the vibrations of the Sun around sunspots shows that beneath the surface of the photosphere, the gases surrounding each sunspot are circulating at high speed—rather like a hurricane as large as the Earth. The circulation of charged gases around the magnetic field holds the fields in place, thus stabilizing the sunspot.

Much about sunspots and solar activity remains mysterious. There are perplexing irregularities in the solar cycle. For example, the overall reversal of the Sun's magnetic field is often piecemeal and haphazard. One pole may reverse polarity long before the other. For several weeks the Sun's surface may have two north magnetic poles and no south magnetic pole at all.

What is more, there seem to be times when all traces of sunspots and the sunspot cycle vanish for many years. For



WEB LINK 16.12

Figure 16-24

Rotation of the Solar Interior This cutaway picture of the Sun shows how the solar rotation period (shown by different colors) varies with depth and latitude. The surface and the convective zone have differential rotation (a short period at the equator and longer periods near the poles). Deeper within the Sun, the radiative zone seems to rotate like a rigid sphere. (Courtesy of K. Libbrecht, Big Bear Solar Observatory) example, virtually no sunspots were seen from 1645 through 1715. Curiously, during these same years Europe experienced record low temperatures, often referred to as the Little Ice Age, whereas the western United States was subjected to severe drought. By contrast, there was apparently a period of increased sunspot activity during the eleventh and twelfth centuries, during which the Earth was warmer than it is today. Thus, variations in solar activity appear to affect climates on the Earth. The origin of this Sun-Earth connection is a topic of ongoing research.

16-10 The Sun's magnetic field heats the corona, produces flares, and causes massive eruptions

Astronomers now understand that the Sun's magnetic field does more than just explain the presence of sunspots. It is also responsible for the existence of spicules, as well as a host of other dramatic phenomena in the chromosphere and corona.

Magnetic Reconnection

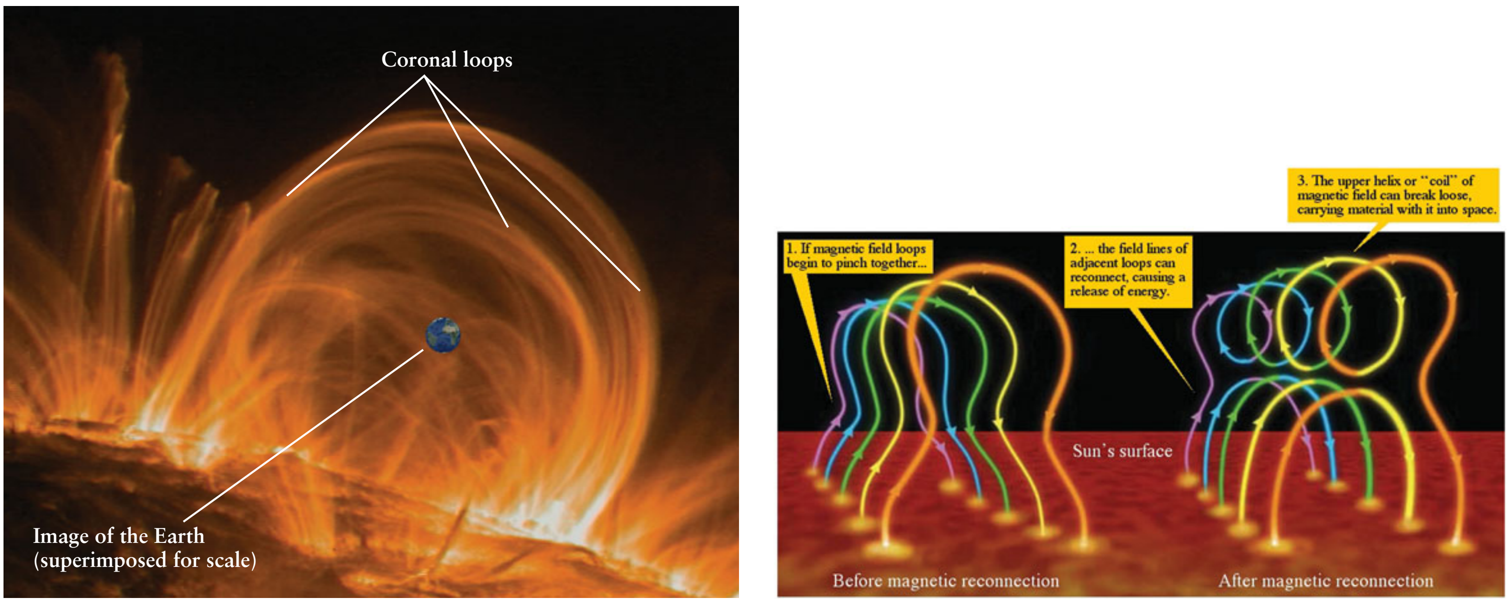
In a plasma, magnetic field lines and the material of the plasma tend to move together. This means that as convection pushes material toward the edge of a supergranule, it pushes magnetic field lines as well. The result is that vertical magnetic field lines pile up around a supergranule. Plasma that "sticks" to these magnetic field lines thus ends up lifted upward, forming a spicule (see Figures 16-12 and 16-13).

The tendency of plasma to follow the Sun's magnetic field can also explain why the temperature of the chromosphere and corona is so high. Spacecraft observations show magnetic field arches extending tens of thousands of kilometers into the corona, with streamers of electrically charged particles moving along each arch (Figure 16-25*a*). If the magnetic fields of two arches come into proximity, their magnetic fields can rearrange in a process called **magnetic reconnection** (Figure 16-25*b*). The tremendous amount of energy stored in the magnetic field is then released into the solar atmosphere. (A single arch contains as much energy as a hydroelectric power plant would generate in a million years.) The amount of energy released in this way appears to be more than enough to maintain the temperatures of the chromosphere and corona.

ANALOGY The idea that a magnetic field can heat gases has applications on Earth as well as on the Sun. In an automobile engine's ignition system an electric current is set up in a coil of wire, which produces a magnetic field. When the current is shut off, the magnetic field collapses and its energy is directed to a spark plug in one of the engine's cylinders. The released energy heats the mixture of air and gasoline around the plug, causing the mixture to ignite. This drives the piston in that cylinder and makes the automobile go.

VIDEO 16.6

Magnetic heating can also explain why the parts of the corona that lie on top of sunspots are often the most prominent in ultraviolet images. (Some examples are the



(a) (b)

WEB LINK 16.13

Figure 16-25 R I V U X G

Magnetic Arches and Magnetic Reconnection **(a)** This false-color ultraviolet image from the *TRACE* spacecraft (*Transition Region and Coronal Explorer*) shows magnetic field loops suspended high above the solar surface. The loops are made visible by the glowing gases trapped within them. **(b)** When the magnetic fields in these loops change their arrangement, a tremendous amount of energy is released and solar material can be ejected upward. (a: Stanford-Lockheed Institute for Space Research; *TRACE;* and NASA)

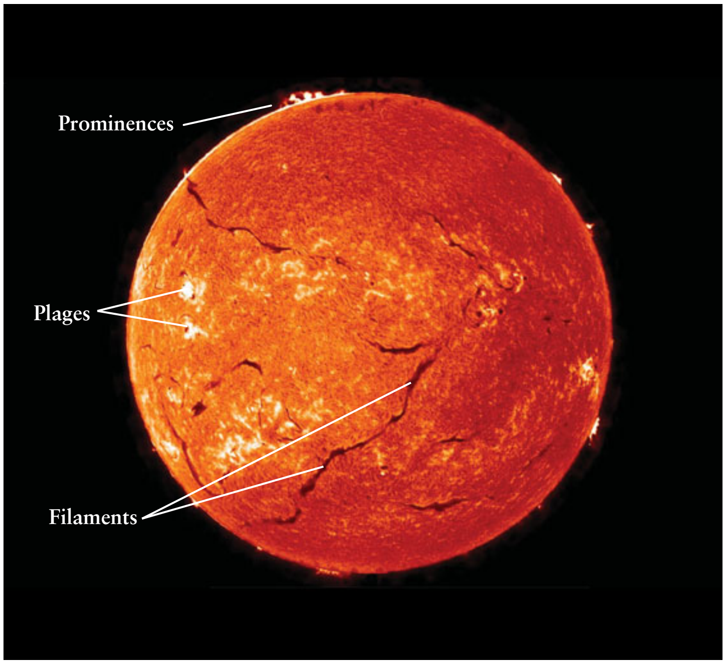


Figure 16-26 R I V U X G

The Active Sun Seen through an H_α Filter This image was made using a red filter that only passes light at a wavelength of 656 nm. The spectrum of the photosphere has an absorption line at this wavelength and so appears dark. Hence, this filter reveals the photosphere and corona. Prominences, plages, and filaments are associated with strong magnetic fields. (NASA)

bright regions in Figure 16-15.) The intense magnetic field of the sunspots helps trap and compress hot coronal gas, giving it such a high temperature that it emits copious amounts of high-energy ultraviolet photons and even more energetic X-ray photons.

Prominences, Flares, ands Coronal Mass Ejections

Spicules and coronal heating occur even when the Sun is quiet. But magnetic fields can also explain many aspects of the active Sun in addition to sunspots. Figure 16-26 is an image of the chromosphere made with an H_α filter during a sunspot maximum. The bright areas are called **plages** (from the French word for "beach"). These are bright, hot regions in the chromosphere that tend to form just before the appearance of new sunspots. They are probably created by magnetic fields that push upward from the Sun's interior, compressing and heating a portion of the chromosphere. The dark streaks, called **filaments,** are relatively cool and dense parts of the chromosphere that have been pulled along with magnetic field lines as they arch to high altitudes.

When seen from the side, so that they are viewed against the dark background of space, filaments appear as bright, arching columns of gas called **prominences** (Figure 16-27). They can extend for tens of thousands of kilometers above the photosphere. Some prominences last for only a few hours, while others persist for many months. The most energetic prominences break free of the magnetic fields that confined them and burst into space.

VIDEO 16.7

Violent, eruptive events on the Sun, called **solar flares,** occur in complex sunspot groups. Within only a few minutes, temperatures in a compact region may soar to 5×10^6 K and vast quantities of particles and radiation—including as much material as is in the prominence shown in Figure 16-27—are blasted out into space. These eruptions can also cause disturbances that spread outward in the solar atmosphere, like the ripples that appear when you drop a rock into a pond.

WEB LINK 16.14

The most energetic flares carry as much as 10^{30} joules of energy, equivalent to 10^{14} one-megaton nuclear weapons being exploded at once! However, the energy of a solar flare does not come from thermonuclear fusion in the solar atmosphere; instead, it appears to be released from the intense magnetic field around a sunspot group.

A solar flare can have as much energy as 100 trillion nuclear bombs

As energetic as solar flares are, they are dwarfed by **coronal mass ejections.** One such event is shown in the image that opens this chapter; Figure 16-28*a* shows another. In a coronal mass ejection, more than 10^{12} kilograms (a billion tons) of high-temperature coronal gas is blasted into space at speeds of hundreds of kilometers per second. A typical coronal mass ejection lasts a few hours. These explosive events seem to be related to large-scale alterations in the Sun's magnetic field, like the magnetic

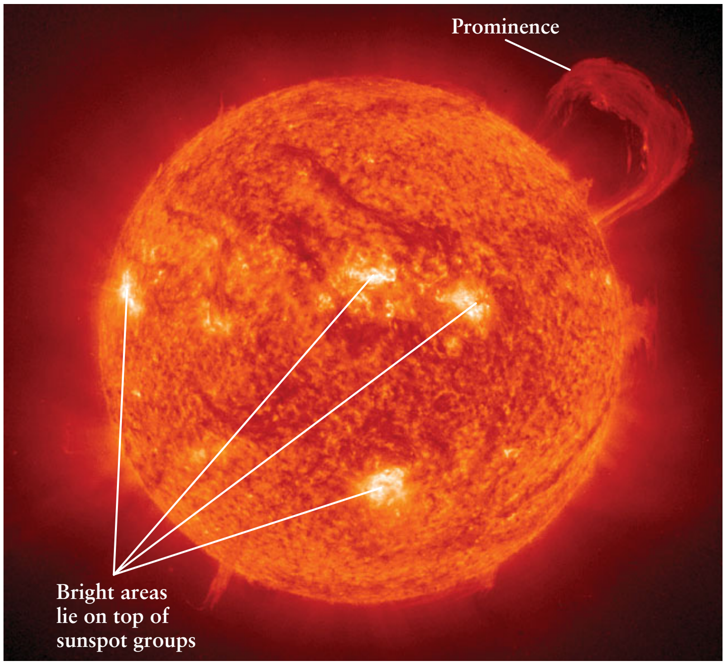
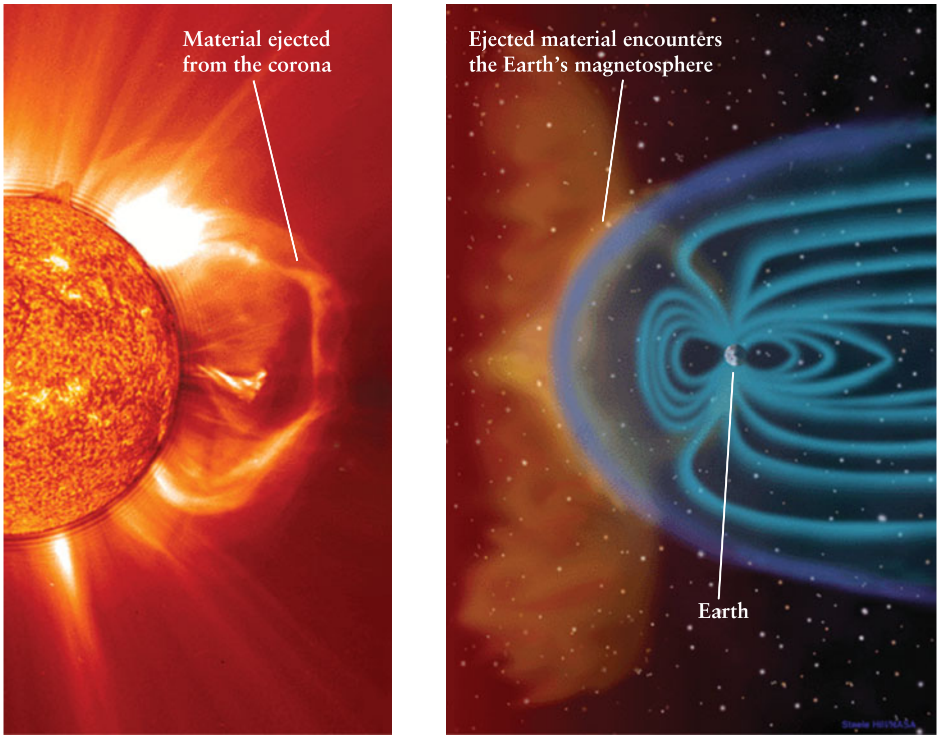


Figure 16-27 R I V U X G

A Solar Prominence A huge prominence arches above the solar surface in this ultraviolet image from the *SOHO* spacecraft. The image was recorded using light at a wavelength of 30.4 nm, emitted by singly ionized helium atoms at a temperature of about 60,000 K. By comparison, the material within the arches in Figure 16-25 reaches temperatures in excess of 2×10^6 K. (*SOHO*/EIT/ESA/NASA)

Figure 16-28 R I V U X G

A Coronal Mass Ejection **(a)** *SOHO* recorded this coronal mass ejection in an X-ray image. (The image of the Sun itself was made at ultraviolet wavelengths.) **(b)** Within two to four days the fastest-moving ejected material reaches a distance of 1 AU from the Sun. Most particles are deflected by the Earth's magnetosphere, but some are able to reach the Earth. (The ejection shown in (a) was not aimed toward the Earth and did not affect us.) (*SOHO*/EIT/LASCO/ESA/NASA)



(a) A coronal mass ejection

(b) Two to four days later

reconnection shown in Figure 16-25*b*. Coronal mass ejections occur every few months; smaller eruptions may occur almost daily.

If a solar flare or coronal mass ejection happens to be aimed toward Earth, a stream of high-energy electrons and nuclei reaches us a few days later (Figure 16-28*b*). When this plasma arrives, it can interfere with satellites, pose a health hazard to astronauts in orbit, and disrupt electrical and communications equipment on the Earth's surface. Telescopes on Earth and on board spacecraft now monitor the Sun continuously to provide warnings of dangerous levels of solar particles.

The numbers of plages, filaments, solar flares, and coronal mass ejections all vary with the same 11-year cycle as sunspots. But unlike sunspots, coronal mass ejections never completely cease, even when the Sun is at its quietest. Astronomers are devoting substantial effort to understanding these and other aspects of our dynamic Sun.

Key Words

The term preceded by an asterisk (*) is discussed in Box 16-1.

22-year solar cycle, p. 424
chromosphere, p. 417
CNO cycle, p. 408
conduction, p. 409
convection, p. 409
convective zone, p. 411
corona, p. 418
coronal hole, p. 419
coronal mass ejection, p. 427
differential rotation, p. 422
filament, p. 427
granulation, p. 415
granule, p. 415
helioseismology, p. 411
hydrogen fusion, p. 406
hydrostatic equilibrium, p. 408
limb darkening, p. 415
luminosity (of the Sun), p. 404
magnetic-dynamo model, p. 424
magnetogram, p. 423
magnetic reconnection, p. 426
negative hydrogen ion, p. 417
neutrino, p. 412
neutrino oscillation, p. 413
photosphere, p. 415
plage, p. 427
plasma, p. 422
*positron, p. 406
prominence, p. 427
proton-proton chain, p. 408
radiative diffusion, p. 409
radiative zone, p. 411
solar flare, p. 427
solar neutrino, p. 412
solar neutrino problem, p. 413
solar wind, p. 419
spicule, p. 417
sunspot, p. 420
sunspot cycle, p. 422
sunspot maximum, p. 422
sunspot minimum, p. 422
supergranule, p. 415
thermal equilibrium, p. 409
thermonuclear fusion, p. 406
Zeeman effect, p. 422

Key Ideas

Hydrogen Fusion in the Sun's Core: The Sun's energy is produced by hydrogen fusion, a sequence of thermonuclear reactions in which four hydrogen nuclei combine to produce a single helium nucleus.

• The energy released in a nuclear reaction corresponds to a slight reduction of mass according to Einstein's equation $E = mc^2$.

• Thermonuclear fusion occurs only at very high temperatures; for example, hydrogen fusion occurs only at temperatures in excess of about 10^7 K. In the Sun, fusion occurs only in the dense, hot core.

Models of the Sun's Interior: A theoretical description of a star's interior can be calculated using the laws of physics.

• The standard model of the Sun suggests that hydrogen fusion takes place in a core extending from the Sun's center to about 0.25 solar radius.

• The core is surrounded by a radiative zone extending to about 0.71 solar radius. In this zone, energy travels outward through radiative diffusion.

• The radiative zone is surrounded by a rather opaque convective zone of gas at relatively low temperature and pressure. In this zone, energy travels outward primarily through convection.

Solar Neutrinos and Helioseismology: Conditions in the solar interior can be inferred from measurements of solar neutrinos and of solar vibrations.

• Neutrinos emitted in thermonuclear reactions in the Sun's core have been detected, but in smaller numbers than expected. Recent neutrino experiments explain why this is so.

• Helioseismology is the study of how the Sun vibrates. These vibrations have been used to infer pressures, densities, chemical compositions, and rotation rates within the Sun.

INTERACTIVE EXERCISE 16.3

The Sun's Atmosphere: The Sun's atmosphere has three main layers: the photosphere, the chromosphere, and the corona. Everything below the solar atmosphere is called the solar interior.

• The visible surface of the Sun, the photosphere, is the lowest layer in the solar atmosphere. Its spectrum is similar to that of a blackbody at a temperature of 5800 K. Convection in the photosphere produces granules.

• Above the photosphere is a layer of less dense but higher-temperature gases called the chromosphere. Spicules extend upward from the photosphere into the chromosphere along the boundaries of supergranules.

• The outermost layer of the solar atmosphere, the corona, is made of very high-temperature gases at extremely low density. Activity in the corona includes coronal mass ejections and coronal holes. The solar corona blends into the solar wind at great distances from the Sun.

The Active Sun: The Sun's surface features vary in an 11-year cycle. This is related to a 22-year cycle in which the surface magnetic field increases, decreases, and then increases again with the opposite polarity.

• Sunspots are relatively cool regions produced by local concentrations of the Sun's magnetic field. The average number of sunspots increases and decreases in a regular cycle of approximately 11 years, with reversed magnetic polarities from one 11-year cycle to the next. Two such cycles make up the 22-year solar cycle.

• The magnetic-dynamo model suggests that many features of the solar cycle are due to changes in the Sun's magnetic field. These changes are caused by convection and the Sun's differential rotation.

• A solar flare is a brief eruption of hot, ionized gases from a sunspot group. A coronal mass ejection is a much larger eruption that involves immense amounts of gas from the corona.

Questions

Review Questions

1. What is meant by the luminosity of the Sun?
2. What is Kelvin-Helmholtz contraction? Why is it ruled out as a source of the present-day Sun's energy?
3. Why is it impossible for the burning of substances like coal to be the source of the Sun's energy?
4. What is hydrogen fusion? Why is hydrogen fusion fundamentally unlike the burning of a log in a fireplace?
5. If the electric force between protons were somehow made stronger, what effect would this have on the temperature required for thermonuclear fusion to take place?
6. Why do thermonuclear reactions occur only in the Sun's core, not in its outer regions?
7. Describe how the net result of the reactions shown in the *Cosmic Connections* figure (Section 16-1) is the conversion of four protons into a single helium nucleus. What other particles are produced in this process? How many of each particle are produced?
8. Give an everyday example of hydrostatic equilibrium. Give an example of thermal equilibrium. Explain how these equilibrium conditions apply to each example.
9. If thermonuclear fusion in the Sun were suddenly to stop, what would eventually happen to the overall radius of the Sun? Justify your answer using the ideas of hydrostatic equilibrium and thermal equilibrium.
10. Give some everyday examples of conduction, convection, and radiative diffusion.
11. Describe the Sun's interior. Include references to the main physical processes that occur at various depths within the Sun.
12. Suppose thermonuclear fusion in the Sun stopped abruptly. Would the intensity of sunlight decrease just as abruptly? Why or why not?
13. Explain how studying the oscillations of the Sun's surface can give important, detailed information about physical conditions deep within the Sun.
14. What is a neutrino? Why is it useful to study neutrinos coming from the Sun? What do they tell us that cannot be learned from other avenues of research?
15. Unlike all other types of telescopes, neutrino detectors are placed deep underground. Why?
16. What was the solar neutrino problem? What solution to this problem was suggested by the results from the Sudbury Neutrino Observatory?
17. Describe the dangers in attempting to observe the Sun. How have astronomers learned to circumvent these observational problems?
18. Briefly describe the three layers that make up the Sun's atmosphere. In what ways do they differ from each other?
19. What is solar granulation? Describe how convection gives rise to granules.
20. High-resolution spectroscopy of the photosphere reveals that absorption lines are blueshifted in the spectrum of the central, bright regions of granules but are redshifted in the spectrum of the dark boundaries between granules. Explain how these observations show that granulation is due to convection.

21. What is the difference between granules and supergranules?
22. What are spicules? Where are they found? How can you observe them? What causes them?
23. How do astronomers know that the temperature of the corona is so high?
24. How do astronomers know when the next sunspot maximum and minimum will occur?
25. Why do astronomers say that the solar cycle is really 22 years long, even though the number of sunspots varies over an 11-year period?
26. Explain how the magnetic-dynamo model accounts for the solar cycle.
27. Describe one explanation for why the corona has a higher temperature than the chromosphere.
28. Why should solar flares and coronal mass ejections be a concern for businesses that use telecommunication satellites?

Advanced Questions

Questions preceded by an asterisk (*) involve the topic discussed in Box 16-1.

> **Problem-solving tips and tools**
>
> You may have to review Wien's law and the Stefan-Boltzmann law, which are the subjects of Section 5-4. Section 5-5 discusses the properties of photons. As we described in Box 5-2, you can simplify calculations by taking ratios, such as the ratio of the flux from a sunspot to the flux from the undisturbed photosphere. When you do this, all the cumbersome constants cancel out. Figure 5-7 shows the various parts of the electromagnetic spectrum. We introduced the Doppler effect in Section 5-9 and Box 5-6. For information about the planets, see Table 7-1.

29. Calculate how much energy would be released if each of the following masses were converted *entirely* into their equivalent energy: (**a**) a carbon atom with a mass of 2×10^{-26} kg, (**b**) 1 kilogram, and (**c**) a planet as massive as the Earth (6×10^{24} kg).
30. Use the luminosity of the Sun (given in Table 16-1) and the answers to the previous question to calculate how long the Sun must shine in order to release an amount of energy equal to that produced by the complete mass-to-energy conversion of (**a**) a carbon atom, (**b**) 1 kilogram, and (**c**) the Earth.
31. Assuming that the current rate of hydrogen fusion in the Sun remains constant, what fraction of the Sun's mass will be converted into helium over the next 5 billion years? How will this affect the overall chemical composition of the Sun?
32. (**a**) Estimate how many kilograms of hydrogen the Sun has consumed over the past 4.56 billion years, and estimate the amount of mass that the Sun has lost as a result. Assume that the Sun's luminosity has remained constant during that time. (**b**) In fact, the Sun's luminosity when it first formed was only about 70% of its present value. With this in mind, explain whether your answers to part (a) are an overestimate or an underestimate.

*33. To convert one kilogram of hydrogen (^{1}H) into helium (^{4}He) as described in Box 16-1, you must start with 1.5 kg of hydrogen. Explain why, and explain what happens to the other 0.5 kg. (*Hint:* How many ^{1}H nuclei are used to make the two ^{3}He nuclei shown in the *Cosmic Connections* figure in Section 16-1? How many of these ^{1}H nuclei end up being incorporated into the ^{4}He nucleus shown in the figure?)

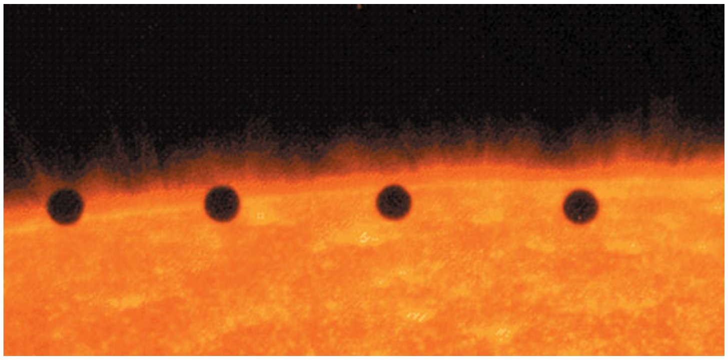
*34. (**a**) A positron has the same mass as an electron (see Appendix 7). Calculate the amount of energy released by the annihilation of an electron and positron. (**b**) The products of this annihilation are two photons, each of equal energy. Calculate the wavelength of each photon, and confirm from Figure 5-7 that this wavelength is the gamma-ray range.

35. Sirius is the brightest star in the night sky. It has a luminosity of 23.5 $L_\odot$, that is, it is 23.5 times as luminous as the Sun and burns hydrogen at a rate 23.5 times greater than the Sun. How many kilograms of hydrogen does Sirius convert into helium each second?
36. (Refer to the preceding question.) Sirius has 2.3 times the mass of the Sun. Do you expect that the lifetime of Sirius will be longer, shorter, or the same length as that of the Sun? Explain your reasoning.
37. (**a**) If the Sun were not in a state of hydrostatic equilibrium, would its diameter remain the same? Explain your reasoning. (**b**) If the Sun were not in a state of thermal equilibrium, would its luminosity remain the same? What about its surface temperature? Explain your reasoning.
38. Using the mass and size of the Sun given in Table 16-1, verify that the average density of the Sun is 1410 kg/m^3. Compare your answer with the average densities of the Jovian planets.
39. Use the data in Table 16-2 to calculate the average density of material within 0.1 $R_\odot$ of the center of the Sun. (You will need to use the mass and radius of the Sun as a whole, given in Table 16-1.) Explain why your answer is not the same as the density at 0.1 $R_\odot$ given in Table 16-2.
40. In a typical solar oscillation, the Sun's surface moves up or down at a maximum speed of 0.1 m/s. An astronomer sets out to measure this speed by detecting the Doppler shift of an absorption line of iron with wavelength 557.6099 nm. What is the maximum wavelength shift that she will observe?
41. Explain why the results from the Sudbury Neutrino Observatory (SNO) only provide an answer to the solar neutrino problem for relatively high-energy neutrinos. (*Hint:* Can SNO detect solar neutrinos of all energies?)
42. The amount of energy required to dislodge the extra electron from a negative hydrogen ion is 1.2×10^{-19} J. (**a**) The extra electron can be dislodged if the ion absorbs a photon of sufficiently short wavelength. (Recall from Section 5-5 that the higher the energy of a photon, the shorter its wavelength.) Find the longest wavelength (in nm) that can accomplish this. (**b**) In what part of the electromagnetic spectrum does this wavelength lie? (**c**) Would a photon of visible light be able to dislodge the extra electron? Explain. (**d**) Explain why the photosphere, which contains negative hydrogen ions, is quite opaque to visible light but is less opaque to light with wavelengths longer than the value you calculated in (a).
43. Astronomers often use an H_α filter to view the chromosphere. Explain why this can also be accomplished with filters that are transparent only to the wavelengths of the H and K lines

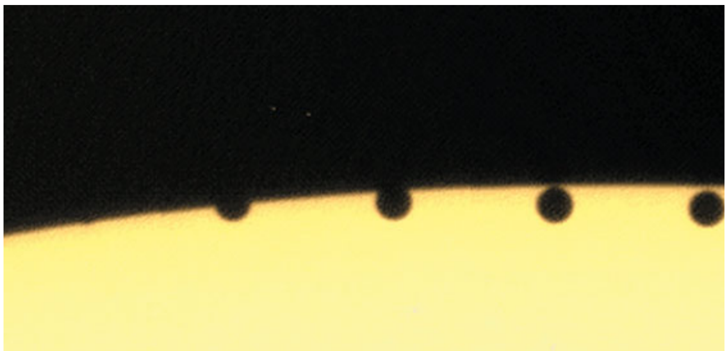
of ionized calcium. (*Hint:* The H and K lines are dark lines in the spectrum of the photosphere.)

44. Calculate the wavelengths at which the photosphere, chromosphere, and corona emit the most radiation. Explain how the results of your calculations suggest the best way to observe these regions of the solar atmosphere. (*Hint:* Treat each part of the atmosphere as a perfect blackbody. Assume average temperatures of 50,000 K and 1.5×10^6 K for the chromosphere and corona, respectively.)

45. On November 15, 1999, the planet Mercury passed in front of the Sun as seen from Earth. The TRACE spacecraft made these time-lapse images of this event using ultraviolet light (top) and visible light (bottom). (Mercury moved from left to right in these images. The time between successive views of Mercury is 6 to 9 minutes.) Explain why the Sun appears somewhat larger in the ultraviolet image than in the visible-light image.



R I V U X G



R I V U X G

(K. Schrijver, Stanford-Lockheed Institute for Space Research, TRACE, and NASA)

46. The moving images on a television set that uses a tube (as opposed to a LCD or plasma flat-screen TV) are made by a fast-moving electron beam that sweeps over the back of the screen. Explain why placing a strong magnet next to the screen distorts the picture. (*Caution:* Don't try this with your television! You can do permanent damage to your set.)

47. Find the wavelength of maximum emission of the umbra of a sunspot and the wavelength of maximum emission of a sunspot's penumbra. In what part of the electromagnetic spectrum do these wavelengths lie?

48. (**a**) Find the ratio of the energy flux from a patch of a sunspot's penumbra to the energy flux from an equally large patch of undisturbed photosphere. Which patch is brighter? (**b**) Find the ratio of the energy flux from a patch of a sunspot's penumbra to the energy flux from an equally large patch of umbra. Again, which patch is brighter?

49. Suppose that you want to determine the Sun's rotation rate by observing its sunspots. Is it necessary to take the Earth's orbital motion into account? Why or why not?

50. (**a**) Using a ruler to make measurements of Figure 16-25, determine how far the arches in that figure extend above the Sun's surface. The diameter of the Earth is 12,756 km. (**b**) In Figure 16-25 (an ultraviolet image) the photosphere appears dark compared to the arches. Explain why.

51. The amount of visible light emitted by the Sun varies only a little over the 11-year sunspot cycle. But the amount of X rays emitted by the Sun can be 10 times greater at solar maximum than at solar minimum. Explain why these two types of radiation should be so different in their variability.

Discussion Questions

52. Discuss the extent to which cultures around the world have worshiped the Sun as a deity throughout history. Why do you suppose there has been such widespread veneration?

53. In the movie *Star Trek IV: The Voyage Home,* the starship *Enterprise* flies on a trajectory that passes close to the Sun's surface. What features should a real spaceship have to survive such a flight? Why?

54. Discuss some of the difficulties in correlating solar activity with changes in the Earth's climate.

55. Describe some of the advantages and disadvantages of observing the Sun (**a**) from space and (**b**) from the Earth's south pole. What kinds of phenomena and issues might solar astronomers want to explore from these locations?

Web/eBook Questions

56. Search the World Wide Web for information about features in the solar atmosphere called *sigmoids*. What are they? What causes them? How might they provide a way to predict coronal mass ejections?

VIDEO 16.1

57. Determining the Lifetime of a Solar Granule. Access and view the video "Granules on the Sun's Surface" in Chapter 16 of the *Universe* Web site or eBook. Your task is to determine the approximate lifetime of a solar granule on the photosphere. Select an area, then slowly and rhythmically repeat "start, stop, start, stop" until you can consistently predict the appearance and disappearance of granules. While keeping your rhythm, move to a different area of the video and continue monitoring the appearance and disappearance of granules. When you are confident you have the timing right, move your eyes (or use a partner) to look at the clock shown in the video. Determine the length of time between the appearance and disappearance of the granules and record your answer.

Activities

Observing Projects

Observing tips and tools

At the risk of repeating ourselves, we remind you to ***never look directly at the Sun, because it can easily cause permanent blindness.*** You can view the Sun safely without a telescope just by using two pieces of white cardboard. First, use a pin to poke a small hole in one piece of cardboard; this will be your "lens," and the other piece of cardboard will be your "viewing screen." Hold the "lens" piece of cardboard so that it is face-on to the Sun and sunlight can pass through the hole. With your other hand, hold the "viewing screen" so that the sunlight from the "lens" falls on it. Adjust the distance between the two pieces of cardboard so that you see a sharp image of the Sun on the "viewing screen." This image is perfectly safe to view. It is actually possible to see sunspots with this low-tech apparatus.

For a better view, use a telescope with a solar filter that fits on the front of the telescope. A standard solar filter is a piece of glass coated with a thin layer of metal to give it a mirrorlike appearance. This coating reflects almost all the sunlight that falls on it, so that only a tiny, safe amount of sunlight enters the telescope. An H_α filter, which looks like a red piece of glass, keeps the light at a safe level by admitting only a very narrow range of wavelengths. (Filters that fit on the back of the telescope are *not* recommended. The telescope focuses concentrated sunlight on such a filter, heating it and making it susceptible to cracking—and if the filter cracks when you are looking through it, your eye will be ruined instantly and permanently.)

To use a telescope with a solar filter, first aim the telescope away from the Sun, then put on the filter. Keep the lens cap on the telescope's secondary wide-angle "finder scope" (if it has one), because the heat of sunlight can fry the finder scope's optics. Next, aim the telescope toward the Sun, using the telescope's shadow to judge when you are pointed in the right direction. You can then safely look through the telescope's eyepiece. When you are done, make sure you point the telescope away from the Sun before removing the filter and storing the telescope.

Note that the amount of solar activity that you can see (sunspots, filaments, flares, prominences, and so on) will depend on where the Sun is in its 11-year sunspot cycle.

58. Use a telescope with a solar filter to observe the surface of the Sun. Do you see any sunspots? Sketch their appearance. Can you distinguish between the umbrae and penumbrae of the sunspots? Can you see limb darkening? Can you see any granulation?

59. If you have access to an H_α filter attached to a telescope especially designed for viewing the Sun safely, use this instrument to examine the solar surface. How does the appearance of the Sun differ from that in white light? What do sunspots look like in H_α? Can you see any prominences? Can you see any filaments? Are the filaments in the H_α image near any sunspots seen in white light? (Note that the amount of activity that you see will be much greater at some times during the solar cycle than at others.)

60. Use the *Starry Night Enthusiast*™ program to measure the Sun's rotation. Display the entire celestial sphere by selecting **Guides > Atlas** in the **Favourites** menu and center on the Sun by double-clicking on **Sun** in the **Find** pane. Using the controls at the right-hand end of the toolbar, zoom in until you can see details on the Sun's surface clearly. In the toolbar, set the **Time Flow Rate** to 1 day. Using the time forward and backward buttons in the toolbar, step through enough time to determine the rotation period of the Sun. Which part of the actual Sun's surface rotates at the rate shown in *Starry Night Enthusiast*™? (Note: The program does not show the Sun's differential rotation.)

61. Use the *Starry Night Enthusiast*™ program to examine the Sun. Open the **Favourites** pane and double-click on **Solar System > Inner Solar System** to display the inner planets surrounding the Sun. Stop **Time Advance** and use the **down** arrow in the toolbar under **Viewing Location** to approach to within about 0.015 AU of the Sun. You can rotate the Sun by placing the mouse cursor over the image and, while holding down the **Shift** key, hold down the mouse button while moving the mouse. (On a two-button mouse, hold down the left mouse button.) (**a**) The Sun's equator lies close to the plane of the ecliptic. Where do most of the sunspots visible on the image of the Sun lie relative to the solar equator? Check Figure 16-18 for more realistic images of sunspots and Figure 16-19 for the latitude distribution of sunspots. (**b**) Based on your observations in (a), does the image in *Starry Night Enthusiast*™ show the Sun near the beginning, middle, or end of the 11-year sunspot cycle? Explain your reasoning. You can see current solar images from both ground and space-based solar telescopes by opening the LiveSky pane if you have an Internet connection on your computer.

Collaborative Exercises

62. Figure 16-19 shows variations in the average latitude of sunspots. Estimate the average latitude of sunspots in the year you were born and estimate the average latitude on your twenty-first birthday. Make rough sketches of the Sun during those years to illustrate your answers.

63. Create a diagram showing a sketch of how limb darkening on the Sun would look different if the Sun had either a thicker or thinner photosphere. Be sure to include a caption explaining your diagram.

64. Solar granules, shown in Figure 16-9, are about 1000 km across. What city is about that distance away from where you are right now? What city is that distance from the birthplace of each group member?

65. Magnetic arches in the corona are shown in Figure 16-25*a*. How many Earths high are these arches, and how many Earths could fit inside one arch?

17

The Nature of the Stars

R I V U X G

WEB LINK 17.1

Some stars in this cluster (called M39) are distinctly blue in color, while others are yellow or red. (Heidi Schweiker/NOAO/AURA/NSF)

To the unaided eye, the night sky is spangled with several thousand stars, each appearing as a bright pinpoint of light. With a pair of binoculars, you can see some 10,000 other, fainter stars; with a 15-cm (6-in.) telescope, the total rises to more than 2 million. Astronomers now know that there are in excess of 100 billion (10^{11}) stars in our Milky Way Galaxy alone.

But what are these distant pinpoints? To the great thinkers of ancient Greece, the stars were bits of light embedded in a vast sphere with the Earth at its center. They thought the stars were composed of a mysterious "fifth element," quite unlike anything found on Earth.

Today, we know that the stars are made of the same chemical elements found on Earth. We know their sizes, their temperatures, their masses, and something of their internal structures. We understand, too, why the stars in the accompanying image come in a range of beautiful colors: Blue stars have high surface temperatures, while the surface temperatures of red and yellow stars are relatively low.

How have we learned these things? How can we know the nature of the stars, objects so distant that their light takes years or centuries to reach us? In this chapter, we will learn about the measurements and calculations that astronomers make to determine the properties of stars. We will also take a first look at the Hertzsprung-Russell diagram, an important tool that helps astronomers systematize the wealth of available information about the stars. In later chapters, we will use this diagram to help us understand how stars are born, evolve, and eventually die.

17-1 Careful measurements of the parallaxes of stars reveal their distances

WEB LINK 17.2

The vast majority of stars are objects very much like the Sun. This understanding followed from the discovery that the stars are tremendously far from us, at distances so great that their light takes years to reach us. Because the stars at night are clearly visible to the naked eye despite these

Learning Goals

By reading the sections of this chapter, you will learn

17-1 How we can measure the distances to the stars

17-2 How we measure a star's brightness and luminosity

17-3 The magnitude scale for brightness and luminosity

17-4 How a star's color indicates its temperature

17-5 How a star's spectrum reveals its chemical composition

17-6 How we can determine the sizes of stars

17-7 How H-R diagrams summarize our knowledge of the stars

17-8 How we can deduce a star's size from its spectrum

17-9 How we can use binary stars to measure the masses of stars

17-10 How we can learn about binary stars in very close orbits

17-11 What eclipsing binaries are and what they tell us about the sizes of stars huge distances, it must be that the **luminosity** of the stars—that is, how much energy they emit into space per second—is comparable to or greater than that of the Sun. Just as for the Sun, the only explanation for such tremendous luminosities is that thermonuclear reactions are occurring within the stars (see Section 16-1).

Clearly, then, it is important to know how distant the stars are. But how do we measure these distances? You might think this is done by comparing how bright different stars appear. Perhaps the star Betelgeuse in the constellation Orion appears bright because it is relatively close, while the dimmer and less conspicuous star Polaris (the North Star, in the constellation Ursa Minor) is farther away.

But this line of reasoning is incorrect: Polaris is actually closer to us than Betelgeuse! How bright a star appears is *not* a good indicator of its distance. If you see a light on a darkened road, it could be a motorcycle headlight a kilometer away or a person holding a flashlight just a few meters away. In the same way, a bright star might be extremely far away but have an unusually high luminosity, and a dim star might be relatively close but have a rather low luminosity. Astronomers must use other techniques to determine the distances to the stars.

Parallax and the Distances to the Stars

The most straightforward way of measuring stellar distances uses an effect called **parallax.** This is the apparent displacement of an object because of a change in the observer's point of view (Figure 17-1). To see how parallax works, hold your arm out straight in front of you. Now look at the hand on your outstretched arm, first with your left eye closed, then with your right eye closed. When you close one eye and open the other, your hand appears to shift back and forth against the background of more distant objects.

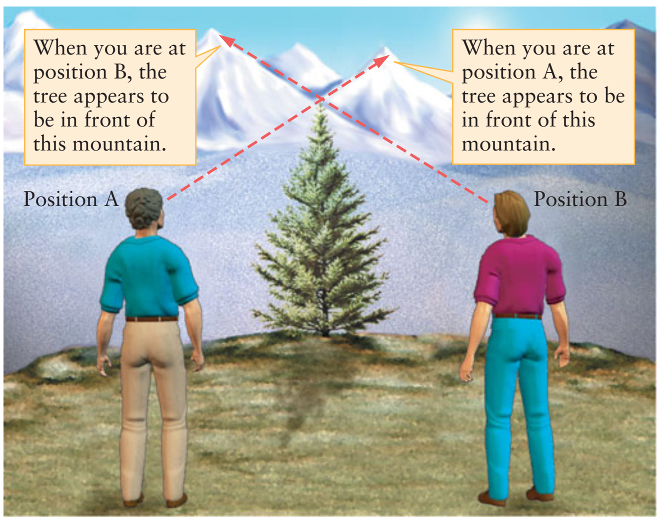


Figure 17-1

Parallax Imagine looking at some nearby object (a tree) against a distant background (mountains). When you move from one location to another, the nearby object appears to shift with respect to the distant background scenery. This familiar phenomenon is called parallax.

The closer the object you are viewing, the greater the parallax shift. To see this, repeat the experiment with your hand held closer to your face. Your brain analyzes such parallax shifts constantly as it compares the images from your left and right eyes, and in this way determines the distances to objects around you. This is the origin of depth perception.

You measure distances around you by comparing the images from your left and right eyes—we find the distances to stars using the same principle

To measure the distance to a star, astronomers measure the parallax shift of the star using two points of view that are as far apart as possible—at opposite sides of the Earth's orbit. The direction from Earth to a nearby star changes as our planet orbits the Sun, and the nearby star appears to move back and forth against the background of more distant stars (Figure 17-2). This motion is called **stellar parallax.** The parallax (p) of a star is equal to half the angle through which the star's apparent position shifts as the Earth moves from one side of its orbit to the other. The larger the parallax p, the smaller the distance d to the star (compare Figure 17-2*a* with Figure 17-2*b*).

It is convenient to measure the distance d in **parsecs.** A star with a parallax angle of 1 second of arc (p = 1 arcsec) is at a distance of 1 parsec (d = 1 pc). (The word "parsec" is a contraction of the phrase "the distance at which a star has a *par*allax of one arc*sec*ond." Recall from Section 1-7 that 1 parsec equals 3.26 light-years, 3.09×10^{13} km, or 206,265 AU; see Figure 1-14.) If the angle p is measured in arcseconds, then the distance d to the star in parsecs is given by the following equation:

Relation between a star's distance and its parallax

$$d = \frac{1}{p}$$

d = distance to a star, in parsecs

p = parallax angle of that star, in arcseconds

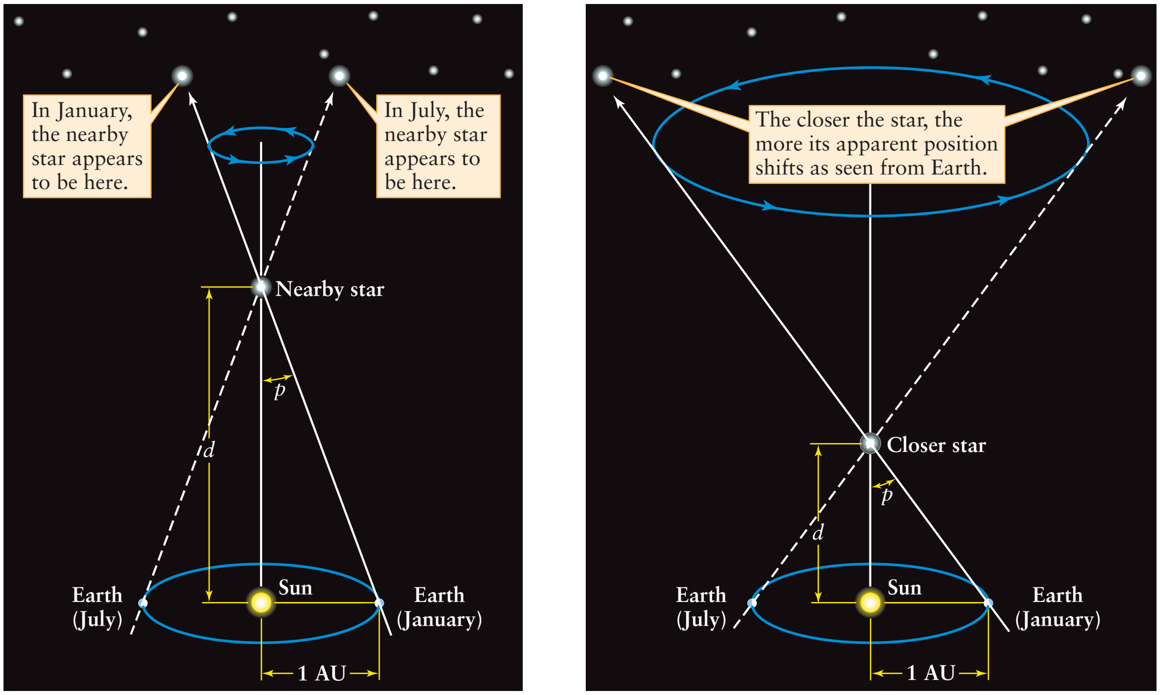
This simple relationship between parallax and distance in parsecs is one of the main reasons that astronomers usually measure cosmic distances in parsecs rather than light-years. For example, a star whose parallax is p = 0.1 arcsec is at a distance d = 1/(0.1) = 10 parsecs from the Earth. Barnard's star, named for the American astronomer Edward E. Barnard, has a parallax of 0.547 arcsec. Hence, the distance to this star is:

$$d = \frac{1}{p} = \frac{1}{0.545} = 1.83 \text{ pc}$$

Because 1 parsec is 3.26 light-years, this can also be expressed as

$$d = 1.83 \text{ pc} \times \frac{3.26 \text{ ly}}{1 \text{ pc}} = 5.98 \text{ ly}$$

All known stars have parallax angles less than one arcsecond. In other words, the closest star is more than 1 parsec away. Such



(a) Parallax of a nearby star

(b) Parallax of an even closer star

ANIMATION 17.1

Figure 17-2

Stellar Parallax (a) As the Earth orbits the Sun, a nearby star appears to shift its position against the background of distant stars. The parallax (p) of the star is equal to the angular radius of the Earth's orbit as seen from the star. **(b)** The closer the star is to us, the greater the parallax angle p. The distance d to the star (in parsecs) is equal to the reciprocal of the parallax angle p (in arcseconds): $d = 1/p$.

small parallax angles are difficult to detect, so it was not until 1838 that the first successful parallax measurements were made by the German astronomer and mathematician Friedrich Wilhelm Bessel. He found the parallax angle of the star 61 Cygni to be just 1/3 arcsec—equal to the angular diameter of a dime at a distance of 11 kilometers, or 7 miles. He thus determined that this star is about 3 pc from the Earth. (Modern measurements give a slightly smaller parallax angle, which means that 61 Cygni is actually more than 3 pc away.) The star Proxima Centauri has the largest known parallax angle, 0.772 arcsec, and hence is the closest known star (other than the Sun); its distance is 1/(0.772) = 1.30 pc.

WEB LINK 17.3 AIMM 17.1

Appendix 4 at the back of this book lists all the stars within 4 pc of the Sun, as determined by parallax measurements. Most of these stars are far too dim to be seen with the naked eye, which is why their names are probably unfamiliar to you. By contrast, the majority of the familiar, bright stars in the nighttime sky (listed in Appendix 5) are so far away that their parallaxes cannot be measured from the Earth's surface. They appear bright not because they are close, but because they are far more luminous than the Sun. The brightest stars in the sky are *not* necessarily the nearest stars!

Measuring Parallax from Space

Parallax angles smaller than about 0.01 arcsec are extremely difficult to measure from the Earth, in part because of the blurring effects of the atmosphere. Therefore, the parallax method used with ground-based telescopes can give fairly reliable distances only for stars nearer than about 1/0.01 = 100 pc. But an observatory in space is unhampered by the atmosphere. Observations made from spacecraft therefore permit astronomers to measure even smaller parallax angles and thus determine the distances to more remote stars.

WEB LINK 17.4

In 1989 the European Space Agency (ESA) launched the satellite *Hipparcos,* an acronym for *High* Precision *Par*allax *Co*llecting *S*atellite (and a commemoration of the ancient Greek astronomer Hipparchus, who created one of the first star charts). Over more than three years of observations, the telescope aboard *Hipparcos* was used to measure the parallaxes of 118,000 stars with an accuracy of 0.001 arcsecond. This has enabled astronomers to determine stellar distances out to several hundred parsecs, and with much greater precision than has been possible with ground-based observations. In the years to come, astronomers will increasingly turn to space-based observations to determine stellar distances.

Unfortunately, most of the stars in the Galaxy are so far away that their parallax angles are too small to measure even with an orbiting telescope. Later in this chapter, we will discuss a technique that can be used to find the distances to these more remote stars. In Chapters 24 and 26 we will learn about other techniques that astronomers use to determine the much larger distances to galaxies beyond the Milky Way. These techniques also

BOX 17-1 Tools of the Astronomer's Trade

Stellar Motions

Stars can move through space in any direction. The **space velocity** of a star describes how fast and in what direction it is moving. As the accompanying figure shows, a star's space velocity v can be broken into components parallel and perpendicular to our line of sight.

The component perpendicular to our line of sight—that is, across the plane of the sky—is called the star's **tangential velocity** (v_t). To determine it, astronomers must know the distance to a star (d) and its **proper motion** (μ, the Greek letter mu), which is the number of arcseconds that the star appears to move per year on the celestial sphere. Proper motion does not repeat itself yearly, so it can be distinguished from the apparent back-and-forth motion due to parallax. In terms of a star's distance and proper motion, its tangential velocity (in km/s) is

$$v_t = 4.74\mu d$$

where μ is in arcseconds per year and d is in parsecs. For example, Barnard's star (Figure 17-3) has a proper motion of 10.358 arcseconds per year and a distance of 1.83 pc. Hence, its tangential velocity is

$$v_t = 4.74(10.358)(1.83) = 89.8 \text{ km/s}$$

The component of a star's motion parallel to our line of sight—that is, either directly toward us or directly away from us—is its **radial velocity** (v_r). It can be determined from measurements of the Doppler shifts of the star's spectral lines (see Section 5-9 and Box 5-6). If a star is approaching us, the wavelengths of all of its spectral lines are decreased (blueshifted); if the star is receding from us, the wavelengths are increased (redshifted). The radial velocity v_r is related to the wavelength shift by the equation

$$\frac{\lambda - \lambda_0}{\lambda_0} = \frac{v_r}{c}$$

In this equation, λ is the wavelength of light coming from the star, λ_0 is what the wavelength would be if the star were not moving, and c is the speed of light. As an illustration, a particular spectral line of iron in the spectrum of Barnard's star has a wavelength (λ) of 516.445 nm. As measured in a laboratory on the Earth, the same spectral line has a wavelength (λ_0) of 516.629 nm. Thus, for Barnard's star, our equation becomes

$$\frac{516.445 \text{ nm} - 516.629 \text{ nm}}{516.629 \text{ nm}} = -0.000356 = \frac{v_r}{c}$$

Solving this equation for the radial velocity v_r, we find

$$v_r = (-0.000356)\ c = (-0.000356)(3.00 \times 10^5 \text{ km/s})$$
$$= -107 \text{ km/s}$$

The negative sign means that Barnard's star is moving toward us. You can check this interpretation by noting that the wavelength $\lambda = 516.445$ nm received from Barnard's star is less than the laboratory wavelength $\lambda_0 = 516.629$ nm; hence, the light from the star is blueshifted, which indeed means that the star is approaching. If the star were receding, its radial velocity would be positive.

The illustration shows that the tangential velocity and radial velocity form two sides of a right triangle. The long side (hypotenuse) of this triangle is the space velocity (v). From the Pythagorean theorem, the space velocity is

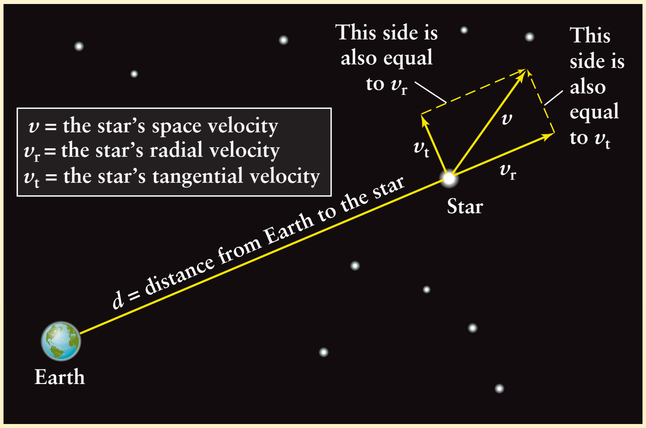
$$v = \sqrt{v_t^2 + v_r^2}$$

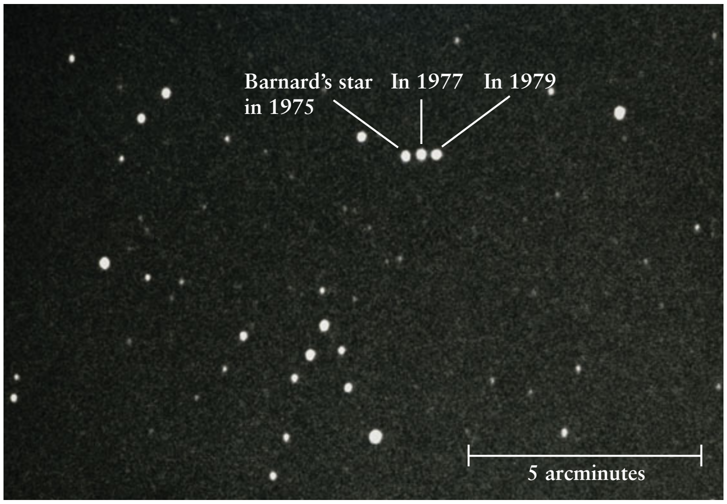
For Barnard's star, the space velocity is

$$v = \sqrt{(89.4 \text{ km/s})^2 + (-107 \text{ km/s})^2} = 140 \text{ km/s}$$

Therefore, Barnard's star is moving through space at a speed of 140 km/s (503,000 km/h, or 312,000 mi/h) relative to the Sun.

Determining the space velocities of stars is essential for understanding the structure of the Galaxy. Studies show that the stars in our local neighborhood are moving in wide orbits around the center of the Galaxy, which lies some 8000 pc (26,000 light-years) away in the direction of the constellation Sagittarius (the Archer). While many of the orbits are roughly circular and lie in nearly the same plane, others are highly elliptical or steeply inclined to the galactic plane. We will see in Chapter 23 how the orbits of stars and gas clouds reveal the Galaxy's spiral structure.





WEB LINK 17.5

Figure 17-3 R I V U X G

The Motion of Barnard's Star Three photographs taken over a four-year period were combined to show the motion of Barnard's star, which lies 1.82 parsecs away in the constellation Ophiuchus. Over this time interval, Barnard's star moved more than 41 arcseconds on the celestial sphere (about 0.69 arcminutes, or 0.012°), more than any other star. (John Sanford/Science Photo Library)

help us understand the overall size, age, and structure of the universe.

The Importance of Parallax Measurements

Because it can be used only on relatively close stars, stellar parallax might seem to be of limited usefulness. But parallax measurements are the cornerstone for all other methods of finding the distances to remote objects. These other methods require a precise and accurate knowledge of the distances to nearby stars, as determined by stellar parallax. Hence, any inaccuracies in the parallax angles for nearby stars can translate into substantial errors in measurement for the whole universe. For this reason, astronomers are continually trying to perfect their parallax-measuring techniques.

Stellar parallax is an *apparent* motion of stars caused by the Earth's orbital motion around the Sun. But stars are not fixed objects and actually do move through space. As a result, stars change their positions on the celestial sphere (Figure 17-3), and they move either toward or away from the Sun. These motions are sufficiently slow, however, that changes in the positions of the stars are hardly noticeable over a human lifetime. Box 17-1 describes how astronomers study these motions and what insights they gain from these studies.

17-2 If a star's distance is known, its luminosity can be determined from its apparent brightness

All the stars you can see in the nighttime sky shine by thermonuclear fusion, just as the Sun does (see Section 16-1). But they are by no means merely identical copies of the Sun. Stars differ in their luminosity (L), the amount of light energy they emit each second. Luminosity is usually measured either in watts (1 watt, or 1 W, is 1 joule per second) or as a multiple of the Sun's luminosity ($L_\odot$, equal to 3.90×10^{26} W). Most stars are less luminous than the Sun, but some blaze forth with a million times the Sun's luminosity. Knowing a star's luminosity is essential for determining the star's history, its present-day internal structure, and its future evolution.

Luminosity, Apparent Brightness, and the Inverse-Square Law

Apparent brightness is a measure of how faint a star looks to us; luminosity is a measure of the star's total light output

To determine the luminosity of a star, we first note that as light energy moves away from its source, it spreads out over increasingly larger regions of space. Imagine a sphere of radius d centered on the light source, as in Figure 17-4. The amount of energy that passes each second through a square meter of the sphere's surface area is the total luminosity of the source (L) divided by the total surface area of the sphere (equal to $4\pi d^2$). This quantity is called the **apparent brightness** of the light, or just **brightness** (b), because how bright a light source appears depends on how much light energy per second enters through the area of a light detector (such as your eye). Apparent brightness is measured in watts per square meter (W/m^2). Written in the form of an equation, the relationship between apparent brightness and luminosity is

Inverse-square law relating apparent brightness and luminosity

$$b = \frac{L}{4\pi d^2}$$

b = apparent brightness of a star's light, in W/m^2

L = star's luminosity, in W

d = distance to star, in meters

This relationship is called the **inverse-square law,** because the apparent brightness of light that an observer can see or measure is inversely proportional to the square of the observer's distance (d) from the source. If you double your distance from a light source, its radiation is spread out over an area 4 times larger, so the apparent brightness you see is decreased by a factor of 4. Similarly, at triple the distance, the apparent brightness is 1/9 as great (see Figure 17-4).

We can apply the inverse-square law to the Sun, which is 1.50×10^{11} m from Earth. Its apparent brightness ($b_\odot$) is

$$b_\odot = \frac{3.90 \times 10^{26}\ \text{W}}{4\pi(1.50 \times 10^{11}\ \text{m})^2} = 1370\ \text{W/m}^2$$

That is, a solar panel with an area of 1 square meter receives 1370 watts of power from the Sun.

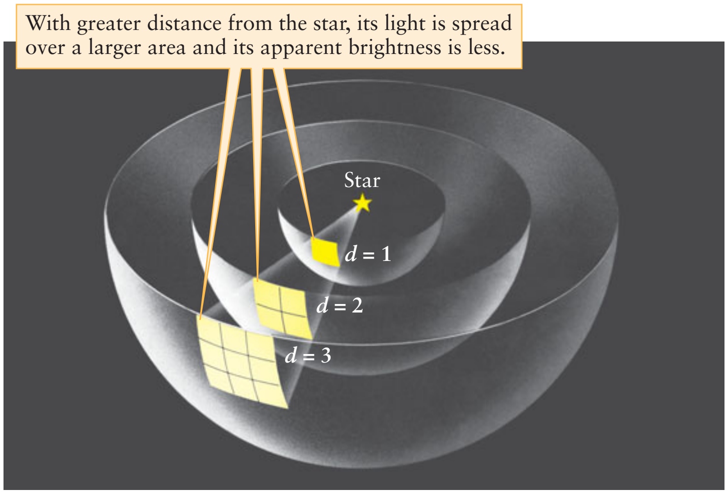


Figure 17-4

The Inverse-Square Law Radiation from a light source illuminates an area that increases as the square of the distance from the source. Hence, the apparent brightness decreases as the square of the distance. The brightness at $d = 2$ is $1/(2^2) = 1/4$ of the brightness at $d = 1$, and the brightness at $d = 3$ is $1/(3^2) = 1/9$ of that at $d = 1$.

Astronomers measure the apparent brightness of a star using a telescope with an attached light-sensitive instrument, similar to the light meter in a camera that determines the proper exposure. Measuring a star's apparent brightness is called **photometry.**

Calculating a Star's Luminosity

The inverse-square law says that we can find a star's luminosity if we know its distance and its apparent brightness. To do this, it is convenient to express this law in a somewhat different form. We first rearrange the above equation:

$$L = 4\pi d^2 b$$

We then apply this equation to the Sun. That is, we write a similar equation relating the Sun's luminosity ($L_\odot$), the distance from the Earth to the Sun ($d_\odot$, equal to 1 AU), and the Sun's apparent brightness ($b_\odot$):

$$L_\odot = 4\pi d_\odot^2 b_\odot$$

If we take the ratio of these two equations, the unpleasant factor of 4π drops out and we are left with the following:

Determining a star's luminosity from its apparent brightness

$$\frac{L}{L_\odot} = \left(\frac{d}{d_\odot}\right)^2 \frac{b}{b_\odot}$$

$L/L_\odot$ = ratio of the star's luminosity to the Sun's luminosity

$d/d_\odot$ = ratio of the star's distance to the Earth-Sun distance

$b/b_\odot$ = ratio of the star's apparent brightness to the Sun's apparent brightness

We need to know just two things to find a star's luminosity: the distance to a star as compared to the Earth-Sun distance (the ratio $d/d_\odot$), and how that star's apparent brightness compares to that of the Sun (the ratio $b/b_\odot$). Then we can use the above equation to find how luminous that star is compared to the Sun (the ratio $L/L_\odot$).

In other words, this equation gives us a general rule relating the luminosity, distance, and apparent brightness of a star:

We can determine the luminosity of a star from its distance and apparent brightness. For a given distance, the brighter the star, the more luminous that star must be. For a given apparent brightness, the more distant the star, the more luminous it must be to be seen at that distance.

Box 17-2 shows how to use the above equation to determine the luminosity of the nearby star ε (epsilon) Eridani, the fifth brightest star in the constellation Eridanus (named for a river in Greek mythology). Parallax measurements indicate that ε Eridani is 3.23 pc away, and photometry shows that the star appears only 6.73×10^{-13} as bright as the Sun; using the above equation, we find that ε Eridani has only 0.30 times the luminosity of the Sun.

The Stellar Population

Calculations of this kind show that stars come in a wide variety of different luminosities, with values that range from about $10^6\ L_\odot$ (a million times the Sun's luminosity) to only about $10^{-4}\ L_\odot$ (a mere ten-thousandth of the Sun's light output). The most luminous star emits roughly 10^{10} times more energy each second than the least luminous! (To put this number in perspective, about 10^{10} human beings have lived on the Earth since our species first evolved.)

As stars go, our Sun is neither extremely luminous nor extremely dim; it is a rather ordinary, garden-variety star. It is somewhat more luminous than most stars, however. Of more than 30 stars within 4 pc of the Sun (see Appendix 4), only three (α Centauri, Sirius, and Procyon) have a greater luminosity than the Sun.

To better characterize a typical population of stars, astronomers count the stars out to a certain distance from the Sun and plot the number of stars that have different luminosities. The resulting graph is called the **luminosity function.** Figure 17-5 on page 440 shows the luminosity function for stars in our part of the Milky Way Galaxy. The curve declines very steeply for the most luminous stars toward the left side of the graph, indicating that they are quite rare. For example, this graph shows that stars like the Sun are about 10,000 times more common than stars like Spica (which has a luminosity of 2100 $L_\odot$).

The exact shape of the curve in Figure 17-5 applies only to the vicinity of the Sun and similar regions in our Milky Way Galaxy. Other locations have somewhat different luminosity functions. In stellar populations in general, however, low-luminosity stars are much more common than high-luminosity ones.

BOX 17-2 Tools of the Astronomer's Trade

Luminosity, Distance, and Apparent Brightness

The inverse-square law (Section 17-2) relates a star's luminosity, distance, and apparent brightness to the corresponding quantities for the Sun:

$$\frac{L}{L_\odot} = \left(\frac{d}{d_\odot}\right)^2 \frac{b}{b_\odot}$$

We can use a similar equation to relate the luminosities, distances, and apparent brightnesses of *any* two stars, which we call star 1 and star 2:

$$\frac{L_1}{L_2} = \left(\frac{d_1}{d_2}\right)^2 \frac{b_1}{b_2}$$

EXAMPLE: The star ε (epsilon) Eridani is 3.23 pc from Earth. As seen from Earth, this star appears only 6.73×10^{-13} as bright as the Sun. What is the luminosity of ε Eridani compared with that of the Sun?

Situation: We are given the distance to ε Eridani (d = 3.23 pc) and this star's brightness compared to that of the Sun ($b/b_\odot = 6.73 \times 10^{-13}$). Our goal is to find the ratio of the luminosity of ε Eridani to that of the Sun, that is, the quantity $L/L_\odot$.

Tools: Since we are asked to compare this star to the Sun, we use the first of the two equations given above, $L/L_\odot = (d/d_\odot)^2(b/b_\odot)$, to solve for $L/L_\odot$.

Answer: Our equation requires the ratio of the star's distance to the Sun's distance, $d/d_\odot$. The distance from the Earth to the Sun is $d_\odot$ = 1 AU. To calculate the ratio $d/d_\odot$, we must express both distances in the same units. There are 206,265 AU in 1 parsec, so we can write the distance to ε Eridani as d = (3.23 pc)(206,265 AU/pc) = 6.66×10^5 AU. Hence, the ratio of distances is $d/d_\odot = (6.66 \times 10^5 \text{ AU})/(1 \text{ AU}) = 6.66 \times 10^5$. Then we find that the ratio of the luminosity of ε Eridani (L) to the Sun's luminosity ($L_\odot$) is

$$\frac{L}{L_\odot} = \left(\frac{d}{d_\odot}\right)^2 \frac{b}{b_\odot} = (6.66 \times 10^5)^2 \times (6.73 \times 10^{-13}) = 0.30$$

Review: This result means that ε Eridani is only 0.30 as luminous as the Sun; that is, its power output is only 30% as great.

EXAMPLE: Suppose star 1 is at half the distance of star 2 (that is, $d_1/d_2 = \frac{1}{2}$) and that star 1 appears twice as bright as star 2 (that is, $b_1/b_2 = 2$). How do the luminosities of these two stars compare?

Situation: For these two stars, we are given the ratio of distances (d_1/d_2) and the ratio of apparent brightnesses (b_1/b_2). Our goal is to find the ratio of their luminosities (L_1/L_2).

Tools: Since we are comparing two stars, neither of which is the Sun, we use the second of the two equations above: $L_1/L_2 = (d_1/d_2)^2(b_1/b_2)$.

Answer: Plugging in to our equation, we find

$$\frac{L_1}{L_2} = \left(\frac{d_1}{d_2}\right)^2 \frac{b_1}{b_2} = \left(\frac{1}{2}\right)^2 \times 2 = \frac{1}{2}$$

Review: This result says that star 1 has only one-half the luminosity of star 2. Despite this, star 1 appears brighter than star 2 because it is closer to us.

The two equations above are also useful in the method of *spectroscopic parallax,* which we discuss in Section 17-8. It turns out that a star's luminosity can be determined simply by analyzing the star's spectrum. If the star's apparent brightness is also known, the star's distance can be calculated. The inverse-square law can be rewritten as an expression for the ratio of the star's distance from the Earth (d) to the Earth-Sun distance ($d_\odot$):

$$\frac{d}{d_\odot} = \sqrt{\frac{(L/L_\odot)}{(b/b_\odot)}}$$

We can also use this formula as a relation between the properties of any two stars, 1 and 2:

$$\frac{d_1}{d_2} = \sqrt{\frac{(L_1/L_2)}{(b_1/b_2)}}$$

EXAMPLE: The star Pleione in the constellation Taurus is 190 times as luminous as the Sun but appears only 3.19×10^{-13} as bright as the Sun. How far is Pleione from Earth?

Situation: We are told the ratio of Pleione's luminosity to that of the Sun ($L/L_\odot$ = 190) and the ratio of their apparent brightnesses ($b/b_\odot = 3.19 \times 10^{-13}$). Our goal is to find the distance d from the Earth to Pleione.

Tools: Since we are comparing Pleione to the Sun, we use the first of the two equations above.

Answer: Our equation tells us the ratio of the Earth-Pleione distance to the Earth-Sun distance:

$$\frac{d}{d_\odot} = \sqrt{\frac{(L/L_\odot)}{(b/b_\odot)}} = \sqrt{\frac{190}{3.19 \times 10^{-13}}}$$

$$= \sqrt{5.95 \times 10^{14}} = 2.44 \times 10^7$$

(continued on the next page)

BOX 17-2 *(continued)*

Hence, the distance from Earth to Pleione is 2.44×10^7 times greater than the distance from Earth to the Sun. The Sun-Earth distance is $d_\odot = 1$ AU and 206,265 AU = 1 pc, so we can express the star's distance as $d = (2.44 \times 10^7 \text{ AU}) \times (1 \text{ pc}/206{,}265 \text{ AU}) = 118$ pc.

Review: We can check our result by comparing with the above example about the star ε Eridani. Pleione has a much greater luminosity than ε Eridani (190 times the Sun's luminosity versus 0.30 times), but Pleione appears dimmer than ε Eridani (3.19×10^{-13} times as bright as the Sun compared to 6.73×10^{-13} times). For this to be true, Pleione must be much farther away from Earth than is ε Eridani. This is just what our results show: $d = 118$ pc for Pleione compared to $d = 3.23$ pc for ε Eridani.

EXAMPLE: The star δ (delta) Cephei, which lies 300 pc from the Earth, is thousands of times more luminous than the Sun. Thanks to this great luminosity, stars like δ Cephei can be seen in galaxies millions of parsecs away. As an example, the Hubble Space Telescope has detected stars like δ Cephei within the galaxy NGC 3351, which lies in the direction of the constellation Leo. These stars appear only 9×10^{-10} as bright as δ Cephei. What is the distance to NGC 3351?

Situation: To determine the distance we want, we need to find the distance to a star within NGC 3351. We are told that certain stars within this galaxy are identical to δ Cephei but appear only 9×10^{-10} as bright.

Tools: We use the equation $d_1/d_2 = \sqrt{(L_1/L_2)/(b_1/b_2)}$ to relate two stars, one within NGC 3351 (call this star 1) and the identical star δ Cephei (star 2). Our goal is to find d_1.

Answer: Since the two stars are identical, they have the same luminosity ($L_1 = L_2$, or $L_1/L_2 = 1$). The brightness ratio is $b_1/b_2 = 9 \times 10^{-10}$, so our equation tells us that

$$\frac{d_1}{d_2} = \sqrt{\frac{(L_1/L_2)}{(b_1/b_2)}} = \sqrt{\frac{1}{9 \times 10^{-10}}} = \sqrt{1.1 \times 10^9} = 33{,}000$$

Hence, NGC 3351 is 33,000 times farther away than δ Cephei, which is 300 pc from Earth. The distance from the Earth to NGC 3351 is therefore (33,000)(300 pc) = 10^7 pc, or 10 megaparsecs (10 Mpc).

Review: This example illustrates one technique that astronomers use to measure extremely large distances. We will learn more about stars like δ Cephei in Chapter 19, and in Chapter 24 we will explore further how they are used to determine the distances to remote galaxies.

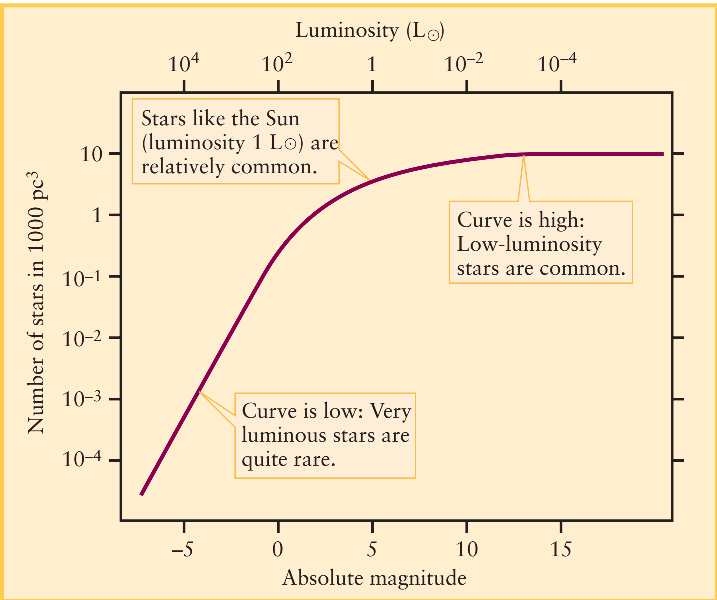


Figure 17-5

The Luminosity Function This graph shows how many stars of a given luminosity lie within a representative 1000 cubic-parsec volume. The scale at the bottom of the graph shows absolute magnitude, an alternative measure of a star's luminosity (described in Section 17-3). (Adapted from J. Bahcall and R. Soneira)

17-3 Astronomers often use the magnitude scale to denote brightness

Because astronomy is among the most ancient of sciences, some of the tools used by modern astronomers are actually many centuries old. One such tool is the **magnitude scale,** which astronomers frequently use to denote the brightness of stars. This scale was introduced in the second century B.C. by the Greek astronomer Hipparchus, who called the brightest stars first-magnitude stars. Stars about half as bright as first-magnitude stars were called second-magnitude stars, and so forth, down to sixth-magnitude stars, the dimmest ones he could see. After telescopes came into use, astronomers extended Hipparchus's magnitude scale to include even dimmer stars.

Apparent Magnitudes

The magnitudes in Hipparchus's scale are properly called **apparent magnitudes,** because they describe how bright an object *appears* to an Earth-based observer. Apparent magnitude is directly related to apparent brightness.

CAUTION! The magnitude scale can be confusing because it works "backward." Keep in mind that the *greater* the apparent magnitude, the *dimmer* the star. A star of apparent magnitude +3 (a third-magnitude star) is dimmer than a star of apparent magnitude +2 (a second-magnitude star).

In the nineteenth century, astronomers developed better techniques for measuring the light energy arriving from a star. These measurements showed that a first-magnitude star is about 100 times brighter than a sixth-magnitude star. In other words, it would take 100 stars of magnitude +6 to provide as much light energy as we receive from a single star of magnitude +1. To make computations easier, the magnitude scale was redefined so that a magnitude difference of 5 corresponds exactly to a factor of 100 in brightness. A magnitude difference of 1 corresponds to a factor of 2.512 in brightness, because

$$2.512 \times 2.512 \times 2.512 \times 2.512 \times 2.512 = (2.512)^5 = 100$$

Thus, it takes 2.512 third-magnitude stars to provide as much light as we receive from a single second-magnitude star.

Figure 17-6 illustrates the modern apparent magnitude scale. The dimmest stars visible through a pair of binoculars have an apparent magnitude of +10, and the dimmest stars that can be photographed in a one-hour exposure with the Keck telescopes (see Section 6-2) or the Hubble Space Telescope have apparent magnitude +30. Modern astronomers also use negative numbers to extend Hipparchus's scale to include very bright objects. For example, Sirius, the brightest star in the sky, has an apparent magnitude of −1.43. The Sun, the brightest object in the sky, has an apparent magnitude of −26.7.

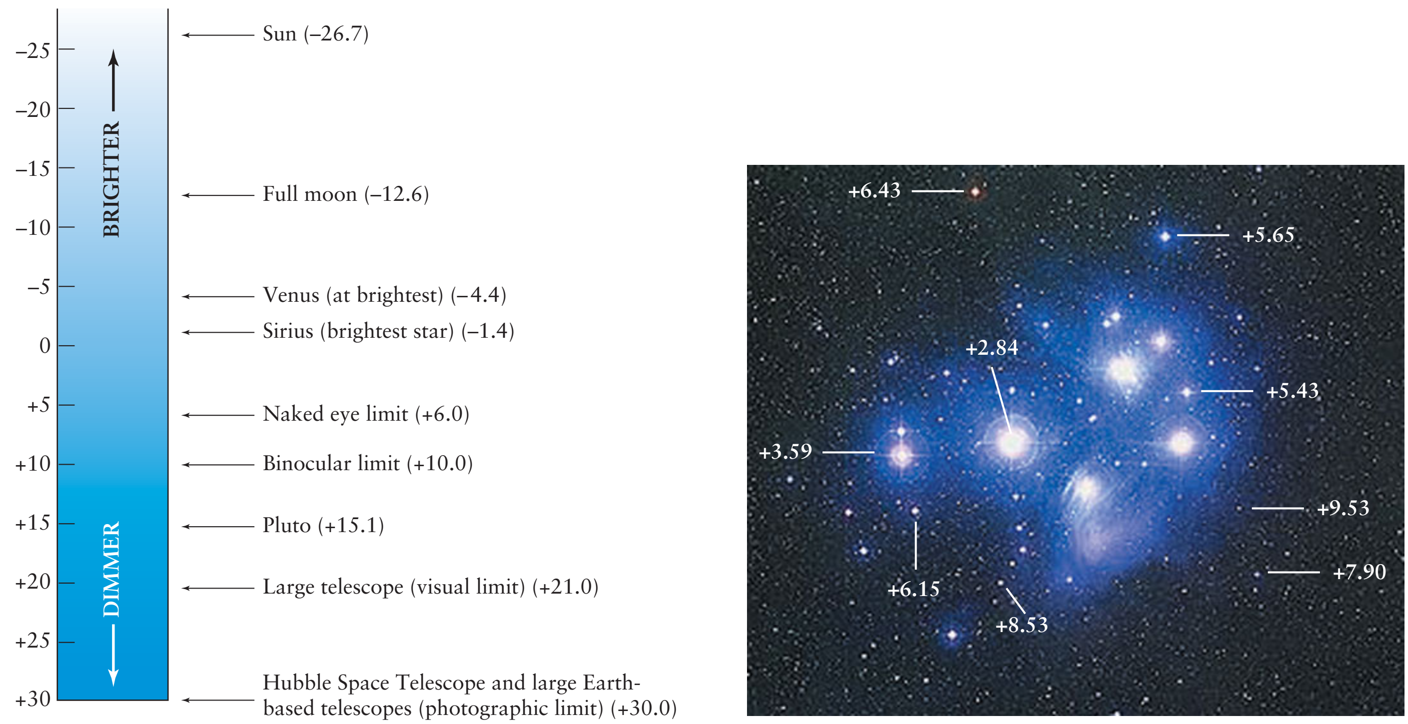
Absolute Magnitudes

Apparent magnitude measures a star's brightness; absolute magnitude measures its luminosity

Apparent magnitude is a measure of a star's apparent brightness as seen from Earth. A related quantity that measures a star's true energy output—that is, its luminosity—is called **absolute magnitude.** This is the apparent magnitude a star would have if it were located exactly 10 parsecs from Earth.

ANALOGY If you wanted to compare the light output of two different lightbulbs, you would naturally place them side by side so that both bulbs were the same distance from you. In the absolute magnitude scale, we imagine doing the same thing with stars to compare their luminosities.

If the Sun were moved to a distance of 10 parsecs from the Earth, it would have an apparent magnitude of +4.8. The absolute magnitude of the Sun is thus +4.8. The absolute magnitudes



(a) Some apparent magnitudes

(b) Apparent magnitudes of stars in the Pleiades

R I V U X G

WEB LINK 17.6

Figure 17-6

The Apparent Magnitude Scale (a) Astronomers denote the apparent brightness of objects in the sky by their apparent magnitudes. The greater the apparent magnitude, the dimmer the object. **(b)** This photograph of the Pleiades cluster, located about 120 pc away in the constellation Taurus, shows the apparent magnitudes of some of its stars. Most are too faint to be seen by the naked eye. (David Malin/Anglo-Australian Observatory) of the stars range from approximately +15 for the least luminous to −10 for the most luminous. (*Note:* Like apparent magnitudes, absolute magnitudes work "backward": The *greater* the absolute magnitude, the *less luminous* the star.) The Sun's absolute magnitude is about in the middle of this range.

We saw in Section 17-2 that we can calculate the luminosity of a star if we know its distance and apparent brightness. There is a mathematical relationship between absolute magnitude and luminosity, which astronomers use to convert one to the other as they see fit. It is also possible to rewrite the inverse-square law, which we introduced in Section 17-2, as a mathematical relationship that allows you to calculate a star's absolute magnitude (a measure of its luminosity) from its distance and apparent magnitude (a measure of its apparent brightness). Box 17-3 describes these relationships and how to use them.

Because the "backward" magnitude scales can be confusing, we will use them only occasionally in this book. We will usually speak of a star's luminosity rather than its absolute magnitude and will describe a star's appearance in terms of apparent brightness rather than apparent magnitude. But if you go on to study more about astronomy, you will undoubtedly make frequent use of apparent magnitude and absolute magnitude.

17-4 A star's color depends on its surface temperature

WEB LINK 17.7

The image that opens this chapter shows that stars come in different colors. You can see these colors even with the naked eye. For example, you can easily see the red color of Betelgeuse, the star in the "armpit" of the constellation Orion, and the blue tint of Bellatrix at Orion's other "shoulder" (see Figure 2-2). Colors are most evident for the brightest stars, because your color vision does not work well at low light levels.

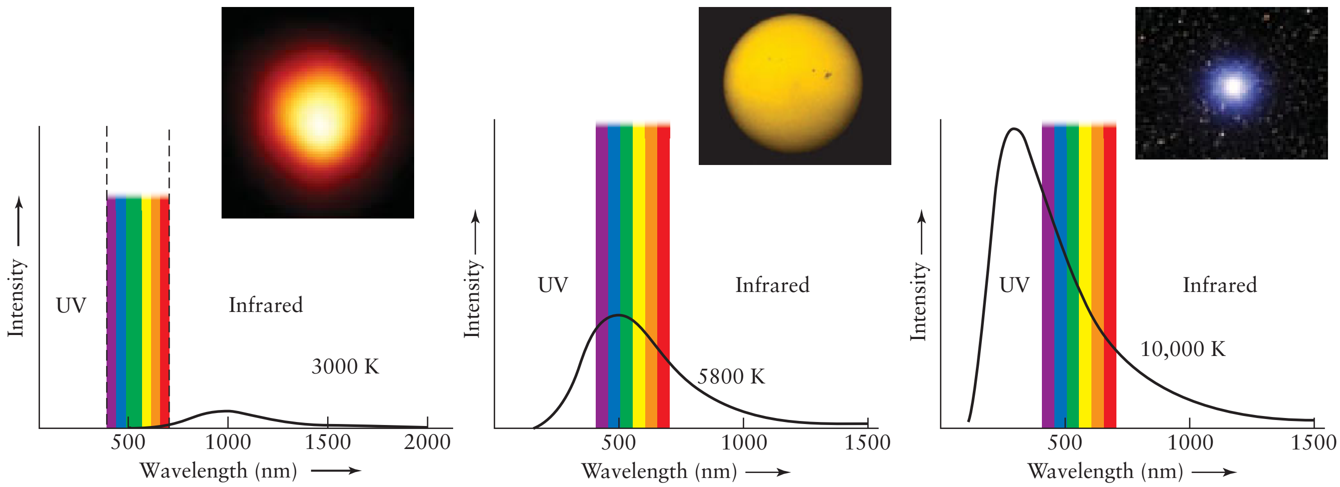
CAUTION! It's true that the light from a star will appear redshifted if the star is moving away from you and blueshifted if it's moving toward you. But for even the fastest stars, these color shifts are so tiny that it takes sensitive instruments to measure them. The red color of Betelgeuse and the blue color of Bellatrix are not due to their motions; they are the actual colors of the stars.

Color and Temperature

We saw in Section 5-3 that a star's color is directly related to its surface temperature. The intensity of light from a relatively cool star peaks at long wavelengths, making the star look red (Figure 17-7*a*). A hot star's intensity curve peaks at shorter wavelengths, so the star looks blue (Figure 17-7*c*). For a star with an intermediate temperature, such as the Sun, the intensity peak is near the middle of the visible spectrum. This gives the star a yellowish color (Figure 17-7*b*). This leads to an important general rule about star colors and surface temperatures:

Red stars are relatively cold, with low surface temperatures; blue stars are relatively hot, with high surface temperatures.

Figure 17-7 shows that astronomers can accurately determine the surface temperature of a star by carefully measuring its color.



(a) A **cool star** with surface temperature 3000 K emits much more red light than blue light, and so appears red.

(b) A **warmer star** with surface temperature 5800 K (like the Sun) emits roughly equal amounts of all visible wavelengths, and so appears yellow-white.

(c) A **hot star** with surface temperature 10,000 K emits much more blue light than red light, and so appears blue.

Figure 17-7

Temperature and Color These graphs show the intensity of light emitted by three hypothetical stars plotted against wavelength (compare with Figure 5-11). The rainbow band indicates the range of visible wavelengths. The star's apparent color depends on whether the intensity curve has larger values at the short-wavelength or long-wavelength end of the visible spectrum. The insets show stars of about these surface temperatures. UV stands for ultraviolet, which extends from 10 to 400 nm. See Figure 3-4 for more on wavelengths of the spectrum. (Inset a: Andrea Dupree/Harvard-Smithsonian CFA, Ronald Gilliland/STScI, NASA and ESA; inset b: NSO/AURA/NSF; inset c: Till Credner, Allthesky.com)

BOX 17-3 Tools of the Astronomer's Trade

Apparent Magnitude and Absolute Magnitude

Astronomers commonly express a star's apparent brightness in terms of apparent magnitude (denoted by a lowercase m), and the star's luminosity in terms of absolute magnitude (denoted by a capital M). While we do not use these quantities extensively in this book, it is useful to know a few simple relationships involving them.

Consider two stars, labeled 1 and 2, with apparent magnitudes m_1 and m_2 and brightnesses b_1 and b_2, respectively. The *ratio* of their apparent brightnesses (b_1/b_2) corresponds to a *difference* in their apparent magnitudes ($m_2 - m_1$). As we learned in Section 17-3, each step in magnitude corresponds to a factor of 2.512 in brightness; we receive 2.512 times more energy per square meter per second from a third-magnitude star than from a fourth-magnitude star. This idea was used to construct the following table:

Apparent magnitude difference ($m_2 - m_1$)	**Ratio of apparent brightness (b_1/b_2)**
1	2.512
2	$(2.512)^2 = 6.31$
3	$(2.512)^3 = 15.85$
4	$(2.512)^4 = 39.82$
5	$(2.512)^5 = 100$
10	$(2.512)^{10} = 10^4$
15	$(2.512)^{15} = 10^6$
20	$(2.512)^{20} = 10^8$

A simple equation relates the difference between two stars' apparent magnitudes to the ratio of their brightnesses:

Magnitude difference related to brightness ratio

$$m_2 - m_1 = 2.5 \log\left(\frac{b_1}{b_2}\right)$$

m_1, m_2 = apparent magnitudes of stars 1 and 2

b_1, b_2 = apparent brightnesses of stars 1 and 2

In this equation, $\log(b_1/b_2)$ is the logarithm of the brightness ratio. The logarithm of $1000 = 10^3$ is 3, the logarithm of $10 = 10^1$ is 1, and the logarithm of $1 = 10^0$ is 0.

EXAMPLE: At their most brilliant, Venus has a magnitude of about -4 and Mercury has a magnitude of about -2. How many times brighter are these than the dimmest stars visible to the naked eye, with a magnitude of $+6$?

Situation: In each case we want to find a ratio of two apparent brightnesses (the brightness of Venus or Mercury compared to that of the dimmest naked-eye stars).

Tools: In each case we will convert a *difference* in apparent magnitude between the planet and the naked-eye star into a *ratio* of their brightnesses.

Answer: The magnitude difference between Venus and the dimmest stars visible to the naked eye is $+6 - (-4) = 10$. From the table, this difference corresponds to a brightness ratio of $(2.512)^{10} = 10^4 = 10{,}000$, so Venus at it most brilliant is 10,000 times brighter than the dimmest naked-eye stars.

The magnitude difference between Mercury and the dimmest naked-eye stars is $+4 - (-4) = 8$. While this value is not in the table, you can see that the corresponding ratio of brightnesses is $(2.512)^8 = (2.512)^{5+3} = (2.512)^5 \times (2.512)^3$. From the table, $(2.512)^5 = 100$ and $(2.512)^3 = 15.85$, so the ratio of brightnesses is $100 \times 15.85 = 1585$. Hence, Mercury at its most brilliant is 1585 times brighter than the dimmest stars visible to the naked eye.

Review: Can you show that when at their most brilliant, Venus is 6.31 times brighter than Mercury? (*Hint:* No multiplication or division is required—just notice the difference in apparent magnitude between Venus and Mercury, and consult the table.)

EXAMPLE: The variable star RR Lyrae in the constellation Lyra (the Harp) periodically doubles its light output. By how much does its apparent magnitude change?

Situation: We are given a ratio of two brightnesses (the star at its maximum is twice as bright as at its minimum). Our goal is to find the corresponding difference in apparent magnitude.

Tools: We let 1 denote the star at its maximum brightness and 2 denote the same star at its dimmest, so the ratio of brightnesses is $b_1/b_2 = 2$. We then use the equation $m_2 - m_1 = 2.5 \log(b_1/b_2)$ to solve for the apparent magnitude difference $m_2 - m_1$.

Answer: Using a calculator, we find $m_2 - m_1 = 2.5 \log(2) = 2.5 \times 0.30 = 0.75$. RR Lyrae therefore varies periodically in brightness by 0.75 magnitude.

Review: Our answer means that at its dimmest, RR Lyrae has an apparent magnitude m_2 that is 0.75 *greater* than its apparent magnitude m_1 when it is brightest. (Remember that a greater value of apparent magnitude means the star is dimmer, not brighter!)

(continued on the next page)

BOX 17-3 *(continued)*

The inverse-square law relating a star's apparent brightness and luminosity can be rewritten in terms of the star's apparent magnitude (m), absolute magnitude (M), and distance from the Earth (d). This can be expressed as an equation:

Relation between a star's apparent magnitude and absolute magnitude

$$m - M = 5 \log d - 5$$

m = star's apparent magnitude

M = star's absolute magnitude

d = distance from the Earth to the star in parsecs

In this expression $m - M$ is called the **distance modulus,** and log d means the logarithm of the distance d in parsecs. For convenience, the following table gives the values of the distance d corresponding to different values of $m - M$.

Distance modulus $m - M$	Distance d (pc)
−4	1.6
−3	2.5
−2	4.0
−1	6.3
0	10
1	16
2	25
3	40
4	63
5	100
10	10^3
15	10^4
20	10^5

This table shows that if a star is less than 10 pc away, its distance modulus $m - M$ is negative. That is, its apparent magnitude (m) is less than its absolute magnitude (M). If the star is more than 10 pc away, $m - M$ is positive and m is greater than M. As an example, the star ε (epsilon) Indi, which is in the direction of the southern constellation Indus, has apparent magnitude $m = +4.7$. It is 3.6 pc away, which is less than 10 pc, so its apparent magnitude is less than its absolute magnitude.

EXAMPLE: Find the absolute magnitude of ε Indi.

Situation: We are given the distance to ε Indi ($d = 3.6$ pc) and its apparent magnitude ($m = +4.7$). Our goal is to find the star's absolute magnitude M.

Tools: We use the formula $m - M = 5 \log d - 5$ to solve for M.

Answer: Since $d = 3.6$ pc, we use a calculator to find log $d = \log 3.6 = 0.56$. Therefore the star's distance modulus is $m - M = 5(0.56) - 5 = -2.2$, and the star's absolute magnitude is $M = m - (-2.2) = +4.7 + 2.2 = +6.9$.

Review: As a check on our calculations, note that this star's distance modulus $m - M = -2.2$ is less than zero, as it should be for a star less than 10 pc away. Note that our Sun has absolute magnitude +4.8; ε Indi has a greater absolute magnitude, so it is less luminous than the Sun.

EXAMPLE: Suppose you were viewing the Sun from a planet orbiting another star 100 pc away. Could you see it without using a telescope?

Situation: We learned in the preceding examples that the Sun has absolute magnitude $M = +4.8$ and that the dimmest stars visible to the naked eye have apparent magnitude $m = +6$. Our goal is to determine whether the Sun would be visible to the naked eye at a distance of 100 pc.

Tools: We use the relationship $m - M = 5 \log d - 5$ to find the Sun's apparent magnitude at $d = 100$ pc. If this is greater than +6, the Sun would not be visible at that distance. (Remember that the greater the apparent magnitude, the dimmer the star.)

Answer: From the table, at $d = 100$ pc the distance modulus is $m - M = 5$. So, as seen from this distant planet, the Sun's apparent magnitude would be $m = M + 5 = +4.8 + 5 = +9.8$. This is greater than the naked-eye limit $m = +6$, so the Sun could not be seen.

Review: The Sun is by far the brightest object in the Earth's sky. But our result tells us that to an inhabitant of a planetary system 100 pc away—a rather small distance in a galaxy that is thousands of parsecs across—our own Sun would be just another insignificant star, visible only through a telescope.

The magnitude system is also used by astronomers to express the colors of stars as seen through different filters, as we describe in Section 17-4. For example, rather than quantifying a star's color by the color ratio b_V/b_B (a star's apparent brightness as seen through a V filter divided by the brightness through a B filter), astronomers commonly use the *color index* B–V, which is the difference in the star's apparent magnitude as measured with these two filters. We will not use this system in this book, however (but see Advanced Questions 53 and 54).

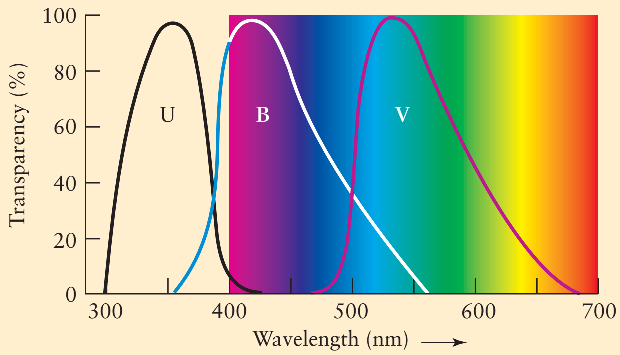


Figure 17-8

U, B, and V Filters This graph shows the wavelengths to which the standard filters are transparent. The U filter is transparent to near-ultraviolet light. The B filter is transparent to violet, blue, and green light, while the V filter is transparent to green and yellow light. By measuring the apparent brightness of a star with each of these filters and comparing the results, an astronomer can determine the star's surface temperature.

Astronomers use a set of filters in their telescopes to measure the surface temperatures of stars

To do this, the star's light is collected by a telescope and passed through one of a set of differently colored filters. The filtered light is then collected by a light-sensitive device such as a CCD (see Section 6-4). The process is then repeated with each of the filters in the set. The star's image will have a different brightness through each colored filter, and by comparing these brightnesses astronomers can find the wavelength at which the star's intensity curve has its peak—and hence the star's temperature.

UBV Photometry

Let's look at this procedure in more detail. The most commonly used filters are called U, B, and V, and the technique that uses them is called **UBV photometry.** Each filter is transparent in a different band of wavelengths: the ultraviolet (U), the blue (B), and the yellow-green (V, for visual) region of the visible spectrum (Figure 17-8). The transparency of the V filter mimics the sensitivity of the human eye.

To determine a star's temperature using UBV photometry, the astronomer first measures the star's brightness through each of the filters individually. This gives three apparent brightnesses for the star, designated b_U, b_B, and b_V. The astronomer then compares the intensity of starlight in neighboring wavelength bands by taking the ratios of these brightnesses: b_V/b_B and b_B/b_U. Table 17-1 gives values for these **color ratios** for several stars with different surface temperatures.

If a star is very hot, its radiation is skewed toward short, ultraviolet wavelengths as in Figure 17-7*c*. This makes the star dim through the V filter, brighter through the B filter, and brightest through the U filter. Hence, for a hot star b_V is less than b_B, which in turn is less than b_U, and the ratios b_V/b_B and b_B/b_U are both less than 1. One such star is Bellatrix (see Table 17-1), which has a surface temperature of 21,500 K.

In contrast, if a star is cool, its radiation peaks at long wavelengths as in Figure 17-7*a*. Such a star appears brightest through the V filter, dimmer through the B filter, and dimmest through the U filter (see Figure 17-8). In other words, for a cool star b_V is greater than b_B, which in turn is greater than b_U. Hence, the ratios b_V/b_B and b_B/b_U will both be greater than 1. The star Betelgeuse (surface temperature 3500 K) is an example.

You can see these differences between hot and cool stars in parts *a* and *c* of Figure 6-27, which show the constellation Orion at ultraviolet wavelengths (a bit shorter than those transmitted by the U filter) and at visible wavelengths that approximate the transmission of a V filter. The hot star Bellatrix is brighter in the ultraviolet image (Figure 6-27*a*) than at visible wavelengths (Figure 6-27*c*). (Figure 6-27*d* shows the names of the stars.) The situation is reversed for the cool star Betelgeuse: It is bright at visible wavelengths, but at ultraviolet wavelengths it is too dim to show up in the image.

Figure 17-9 graphs the relationship between a star's b_V/b_B color ratio and its temperature. If you know the value of the b_V/b_B color ratio for a given star, you can use this graph to find the star's surface temperature. As an example, for the Sun b_V/b_B

Table 17-1 Colors of Selected Stars

Star	Surface temperature (K)	b_V/b_B	b_B/b_U	Apparent color
Bellatrix (γ Orionis)	21,500	0.81	0.45	Blue
Regulus (α Leonis)	12,000	0.90	0.72	Blue-white
Sirius (α Canis Majoris)	9400	1.00	0.96	Blue-white
Megrez (δ Ursae Majoris)	8630	1.07	1.07	White
Altair (α Aquilae)	7800	1.23	1.08	Yellow-white
Sun	5800	1.87	1.17	Yellow-white
Aldebaran (α Tauri)	4000	4.12	5.76	Orange
Betelgeuse (α Orionis)	3500	5.55	6.66	Red

Source: J.-C. Mermilliod, B. Hauck, and M. Mermilliod, University of Lausanne

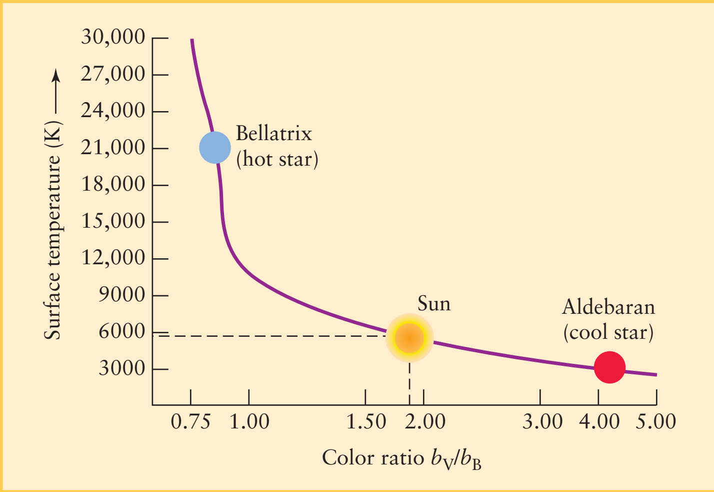


Figure 17-9

Temperature, Color, and Color Ratio The b_V/b_B color ratio is the ratio of a star's apparent brightnesses through a V filter and through a B filter. This ratio is small for hot, blue stars but large for cool, red stars. After measuring a star's brightness with the B and V filters, an astronomer can estimate the star's surface temperature from a graph like this one.

equals 1.87, which corresponds to a surface temperature of 5800 K.

CAUTION! As we will see in Chapter 18, tiny dust particles that pervade interstellar space cause distant stars to appear redder than they really are. (In the same way, particles in the Earth's atmosphere make the setting Sun look redder; see Box 5-4.) Astronomers must take this reddening into account whenever they attempt to determine a star's surface temperature from its color ratios. A star's spectrum provides a more precise measure of a star's surface temperature, as we will see next. But it is quicker and easier to observe a star's colors with a set of U, B, and V filters than it is to take the star's spectrum with a spectrograph.

17-5 The spectra of stars reveal their chemical compositions as well as their surface temperatures

We have seen how the color of a star's light helps astronomers determine its surface temperature. To determine the other properties of a star, astronomers must analyze the spectrum of its light in more detail. This technique of *stellar spectroscopy* began in 1817 when Joseph Fraunhofer, a German instrument maker, attached a spectroscope to a telescope and pointed it toward the stars. Fraunhofer had earlier observed that the Sun has an absorption line spectrum—that is, a continuous spectrum with dark absorption lines (see Section 5-6). He found that stars have the same kind of spectra, which reinforces the idea that our Sun is a rather typical star. But Fraunhofer also found that the pattern of absorption lines is different for different stars.

Classifying Stars: Absorption Line Spectra and Spectral Classes

We see an absorption line spectrum when a cool gas lies between us and a hot, glowing object (recall Figure 5-16). The light from the hot, glowing object itself has a continuous spectrum. In the case of a star, light with a continuous spectrum is produced at low-lying levels of the star's atmosphere where the gases are hot and dense. The absorption lines are created when this light flows outward through the upper layers of the star's atmosphere. Atoms in these cooler, less dense layers absorb radiation at specific wavelengths, which depend on the specific kinds of atoms present—hydrogen, helium, or other elements—and on whether or not the atoms are ionized. Absorption lines in the Sun's spectrum are produced in this same way (see Section 16-5).

WEB LINK 17.8

Some stars have spectra in which the Balmer absorption lines of hydrogen are prominent. But in the spectra of other stars, including the Sun, the Balmer lines are nearly absent and the dominant absorption lines are those of heavier elements such as calcium, iron, and sodium. Still other stellar spectra are dominated by broad absorption lines caused by molecules, such as titanium oxide, rather than single atoms. To cope with this diversity, astronomers group similar-appearing stellar spectra into **spectral classes.** In a popular classification scheme that emerged in the late 1890s, a star was assigned a letter from A through O according to the strength or weakness of the hydrogen Balmer lines in the star's spectrum.

Nineteenth-century science could not explain why or how the spectral lines of a particular chemical are affected by the temperature and density of the gas. Nevertheless, a team of astronomers at the Harvard College Observatory forged ahead with a monumental project of examining the spectra of hundreds of thousands of stars. Their goal was to develop a system of spectral classification in which all spectral features, not just Balmer lines, change smoothly from one spectral class to the next.

Deciphering the information in starlight took the painstaking work of generations of astronomers

The Harvard project was financed by the estate of Henry Draper, a wealthy New York physician and amateur astronomer who in 1872 became the first person to photograph stellar absorption lines. Researchers on the project included Edward C. Pickering, Williamina Fleming, Antonia Maury, and Annie Jump Cannon (Figure 17-10). As a result of their efforts, many of the original A-through-O classes were dropped and others were consolidated. The remaining spectral classes were reordered in the sequence **OBAFGKM.** You can remember this sequence with the mnemonic: "*O*h, *B*e *A* *F*ine *G*irl (or Guy), *K*iss *M*e!"

Refining the Classification: Spectral Types

Cannon refined the original OBAFGKM sequence into smaller steps called **spectral types.** These steps are indicated by attaching an integer from 0 through 9 to the original letter. For example, the spectral class F includes spectral types F0, F1, F2, . . . , F8, F9, which are followed by the spectral types G0, G1, G2, . . . , G8, G9, and so on.

WEB LINK 17.9

Figure 17-10 RIVUXG

Classifying the Spectra of the Stars The modern scheme of classifying stars by their spectra was developed at the Harvard College Observatory in the late nineteenth century. A team of women astronomers led by Edward C. Pickering and Williamina Fleming (standing) analyzed hundreds of thousands of spectra. Social conventions of the time prevented most women astronomers from using research telescopes or receiving salaries comparable to men's. (Harvard College Observatory)

Figure 17-11 shows representative spectra of several spectral types. The strengths of spectral lines change gradually from one spectral type to the next. For example, the Balmer absorption lines of hydrogen become increasingly prominent as you go from spectral type B0 to A0. From A0 onward through the F and G classes, the hydrogen lines weaken and almost fade from view. The Sun, whose spectrum is dominated by calcium and iron, is a G2 star.

The Harvard project culminated in the *Henry Draper Catalogue,* published between 1918 and 1924. It listed 225,300 stars, each of which Cannon had personally classified. Meanwhile, physicists had been making important discoveries about the structure of atoms. Ernest Rutherford had shown that atoms have nuclei (recall Figure 5-19), and Niels Bohr made the remarkable hypothesis that electrons circle atomic nuclei along discrete orbits (see Figure 5-22). These advances gave scientists the conceptual and mathematical tools needed to understand stellar spectra.

In the 1920s, the Harvard astronomer Cecilia Payne and the Indian physicist Meghnad Saha demonstrated that the OBAFGKM spectral sequence is actually a sequence in temperature. The hottest stars are O stars. Their absorption lines can occur only if these stars have surface temperatures above 25,000 K. M stars are the coolest stars. The spectral features of M stars are consistent with stellar surface temperatures of about 3000 K.

Why Surface Temperature Affects Stellar Spectra

To see why the appearance of a star's spectrum is profoundly affected by the star's surface temperature, consider the Balmer lines of hydrogen. Hydrogen is by far the most abundant element in the universe, accounting for about three-quarters of the mass of a typical star. Yet the Balmer lines do not necessarily show up in a star's spectrum. As we saw in Section 5-8, Balmer absorption lines are produced when an electron in the $n = 2$ orbit of hydrogen is lifted into a higher orbit by absorbing a photon with the right amount of energy (see Figure 5-24). If the star is much hotter than 10,000 K, the photons pouring out of the star's interior have such high energy that they easily knock electrons out of hydrogen atoms in the star's atmosphere. This process ionizes the gas. With its only electron torn away, a hydrogen atom cannot produce absorption lines. Hence, the Balmer lines will be relatively weak in the spectra of such hot stars, such as the hot O and B2 stars in Figure 17-11.

Conversely, if the star's atmosphere is much cooler than 10,000 K, almost all the hydrogen atoms are in the lowest ($n = 1$) energy state. Most of the photons passing through the star's atmosphere possess too little energy to boost electrons up from the $n = 1$ to the $n = 2$ orbit of the hydrogen atoms. Hence, very few of these atoms will have electrons in the $n = 2$ orbit, and only these few can absorb the photons characteristic of the Balmer lines. As a result, these lines are nearly absent from the spectrum of a cool star. (You can see this in the spectra of the cool M0 and M2 stars in Figure 17-11.)

For the Balmer lines to be prominent in a star's spectrum, the star must be hot enough to excite the electrons out of the ground state but not so hot that all the hydrogen atoms become ionized. A stellar surface temperature of about 9000 K produces the strongest hydrogen lines; this is the case for the stars of spectral types A0 and A5 in Figure 17-11.

LOOKING DEEPER 17.1

Every other type of atom or molecule also has a characteristic temperature range in which it produces prominent absorption lines in the observable part of the spectrum. Figure 17-12 shows the relative strengths of absorption lines produced by different chemicals. By measuring the details of these lines in a given star's spectrum, astronomers can accurately determine that star's surface temperature.

For example, the spectral lines of neutral (that is, un-ionized) helium are strong around 25,000 K. At this temperature, photons have enough energy to excite helium atoms without tearing away the electrons altogether. In stars hotter than about 30,000 K, helium atoms become singly ionized, that is, they lose one of their two electrons. The remaining electron produces a set of spectral lines that is recognizably different from those of neutral helium. Hence, when the spectral lines of singly ionized helium appear in a star's spectrum, we know that the star's surface temperature is greater than 30,000 K.

Astronomers use the term **metals** to refer to all elements other than hydrogen and helium. (This idiosyncratic use of the term "metal" is quite different from the definition used by chemists and other scientists. To a chemist, sodium and iron are metals but carbon and oxygen are not; to an astronomer, all of these substances are metals.) In this terminology, metals dominate the

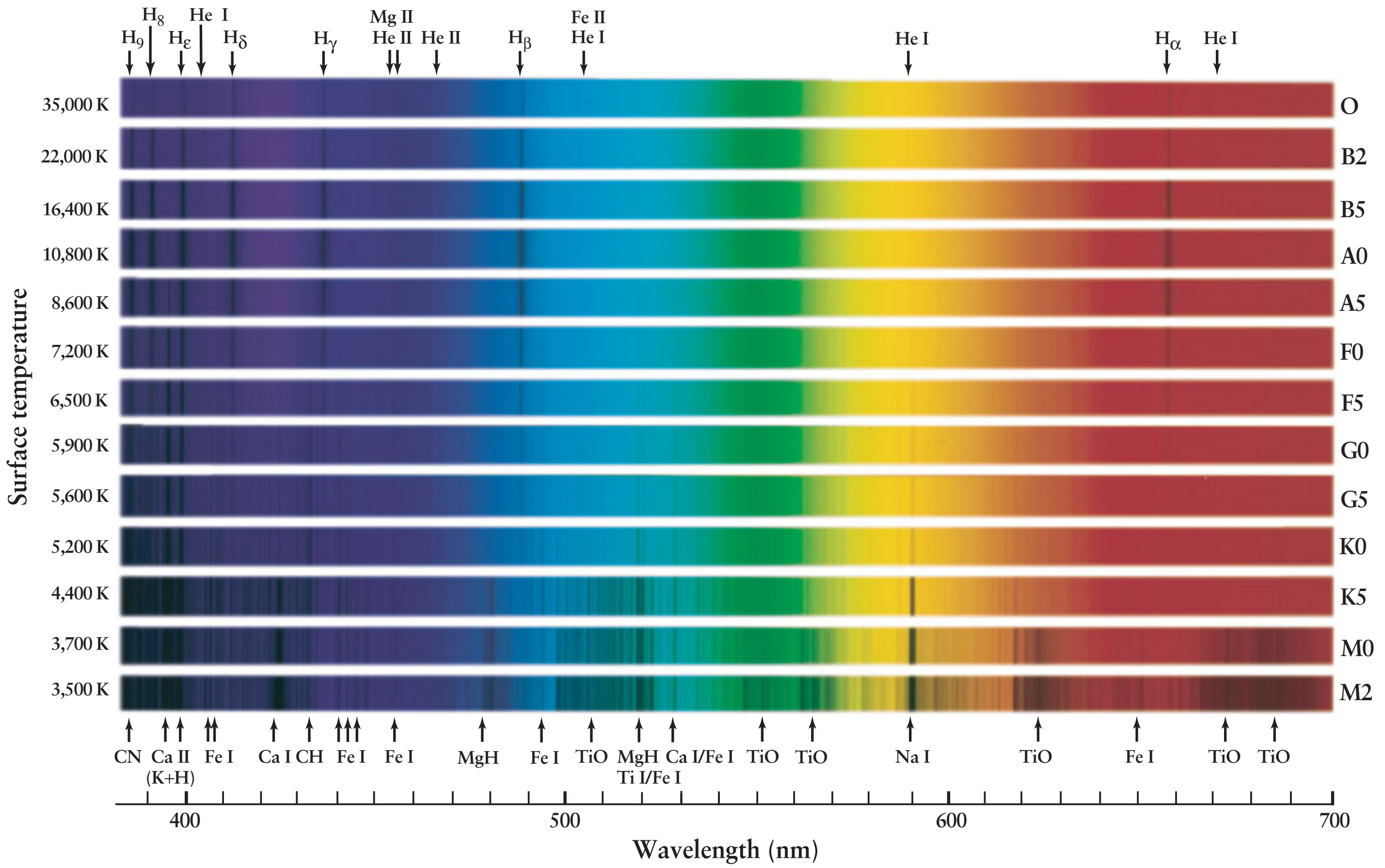
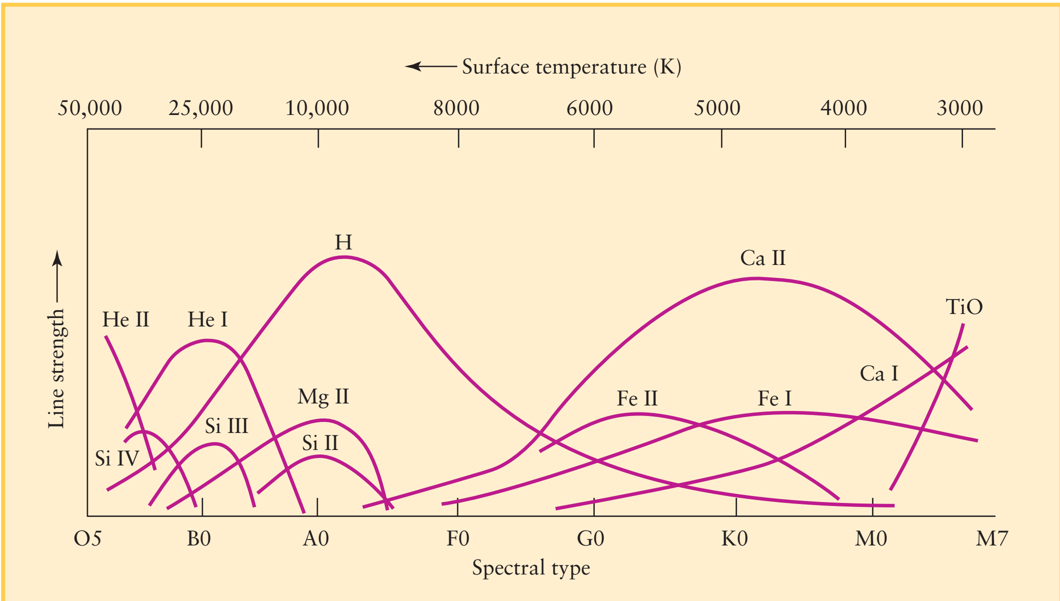


Figure 17-11 R I V U X G

Principal Types of Stellar Spectra Stars of different spectral classes and different surface temperatures have spectra dominated by different absorption lines. Notice how the Balmer lines of hydrogen (H_α, H_β, H_γ, and H_δ) are strongest for hot stars of spectral class A, while absorption lines due to calcium (Ca) are strongest in cooler K and M stars. The spectra of M stars also have broad, dark bands caused by molecules of titanium oxide (TiO), which can only exist at relatively low temperatures. A roman numeral after a chemical symbol shows whether the absorption line is caused by un-ionized atoms (roman numeral I) or by atoms that have lost one electron (roman numeral II). (R. Bell, University of Maryland, and M. Briley, University of Wisconsin at Oshkosh)

Figure 17-12

The Strengths of Absorption Lines Each curve in this graph peaks at the stellar surface temperature for which that chemical's absorption line is strongest. For example, hydrogen (H) absorption lines are strongest in A stars with surface temperatures near 10,000 K. Roman numeral I denotes neutral, un-ionized atoms; II, III, and IV denote atoms that are singly, doubly, or triply ionized (that is, have lost one, two, or three electrons).



spectra of stars cooler than 10,000 K. Ionized metals are prominent for surface temperatures between 6000 and 8000 K, while neutral metals are strongest between approximately 5500 and 4000 K.

Below 4000 K, certain atoms in a star's atmosphere combine to form molecules. (At higher temperatures atoms move so fast that when they collide, they bounce off each other rather than "sticking together" to form molecules.) As these molecules vibrate and rotate, they produce bands of spectral lines that dominate the star's spectrum. Most noticeable are the lines of titanium oxide (TiO), which are strongest for surface temperatures of about 3000 K.

Spectral Classes for Brown Dwarfs

Since 1995 astronomers have found a number of stars with surface temperatures even lower than those of spectral class M. Strictly speaking, these are not stars but **brown dwarfs,** which we introduced in Section 8-6. Brown dwarfs are too small to sustain thermonuclear fusion in their cores. Instead, these "substars" glow primarily from the heat released by Kelvin-Helmholtz contraction, which we described in Section 16-1. (They do undergo fusion reactions for a brief period during their evolution.) Brown dwarfs are so cold that they are best observed with infrared telescopes (see Figure 17-13). Such observations reveal that brown dwarf spectra have a rich variety of absorption lines due to molecules. Some of these molecules actually form into solid grains in a brown dwarf's atmosphere.

To describe brown dwarf spectra, astronomers have defined two new spectral classes, L and T. Thus, the modern spectral sequence of stars and brown dwarfs from hottest to coldest surface temperature is OBAFGKMLT. (Can you think of a new mnemonic that includes L and T?) For example, Figure 17-13 shows a star of spectral class K and a brown dwarf of spectral class T. Table 17-2 summarizes the relationship between the temperature and spectra of stars and brown dwarfs.

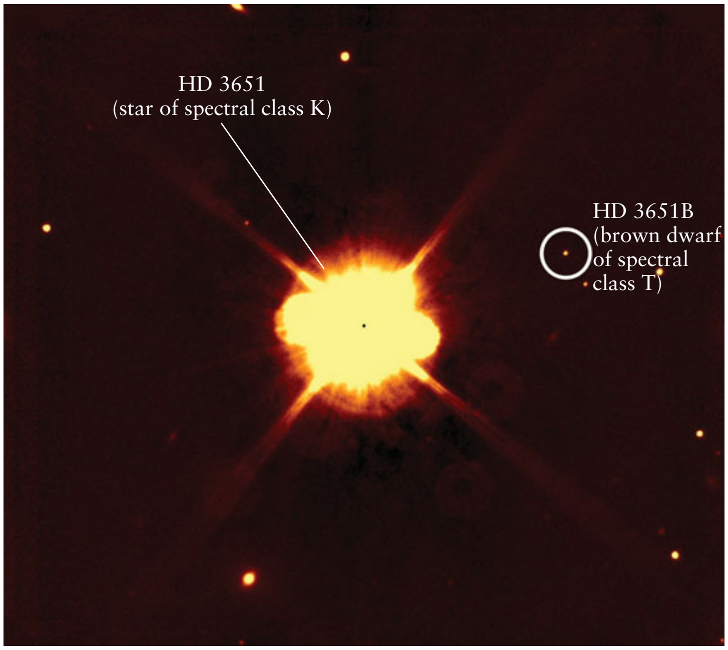


Figure 17-13 R I V U X G

An Infrared Image of Brown Dwarf HD 3651B The star HD 3561 is of spectral class K, with a surface temperature of about 5200 K. ("HD" refers to the *Henry Draper Catalogue*.) In 2006 it was discovered that HD 3651 is orbited by a brown dwarf named HD 3651B with a surface temperature between 800 and 900 K and a luminosity just 1/300,000 that of the Sun. The brown dwarf emits most of its light at infrared wavelengths, so an infrared telescope was used to record this image. The hotter and more luminous star HD 3651 is greatly overexposed in this image and appears much larger than its true size. HD 3651 and HD 3651B are both 11 pc (36 ly) from Earth in the constellation Pisces (the Fish); the other stars in this image are much farther away. (M. Mugrauer and R. Neuhäuser, U. of Jena; A. Seifahrt, ESO; and T. Mazeh, Tel Aviv U.)

Table 17-2 The Spectral Sequence

Spectral class	Color	Temperature (K)	Spectral lines	Examples
O	Blue-violet	30,000–50,000	Ionized atoms, especially helium	Naos (ζ Puppis), Mintaka (δ Orionis)
B	Blue-white	11,000–30,000	Neutral helium, some hydrogen	Spica (α Virginis), Rigel (β Orionis)
A	White	7500–11,000	Strong hydrogen, some ionized metals	Sirius (α Canis Majoris), Vega (α Lyrae)
F	Yellow-white	5900–7500	Hydrogen and ionized metals such as calcium and iron	Canopus (α Carinae), Procyon (α Canis Minoris)
G	Yellow	5200–5900	Both neutral and ionized metals, especially ionized calcium	Sun, Capella (α Aurigae)
K	Orange	3900–5200	Neutral metals	Arcturus (α Boötis), Aldebaran (α Tauri)
M	Red-orange	2500–3900	Strong titanium oxide and some neutral calcium	Antares (α Scorpii), Betelgeuse (α Orionis)
L	Red	1300–2500	Neutral potassium, rubidium, and cesium, and metal hydrides	Brown dwarf Teide 1
T	Red	below 1300	Strong neutral potassium and some water (H_2O)	Brown dwarfs Gliese 229B, HD 3651B

When the effects of temperature are accounted for, astronomers find that *all* stars have essentially the same chemical composition. We can state the results as a general rule:

By mass, almost all stars (including the Sun) and brown dwarfs are about three-quarters hydrogen, one-quarter helium, and 1% or less metals.

Our Sun is about 1% metals by mass, as are most of the stars you can see with the naked eye. But some stars have an even lower percentage of metals. We will see in Chapter 19 that these seemingly minor differences tell an important tale about the life stories of stars.

17-6 Stars come in a wide variety of sizes

With even the best telescopes, stars appear as nothing more than bright points of light. On a photograph or CCD image, brighter stars appear larger than dim ones (see Figures 17-3, 17-6*b*, and 17-13), but these apparent sizes give no indication of the star's actual size. To determine the size of a star, astronomers combine information about its luminosity (determined from its distance and apparent brightness) and its surface temperature (determined from its spectral type). In this way, they find that some stars are quite a bit smaller than the Sun, while others are a thousand times larger.

A star's radius can be calculated if we know its luminosity and surface temperature

Calculating the Radii of Stars

The key to finding a star's radius from its luminosity and surface temperature is the Stefan-Boltzmann law (see Section 5-4). This law says that the amount of energy radiated per second from a square meter of a blackbody's surface—that is, the energy flux (F)—is proportional to the fourth power of the temperature of that surface (T), as given by the equation $F = \sigma T^4$. This equation applies very well to stars, whose spectra are quite similar to that of a perfect blackbody. (Absorption lines, while important for determining the star's chemical composition and surface temperature, make only relatively small modifications to a star's blackbody spectrum.)

A star's luminosity is the amount of energy emitted per second from its entire surface. This equals the energy flux F multiplied by the total number of square meters on the star's surface (that is, the star's surface area). We expect that most stars are nearly spherical, like the Sun, so we can use the formula for the surface area of a sphere. This is $4\pi R^2$, where R is the star's

BOX 17-4 Tools of the Astronomer's Trade

Stellar Radii, Luminosities, and Surface Temperatures

Because stars emit light in almost exactly the same fashion as blackbodies, we can use the Stefan-Boltzmann law to relate a star's luminosity (L), surface temperature (T), and radius (R). The relevant equation is

$$L = 4\pi R^2 \sigma T^4$$

As written, this equation involves the Stefan-Boltzmann constant σ, which is equal to 5.67×10^{-8} W m^{-2} K^{-4}. In many calculations, it is more convenient to relate everything to the Sun, which is a typical star. Specifically, for the Sun we have $L_\odot = 4\pi R_\odot^2 \sigma T_\odot^4$, where $L_\odot$ is the Sun's luminosity, $R_\odot$ is the Sun's radius, and $T_\odot$ is the Sun's surface temperature (equal to 5800 K). Dividing the general equation for L by this specific equation for the Sun, we obtain

$$\frac{L}{L_\odot} = \left(\frac{R}{R_\odot}\right)^2 \left(\frac{T}{T_\odot}\right)^4$$

This is an easier formula to use because the constant σ has cancelled out. We can also rearrange terms to arrive at a useful equation for the radius (R) of a star:

Radius of a star related to its luminosity and surface temperature

$$\frac{R}{R_\odot} = \left(\frac{T_\odot}{T}\right)^2 \sqrt{\frac{L}{L_\odot}}$$

$R/R_\odot$ = ratio of the star's radius to the Sun's radius

$T_\odot/T$ = ratio of the Sun's surface temperature to the star's surface temperature

$L/L_\odot$ = ratio of the star's luminosity to the Sun's luminosity

EXAMPLE: The bright reddish star Betelgeuse in the constellation Orion (see Figure 2-2 or Figure 6-30*c*) is 60,000 times more luminous than the Sun and has a surface temperature of 3500 K. What is its radius?

Situation: We are given the star's luminosity $L = 60{,}000\ L_\odot$ and its surface temperature $T = 3500$ K. Our goal is to find the star's radius R.

radius (the distance from its center to its surface). Multiplying together the formulas for energy flux and surface area, we can write the star's luminosity as follows:

Relationship between a star's luminosity, radius, and surface temperature

$$L = 4\pi R^2 \sigma T^4$$

L = star's luminosity, in watts

R = star's radius, in meters

σ = Stefan-Boltzmann constant = 5.67×10^{-8} W m^{-2} K^{-4}

T = star's surface temperature, in kelvins

This equation says that a relatively cool star (low surface temperature T), for which the energy flux is quite low, can nonetheless be very luminous if it has a large enough radius R. Alternatively, a relatively hot star (large T) can have a very low luminosity if the star has only a little surface area (small R).

Box 17-4 describes how to use the above equation to calculate a star's radius if its luminosity and surface temperature are known. We can express the idea behind these calculations in terms of the following general rule:

We can determine the radius of a star from its luminosity and surface temperature. For a given luminosity, the greater the surface temperature, the smaller the radius must be. For a given surface temperature, the greater the luminosity, the larger the radius must be.

ANALOGY In a similar way, a roaring campfire can emit more light than a welder's torch. The campfire is at a lower temperature than the torch, but has a much larger surface area from which it emits light.

The Range of Stellar Radii

Using this general rule as shown in Box 17-4, astronomers find that stars come in a wide range of sizes. The smallest stars visible through ordinary telescopes, called *white dwarfs,* are about the same size as the Earth. Although their surface temperatures can be very high (25,000 K or more), white dwarfs have so little surface area that their luminosities are very low (less than 0.01 $L_\odot$). The largest stars, called *supergiants,* are a thousand times larger in radius than the Sun and 10^5 times larger than white dwarfs. If our own Sun were replaced by one of these supergiants, the Earth's orbit would lie completely inside the star!

Figure 17-14 summarizes how astronomers determine the distance from Earth, luminosity, surface temperature, chemical

Tools: We use the above equation to find the ratio of the star's radius to the radius of the Sun, $R/R_\odot$. Note that we also know the Sun's surface temperature, $T_\odot$ = 5800 K.

Answer: Substituting these data into the above equation, we get

$$\frac{R}{R_\odot} = \left(\frac{5800\ \text{K}}{3500\ \text{K}}\right)^2 \sqrt{6 \times 10^4} = 670$$

Review: Our result tells us that Betelgeuse's radius is 670 times larger than that of the Sun. The Sun's radius is 6.96×10^5 km, so we can also express the radius of Betelgeuse as $(670)(6.96 \times 10^5\ \text{km}) = 4.7 \times 10^8$ km. This is more than 3 AU. If Betelgeuse were located at the center of our solar system, it would extend beyond the orbit of Mars!

EXAMPLE: Sirius, the brightest star in the sky, is actually two stars orbiting each other (a binary star). The less luminous star, Sirius B, is a white dwarf that is too dim to see with the naked eye. Its luminosity is 0.0025 $L_\odot$ and its surface temperature is 10,000 K. How large is Sirius B compared to the Earth?

Situation: Again we are asked to find a star's radius from its luminosity and surface temperature.

Tools: We use the same equation as in the preceding example.

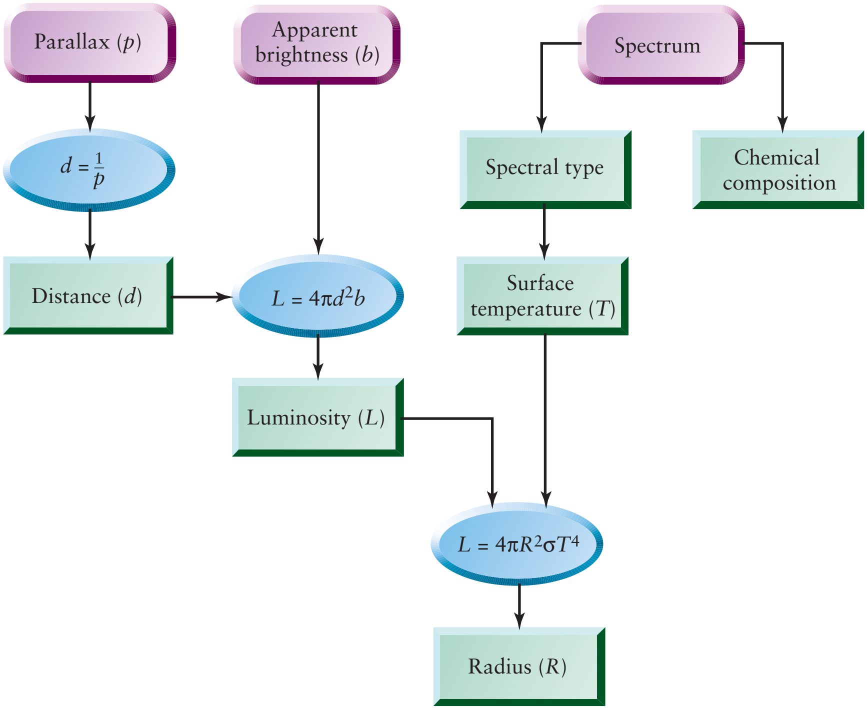
Answer: The ratio of the radius of Sirius B to the Sun's radius is

$$\frac{R}{R_\odot} = \left(\frac{5800\ \text{K}}{10{,}000\ \text{K}}\right)^2 \sqrt{0.0025} = 0.017$$

Since the Sun's radius is $R_\odot = 6.96 \times 10^5$ km, the radius of Sirius B is $(0.017)(6.96 \times 10^5\ \text{km})$ = 12,000 km. From Table 7-1, the Earth's radius (half its diameter) is 6378 km. Hence, this star is only about twice the radius of the Earth.

Review: Sirius B's radius would be large for a terrestrial planet, but it is minuscule for a star. The name *dwarf* is well deserved!

The radii of some stars have been measured with other techniques (see Section 17-11). These other methods yield values consistent with those calculated by the methods we have just described.



INTERACTIVE EXERCISE 17.1

Figure 17-14

Finding Key Properties of a Nearby Star This flowchart shows how astronomers determine the properties of a relatively nearby star (one close enough that its parallax can be measured). The rounded purple boxes show the measurements that must be made of the star, the blue ovals show the key equations that are used (from Sections 17-2, 17-5, and 17-6), and the green rectangles show the inferred properties of the stars. A different procedure is followed for more distant stars (see Section 17-8, especially Figure 17-17).

composition, and radius of a star close enough to us so that its parallax can be measured. Remarkably, all of these properties can be deduced from just a few measured quantities: the star's parallax angle, apparent brightness, and spectrum.

17-7 Hertzsprung-Russell (H-R) diagrams reveal the different kinds of stars

Astronomers have collected a wealth of data about the stars, but merely having tables of numerical data is not enough. Like all scientists, astronomers want to analyze their data to look for trends and underlying principles. One of the best ways to look for trends in any set of data, whether it comes from astronomy, finance, medicine, or meteorology, is to create a graph showing how one quantity depends on another. For example, investors consult graphs of stock market values versus dates, and weather forecasters make graphs of temperature versus altitude to determine whether thunderstorms will form. Astronomers have found that a particular graph of stellar properties shows that stars fall naturally into just a few categories. This graph, one of the most important in all astronomy, will in later chapters help us understand how stars form, evolve, and eventually die.

H-R Diagrams

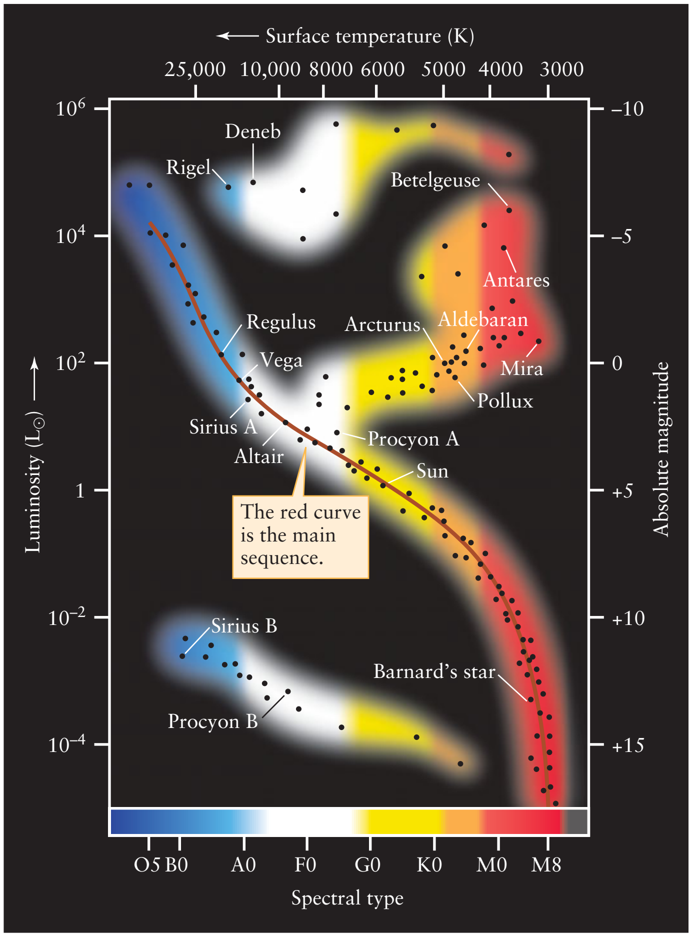
Which properties of stars should we include in a graph? Most stars have about the same chemical composition, but two properties of stars—their luminosities and surface temperatures—differ substantially from one star to another. Stars also come in a wide range of radii, but a star's radius is a secondary property that can be found from the luminosity and surface temperature (as we saw in Section 17-6 and Box 17-4). We also relegate the positions and space velocities of stars to secondary importance. (In a similar way, a physician is more interested in your weight and blood pressure than in where you live or how fast you drive.) We can then ask the following question: What do we learn when we graph the luminosities of stars versus their surface temperatures?

A graph of luminosity versus surface temperature reveals that stars fall into a few basic categories

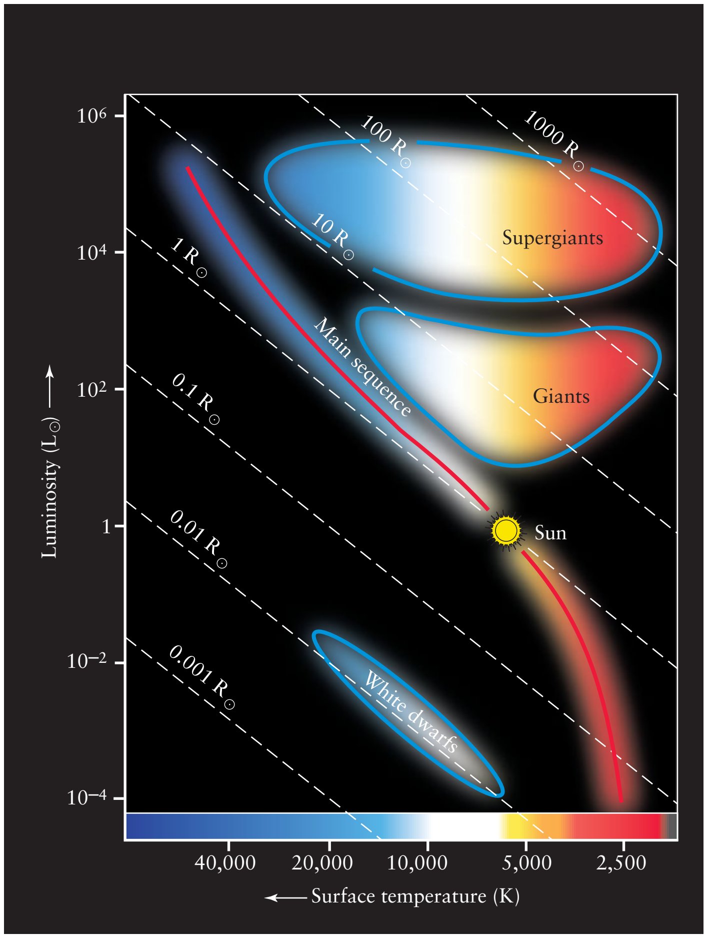
The first answer to this question was given in 1911 by the Danish astronomer Ejnar Hertzsprung. He pointed out that a regular pattern appears when the absolute magnitudes of stars (which measure their luminosities) are plotted against their colors (which measure their surface temperatures). Two years later, the American astronomer Henry Norris Russell independently discovered a similar regularity in a graph using spectral types (another measure of surface temperature) instead of colors. In recognition of their originators, graphs of this kind are today known as **Hertzsprung-Russell diagrams,** or **H-R diagrams** (Figure 17-15).

Figure 17-15*a* is a typical Hertzsprung-Russell diagram. Each dot represents a star whose spectral type and luminosity have been determined. The most luminous stars are near the top of the diagram, the least luminous stars near the bottom. Hot stars of spectral classes O and B are toward the left side of the graph and cool stars of spectral class M are toward the right.

CAUTION! You are probably accustomed to graphs in which the numbers on the horizontal axis increase as you move to the right. (For example, the business section of a newspaper includes a graph of stock market values versus dates, with later



(a) A Hertzsprung-Russell (H-R) diagram



(b) The sizes of stars on an H-R diagram

INTERACTIVE EXERCISE 17.2

Figure 17-15

Hertzsprung-Russell (H-R) Diagrams On an H-R diagram, the luminosities (or absolute magnitudes) of stars are plotted against their spectral types (or surface temperatures). **(a)** The data points are grouped in just a few regions on the graph, showing that luminosity and spectral type are correlated. Most stars lie along the red curve called the main sequence. Giants like Arcturus as well as supergiants like Rigel and Betelgeuse are above the main sequence, and white dwarfs like Sirius B are below it. **(b)** The blue curves on this H-R diagram enclose the regions of the diagram in which different types of stars are found. The dashed diagonal lines indicate different stellar radii. For a given stellar radius, as the surface temperature increases (that is, moving from right to left in the diagram), the star glows more intensely and the luminosity increases (that is, moving upward in the diagram). Note that the Sun is intermediate in luminosity, surface temperature, and radius.

dates to the right of earlier ones.) But on an H-R diagram the temperature scale on the horizontal axis increases toward the *left*. This practice stems from the original diagrams of Hertzsprung and Russell, who placed hot O stars on the left and cool M stars on the right. This arrangement is a tradition that no one has seriously tried to change.

Star Varieties: Main-Sequence Stars, Giants, Supergiants, White Dwarfs, and Brown Dwarfs

The most striking feature of the H-R diagram is that the data points are not scattered randomly over the graph but are grouped in a few distinct regions. The luminosities and surface temperatures of stars do *not* have random values; instead, these two quantities are related!

The band stretching diagonally across the H-R diagram includes about 90% of the stars in the night sky. This band, called the **main sequence,** extends from the hot, luminous, blue stars in the upper left corner of the diagram to the cool, dim, red stars in the lower right corner. A star whose properties place it in this region of an H-R diagram is called a **main-sequence star.** The Sun (spectral type G2, luminosity 1 $L_\odot$, absolute magnitude +4.8) is such a star. We will find that all main-sequence stars are like the Sun in that *hydrogen fusion*—thermonuclear reactions that convert hydrogen into helium (see Section 16-1)—is taking place in their cores.

The upper right side of the H-R diagram shows a second major grouping of data points. Stars represented by these points are both luminous and cool. From the Stefan-Boltzmann law, we know that a cool star radiates much less light per unit of surface area than a hot star. In order for these stars to be as luminous as they are, they must be huge (see Section 17-6), and so they are called **giants.** These stars are around 10 to 100 times larger than the Sun. You can see this in Figure 17-15*b*, which is an H-R diagram to which dashed lines have been added to represent stellar radii. Most giant stars are around 100 to 1000 times more luminous than the Sun and have surface temperatures of about 3000 to 6000 K. Cooler members of this class of stars (those with surface temperatures from about 3000 to 4000 K) are often called **red giants** because they appear reddish. The bright yellowish stars in the image that opens this chapter are red giants. (Notice how they outshine the blue stars in this image, which are at the same distance from Earth as the red giants but are smaller and less luminous.) A number of red giants can easily be seen with the naked eye, including Aldebaran in the constellation Taurus and Arcturus in Boötes.

A few rare stars are considerably bigger and brighter than typical red giants, with radii up to 1000 $R_\odot$. Appropriately enough, these superluminous stars are called **supergiants.** Betelgeuse in Orion (see Box 17-4) and Antares in Scorpius are two supergiants you can find in the nighttime sky. Together, giants and supergiants make up about 1% of the stars in the sky.

Both giants and supergiants have thermonuclear reactions occurring in their interiors, but the character of those reactions and where in the star they occur can be quite different than for a main-sequence star like the Sun. We will study these stars in more detail in Chapters 21 and 22.

The remaining 9% of stars form a distinct grouping of data points toward the lower left corner of the Hertzsprung-Russell diagram. Although these stars are hot, their luminosities are quite low; hence, they must be small. They are appropriately called **white dwarfs.** These stars, which are so dim that they can be seen only with a telescope, are approximately the same size as the Earth. As we will learn in Chapter 22, no thermonuclear reactions take place within white dwarf stars. Rather, like embers left from a fire, they are the still-glowing remnants of what were once giant stars.

By contrast, *brown* dwarfs (which lie at the extreme lower right of the main sequence, off the bottom and right-hand edge of Figure 17-15*a* or Figure 17-15*b*) are objects that will never become stars. They are comparable in radius to the planet Jupiter (that is, intermediate in size between the Earth and the Sun; see Figure 7-2). The study of brown dwarfs is still in its infancy, but it appears that there may be twice as many brown dwarfs as there are "real" stars.

ANALOGY You can think of white dwarfs as "has-been" stars whose days of glory have passed. In this analogy, a brown dwarf is a "never-will-be."

The existence of fundamentally different types of stars is the first important lesson to come from the H-R diagram. In later chapters we will find that these different types represent various stages in the lives of stars. We will use the H-R diagram as an essential tool for understanding how stars evolve.

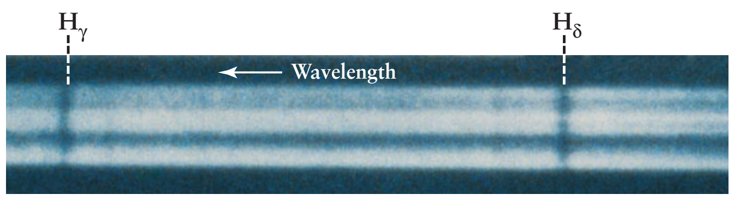
17-8 Details of a star's spectrum reveal whether it is a giant, a white dwarf, or a main-sequence star

A star's surface temperature largely determines which lines are prominent in its spectrum. Therefore, classifying stars by spectral type is essentially the same as categorizing them by surface temperature. But as the H-R diagram in Figure 17-15*b* shows, stars of the same surface temperature can have very different luminosities. As an example, a star with surface temperature 5800 K could be a white dwarf, a main-sequence star, a giant, or a supergiant, depending on its luminosity. By examining the details of a star's spectrum, however, astronomers can determine to which of these categories a star belongs. This gives astronomers a tool to determine the distances to stars millions of parsecs away, far beyond the maximum distance that can be measured using stellar parallax.

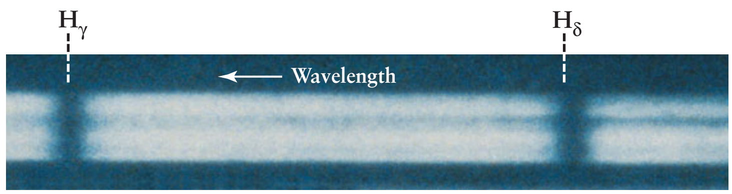
Determining a Star's Size from Its Spectrum

Figure 17-16 compares the spectra of two stars of the same spectral type but different luminosity (and hence different size): a B8 supergiant and a B8 main-sequence star. Note that the Balmer lines of hydrogen are narrow in the spectrum of the very large, very luminous supergiant but quite broad in the spectrum of the small, less luminous main-sequence star. In general, for stars of spectral types B through F, the larger and more luminous the star, the narrower its hydrogen lines.

The smaller a star and the denser its atmosphere, the broader the absorption lines in its spectrum



(a) A supergiant star has a low-density, low-pressure atmosphere: its spectrum has narrow absorption lines



(b) A main-sequence star has a denser, higher-pressure atmosphere: its spectrum has broad absorption lines

Figure 17-16 R I **V** U X G

How a Star's Size Affects Its Spectrum These are the spectra of two stars of the same spectral type (B8) and surface temperature (13,400 K) but different radii and luminosities: **(a)** the B8 supergiant Rigel (luminosity 58,000 $L_\odot$) in Orion, and **(b)** the B8 main-sequence star Algol (luminosity 100 $L_\odot$) in Perseus. (From W. W. Morgan, P. C. Keenan, and E. Kellman, *An Atlas of Stellar Spectra*)

Fundamentally, these differences between stars of different luminosity are due to differences between the stars' atmospheres, where absorption lines are produced. Hydrogen lines in particular are affected by the density and pressure of the gas in a star's atmosphere. The higher the density and pressure, the more frequently hydrogen atoms collide and interact with other atoms and ions in the atmosphere. These collisions shift the energy levels in the hydrogen atoms and thus broaden the hydrogen spectral lines.

In the atmosphere of a luminous giant star, the density and pressure are quite low because the star's mass is spread over a huge volume. Atoms and ions in the atmosphere are relatively far apart; hence, collisions between them are sufficiently infrequent that hydrogen atoms can produce narrow Balmer lines. A main-sequence star, however, is much more compact than a giant or supergiant. In the denser atmosphere of a main-sequence star, frequent interatomic collisions perturb the energy levels in the hydrogen atoms, thereby producing broader Balmer lines.

Luminosity Classes

In the 1930s, W. W. Morgan and P. C. Keenan of the Yerkes Observatory of the University of Chicago developed a system of **luminosity classes** based upon the subtle differences in spectral lines. When these luminosity classes are plotted on an H-R diagram (Figure 17-17), they provide a useful subdivision of the star types in the upper right of the diagram. Luminosity classes Ia and Ib are composed of supergiants; luminosity class V includes all the main-sequence stars. The intermediate classes distinguish giant stars of various luminosities. Note that for stars of a given surface temperature (that is, a given spectral type), the *higher* the number of the luminosity class, the *lower* the star's luminosity.

As we will see in Chapters 19 and 20, different luminosity classes represent different stages in the evolution of a star. White dwarfs are not given a luminosity class of their own; as we mentioned in Section 17-7, they represent a final stage in stellar evolution in which no thermonuclear reactions take place.

Astronomers commonly use a shorthand description that combines a star's spectral type and its luminosity class. For example, the Sun is said to be a G2 V star. The spectral type indicates the star's surface temperature, and the luminosity class indicates its luminosity. Thus, an astronomer knows immediately that any G2 V star is a main-sequence star with a luminosity of about 1 $L_\odot$ and a surface temperature of about 5800 K. Similarly, a description of Aldebaran as a K5 III star tells an astronomer that it is a red giant with a luminosity of around 370 $L_\odot$ and a surface temperature of about 4000 K.

Spectroscopic Parallax

A star's spectral type and luminosity class, combined with the information on the H-R diagram, enable astronomers to estimate the star's distance from the Earth. As an example, consider the star Pleione in the constellation Taurus. Its spectrum reveals Pleione to be a B8 V star (a hot, blue, main-sequence star, like the one in Figure 17-16*b*). Using Figure 17-17, we can read off that such a star's luminosity is 190 $L_\odot$. Given the star's luminosity and its apparent brightness—in the case of Pleione, 3.9×10^{-13} of the apparent brightness of the Sun—we can use the inverse-square law to determine its distance from the Earth. The mathematical details are worked out in Box 17-2.

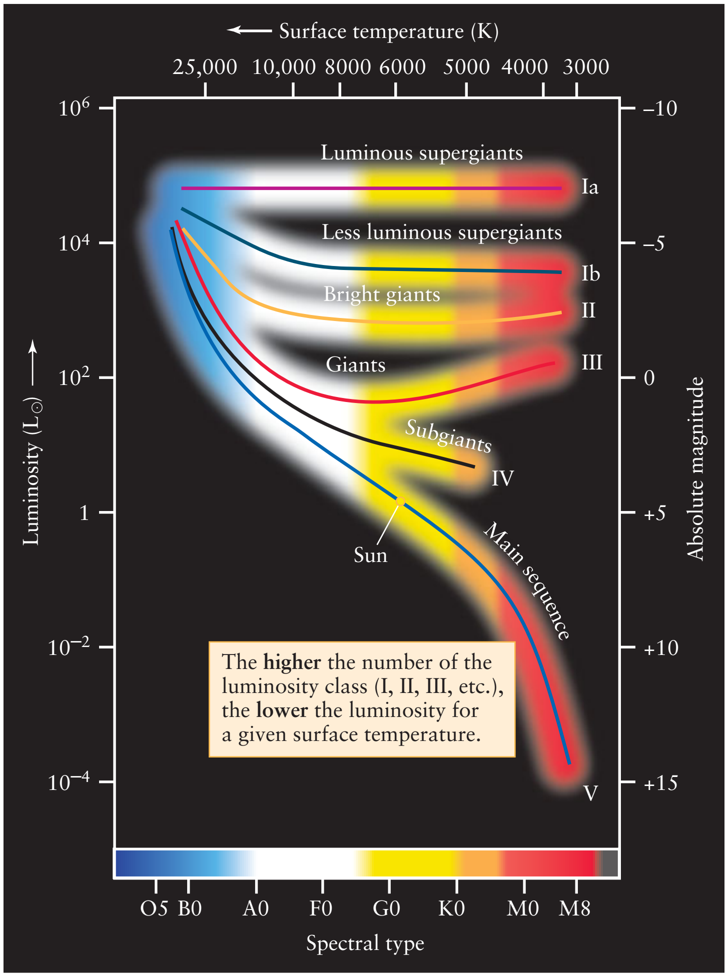


Figure 17-17

Luminosity Classes The H-R diagram is divided into regions corresponding to stars of different luminosity classes. (White dwarfs do not have their own luminosity class.) A star's spectrum reveals both its spectral type and its luminosity class; from these, the star's luminosity can be determined.

This method for determining distance, in which the luminosity of a star is found using spectroscopy, is called **spectroscopic parallax.** Figure 17-18 summarizes the method of spectroscopic parallax.

CAUTION! The name "spectroscopic parallax" is a bit misleading, because no parallax angle is involved! The idea is that measuring the star's spectrum takes the place of measuring its parallax as a way to find the star's distance. A better name for this method, although not the one used by astronomers, would be "spectroscopic distance determination."

Spectroscopic parallax is an incredibly powerful technique. No matter how remote a star is, this technique allows astronomers

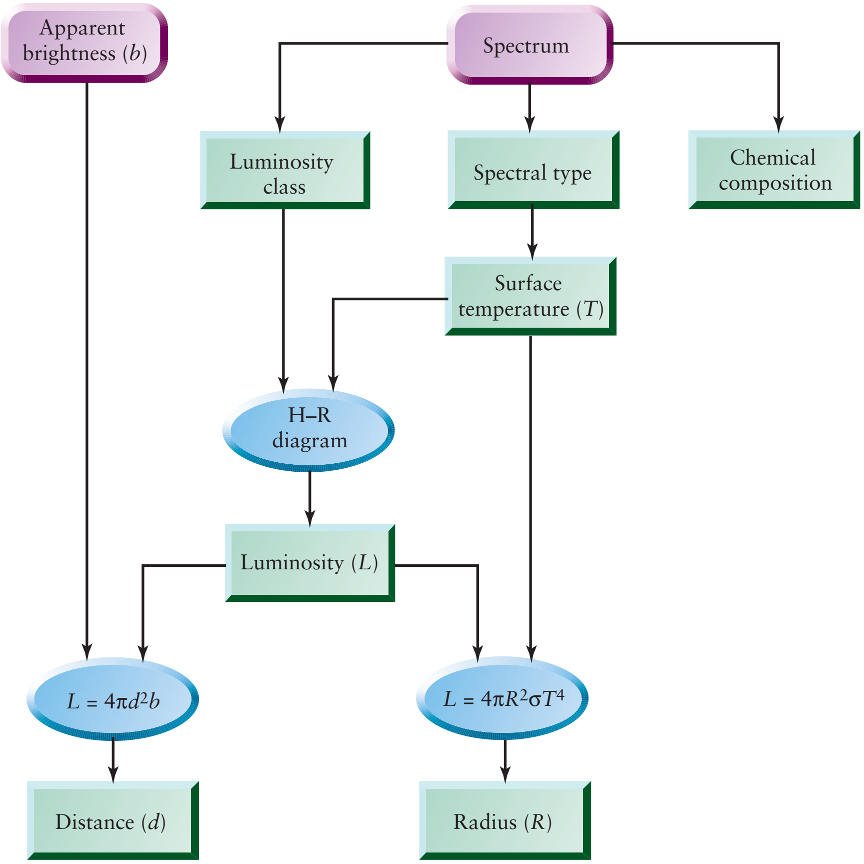


Figure 17-18

The Method of Spectroscopic Parallax If a star is too far away, its parallax angle is too small to allow a direct determination of its distance. This flowchart shows how astronomers deduce the properties of such a distant star. Note that the H-R diagram plays a central role in determining the star's luminosity from its spectral type and luminosity class. Just as for nearby stars (see Figure 17-14), the star's chemical composition is determined from its spectrum, and the star's radius is calculated from the luminosity and surface temperature.

to determine its distance, provided only that its spectrum and apparent brightness can be measured. Box 17-2 gives an example of how spectroscopic parallax has been used to find the distance to stars in other galaxies tens of millions of parsecs away. By contrast, we saw in Section 17-1 that "real" stellar parallaxes can be measured only for stars within a few hundred parsecs.

Unfortunately, spectroscopic parallax has its limitations; distances to individual stars determined using this method are only accurate to at best 10%. The reason is that the luminosity classes shown in Figure 17-17 are not thin lines on the H-R diagram but are moderately broad bands. Hence, even if a star's spectral type and luminosity class are known, there is still some uncertainty in the luminosity that we read off an H-R diagram. Nonetheless, spectroscopic parallax is often the only means that an astronomer has to estimate the distance to remote stars.

What has been left out of this discussion is *why* different stars have different spectral types and luminosities. One key factor, as we shall see, turns out to be the mass of the star.

17-9 Observing binary star systems reveals the masses of stars

We now know something about the sizes, temperatures, and luminosities of stars. To complete our picture of the physical properties of stars, we need to know their masses. In this section, we will see that stars come in a wide range of masses. We will also discover an important relationship between the mass and luminosity of main-sequence stars. This relationship is crucial to understanding why some main-sequence stars are hot and luminous, while others are cool and dim. It will also help us understand what happens to a star as it ages and evolves.

For main-sequence stars, there is a direct correlation between mass and luminosity

Determining the masses of stars is not trivial, however. The problem is that there is no practical, direct way to measure the mass of an isolated star. Fortunately for astronomers, about half of the visible stars in the night sky are not isolated individuals. Instead, they are *multiple-star systems,* in which two or more stars orbit each other. By carefully observing the motions of these stars, astronomers can glean important information about their masses.

Binary Stars

A pair of stars located at nearly the same position in the night sky is called a **double star.** The Anglo-German astronomer William Herschel made the first organized search for such pairs. Between 1782 and 1821, he published three catalogs listing more than 800 double stars. Late in the nineteenth century, his son, John Herschel, discovered 10,000 more doubles. Some of these double stars are **optical double stars,** which are two stars that lie along nearly the same line of sight but are actually at very different distances from us. But many double stars are true **binary stars,** or

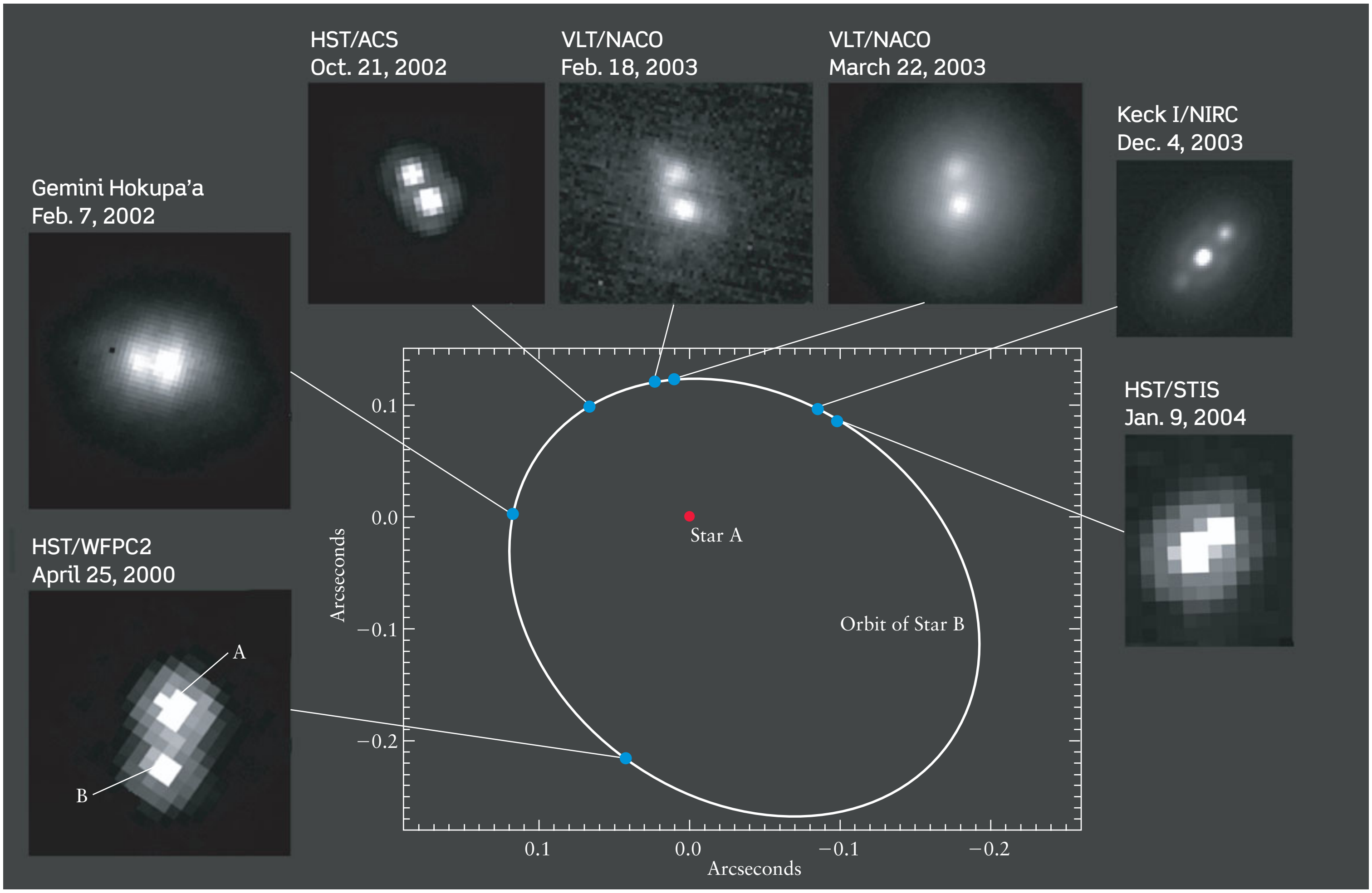


Figure 17-19 R I V U X G

A Binary Star System As seen from Earth, the two stars that make up the binary system called 2MASSW J0746425+2000321 are separated by less than 1/3 arcsecond. The images surrounding the center diagram show the relative positions of the two stars over a four-year period. These images were made by the Hubble Space Telescope (HST), the European Southern Observatory's Very Large Telescope (VLT), and Keck I and Gemini North in Hawaii (see Figure 6-16). For simplicity, the diagram shows one star as remaining stationary; in reality, both stars move around their common center of mass. (H. Bouy et al., MPE and ESO)

binaries—pairs of stars that actually orbit each other. Figure 17-19 shows an example of this orbital motion.

When astronomers can actually see the two stars orbiting each other, a binary is called a **visual binary.** By observing the binary over an extended period, astronomers can plot the orbit that one star appears to describe around the other, as shown in the center diagram in Figure 17-19.

In fact, *both* stars in a binary system are in motion: They orbit each other because of their mutual gravitational attraction, and their orbital motions obey Kepler's third law as formulated by Isaac Newton (see Section 4-7 and Box 4-4). This law can be written as follows:

Kepler's third law for binary star systems

$$M_1 + M_2 = \frac{a^3}{P^2}$$

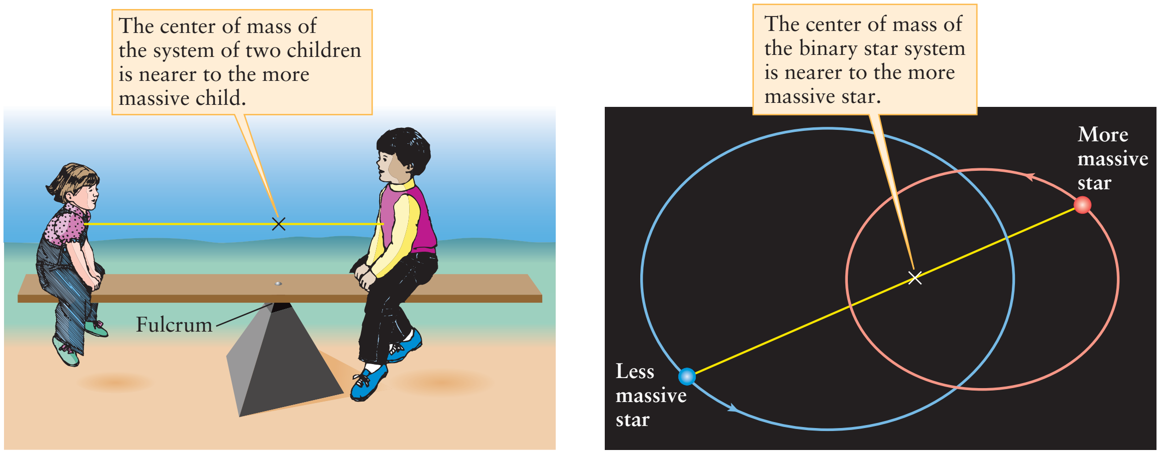
M_1, M_2 = masses of two stars in binary system, in solar masses

a = semimajor axis of one star's orbit around the other, in AU

P = orbital period, in years

Here a is the semimajor axis of the elliptical orbit that one star appears to describe around the other, plotted as in the center diagram in Figure 17-19. As this equation indicates, if we can measure this semimajor axis (a) and the orbital period (P), we can learn something about the masses of the two stars.

In principle, the orbital period of a visual binary is easy to determine. All you have to do is see how long it takes for the two stars to revolve once about each other. The two stars shown in Figure 17-19 are relatively close, about 2.5 AU on average, and their orbital period is only 10 years. Many binary systems have



(a) A "binary system" of two children

(b) A binary star system

ANIMATION 17.3

Figure 17-20

Center of Mass in a Binary Star System **(a)** A seesaw balances if the fulcrum is at the center of mass of the two children. **(b)** The members of a binary star system orbit around the center of mass of the two stars. Although their elliptical orbits cross each other, the two stars are always on opposite sides of the center of mass and thus never collide.

much larger separations, however, and the period may be so long that more than one astronomer's lifetime is needed to complete the observations.

Determining the semimajor axis of an orbit can also be a challenge. The *angular* separation between the stars can be determined by observation. To convert this angle into a physical distance between the stars, we need to know the distance between the binary and the Earth. This can be found from parallax measurements or by using spectroscopic parallax. The astronomer must also take into account how the orbit is tilted to our line of sight.

Once both P and a have been determined, Kepler's third law can be used to calculate $M_1 + M_2$, the sum of the masses of the two stars in the binary system. But this analysis tells us nothing about the *individual* masses of the two stars. To obtain these, more information about the motions of the two stars is needed.

Each of the two stars in a binary system actually moves in an elliptical orbit about the **center of mass** of the system. Imagine two children sitting on opposite ends of a seesaw (Figure 17-20*a*). For the seesaw to balance properly, they must position themselves so that their center of mass—an imaginary point that lies along a line connecting their two bodies—is at the fulcrum, or pivot point of the seesaw. If the two children have the same mass, the center of mass lies midway between them, and they should sit equal distances from the fulcrum. If their masses are different, the center of mass is closer to the heavier child.

Just as the seesaw naturally balances at its center of mass, the two stars that make up a binary system naturally orbit around their center of mass (Figure 17-20*b*). The center of mass always lies along the line connecting the two stars and is closer to the more massive star.

The center of mass of a visual binary is located by plotting the separate orbits of the two stars, as in Figure 17-20*b*, using the background stars as reference points. The center of mass lies at the common focus of the two elliptical orbits. Comparing the relative sizes of the two orbits around the center of mass yields the ratio of the two stars' masses, M_1/M_2. The sum $M_1 + M_2$ is already known from Kepler's third law, so the individual masses of the two stars can then be determined.

Main-Sequence Masses and the Mass-Luminosity Relation

Years of careful, patient observations of binaries have slowly yielded the masses of many stars. As the data accumulated, an important trend began to emerge: For main-sequence stars, there is a direct correlation between mass and luminosity. The more massive a main-sequence star, the more luminous it is. Figure 17-21 depicts this **mass-luminosity relation** as a graph. The range of stellar masses extends from less than 0.1 of a solar mass to more than 50 solar masses. The Sun's mass lies between these extremes.

The *Cosmic Connections* figure on the next page depicts the mass-luminosity relation for main-sequence stars on an H-R diagram. This figure shows the main sequence on an H-R diagram is a progression in mass as well as in luminosity and surface temperature. The hot, bright, bluish stars in the upper left corner of an H-R diagram are the most massive main-sequence stars. Likewise, the dim, cool, reddish stars in the lower right corner of an H-R diagram are the least massive. Main-sequence stars of intermediate temperature and luminosity also have intermediate masses.

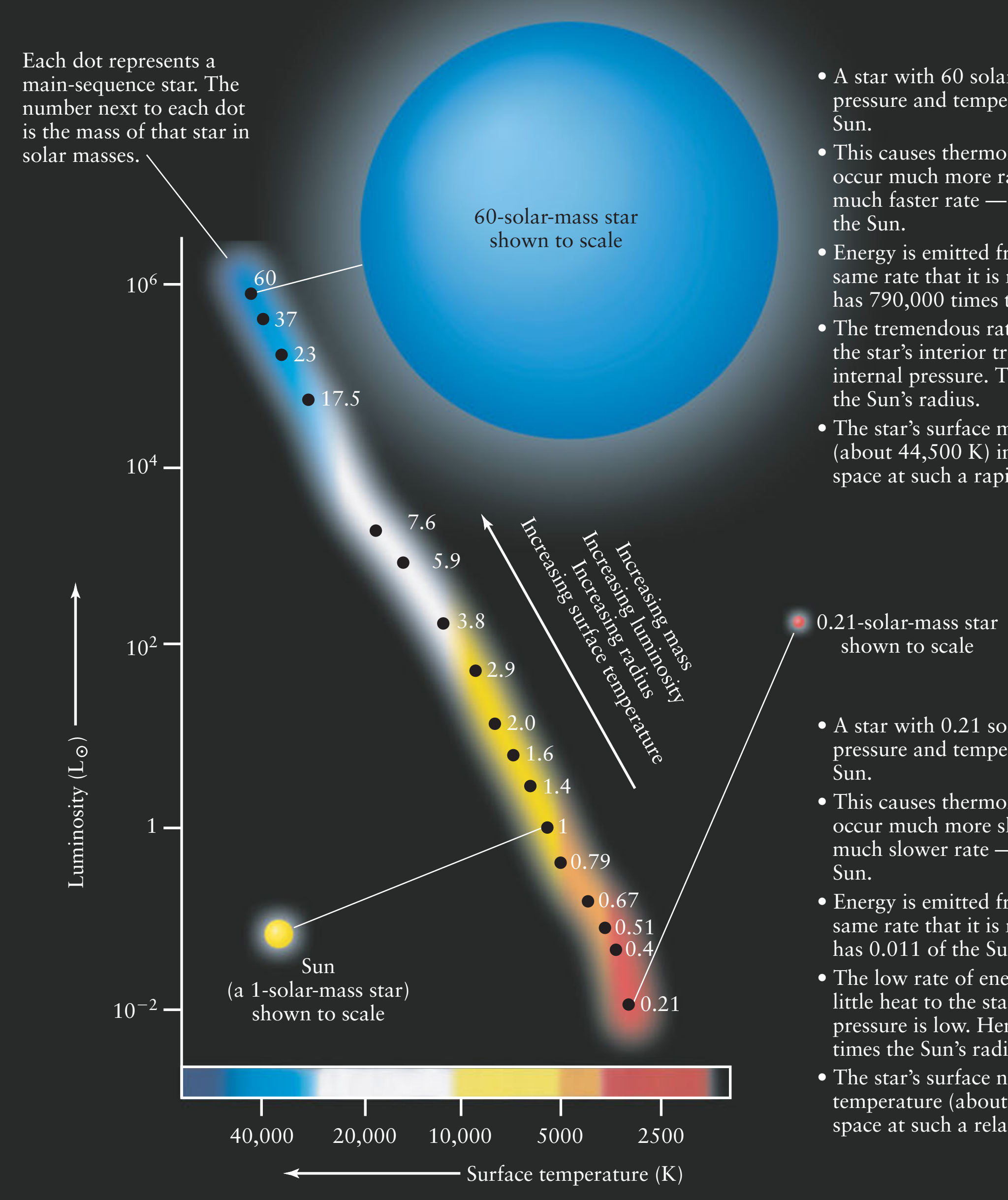
The mass of a main-sequence star also helps determine its radius. Referring back to Figure 17-15*b*, we see that if we go along the main sequence from low luminosity to high luminosity, the radius of the star increases. Thus, we have the following general rule for main-sequence stars:

The greater the mass of a main-sequence star, the greater its luminosity, its surface temperature, and its radius.

COSMIC CONNECTIONS

The Main Sequence and Masses

The main sequence is an arrangement of stars according to their mass. The most massive main-sequence stars have the greatest luminosity, greatest radius, and greatest surface temperature. This is a consequence of the behavior of thermonuclear reactions at the core of a main-sequence star.



- A star with 60 solar masses has much higher pressure and temperature at its core than does the Sun.
- This causes thermonuclear reactions in the core to occur much more rapidly and release energy at a much faster rate — 790,000 times faster than in the Sun.
- Energy is emitted from the star's surface at the same rate that it is released in the core, so the star has 790,000 times the Sun's luminosity.
- The tremendous rate of energy release also heats the star's interior tremendously, increasing the star's internal pressure. This inflates the star to 15 times the Sun's radius.
- The star's surface must be at a high temperature (about 44,500 K) in order for it to emit energy into space at such a rapid rate.

- A star with 0.21 solar mass has much lower pressure and temperature at its core than does the Sun.
- This causes thermonuclear reactions in the core to occur much more slowly and release energy at a much slower rate — 0.011 times as fast as in the Sun.
- Energy is emitted from the star's surface at the same rate that it is released in the core, so the star has 0.011 of the Sun's luminosity.
- The low rate of energy release supplies relatively little heat to the star's interior, so the star's internal pressure is low. Hence the star's radius is only 0.33 times the Sun's radius.
- The star's surface need be at only a low temperature (about 3200 K) to emit energy into space at such a relatively slow rate.

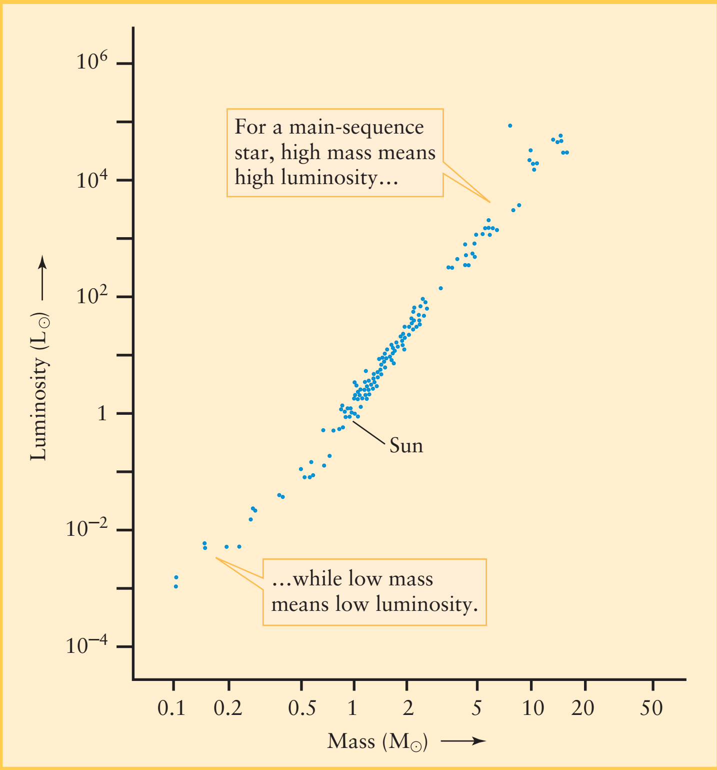


Figure 17-21

The Mass-Luminosity Relation For main-sequence stars, there is a direct correlation between mass and luminosity—the more massive a star, the more luminous it is. A main-sequence star of mass 10 $M_\odot$ (that is, 10 times the Sun's mass) has roughly 3000 times the Sun's luminosity (3000 $L_\odot$); one with 0.1 $M_\odot$ has a luminosity of only about 0.001 $L_\odot$.

Mass and Main-Sequence Stars

Why is mass the controlling factor in determining the properties of a main-sequence star? The answer is that all main-sequence stars are objects like the Sun, with essentially the same chemical composition as the Sun but with different masses. Like the Sun, all main-sequence stars shine because thermonuclear reactions at their cores convert hydrogen to helium and release energy. The greater the total mass of the star, the greater the pressure and temperature at the core, the more rapidly thermonuclear reactions take place in the core, and the greater the energy output—that is, the luminosity—of the star. In other words, the greater the mass of a main-sequence star, the greater its luminosity. This statement is just the mass-luminosity relation, which we can now recognize as a natural consequence of the nature of main-sequence stars.

Like the Sun, main-sequence stars are in a state of both hydrostatic equilibrium and thermal equilibrium (see Section 16-2). Calculations using models of a main-sequence star's interior (like the solar models we discussed in Section 16-2) show that to maintain equilibrium, a more massive star must have a larger radius and a higher surface temperature. This is just what we see when we plot the curve of the main sequence on an H-R diagram (see Figure 17-15*b*). As you move up the main sequence from less massive stars (at the lower right in the H-R diagram) to more massive stars (at the upper left), the radius and surface temperature both increase.

Calculations using hydrostatic and thermal equilibrium also show that if a star's mass is less than about 0.08$M_\odot$, the core pressure and temperature are too low for thermonuclear reactions to take place. The "star" is then a brown dwarf. Brown dwarfs also obey a mass-luminosity relation: The greater the mass, the faster the brown dwarf contracts because of its own gravity, the more rapidly it radiates energy into space, and, hence, the more luminous the brown dwarf is.

CAUTION! The mass-luminosity relation we have discussed applies to main-sequence stars only. There are *no* simple mass-luminosity relations for giant, supergiant, or white dwarf stars. Why these stars lie where they do on an H-R diagram will become apparent when we study the evolution of stars in Chapters 19 and 20. We will find that main-sequence stars evolve into giant and supergiant stars, and that some of these eventually end their lives as white dwarfs.

17-10 Spectroscopy makes it possible to study binary systems in which the two stars are close together

We have described how the masses of stars can be determined from observations of visual binaries, in which the two stars can be distinguished from each other. But if the two stars in a binary system are too close together, the images of the two stars can blend to produce the semblance of a single star. Happily, in many cases we can use spectroscopy to decide whether a seemingly single star is in fact a binary system. Spectroscopic observations of binaries provide additional useful information about star masses.

Some binaries are discovered when the spectrum of a star shows incongruous spectral lines. For example, the spectrum of what appears to be a single star may include both strong hydrogen lines (characteristic of a type A star) and strong absorption bands of titanium oxide (typical of a type M star). Because a single star cannot have the differing physical properties of these two spectral types, such a star must actually be a binary system that is too far away for us to resolve its individual stars. A binary system detected in this way is called a **spectrum binary.**

Other binary systems can be detected using the Doppler effect. If a star is moving toward the Earth, its spectral lines are displaced toward the short-wavelength (blue) end of the spectrum. Conversely, the spectral lines of a star moving away from us are shifted toward the long-wavelength (red) end of the spectrum. The upper portion of Figure 17-22 applies these ideas to a hypothetical binary star system with an orbital plane that is edge-on to our line of sight.

As the two stars move around their orbits, they periodically approach and recede from us. Hence, the spectral lines of the two stars are alternately blueshifted and redshifted. The two stars in this hypothetical system are so close together that they appear through a telescope as a single star with a single spectrum.

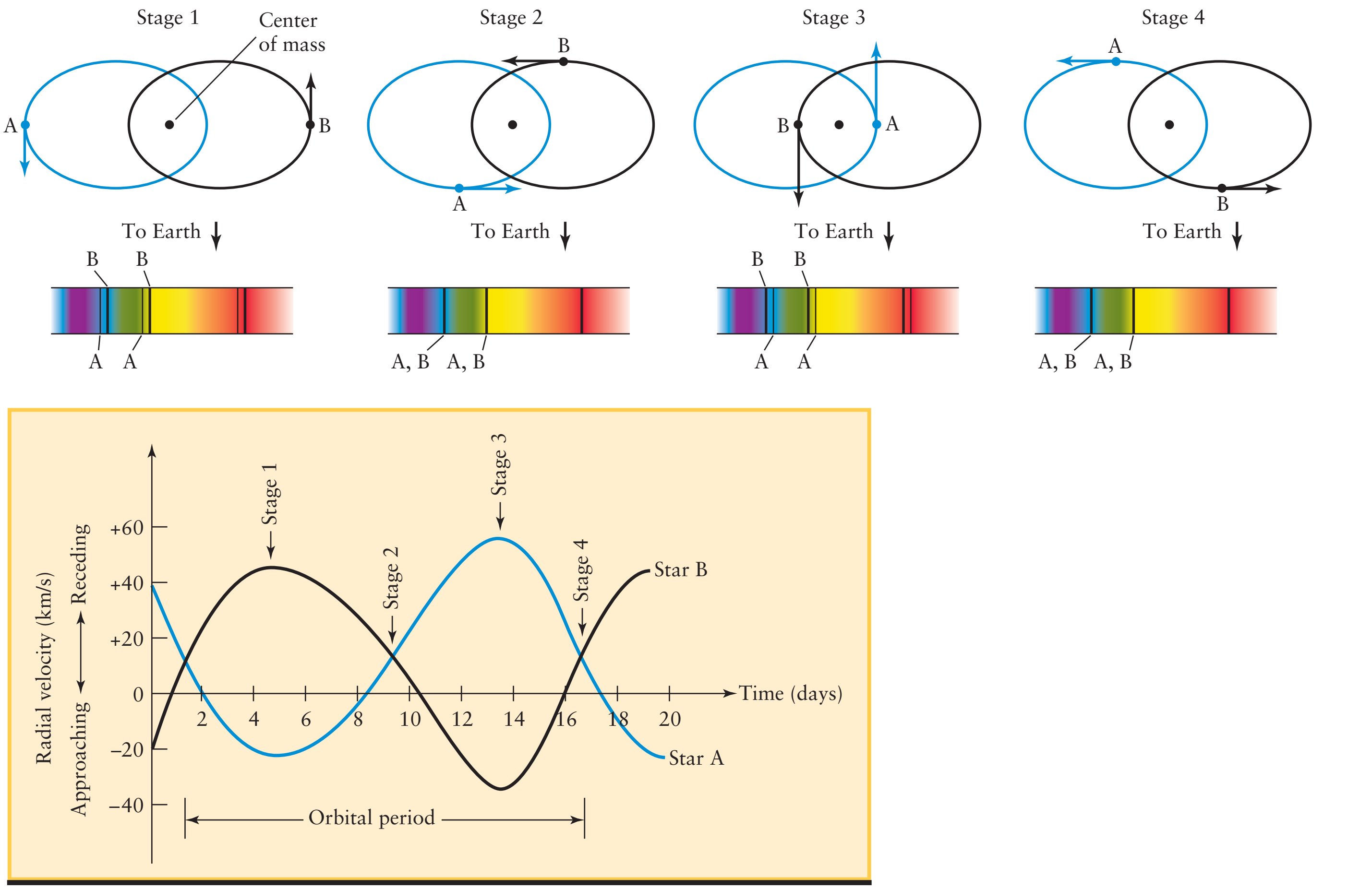


Figure 17-22

Radial Velocity Curves The lower graph displays the radial velocity curves of the binary system HD 171978. The drawings at the top indicate the positions of the stars (labeled A and B) and the spectra of the binary at four selected moments (stages 1, 2, 3, and 4) during an orbital period. Note that at stages 1 and 3, the Doppler effect splits apart the absorption lines from stars A and B; compare with Figure 17-23. The entire binary star system is moving away from us at 12 km/s, which is why the entire pattern of radial velocity curves is displaced upward from the zero-velocity line.

Because one star shows a blueshift while the other is showing a redshift, the spectral lines of the binary system appear to split apart and rejoin periodically. Stars whose binary character is revealed by such shifting spectral lines are called **spectroscopic binaries.**

Exploring Spectroscopic Binary Stars

To analyze a spectroscopic binary, astronomers measure the wavelength shift of each star's spectral lines and use the Doppler shift formula (introduced in Section 5-9 and Box 5-6) to determine the **radial velocity** of each star—that is, how fast and in what direction it is moving along our line of sight. The lower portion of Figure 17-22 shows a graph of the radial velocity versus time, called a **radial velocity curve,** for the binary system HD 171978. Each of the two stars alternately approaches and recedes as it orbits around the center of mass. The pattern of the curves repeats every 15 days, which is the orbital period of the binary.

> **Stars in close binary systems move so rapidly that we can detect their motion using the Doppler effect**

Figure 17-23 shows two spectra of the spectroscopic binary κ (kappa) Arietis taken a few days apart. In Figure 17-23*a*, two sets of spectral lines are visible, offset slightly in opposite directions from the normal positions of these lines. This corresponds to stage 1 or stage 3 in Figure 17-22; one of the orbiting stars is moving toward the Earth and has its spectral lines blueshifted, and the other star is moving away from the Earth and has its lines redshifted. A few days later, the stars have progressed along their orbits so that neither star is moving toward or away from the Earth, corresponding to stage 2 or stage 4 in Figure 17-22. At this time there are no Doppler shifts, and the spectral lines of both stars are at the same positions. That is why only one set of spectral lines appears in Figure 17-23*b*.

It is important to emphasize that the Doppler effect applies only to motion along the line of sight. Motion perpendicular to

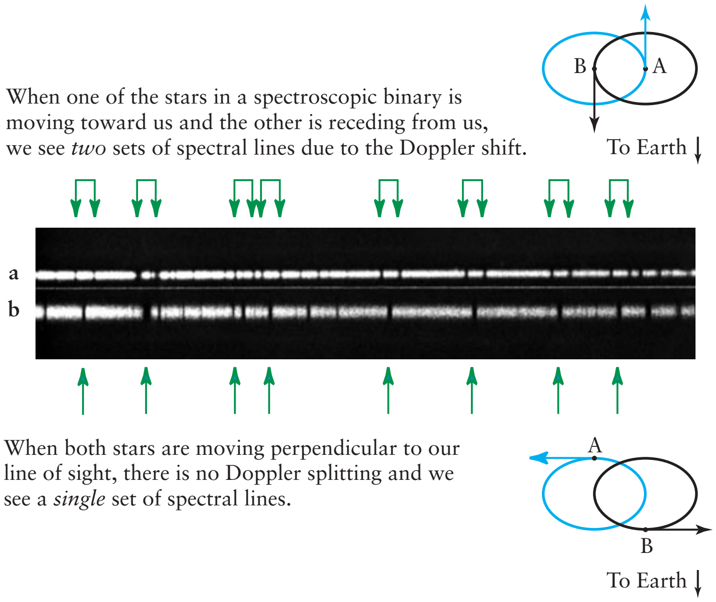


Figure 17-23 R I V U X G

A Spectroscopic Binary The visible-light spectrum of the double-line spectroscopic binary κ (kappa) Arietis has spectral lines that shift back and forth as the two stars revolve about each other. (Lick Observatory)

the line of sight does not affect the observed wavelengths of spectral lines. Hence, the ideal orientation for a spectroscopic binary is to have the stars orbit in a plane that is edge-on to our line of sight. (By contrast, a *visual* binary is best observed if the orbital plane is face-on to our line of sight.) For the Doppler shifts to be noticeable, the orbital speeds of the two stars should be at least a few kilometers per second.

The binaries depicted in Figures 17-22 and 17-23 are called *double-line* spectroscopic binaries, because the spectral lines of both stars in the binary system can be seen. Most spectroscopic binaries, however, are *single-line* spectroscopic binaries: One of the stars is so dim that its spectral lines cannot be detected. The star is obviously a binary, however, because its spectral lines shift back and forth, thereby revealing the orbital motions of two stars about their center of mass.

As for visual binaries, spectroscopic binaries allow astronomers to learn about stellar masses. From a radial velocity curve, one can find the *ratio* of the masses of the two stars in a binary. The *sum* of the masses is related to the orbital speeds of the two stars by Kepler's laws and Newtonian mechanics. If both the ratio of the masses and their sum are known, the individual masses can be determined using algebra. However, determining the sum of the masses requires that we know how the binary orbits are tilted from our line of sight. This is because the Doppler shifts reveal only the radial velocities of the stars rather than their true orbital speeds. This tilt is often impossible to determine, because we cannot see the individual stars in the binary. Thus, the masses of stars in spectroscopic binaries tend to be uncertain.

There is one important case in which we can determine the orbital tilt of a spectroscopic binary. If the two stars are observed to eclipse each other periodically, then we must be viewing the orbit nearly edge-on. As we will see next, individual stellar masses—as well as other useful data—can be determined if a spectroscopic binary also happens to be such an *eclipsing* binary.

17-11 Light curves of eclipsing binaries provide detailed information about the two stars

Some binary systems are oriented so that the two stars periodically eclipse each other as seen from the Earth. These **eclipsing binaries** can be detected even when the two stars cannot be resolved visually as two distinct images in the telescope. The apparent brightness of the image of the binary dims briefly each time one star blocks the light from the other.

Using a sensitive detector at the focus of a telescope, an astronomer can measure the incoming light intensity quite accurately and create a **light curve** (Figure 17-24). The shape of the light curve for an eclipsing binary reveals at a glance whether the eclipse is partial or total (compare Figures 17-24*a* and 17-24*b*). Figure 17-24*d* shows an observation of a binary system undergoing a total eclipse.

Eclipsing binaries can reveal the sizes and shapes of stars

In fact, the light curve of an eclipsing binary can yield a surprising amount of information. For example, the ratio of the surface temperatures can be determined from how much their combined light is diminished when the stars eclipse each other. Also, the duration of a mutual eclipse tells astronomers about the relative sizes of the stars and their orbits.

If the eclipsing binary is also a double-line spectroscopic binary, an astronomer can calculate the mass and radius of each star from the light curves and the velocity curves. Unfortunately, very few binary stars are of this ideal type. Stellar radii determined in this way agree well with the values found using the Stefan-Boltzmann law, as described in Section 17-6.

The shape of a light curve can reveal many additional details about a binary system. In some binaries, for example, the gravitational pull of one star distorts the other, much as the Moon distorts the Earth's oceans in producing tides (see Figure 4-26). Figure 17-24*c* shows how such tidal distortion gives the light curve a different shape than in Figure 17-24*b*.

Information about stellar atmospheres can also be derived from light curves. Suppose that one star of a binary is a luminous main-sequence star and the other is a bloated red giant. By observing exactly how the light from the bright main-sequence star is gradually cut off as it moves behind the edge of the red giant during the beginning of an eclipse, astronomers can infer the pressure and density in the upper atmosphere of the red giant.

Binary systems are tremendously important because they enable astronomers to measure stellar masses as well as other key properties of stars. In the next several chapters, we will use this information to help us piece together the story of *stellar evolution*—how stars are born, evolve, and eventually die.

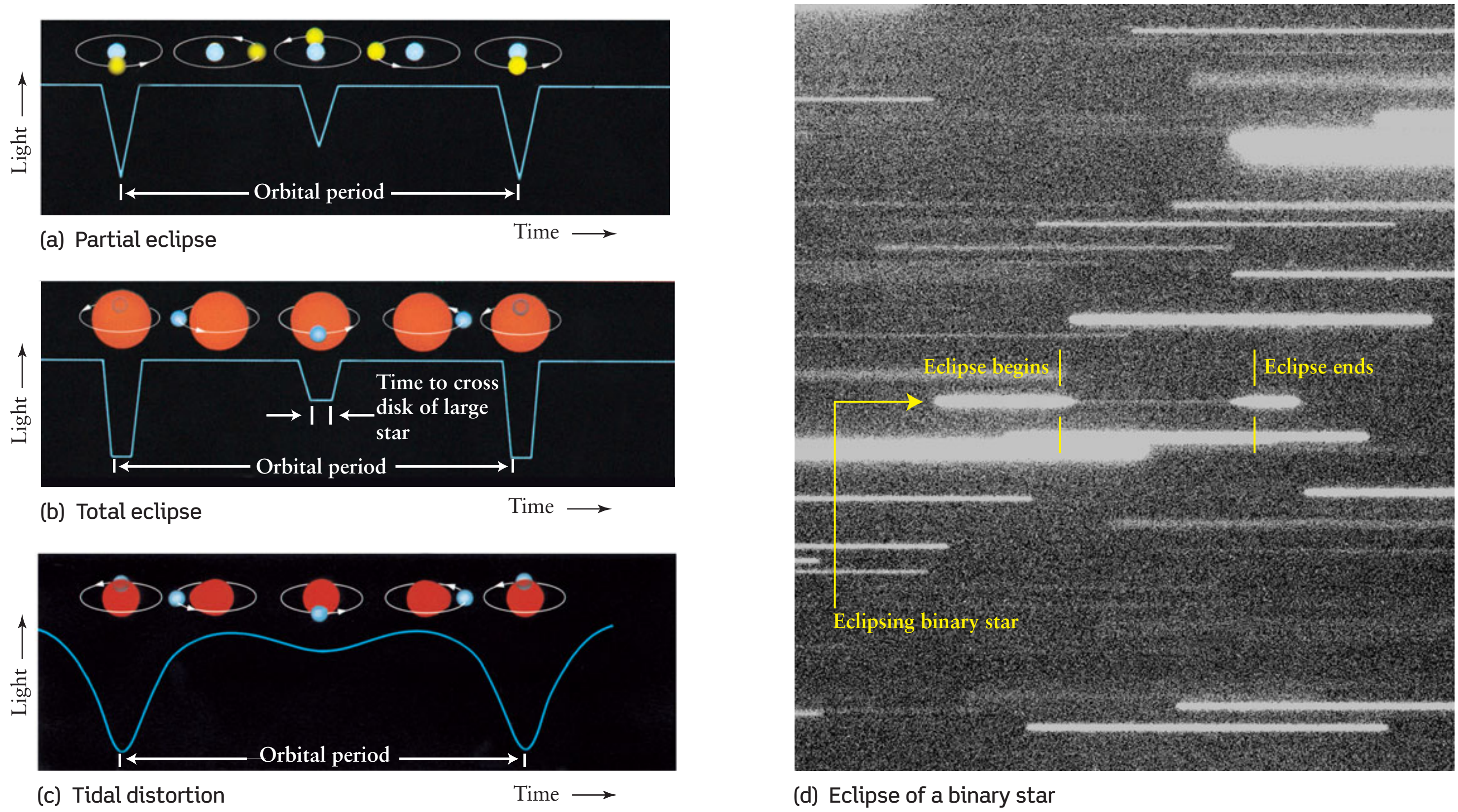


Figure 17-24 R I V U X G

Representative Light Curves of Eclipsing Binaries (a), (b), (c) The shape of the light curve of an eclipsing binary can reveal many details about the two stars that make up the binary. **(d)** This image shows the binary star NN Serpens (indicated by the arrow) undergoing a total eclipse. The telescope was moved during the exposure so that the sky drifted slowly from left to right across the field of view. During the 10.5-minute duration of the eclipse, the dimmer star of the binary system (an M6 main-sequence star) passed in front of the other, more luminous star (a white dwarf). The binary became so dim that it almost disappeared. (European Southern Observatory)

Key Words

Terms preceded by an asterisk (*) are discussed in the Boxes.

absolute magnitude, p. 441
apparent brightness (brightness), p. 437
apparent magnitude, p. 440
binary star (binary), p. 456
brown dwarf, p. 449
center of mass, p. 458
color ratio, p. 445
*distance modulus, p. 444
double star, p. 456
eclipsing binary, p. 462
giant, p. 454
Hertzsprung-Russell diagram (H-R diagram), p. 452
inverse-square law, p. 437
light curve, p. 462
luminosity, p. 434
luminosity class, p. 455
luminosity function, p. 438
magnitude scale, p. 440
main sequence, p. 453
main-sequence star, p. 453
mass-luminosity relation, p. 458
metals, p. 447
OBAFGKM, p. 446
optical double star, p. 456
parallax, p. 434
parsec, p. 434
photometry, p. 438
*proper motion, p. 436
radial velocity, p. 461
radial velocity curve, p. 461
red giant, p. 454
*space velocity, p. 436
spectral classes, p. 446
spectral types, p. 446
spectroscopic binary, p. 461
spectroscopic parallax, p. 455
spectrum binary, p. 460
stellar parallax, p. 434
supergiant, p. 454
*tangential velocity, p. 436
UBV photometry, p. 445
visual binary, p. 457
white dwarf, p. 454

Key Ideas

Measuring Distances to Nearby Stars: Distances to the nearer stars can be determined by parallax, the apparent shift of a star against the background stars observed as the Earth moves along its orbit.

• Parallax measurements made from orbit, above the blurring effects of the atmosphere, are much more accurate than those made with Earth-based telescopes.

• Stellar parallaxes can only be measured for stars within a few hundred parsecs.

The Inverse-Square Law: A star's luminosity (total light output), apparent brightness, and distance from the Earth are related by the inverse-square law. If any two of these quantities are known, the third can be calculated.

The Population of Stars: Stars of relatively low luminosity are more common than more luminous stars. Our own Sun is a rather average star of intermediate luminosity.

The Magnitude Scale: The apparent magnitude scale is an alternative way to measure a star's apparent brightness.

• The absolute magnitude of a star is the apparent magnitude it would have if viewed from a distance of 10 parsecs. A version of the inverse-square law relates a star's absolute magnitude, apparent magnitude, and distance.

Photometry and Color Ratios: Photometry measures the apparent brightness of a star. The color ratios of a star are the ratios of brightness values obtained through different standard filters, such as the U, B, and V filters. These ratios are a measure of the star's surface temperature.

Spectral Types: Stars are classified into spectral types (subdivisions of the spectral classes O, B, A, F, G, K, and M), based on the major patterns of spectral lines in their spectra. The spectral class and type of a star is directly related to its surface temperature: O stars are the hottest and M stars are the coolest.

• Most brown dwarfs are in even cooler spectral classes called L and T. Unlike true stars, brown dwarfs are too small to sustain thermonuclear fusion.

Hertzsprung-Russell Diagram: The Hertzsprung-Russell (H-R) diagram is a graph plotting the absolute magnitudes of stars against their spectral types—or, equivalently, their luminosities against surface temperatures.

• The positions on the H-R diagram of most stars are along the main sequence, a band that extends from high luminosity and high surface temperature to low luminosity and low surface temperature.

• On the H-R diagram, giant and supergiant stars lie above the main sequence, while white dwarfs are below the main sequence.

• By carefully examining a star's spectral lines, astronomers can determine whether that star is a main-sequence star, giant, supergiant, or white dwarf. Using the H-R diagram and the inverse-square law, the star's luminosity and distance can be found without measuring its stellar parallax.

Binary Stars: Binary stars, in which two stars are held in orbit around each other by their mutual gravitational attraction, are surprisingly common. Those that can be resolved into two distinct star images by an Earth-based telescope are called visual binaries.

• Each of the two stars in a binary system moves in an elliptical orbit about the center of mass of the system.

• Binary stars are important because they allow astronomers to determine the masses of the two stars in a binary system. The masses can be computed from measurements of the orbital period and orbital dimensions of the system.

Mass-Luminosity Relation for Main-Sequence Stars: Main-sequence stars are stars like the Sun but with different masses.

• The mass-luminosity relation expresses a direct correlation between mass and luminosity for main-sequence stars. The greater the mass of a main-sequence star, the greater its luminosity (and also the greater its radius and surface temperature).

Spectroscopic Observations of Binary Stars: Some binaries can be detected and analyzed, even though the system may be so distant or the two stars so close together that the two star images cannot be resolved.

• A spectrum binary appears to be a single star but has a spectrum with the absorption lines for two distinctly different spectral types.

• A spectroscopic binary has spectral lines that shift back and forth in wavelength. This is caused by the Doppler effect, as the orbits of the stars carry them first toward then away from the Earth.

• An eclipsing binary is a system whose orbits are viewed nearly edge-on from the Earth, so that one star periodically eclipses the other. Detailed information about the stars in an eclipsing binary can be obtained from a study of the binary's radial velocity curve and its light curve.

Questions

Review Questions

1. Explain the difference between a star's apparent brightness and its luminosity.
2. Describe how the parallax method of finding a star's distance is similar to binocular (two-eye) vision in humans.
3. Why does it take at least six months to make a measurement of a star's parallax?
4. Why are measurements of stellar parallax difficult to make? What are the advantages of making these measurements from orbit?
5. What is the inverse-square law? Use it to explain why an ordinary lightbulb can appear brighter than a star, even though the lightbulb emits far less light energy per second.
6. Briefly describe how you would determine the luminosity of a nearby star. Of what value is knowing the luminosity of various stars?
7. Which are more common, stars more luminous than the Sun or stars less luminous than the Sun?
8. Why is the magnitude scale called a "backward" scale? What is the difference between apparent magnitude and absolute magnitude?
9. The star Zubenelgenubi (from the Arabic for "scorpion's southern claw") has apparent magnitude +2.75, while the star Sulafat (Arabic for "tortoise") has apparent magnitude +3.25. Which star appears brighter? From this information alone, what can you conclude about the luminosities of these stars? Explain.
10. Explain why the color ratios of a star are related to the star's surface temperature.
11. Would it be possible for a star to appear bright when viewed through a U filter or a V filter, but dim when viewed through a B filter? Explain.

12. Which gives a more accurate measure of a star's surface temperature, its color ratios or its spectral lines? Explain.
13. Menkalinan (Arabic for "shoulder of the rein-holder") is an A2 star in the constellation Auriga (the Charioteer). What is its spectral class? What is its spectral type? Which gives a more precise description of the spectrum of Menkalinan?
14. What are the most prominent absorption lines you would expect to find in the spectrum of a star with a surface temperature of (**a**) 35,000 K, (**b**) 2800 K, and (**c**) 5800 K (like the Sun)? Briefly describe why these stars have such different spectra even though they have essentially the same chemical composition.
15. A fellow student expresses the opinion that since the Sun's spectrum has only weak absorption lines of hydrogen, this element cannot be a major constituent of the Sun. How would you enlighten this person?
16. If a red star and a blue star both have the same radius and both are the same distance from the Earth, which one looks brighter in the night sky? Explain why.
17. If a red star and a blue star both appear equally bright and both are the same distance from the Earth, which one has the larger radius? Explain why.
18. If a red star and a blue star both have the same radius and both appear equally bright, which one is farther from Earth? Explain why.
19. Sketch a Hertzsprung-Russell diagram. Indicate the regions on your diagram occupied by (**a**) main-sequence stars, (**b**) red giants, (**c**) supergiants, (**d**) white dwarfs, and (**e**) the Sun.
20. Most of the bright stars in the night sky (see Appendix 5) are giants and supergiants. How can this be, if giants and supergiants make up only 1% of the population of stars?
21. Explain why the dashed lines in Figure 17-15*b* slope down and to the right.
22. Some giant and supergiant stars are of the same spectral type (G2) as the Sun. What aspects of the spectrum of a G2 star would you concentrate on to determine the star's luminosity class? Explain what you would look for.
23. Briefly describe how you would determine the distance to a star whose parallax is too small to measure.
24. What information about stars do astronomers learn from binary systems that cannot be learned in any other way? What measurements do they make of binary systems to garner this information?
25. Suppose that you want to determine the temperature, diameter, and luminosity of an isolated star (not a member of a binary system). Which of these physical quantities require you to know the distance to the star? Explain.
26. What is the mass-luminosity relation? Does it apply to stars of all kinds?
27. Use Figure 17-21 to (**a**) estimate the mass of a main-sequence star that is 1000 times as luminous as the Sun, and (**b**) estimate the luminosity of a main-sequence star whose mass is one-fifth that of the Sun. Explain your answers.
28. Which is more massive, a red main-sequence star or a blue main-sequence star? Which has the greater radius? Explain your answers.
29. How do white dwarfs differ from brown dwarfs? Which are more massive? Which are larger in radius? Which are denser?
30. Sketch the radial velocity curves of a binary consisting of two identical stars moving in circular orbits that are (**a**) perpendicular to and (**b**) parallel to our line of sight.
31. Give two reasons why a visual binary star is unlikely to also be a spectroscopic binary star.
32. Sketch the light curve of an eclipsing binary consisting of two identical stars in highly elongated orbits oriented so that (**a**) their major axes are pointed toward the Earth and (**b**) their major axes are perpendicular to our line of sight.

Advanced Questions

Questions preceded by an asterisk (*) involve topics discussed in the Boxes.

Problem-solving tips and tools

Look carefully at the worked examples in Boxes 17-1, 17-2, 17-3, and 17-4 before attempting these exercises. For data on the planets, see Table 7-1 or Appendices 1 and 2 at the back of this book. Remember that a telescope's light-gathering power is proportional to the area of its objective or primary mirror. The volume of a sphere of radius r is $4\pi r^3/3$. Make use of the H-R diagrams in this chapter to answer questions involving spectroscopic parallax. As Box 17-3 shows, some of the problems concerning magnitudes may require facility with logarithms.

33. Find the average distance from the Sun to Neptune in parsecs. Compared to Neptune, how many times farther away from the Sun is Proxima Centauri?
34. Suppose that a dim star were located 2 million AU from the Sun. Find (**a**) the distance to the star in parsecs and (**b**) the parallax angle of the star. Would this angle be measurable with present-day techniques?
35. The star GJ 1156 has a parallax angle of 0.153 arcsec. How far away is the star?

*36. Kapteyn's star (named after the Dutch astronomer who found it) has a parallax of 0.255 arcsec, a proper motion of 8.67 arcsec per year, and a radial velocity of +246 km/s. (**a**) What is the star's tangential velocity? (**b**) What is the star's actual speed relative to the Sun? (**c**) Is Kapteyn's star moving toward the Sun or away from the Sun? Explain.

*37. How far away is a star that has a proper motion of 0.08 arcseconds per year and a tangential velocity of 40 km/s? For a star at this distance, what would its tangential velocity have to be in order for it to exhibit the same proper motion as Barnard's star (see Box 17-1)?

*38. The space velocity of a certain star is 120 km/s and its radial velocity is 72 km/s. Find the star's tangential velocity.

*39. In the spectrum of a particular star, the Balmer line H_α has a wavelength of 656.15 nm. The laboratory value for the wavelength of H_α is 656.28 nm. (**a**) Find the star's radial velocity. (**b**) Is this star approaching us or moving away? Explain. (**c**) Find the wavelength at which you would expect to find H_α in the spectrum of this star, given that the laboratory wavelength of H_α is 486.13 nm. (**d**) Do your answers depend on the distance from the Sun to this star? Why or why not?

*40. Derive the equation given in Box 17-1 relating proper motion and tangential velocity. (*Hint:* See Box 1-1.)

41. How much dimmer does the Sun appear from Neptune than from Earth? (*Hint:* The average distance between a planet and the Sun equals the semimajor axis of the planet's orbit.)

42. Stars *A* and *B* are both equally bright as seen from Earth, but *A* is 120 pc away while *B* is 24 pc away. Which star has the greater luminosity? How many times greater is it?

43. Stars *C* and *D* both have the same luminosity, but *C* is 32 pc from Earth while *D* is 128 pc from Earth. Which star appears brighter as seen from Earth? How many times brighter is it?

44. Suppose two stars have the same apparent brightness, but one star is 8 times farther away than the other. What is the ratio of their luminosities? Which one is more luminous, the closer star or the farther star?

45. The *solar constant,* equal to 1370 W/m^2, is the amount of light energy from the Sun that falls on 1 square meter of the Earth's surface in 1 second (see Section 17-2). What would the distance between the Earth and the Sun have to be in order for the solar constant to be 1 watt per square meter (1 W/m^2)?

46. The star Procyon in Canis Minor (the Small Dog) is a prominent star in the winter sky, with an apparent brightness 1.3×10^{-11} that of the Sun. It is also one of the nearest stars, being only 3.50 parsecs from Earth. What is the luminosity of Procyon? Express your answer as a multiple of the Sun's luminosity.

47. The star HIP 92403 (also called Ross 154) is only 2.97 parsecs from Earth but can be seen only with a telescope, because it is 60 times dimmer than the dimmest star visible to the unaided eye. How close to us would this star have to be in order for it to be visible without a telescope? Give your answer in parsecs and in AU. Compare with the semimajor axis of Pluto's orbit around the Sun.

*48. The star HIP 72509 has an apparent magnitude of +12.1 and a parallax angle of 0.222 arcsecond. (**a**) Determine its absolute magnitude. (**b**) Find the approximate ratio of the luminosity of HIP 72509 to the Sun's luminosity.

*49. Suppose you can just barely see a twelfth-magnitude star through an amateur's 6-inch telescope. What is the magnitude of the dimmest star you could see through a 60-inch telescope?

*50. A certain type of variable star is known to have an average absolute magnitude of 0.0. Such stars are observed in a particular star cluster to have an average apparent magnitude of +14.0. What is the distance to that star cluster?

*51. (**a**) Find the absolute magnitudes of the brightest and dimmest of the labeled stars in Figure 17-6*b*. Assume that all of these stars are 110 pc from Earth. (**b**) If a star in the Pleiades cluster is just bright enough to be seen from Earth with the naked eye, what is its absolute magnitude? Is such a star more or less luminous than the Sun? Explain.

52. (**a**) On a copy of Figure 17-8, sketch the intensity curve for a blackbody at a temperature of 3000 K. Note that this figure shows a smaller wavelength range than Figure 17-7*a*. (**b**) Repeat part (a) for a blackbody at 12,000 K (see Figure 17-7*c*). (**c**) Use your sketches from parts (a) and (b) to explain why the color ratios b_V/b_B and b_B/b_U are less than 1 for very hot stars but greater than 1 for very cool stars.

*53. Astronomers usually express a star's color using apparent magnitudes. The star's apparent magnitude as viewed through a B filter is called m_B, and its apparent magnitude as viewed through a V filter is m_V. The difference $m_B - m_V$ is called the *B–V color index* ("B minus V"). Is the B–V color index positive or negative for very hot stars? What about very cool stars? Explain your answers.

*54. (See Question 53.) The B–V color index is related to the color ratio b_V/b_B by the equation

$$m_B = m_V = 2.5 \log\left(\frac{b_V}{b_B}\right)$$

(**a**) Explain why this equation is correct. (**b**) Use the data in Table 17-1 to calculate the B–V color indices for Bellatrix, the Sun, and Betelgeuse. From your results, describe a simple rule that relates the value of the B–V color index to a star's color.

55. The bright star Rigel in the constellation Orion has a surface temperature about 1.6 times that of the Sun. Its luminosity is about 64,000 $L_\odot$. What is Rigel's radius compared to the radius of the Sun?

56. (See Figure 17-12.) What temperature and spectral classification would you give to a star with equal line strengths of hydrogen (H) and neutral helium (He I)? Explain.

57. The Sun's surface temperature is 5800 K. Using Figure 17-12, arrange the following absorption lines in the Sun's spectrum from the strongest to the weakest, and explain your reasoning: (i) neutral calcium; (ii) singly ionized calcium; (iii) neutral iron; (iv) singly ionized iron.

58. Star *P* has one-half the radius of star *Q*. Stars *P* and *Q* have surface temperatures 4000 K and 8000 K, respectively. Which star has the greater luminosity? How many times greater is it?

59. Star *X* has 12 times the luminosity of star *Y*. Stars *X* and *Y* have surface temperatures 3500 K and 7800 K, respectively. Which star has the larger radius? How many times larger is it?

60. Suppose a star experiences an outburst in which its surface temperature doubles but its average density (its mass divided by its volume) decreases by a factor of 8. The mass of the star stays the same. By what factors do the star's radius and luminosity change?

61. The Sun experiences solar flares (see Section 16-10). The amount of energy radiated by even the strongest solar flare is not enough to have an appreciable effect on the Sun's luminosity. But when a flare of the same size occurs on a main-sequence star of spectral class M, the star's brightness can increase by as much as a factor of 2. Why should there be an appreciable increase in brightness for a main-sequence M star but not for the Sun?

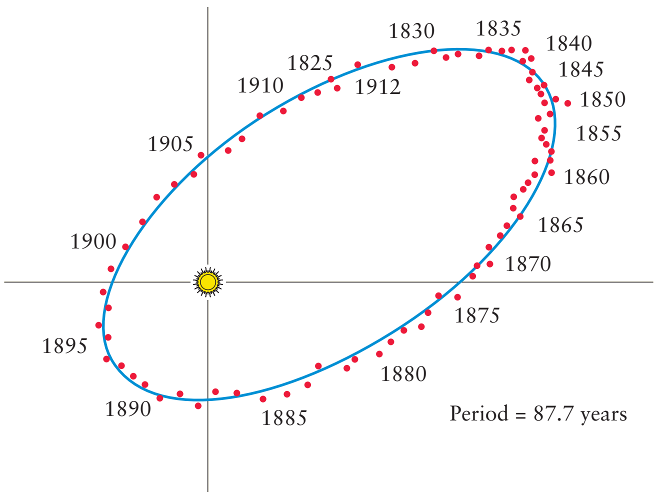
62. The bright star Zubeneschmali (β Librae) is of spectral type B8 and has a luminosity of 130 $L_\odot$. What is the star's approximate surface temperature? How does its radius compare to that of the Sun?

63. Castor (α Geminorum) is an A1 V star with an apparent brightness of 4.4×10^{-12} that of the Sun. Determine the approximate distance from the Earth to Castor (in parsecs).

64. A brown dwarf called CoD–33°7795 B has a luminosity of $0.0025 L_{\odot}$. It has a relatively high surface temperature of 2550 K, which suggests that it is very young and has not yet had time to cool down by emitting radiation. (**a**) What is this brown dwarf's spectral class? (**b**) Find the radius of CoD–33°7795 B. Express your answer in terms of the Sun's radius and in kilometers. How does this compare to the radius of Jupiter? Is the name "dwarf" justified?

65. The star HD 3651 shown in Figure 17-13 has a mass of $0.79\ M_{\odot}$. Its brown dwarf companion, HD 3651B, has about 40 times the mass of Jupiter. The average distance between the two stars is about 480 AU. How long does it take the two stars to complete one orbit around each other?

66. The visual binary 70 Ophiuchi (see the accompanying figure) has a period of 87.7 years. The parallax of 70 Ophiuchi is 0.2 arcsec, and the apparent length of the semimajor axis as seen through a telescope is 4.5 arcsec. (**a**) What is the distance to 70 Ophiuchi in parsecs? (**b**) What is the actual length of the semimajor axis in AU? (**c**) What is the sum of the masses of the two stars? Give your answer in solar masses.



67. An astronomer observing a binary star finds that one of the stars orbits the other once every 5 years at a distance of 10 AU. (**a**) Find the sum of the masses of the two stars. (**b**) If the mass ratio of the system is $M_1/M_2 = 0.25$, find the individual masses of the stars. Give your answers in terms of the mass of the Sun.

Discussion Questions

68. From its orbit around the Earth, the *Hipparcos* satellite could measure stellar parallax angles with acceptable accuracy only if the angles were larger than about 0.002 arcsec. Discuss the advantages or disadvantages of making parallax measurements from a satellite in a large solar orbit, say at the distance of Jupiter from the Sun. If this satellite can also measure parallax angles of 0.002 arcsec, what is the distance of the most remote stars that can be accurately determined? How much bigger a volume of space would be covered compared to the Earth-based observations? How many more stars would you expect to be contained in that volume?

***69.** As seen from the starship *Enterprise* in the *Star Trek* television series and movies, stars appear to move across the sky due to the starship's motion. How fast would the *Enterprise* have to move in order for a star 1 pc away to appear to move 1° per second? (*Hint:* The speed of the star as seen from the *Enterprise* is the same as the speed of the *Enterprise* relative to the star.) How does this compare with the speed of light? Do you think the stars appear to move as seen from an orbiting space shuttle, which moves at about 8 km/s?

70. It is desirable to be able to measure the radial velocity of stars (using the Doppler effect) to an accuracy of 1 km/s or better. One complication is that radial velocities refer to the motion of the star relative to the Sun, while the observations are made using a telescope on the Earth. Is it important to take into account the motion of the Earth around the Sun? Is it important to take into account the Earth's rotational motion? To answer this question, you will have to calculate the Earth's orbital speed and the speed of a point on the Earth's equator (the part of the Earth's surface that moves at the greatest speed because of the planet's rotation). If one or both of these effects are of importance, how do you suppose astronomers compensate for them?

Web/eBook Questions

71. Search the World Wide Web for information about *Gaia*, a European Space Agency (ESA) spacecraft planned to extend the work carried out by *Hipparcos*. When is the spacecraft planned to be launched? How will *Gaia* compare to *Hipparcos*? For how many more stars will it be able to measure parallaxes? What other types of research will it carry out?

72. Search the World Wide Web for recent discoveries about brown dwarfs. Are all brown dwarfs found orbiting normal stars, or are they also found orbiting other brown dwarfs? Are any found in isolation (that is, not part of a binary system)? The Sun experiences flares (see Section 16-10), as do other normal stars; is there any evidence that brown dwarfs also experience flares? If so, is there anything unusual about these flares?

AIMM 17.2 **73.** **Distances to Stars Using Parallax.** Access the Active Integrated Media Module "Using Parallax to Determine Distance" in Chapter 17 of the *Universe* Web site or eBook. Use this to determine the distance in parsecs and in light-years to each of the following stars: (**a**) Betelgeuse (parallax $p = 0.00763$ arcsecond); (**b**) Vega ($p = 0.129$ arcsecond); (**c**) Antares ($p = 0.00540$ arcsecond); (**d**) Sirius ($p = 0.379$ arcsecond).

AIMM 17.3 ***74.** **Finding Absolute Magnitudes.** Access the Active Integrated Media Module "The Distance-Magnitude Relationship" in Chapter 17 of the *Universe* Web site or eBook. Use this to determine the absolute magnitudes of stars with the following properties: (**a**) apparent magnitude +6.3, distance = 125 pc; (**b**) apparent magnitude +11.4, distance = 48 pc; (**c**) apparent magnitude +9.8, distance = 70 pc. (**d**) Rank the three stars from parts (a), (b), and (c) in order from highest to lowest luminosity.

Activities

Observing Projects

Observing tips and tools

Even through a telescope, the colors of stars are sometimes subtle and difficult to see. To give your eye the best chance of seeing color, use the "averted vision" trick: When looking through the telescope eyepiece, direct your vision a little to one side of the star you are most interested in. This places the light from that star on a more sensitive part of your eye's retina.

75. The table below lists five well-known red stars. It includes their right ascension and declination (celestial coordinates described in Box 2-1), apparent magnitudes, and color ratios. As their apparent magnitudes indicate, all these stars are somewhat variable. Observe at least two of these stars both by eye and through a small telescope. Is the reddish color of the stars readily apparent, especially in contrast to neighboring stars? (The Jesuit priest and astronomer Angelo Secchi named Y Canum Venaticorum "La Superba," and μ Cephei is often called William Herschel's "Garnet Star.")

Star	Right ascension	Declination	Apparent magnitude	b_V/b_B
Betelgeuse	$5^h\ 55.2^m$	+7° 24′	0.4–1.3	5.5
Y Canum Venaticorum	12 45.1	+45 26	5.5–6.0	10.4
Antares	16 29.4	−26 26	0.9–1.8	5.4
μ Cephei	21 43.5	+58 47	3.6–5.1	8.7
TX Piscium	23 46.4	+3 29	5.3–5.8	11.0

Note: The right ascensions and declinations are given for epoch 2000.

76. The table of double stars shown below includes vivid examples of contrasting star colors. The table lists the angular separation between the stars of each double. Observe at least four of these double stars through a telescope. Use the spectral types listed to estimate the difference in surface temperature of the stars in each pair you observe. Does the double with the greatest difference in temperature seem to present the greatest color contrast? From what you see through the telescope and on what you know about the H-R diagram, explain why all the cool stars (spectral types K and M) listed are probably giants or supergiants.

77. Observe the eclipsing binary Algol (β Persei), using nearby stars to judge its brightness during the course of an eclipse. Algol has an orbital period of 2.87 days, and, with the onset of primary eclipse, its apparent magnitude drops from 2.1 to 3.4. It remains this faint for about 2 hours. The entire eclipse, from start to finish, takes about 10 hours. Consult the "Celestial Calendar" section of the current issue of *Sky & Telescope* for the predicted dates and times of the minima of Algol. Note that the schedule is given in Universal Time (the same as Greenwich Mean Time), so you will have to convert the time to that of your own time zone. Algol is normally the second brightest star in the constellation of Perseus. Because of its position on the celestial sphere (R.A. = $3^h\ 08.2^m$, Decl. = 40° 57′), Algol is readily visible from northern latitudes during the fall and winter months.

78. Use the *Starry Night Enthusiast™* program to investigate the brightest stars. Click on **Home** to show the sky as seen from your location. Set the date to today's date and the time to midnight (12:00:00 A.M.). Open the **Options** pane, expand the **Constellations** layer and click the **Boundaries** and **Labels** options to turn these displays on. You will now see the boundaries of the constellations. Open the **Info** pane and expand the **Position in Space** and **Other Data** lists. (**a**) Scroll around the sky and identify at least five of the brighter stars (shown as larger dots) and click on them to reveal relevant data in the **Info** pane. Make a list of these stars and record **Luminosity** and **Distance from Sun** from the **Info** pane. Which stars did you select? In which constellation does each of these stars lie? Which of these stars are listed in Appendix 5? Of these, which is the most luminous? Which is the most distant? (**b**) Set the

Star	Right ascension	Declination	Apparent magnitudes	Angular separation (arcseconds)	Spectral types
55 Piscium	$0^h\ 39.9^m$	+21° 26′	5.4 and 8.7	6.5	K0 and F3
γ Andromedae	2 03.9	+42 20	2.3 and 4.8	9.8	K3 and A0
32 Eridani	3 54.3	−2 57	4.8 and 6.1	6.8	G5 and A2
ι Cancri	8 46.7	+28 46	4.2 and 6.6	30.5	G5 and A5
γ Leonis	10 20.0	+19 51	2.2 and 3.5	4.4	K0 and G7
24 Coma Berenicis	12 35.1	+18 23	5.2 and 6.7	20.3	K0 and A3
ν Boötis	14 45.0	+27 04	2.5 and 4.9	2.8	K0 and A0
α Herculis	17 14.6	+14 23	3.5 and 5.4	4.7	M5 and G5
59 Serpentis	18 27.2	+0 12	5.3 and 7.6	3.8	G0 and A6
β Cygni	19 30.7	+27 58	3.1 and 5.1	34.3	K3 and B8
δ Cephei	22 29.2	+58 25	4* and 7.5	20.4	F5 and A0

Note: The right ascensions and declinations are given for epoch 2000.

**The brighter star in the δ Cephei binary system is a variable star of approximately the fourth magnitude.*

date to six months from today, and again set the time to 12:00:00 A.M. Which of the stars that you selected in part (a) are visible? (You can use the **Find** pane to attempt to locate your selected stars.) Which are not? Explain why the passage of six months should make a difference.

STARRY NIGHT

79. Use the *Starry Night Enthusiast™* program to examine the nearby stars. Click on **Favourites > Stars > Local Neighborhood** and **Stop time.** Select **View > Feet** to hide the spacesuit image. Center this view upon the Sun by opening the **Find** pane and double-clicking on Sun. You are now 16.41 light years from the Sun, looking at the labeled nearby stars. **Increase current elevation** to about 70,000 light-years using the button on the toolbar below the Viewing Location box (an upward-pointing triangle) to see these nearby stars within the Milky Way Galaxy. You can rotate the galaxy by placing the mouse cursor over the image and holding down the **Shift** key while holding down the mouse button and moving the mouse. (On a two-button mouse, hold down the left mouse button). **Decrease current elevation** to a distance of about 100 light-years from the Sun to return to the solar neighborhood. Again, you can rotate this swarm of stars by holding down the **Shift** key while holding down the mouse button and moving the mouse. Open the **Info** pane. If you click the mouse while the cursor is over a star, you will see the star's apparent magnitude as seen from Earth in the **Other Data** layer and its distance from the Sun in the **Position in Space** layer of the **Info** pane. (**a**) Select at least 5 stars within 50 light-years of the Sun and note their names, apparent magnitudes, luminosities, and distances from the Sun in a list. Which of these stars would be visible from Earth with the naked eye from a dark location? Which are visible with the naked eye from a brightly lit city? (*Hint:* The naked eye can see stars as faint as apparent magnitude $m = +6$ from a dark location, but only as faint as $m = +4$ from an inner city.) (**b**) **Increase current elevation** once more to about 1000 light-years from Earth and locate at least 5 stars that are further than 500 light-years from the Sun, making a list of these stars, their names, apparent magnitudes, luminosities and distances from the Sun. Which of these stars are visible from Earth with the naked eye from a dark location? Are the naked-eye stars more likely to be giants or supergiants, or are they more likely to be main-sequence stars? Explain your answer.

This page intentionally left blank

18 The Birth of Stars

R I V U X G

A region of star formation about 1400 pc (4000 ly) from Earth in the southern constellation Ara (the Altar). (European Southern Observatory)

The stars that illuminate our nights seem eternal and unchanging. But this permanence is an illusion. Each of the stars visible to the naked eye shines due to thermonuclear reactions and has only a finite amount of fuel available for these reactions. Hence, stars cannot last forever: They form from material in interstellar space, evolve over millions or billions of years, and eventually die. In this chapter our concern is with how stars are born and become part of the main sequence.

Stars form within cold, dark clouds of gas and dust that are scattered abundantly throughout our Galaxy. One such cloud appears as a dark area on the far right-hand side of the photograph at the top of this page. Perhaps a dark cloud like this encounters one of the Galaxy's spiral arms, or perhaps a supernova detonates nearby. From the shock of events like these, the cloud begins to contract under the pull of gravity, forming protostars—the fragments that will one day become stars. As a protostar develops, its internal pressure builds and its temperature rises. In time, hydrogen fusion begins, and a star is born. The hottest, bluest, and brightest young stars, like those in the accompanying image, emit ultraviolet radiation that excites the surrounding interstellar gas. The result is a beautiful glowing nebula, which typically has the red color characteristic of excited hydrogen (as shown in the photograph).

In Chapters 19 and 20 we will see how stars mature and grow old. Some even blow themselves apart in death throes that enrich interstellar space with the material for future generations of stars. Thus, like the mythical phoenix, new stars arise from the ashes of the old.

18-1 Understanding how stars evolve requires observation as well as ideas from physics

Over the past several decades, astronomers have labored to develop an understanding of **stellar evolution,** that is, how stars are born, live their lives, and finally die. Our own Sun provides evidence that stars are not permanent. The energy radiated by the Sun comes from thermonuclear reactions in its core, which consume 6×10^{11} kg of hydrogen each second and convert it into helium (see Section 16-1). While the amount of hydrogen in the Sun's core is vast, it is not infinite; therefore, the Sun cannot always have been shining, nor can it continue to shine forever. The same is true for all other main-sequence stars, which are fundamentally the same kinds of objects as the Sun but with different masses (see Section 17-9). Thus, stars must have a beginning as well as an end.

Stars consume the material of which they are made, and so cannot last forever

Learning Goals

By reading the sections of this chapter, you will learn

18-1 How astronomers have pieced together the story of stellar evolution

18-2 What interstellar nebulae are and what they are made of

18-3 What happens as a star begins to form

18-4 The stages of growth from young protostars to main-sequence stars

18-5 How stars gain and lose mass during their growth

18-6 What insights star clusters add to our understanding of stellar evolution

18-7 Where new stars form within galaxies

18-8 How the death of old stars can trigger the birth of new stars

Stars last very much longer than the lifetime of any astronomer—indeed, far longer than the entire history of human civilization. Thus, it is impossible to watch a single star go through its formation, evolution, and eventual demise. Rather, astronomers have to piece together the evolutionary history of stars by studying different stars at different stages in their life cycles.

ANALOGY To see the magnitude of this task, imagine that you are a biologist from another planet who sets out to understand the life cycles of human beings. You send a spacecraft to fly above the Earth and photograph humans in action. Unfortunately, the spacecraft fails after collecting only 20 seconds of data, but during that time its sophisticated equipment sends back observations of thousands of different humans. From this brief snapshot of life on the Earth—only 10^{-8} (a hundred-millionth) of a typical human lifetime—how would you decide which were the young humans and which were the older ones? Without a look inside our bodies to see the biological processes that shape our lives, could you tell how humans are born and how they die? And how could you deduce the various biological changes that humans undergo as they age?

Astronomers, too, have data spanning only a tiny fraction of any star's lifetime. A star like the Sun can last for about 10^{10} years, whereas astronomers have been observing stars in detail for only about a century—as in our analogy, roughly 10^{-8} of the life span of a typical star. Astronomers are also frustrated by being unable to see the interiors of stars. For example, we cannot see the thermonuclear reactions that convert hydrogen into helium. But astronomers have an advantage over the biologist in our story. Unlike humans, stars are made of relatively simple substances, primarily hydrogen and helium, that are found almost exclusively in the form of gases. Of the three phases of matter—gas, liquid, and solid—gases are by far the simplest to understand.

Astronomers use our understanding of gases to build theoretical models of the interiors of stars, like the model of the Sun we saw in Section 16-2. Models help to complete the story of stellar evolution. In fact, like all great dramas, the story of stellar evolution can be regarded as a struggle between two opposing and unyielding forces: Gravity continually tries to make a star shrink, while the star's internal pressure tends to make the star expand. When these two opposing forces are in balance, the star is in a state of hydrostatic equilibrium (see Figure 16-2).

But what happens when changes within the star cause either pressure or gravity to predominate? The star must then either expand or contract until it reaches a new equilibrium. In the process, it will change not only in size but also in luminosity and color.

In the following chapters, we will find that giant and supergiant stars are the result of pressure gaining the upper hand over gravity. Both giants and supergiants turn out to be aging stars that have become tremendously luminous and ballooned to hundreds or thousands of times their previous size. White dwarfs, by contrast, are the result of the balance tipping in gravity's favor. These are even older stars that have collapsed to a fraction of the size they had while on the main sequence. In this chapter, however, we will see how the combat between gravity and pressure explains the birth of stars. We start our journey within the diffuse clouds of gas and dust that permeate our galaxy.

18-2 Interstellar gas and dust pervade the galaxy

Different types of nebulae emit, absorb, or reflect light

Where do stars come from? As we saw in Section 8-4, our Sun condensed from a solar nebula, a collection of gas and dust in interstellar space. Observations suggest that other stars originate in a similar way (see Figure 8-8). To understand the formation of stars, we must first understand the nature of the interstellar matter from which the stars form.

Nebulae and the Interstellar Medium

At first glance, the space between the stars seems to be empty. On closer inspection, we find that it is filled with a thin gas laced with microscopic dust particles. This combination of gas and dust is called the **interstellar medium.** Evidence for this medium includes interstellar clouds of various types, curious lines in the spectra of binary star systems, and an apparent dimming and reddening of distant stars.

You can see evidence for the interstellar medium with the naked eye. Look carefully at the constellation Orion (Figure 18-1*a*), visible on winter nights in the northern hemisphere and summer nights in the southern hemisphere. While most of the stars in the constellation appear as sharply defined points of light, the middle "star" in Orion's sword has a fuzzy appearance. This becomes more obvious with binoculars or a telescope. As Figure 18-1*b* shows, this "star" is actually not a star at all, but the Orion Nebula—a cloud in interstellar space. Any interstellar cloud is called a **nebula** (plural **nebulae**) or **nebulosity.**

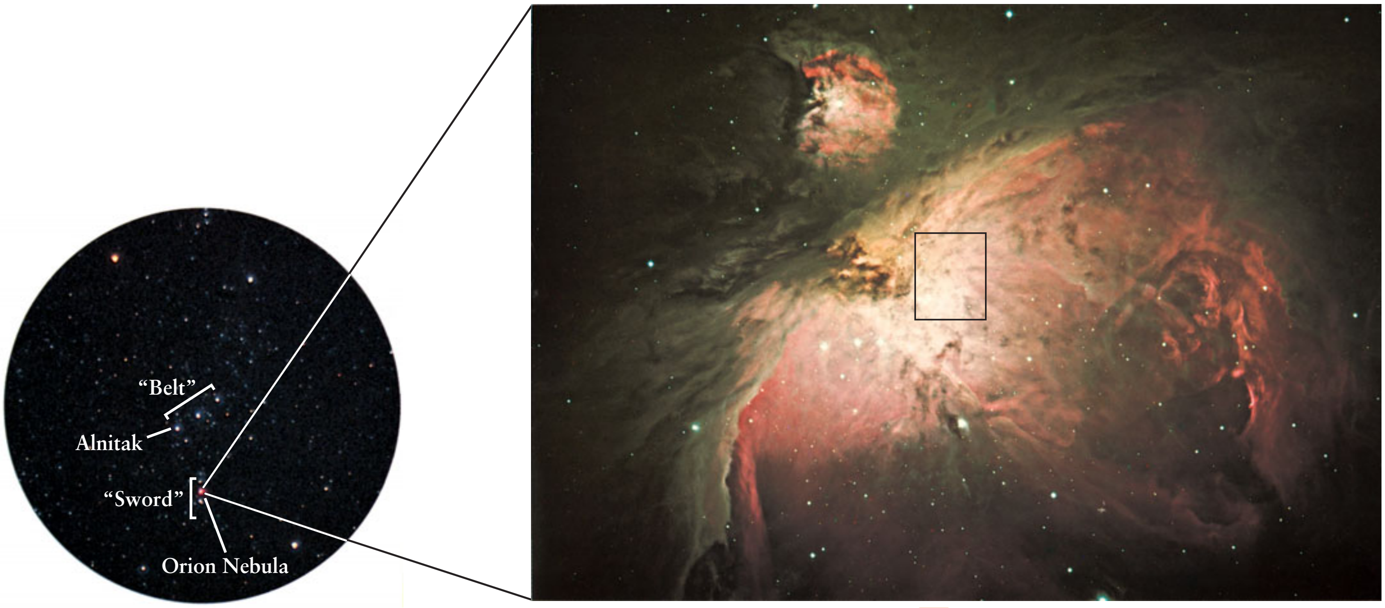
Emission Nebulae: Clouds of Excited Gas

The Orion Nebula emits its own light, with the characteristic emission line spectrum of a hot, thin gas. For this reason it is called an **emission nebula.** Many emission nebulae can be seen with a small telescope. Figure 18-2 shows some of these nebulae in a different part of the constellation Orion. Emission nebulae are direct evidence of gas atoms in the interstellar medium.

Typical emission nebulae have masses that range from about 100 to about 10,000 solar masses. Because this mass is spread over a huge volume that is light-years across, the density is quite low by Earth standards, only a few thousand hydrogen atoms per cubic centimeter. (By comparison, the air you are breathing contains more than 10^{19} atoms per cm^3.)

Emission nebulae are found near hot, luminous stars of spectral types O and B. Such stars emit copious amounts of ultraviolet radiation. When atoms in the nearby interstellar gas absorb these energetic ultraviolet photons, the atoms become ionized. Indeed, emission nebulae are composed primarily of ionized hydrogen atoms, that is, free protons (hydrogen nuclei) and electrons. Astronomers use the notation H I for neutral, un-ionized hydrogen atoms and H II for ionized hydrogen atoms, which is why emission nebulae are also called **H II regions.**

H II regions emit visible light when some of the free protons and electrons get back together to form hydrogen atoms, a process



(a) A wide-angle view of Orion (b) A closeup of the Orion Nebula R I V U X G

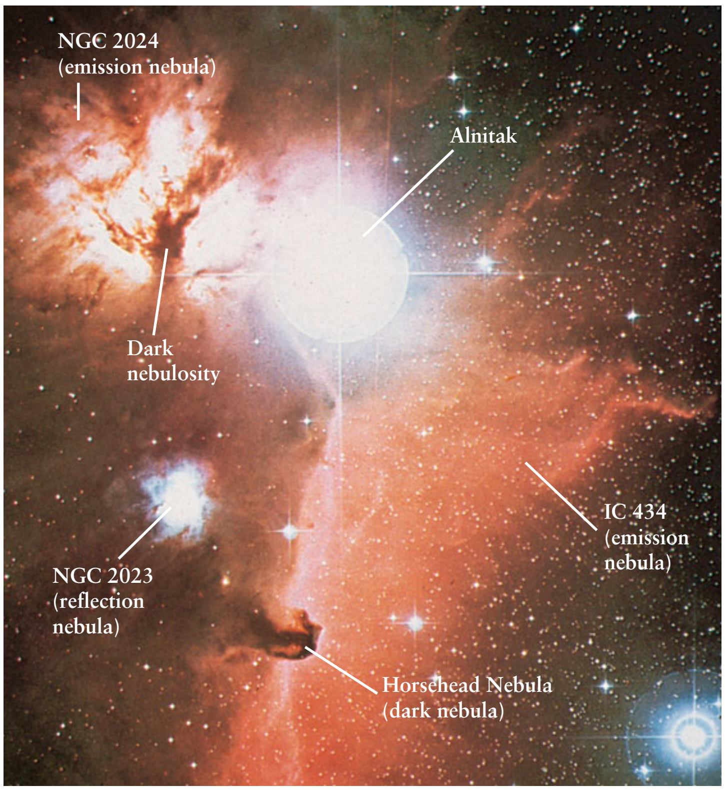
WEB LINK 18.2

Figure 18-1

The Orion Nebula (a) The middle "star" of the three that make up Orion's sword is actually an interstellar cloud called the Orion Nebula. **(b)** The nebula is about 450 pc (1500 ly) from Earth and contains about 300 solar masses of material. Within the area shown by the box are four hot, massive stars called the Trapezium. They produce the ultraviolet light that makes the nebula glow. (a: R. C. Mitchell, Central Washington University; b: Anglo-Australian Observatory)

called **recombination** (Figure 18-3). When an atom forms by recombination, the electron is typically captured into a high-energy orbit. As the electron cascades downward through the atom's energy levels toward the ground state, the atom emits photons with lower energy and longer wavelength than the photons that originally caused the ionization. Particularly important is the transition from $n = 3$ to $n = 2$. It produces H_α photons with a wavelength of 656 nm, in the red portion of the visible spectrum (see Section 5-8, especially Figure 5-23*b*). These photons give H II regions their distinctive reddish color.

For each high-energy, ultraviolet photon absorbed by a hydrogen atom to ionize it, several photons of lower energy are emitted when a proton and electron recombine. As Box 18-1 describes, a similar effect takes place in a fluorescent lightbulb. In this sense, H II regions are immense fluorescent light fixtures!



WEB LINK 18.3

Figure 18-2 R I V U X G

Emission, Reflection, and Dark Nebulae in Orion A variety of different nebulae appear in the sky around Alnitak, the easternmost star in Orion's belt (see Figure 20-1a). All the nebulae lie approximately 500 pc (1600 ly) from Earth. They are actually nowhere near Alnitak, which is only 250 pc (820 ly) distant. This photograph shows an area of the sky about 1.5° across. (Royal Observatory, Edinburgh)

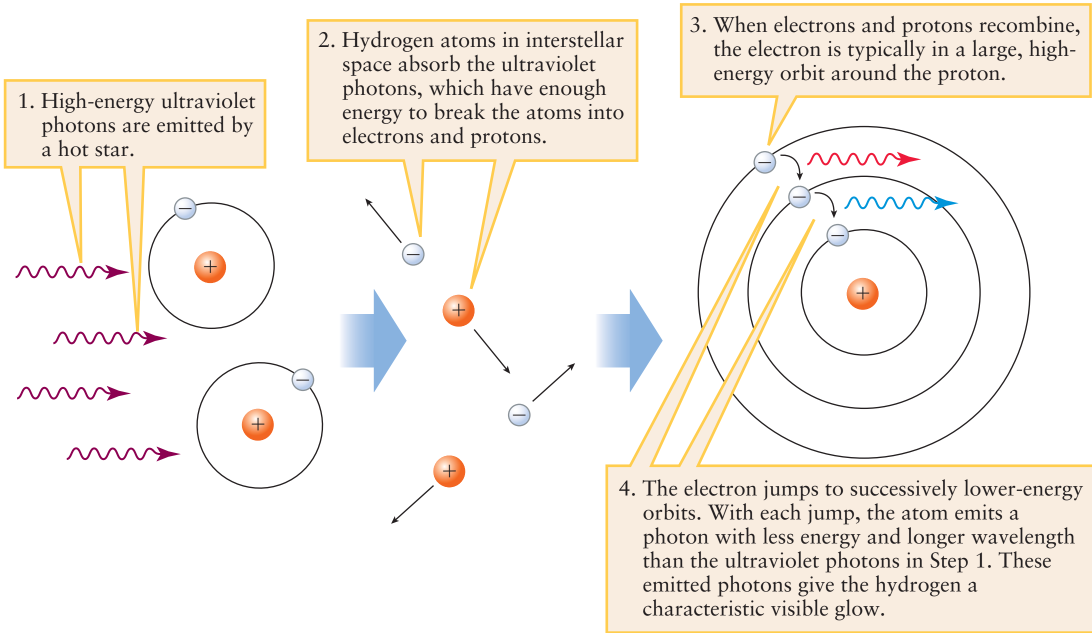


Figure 18-3

Ionization and Recombination The characteristic glow of emission nebulae (like those shown in Figure 18-1 and Figure 18-2) comes from gas atoms that are excited by ultraviolet radiation from nearby hot stars.

Further evidence of interstellar gas comes from the spectra of binary star systems. As the two stars that make up a binary system orbit their common center of mass, their spectral lines shift back and forth (see Figure 17-23). But certain calcium and sodium lines are found to remain at fixed wavelengths. These **stationary absorption lines** are therefore not associated with the binary star. Instead, they must be caused by interstellar gas between us and the binary system.

BOX 18-1 — The Heavens on the Earth

Fluorescent Lights

The light that comes from glowing interstellar clouds is, quite literally, otherworldly. But the same principles that explain how such clouds emit light are also at the heart of light phenomena that we see here on the Earth.

A fluorescent lamp produces light in a manner not too different from an emission nebula (H II region). In both cases, the physical effect is called **fluorescence:** High-energy ultraviolet photons are absorbed, and the absorbed energy is reradiated as lower-energy photons of visible light.

Within the glass tube of a fluorescent lamp is a small amount of the element mercury. When you turn on the lamp, an electric current passes through the tube, vaporizing the mercury and exciting its atoms. This excited mercury vapor radiates light with an emission-line spectrum, including lines in the ultraviolet. The white fluorescent coating on the inside of the glass tube absorbs these ultraviolet photons, exciting the coating's molecules to high energy levels.

The molecules then cascade down through a number of lower levels before reaching the ground state. During this cascade, visible-light photons of many different wavelengths are emitted, giving an essentially continuous spectrum and a very white light. (By comparison, the hydrogen atoms in an H II region emit at only certain discrete wavelengths, because the spectrum of hydrogen is much simpler than that of the fluorescent coating's molecules. Another difference is that the molecules in the fluorescent tube never become ionized.)

Many common materials display fluorescence. Among them are teeth, fingernails, and certain minerals. When illuminated with ultraviolet light, these materials glow with a blue or green color. (Most natural history museums and science museums have an exhibit showing fluorescent minerals.) Laundry detergent also contains fluorescent material. After washing, your laundry absorbs ultraviolet light from the Sun and glows faintly, making it appear "whiter than white."

Dark Nebulae and Reflection Nebulae: Abodes of Dust

In addition to the presence of gas atoms in H II regions, Figure 18-2 also shows two kinds of evidence for larger bits of matter, called **dust grains,** in the interstellar medium. These dust grains make their appearance in *dark nebulae* and *reflection nebulae*.

A **dark nebula** is so opaque that it blocks any visible light coming from stars that lie behind it. One such dark nebula, called the Horsehead Nebula for its shape, protrudes in front of one of the H II regions in Figure 18-2. Figure 18-4 shows another dark nebula, called Barnard 86. These nebulae have a relatively dense concentration of microscopic dust grains, which scatter and absorb light much more efficiently than single atoms. Typical dark nebulae have temperatures between 10 K to 100 K, which is low enough for hydrogen atoms to form molecules. Such nebulae contain from 10^4 to 10^9 particles (atoms, molecules, and dust grains) per cubic centimeter. Although tenuous by Earth standards, dark nebulae are large enough—typically many light-years deep—that they block the passage of light. In the same way, a sufficient depth of haze or smoke in our atmosphere can make it impossible to see distant mountains.

The other evidence for dust in Figure 18-2 is the bluish haze surrounding the star immediately above and to the left of the Horsehead Nebula. Figure 18-5 shows a similar haze around a different set of stars. A haze of this kind, called a **reflection nebula,** is caused by fine grains of dust in a lower concentration than that found in dark nebulae. The light we see coming from the nebula is starlight that has been scattered and reflected by these dust grains. The grains are only about 500 nm across, no larger than a typical wavelength of visible light, and they scatter short-wavelength blue light more efficiently than long-wavelength red light. Hence, reflection nebulae have a characteristic blue color. Box 5-4 explains how a similar process—the scattering of sunlight in our atmosphere—gives rise to the blue color of the sky.

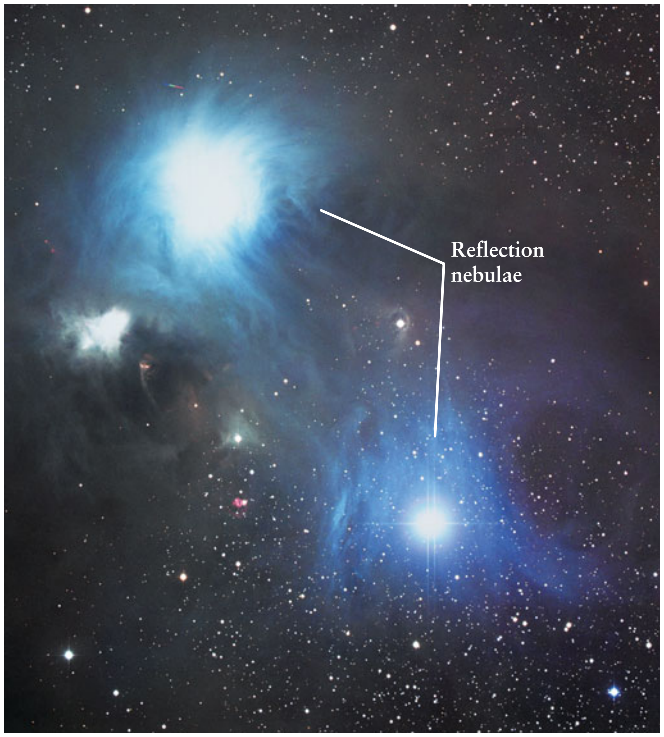


Figure 18-5 R I V U X G

Reflection Nebulae Wispy reflection nebulae called NGC 6726-27-29 surround several stars in the constellation Corona Australis (the Southern Crown). Unlike emission nebulae, reflection nebulae do not emit their own light, but scatter and reflect light from the stars that they surround. This scattered starlight is quite blue in color. The region shown here is about 23 arcminutes across. (David Malin/Anglo-Australian Observatory)

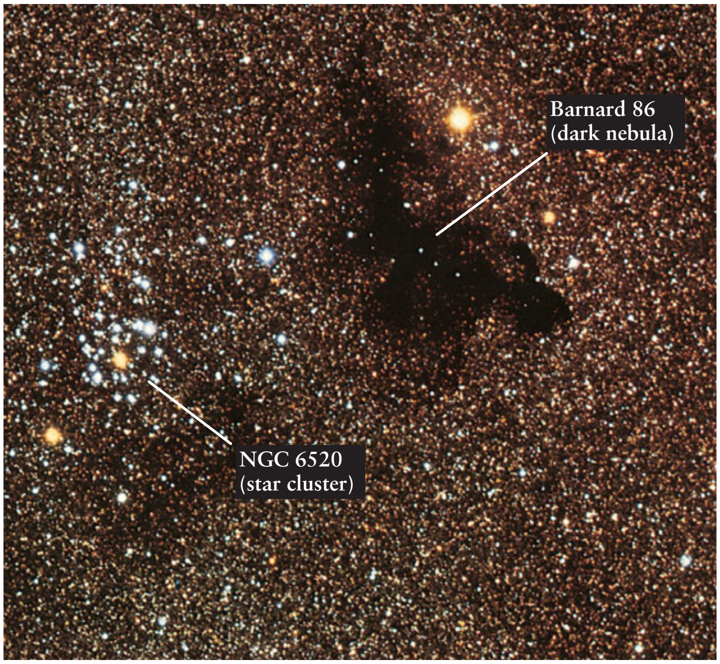
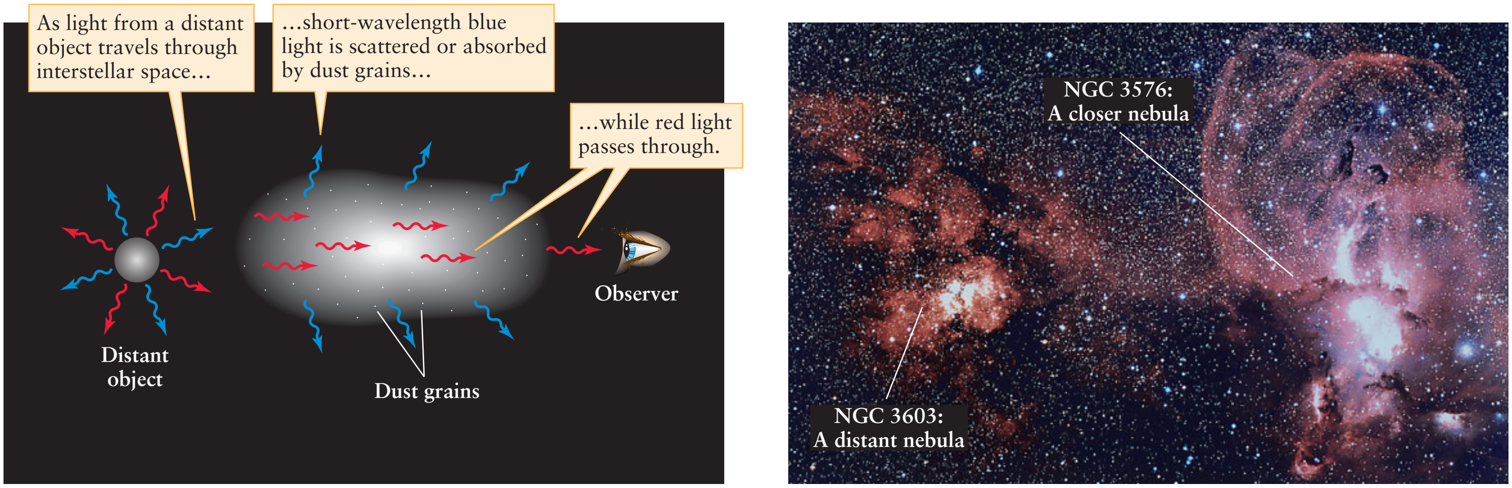


Figure 18-4 R I V U X G

A Dark Nebula When first discovered in the late 1700s, dark nebulae were thought to be "holes in the heavens" where very few stars are present. In fact, they are opaque regions that block out light from the stars beyond them. The few stars that appear to be within Barnard 86 lie between us and the nebula. Barnard 86 is in the constellation Sagittarius and has an angular diameter of 4 arcminutes, about 1/7 the angular diameter of the full moon. (Anglo-Australian Observatory)

Interstellar Extinction and Reddening

In the 1930s, the American astronomer Robert Trumpler discovered two other convincing pieces of evidence for the existence of interstellar matter—*interstellar extinction* and *interstellar reddening*. While studying the brightness and distances of certain star clusters, Trumpler noticed that remote clusters seem to be dimmer than would be expected from their distance alone. His observations demonstrated that the intensity of light from remote stars is reduced as the light passes through material in interstellar space. This process is called **interstellar extinction.** (In the same way, the headlights of oncoming cars appear dimmer when there is smoke or fog in the air.) The light from remote stars is also reddened as it passes through the interstellar medium, because the blue



(a) How dust causes interstellar reddening

(b) Reddening depends on distance R I V U X G

Figure 18-6

Interstellar Reddening **(a)** Dust grains in interstellar space scatter or absorb blue light more than red light. Thus light from a distant object appears redder than it really is. **(b)** The emission nebulae NGC 3603 and NGC 3576 are different distances from Earth. Light from the more distant nebula must pass through more interstellar dust to reach us, so more interstellar reddening occurs and NGC 3603 is a ruddier shade of red. The two nebulae are about 1° apart in the sky. (Anglo-Australian Observatory)

component of their starlight is scattered and absorbed by interstellar dust. This effect, shown in Figure 18-6, is called **interstellar reddening.** The same effect makes the setting Sun look red (see Box 5-4).

CAUTION! It is important to understand the distinction between interstellar reddening and the Doppler effect. If an object is moving away from us, the Doppler effect *shifts* all of that object's light toward longer wavelengths (a redshift). Interstellar reddening, by contrast, makes objects appear red not by shifting wavelengths but by filtering out short wavelengths. The effect is the same as if we looked at an object through red-colored glasses, which let red light pass but block out blue light. The one similarity between Doppler shifts and interstellar reddening is that neither has any discernible effect on what you see with the naked eye. Both of these effects cause only subtle color changes that your eye (which does a poor job of seeing colors in faint objects like stars) cannot detect. If you see a star with a red color, the reason is *not* interstellar reddening or the Doppler effect; it is because the star really is red due to its low surface temperature (see Figure 17-7).

Interstellar Gas and Dust in Spiral Galaxies

Observations indicate that interstellar gas and dust are largely confined to the plane of the Milky Way—that faint, hazy band of stars you can see stretching across the sky on a dark, moonless night. Figure 18-7 shows glowing emission nebulae and dark lanes of dust along the Milky Way.

As we will see in Chapter 23, the band of the Milky Way is actually our inside view of our Galaxy, which is a flat, disk-shaped collection of several hundred billion stars about 50,000 pc (160,000 ly) in diameter. The Sun is located about 8000 pc (26,000 ly) from our Galaxy's center. Astronomers know these dimensions because they can map our Galaxy from the locations of bright stars and nebulae and also by using radio telescopes. Observations indicate that bright stars, gas, and glowing nebulae are mostly located within a few hundred parsecs of the midplane of our Galaxy, and that they are concentrated along arching spiral arms that wind outward from the Galaxy's center. If we could view our Galaxy from a great distance, it would look somewhat like one of the galaxies shown in Figure 18-8.

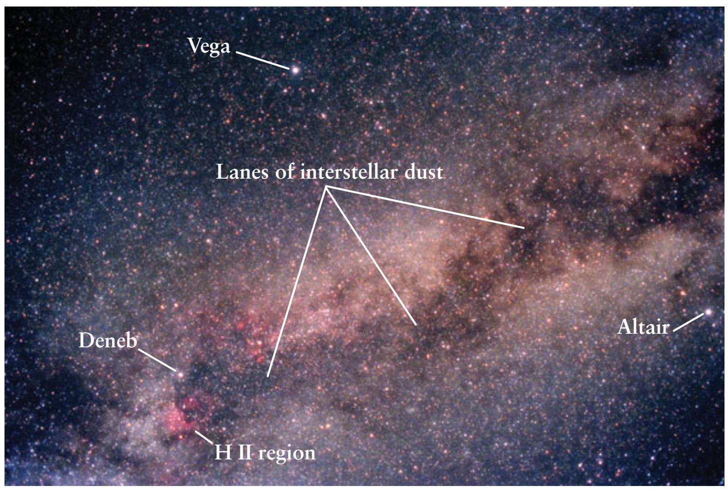
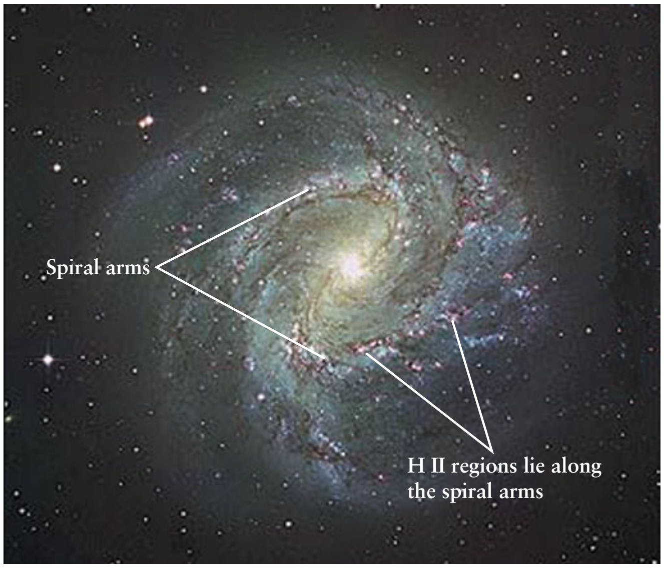
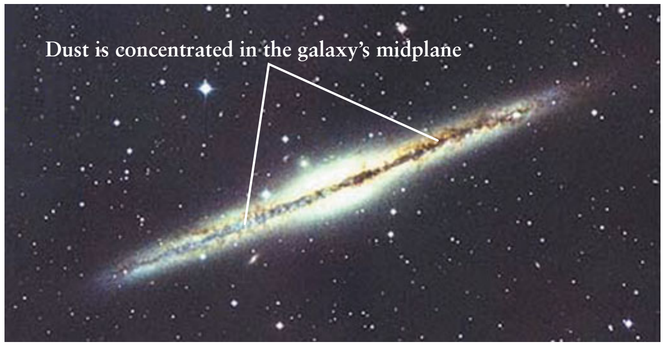


Figure 18-7 R I V U X G

Gas and Dust in the Milky Way Glowing gas clouds (emission nebulae or H II regions) and dark, dusty regions are concentrated close to the midplane of the Milky Way Galaxy, of which our Sun is part. This wide-angle photograph also shows the three bright stars that make up the "summer triangle" (see Figure 2-8). (© Jerry Lodriguss/Photo Researchers).



(a) We see spiral galaxy M83 nearly face-on



(b) We see spiral galaxy NGC 891 nearly edge-on

WEB LINK 18.4

Figure 18-8 R I V U X G

Spiral Galaxies Spiral galaxies, like our own Milky Way Galaxy, consist of stars, gas, and dust that are largely confined to a flattened, rotating disk. **(a)** This face-on view of M83 shows luminous stars and H II regions along the spiral arms. **(b)** This edge-on view of NGC 891 shows a dark band caused by dust in this galaxy's interstellar medium. Although in different parts of the sky, both galaxies are about 7 million pc (23 million ly) from Earth and have angular diameters of about 13 arcminutes. (a: David Malin/Anglo-Australian Observatory; b: Instituto de Astrofísica de Canarias/Royal Greenwich Observatory/David Malin)

Interstellar gas and dust are the raw material from which stars are made. The disk of our Galaxy, where most of this matter is concentrated, is therefore the site of ongoing star formation. Our next goal is to examine the steps by which stars are formed.

18-3 Protostars form in cold, dark nebulae

How do stars form in the interstellar medium? In this section, we discuss how a contracting cloud gives birth to a clump called a *protostar*—a future main-sequence star.

Protostars: Initial Gravitational Collapse

In order for interstellar material to condense and form a star, the force of gravity—which tends to draw interstellar material together—must overwhelm the internal pressure pushing the material apart. This means that stars will most easily form in regions where the interstellar material is relatively dense, so that atoms and dust grains are close together and gravitational attraction is enhanced.

To assist star formation, the pressure of the interstellar medium should be relatively low. This means that the interstellar medium should be as cold as possible, because the pressure of a gas goes down as the gas temperature decreases. (Conversely, increasing the temperature makes the pressure increase. That's why the air pressure inside automobile tires is higher after the auto has been driven a while and the tires have warmed up.)

The only parts of the interstellar medium with high enough density and low enough temperature for stars to form are the dark nebulae. Many of these, such as the Horsehead Nebula (see Figure 18-2) and Barnard 86 (see Figure 18-4), were discovered and catalogued around 1900 by Edward Barnard and are known as **Barnard objects.** (The Horsehead Nebula is also known as Barnard 33.) Other relatively small, nearly spherical dark nebulae are known as **Bok globules,** after the Dutch-American astronomer Bart Bok, who first called attention to them during the 1940s (Figure 18-9). A Bok globule resembles the inner core of a Barnard object with the outer, less dense portions stripped away. The density of the gas and dust within a Barnard object or Bok globule is indeed quite high by cosmic standards, in the range from 100 to 10,000 particles per cm^3; by comparison, most of the interstellar medium contains only 0.1 to 20 particles per cm^3. Radio emissions from molecules within these clouds indicate that their internal temperatures are very low, only about 10 K.

A typical Barnard object contains a few thousand solar masses of gas and dust spread over a volume roughly 10 pc (30 ly) across; a typical Bok globule is about one-tenth as large. The chemical composition of this material is the standard "cosmic abundance" of about 74% (by mass) hydrogen, 25% helium, and 1% heavier elements (review Figure 8-4). Within these clouds, the densest portions can contract under their own mutual gravitational attraction and form clumps called **protostars.** Each protostar will eventually evolve into a main-sequence star. Because dark nebulae contain many solar masses of material, it is possible for a large number of protostars to form

In the first stage of star formation, a protostar coalesces and contracts due to the mutual gravitational attraction of its parts

Figure 18-9 R I V U X G

Bok Globules The dark blobs in this photograph of a glowing H II region are clouds of dust and gas called Bok globules. A typical Bok globule is a parsec or less in size and contains from one to a thousand solar masses of material. The Bok globules and H II region in this image are part of a much larger star-forming region called NGC 281, which lies about 9500 ly (2900 pc) from Earth in the constellation Cassiopeia. The image shows an area about 8.8 ly (2.7 pc) across. (NASA, ESA, and The Hubble Heritage Team (STScI/AURA))

out of a single such nebula. Thus, we can think of dark nebulae as "stellar nurseries."

The Evolution of a Protostar

In the 1950s, astrophysicists such as Louis Henyey in the United States and C. Hayashi in Japan performed calculations that enabled them to describe the earliest stages of a protostar. At first, a protostar is merely a cool blob of gas several times larger than our solar system. The pressure inside the protostar is too low to support all this cool gas against the mutual gravitational attraction of its parts, and so the protostar collapses. As the protostar collapses, gravitational energy is converted into thermal energy, making the gases heat up and start glowing. (We discussed this process, called Kelvin-Helmholtz contraction, in Section 8-4 and Section 16-1.) Energy from the interior of the protostar is transported outward by convection, warming its surface.

After only a few thousand years of gravitational contraction, the protostar's surface temperature reaches 2000 to 3000 K. At this point the protostar is still quite large, so its glowing gases produce substantial luminosity. (The greater a star's radius and surface temperature, the greater its luminosity. You can review the details in Section 17-6 and Box 17-4.) For example, after only 1000 years of contraction, a protostar of 1 solar mass (1 $M_\odot$) is 20 times larger in radius than the Sun and has 100 times the Sun's luminosity. Unlike the Sun, however, the luminosity of a young protostar is not the result of thermonuclear fusion, because the protostar's core is not yet hot enough for fusion reactions to begin. Instead, the radiated energy comes exclusively from the heating of the protostar as it contracts.

In order to determine the conditions inside a contracting protostar, astrophysicists use computers to solve equations similar to those used for calculating the structure of the Sun (see Section 18-2). The results tell how the protostar's luminosity and surface temperature change at various stages in its contraction. This information, when plotted on a Hertzsprung-Russell diagram, provides a protostar's **evolutionary track.** The track shows us how the protostar's appearance changes because of changes in its interior. These theoretical tracks agree quite well with actual observations of protostars.

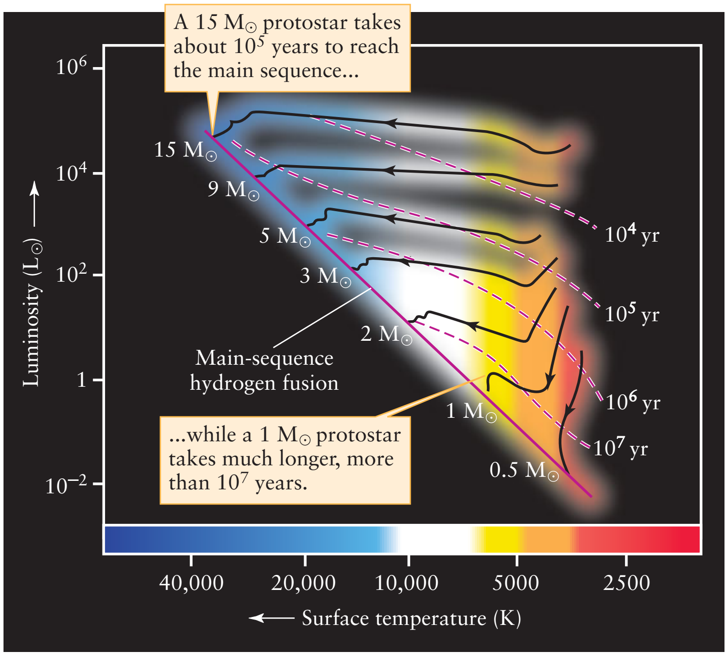
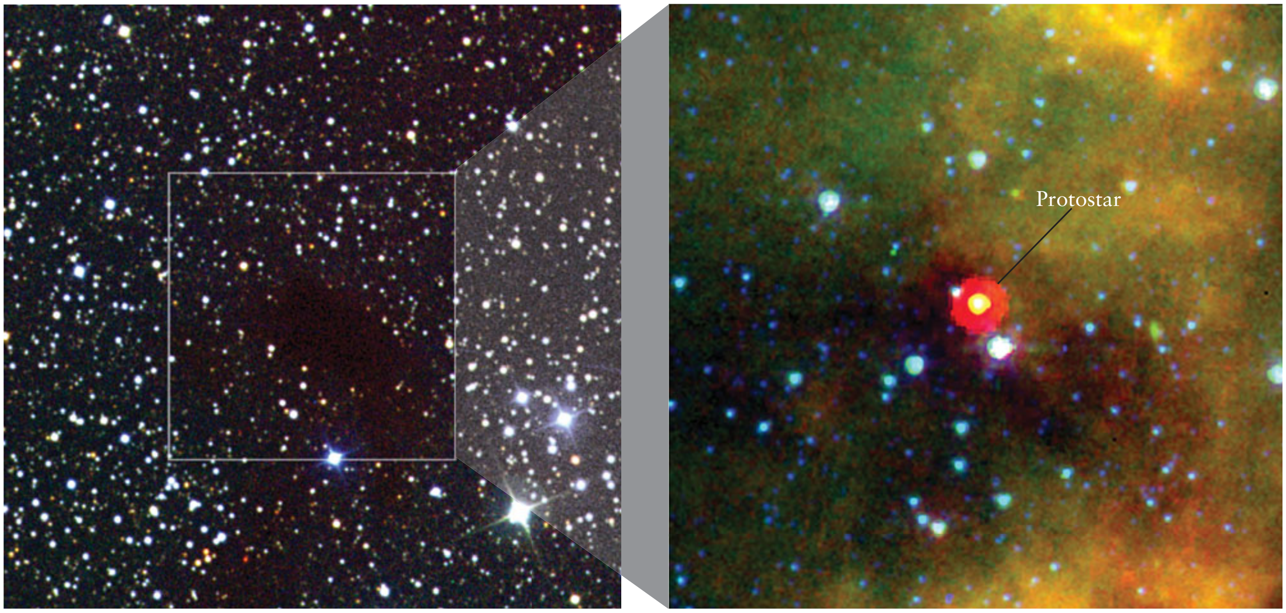


Figure 18-10

Pre-Main-Sequence Evolutionary Tracks As a protostar evolves, its luminosity and surface temperature both change. The tracks shown here depict these changes for protostars of seven different masses. Each dashed red line shows the age of a protostar when its evolutionary track crosses that line. (We will see in Section 18-5 that protostars lose quite a bit of mass as they evolve: The mass shown for each track is the value when the protostar finally settles down as a main-sequence star.)



(a) A dark nebula R I V U X G

(b) A hidden protostar within the dark nebula R I V U X G

Figure 18-11

Revealing a Hidden Protostar **(a)** This visible-light view shows a dark nebula called L1014 in the constellation Cygnus (the Swan). No stars are visible within the nebula. **(b)** The Spitzer Space Telescope was used to make this false-color infrared image of the outlined area in (a). The bright red-yellow spot is a protostar within the dark nebula. (a: Deep Sky Survey; b: NASA/JPL-Caltech/N. Evans, Univ. of Texas at Austin)

CAUTION! When astronomers refer to a star "following an evolutionary track" or "moving on the H-R diagram," they mean that the star's luminosity, surface temperature, or both change. Hence, the point that represents the star on an H-R diagram changes its position. This is completely unrelated to how the star physically moves through space!

Figure 18-10 shows evolutionary tracks for protostars of seven different masses, ranging from 0.5 $M_\odot$ to 15 $M_\odot$. Because protostars are relatively cool when they begin to shine at visible wavelengths, these tracks all begin near the right (low-temperature) side of the H-R diagram. The subsequent evolution is somewhat different, however, depending on the protostar's mass. Note that the greater a star's mass, the more rapidly it moves along its evolutionary track: a 15-$M_\odot$ protostar takes only about 10^5 years to reach the main sequence, while a 1-$M_\odot$ protostar takes over a hundred times longer (more than 10^7 years).

Observing Protostars

Observing the evolution of a protostar can be quite a challenge. Indeed, while Figure 18-10 shows that young protostars are much more luminous than the Sun, it is quite unlikely that you have ever seen a protostar shining in the night sky. The reason is that protostars form within clouds that contain substantial amounts of interstellar dust. The dust in a protostar's immediate surroundings, called its **cocoon nebula,** absorbs the vast amounts of visible light emitted by the protostar and makes it very hard to detect using visible wavelengths.

Protostars can be seen, however, using infrared wavelengths. Because it absorbs so much energy from the protostar that it surrounds, a dusty cocoon nebula becomes heated to a few hundred kelvins. The warmed dust then reradiates its thermal energy at infrared wavelengths, to which the cocoon nebula is relatively transparent. So, by using infrared telescopes, astronomers can see protostars within the "stellar nursery" of a dark nebula.

Figure 18-11 shows visible-light and infrared views of one such stellar nursery. The visible-light view (Figure 18-11*a*) shows a dark, dusty nebula that appears completely opaque. The infrared image (Figure 18-11*b*) allows us to see through the dust, revealing a newly formed protostar within the dark nebula. The properties of this protostar agree well with the theoretical ideas outlined in this section.

18-4 Protostars evolve into main-sequence stars

Evolutionary tracks show how a protostar matures into a star as its gases contract. The details of this evolution depend on the star's mass. To follow these details, you should keep in mind the basic principle that a star's luminosity, radius, and surface temperature are intimately related: luminosity is proportional to the square of radius and to the fourth power of surface temperature (see Section 17-6).

The course of a protostar's evolution depends on its mass

A One-Solar-Mass Protostar

For a protostar with the same mass as our Sun (1 $M_\odot$), the outer layers are cool and quite opaque (for the same reasons that the Sun's photosphere is opaque; see Section 16-5). This means that energy released from the shrinking inner layers in the form of radiation cannot reach the surface. Instead, energy flows outward by the slower and less effective method of convection. Hence, the surface temperature of the contracting protostar stays roughly constant, the luminosity decreases as the radius decreases, and the evolutionary track moves downward on the H-R diagram in Figure 18-10.

Although its surface temperature changes relatively little, the internal temperature of the shrinking protostar increases. After a time, the interior becomes ionized, which makes it less opaque. Energy is then conveyed outward by radiation in the interior and by convection in the opaque outer layers, just as in the present-day Sun (see Section 16-2, especially Figure 16-4). This makes it easier for energy to escape from the protostar, so the luminosity—the rate at which energy is emitted from the protostar's surface—increases. As a result, the evolutionary track in Figure 18-10 bends upward (higher luminosity) and to the left (higher surface temperature, caused by the increased energy flow).

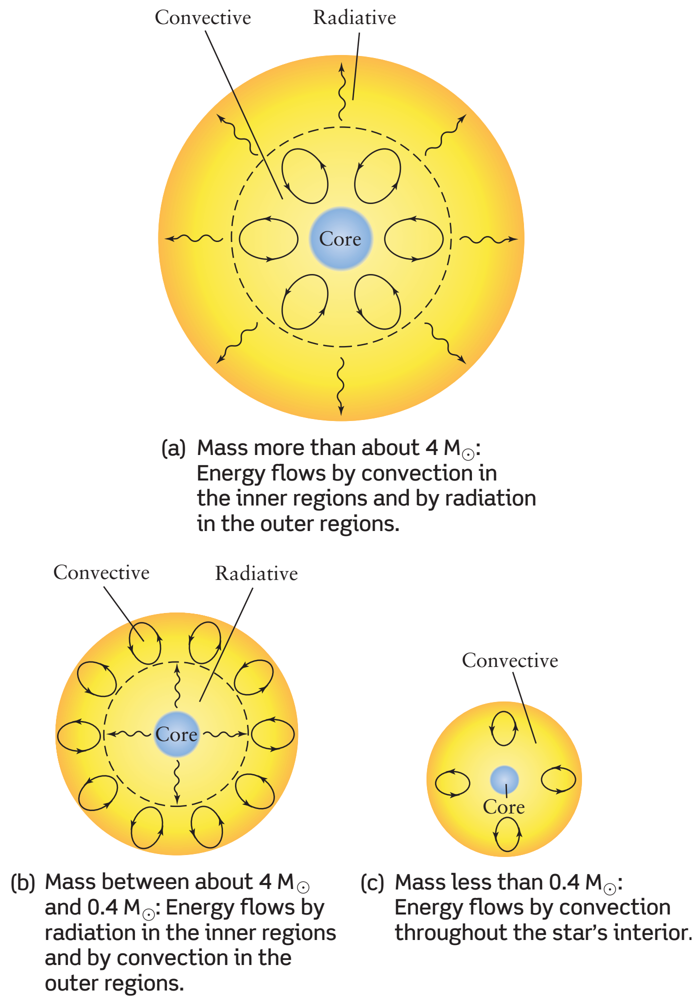
In time, the 1-$M_\odot$ protostar's interior temperature reaches a few million kelvins, hot enough for thermonuclear reactions to begin converting hydrogen into helium. As we saw in Section 16-1 and Box 16-1, these reactions release enormous amounts of energy. Eventually, these reactions provide enough heat and internal pressure to stop the star's gravitational contraction, and hydrostatic equilibrium is reached. The protostar's evolutionary track has now led it to the main sequence, and the protostar has become a full-fledged main-sequence star.

High-Mass and Low-Mass Protostars

More massive protostars evolve a bit differently. If its mass is more than about 4 $M_\odot$, a protostar contracts and heats more rapidly, and hydrogen fusion begins quite early. As a result, the luminosity quickly stabilizes at nearly its final value, while the surface temperature continues to increase as the star shrinks. Thus, the evolutionary tracks of massive protostars traverse the H-R diagram roughly horizontally (signifying approximately constant luminosity) in the direction from right to left (from low to high surface temperature). You can see this most easily for the 9-$M_\odot$ and 15-$M_\odot$ evolutionary tracks in Figure 18-10.

Greater mass means greater pressure and temperature in the interior, which means that a massive star has an even larger temperature difference between its core and its outer layers than the Sun. This causes convection deep in the interior of a massive star (Figure 18-12). By contrast, a massive star's outer layers are of such low density that energy flows through them more easily by radiation than by convection. Therefore, main-sequence stars with masses more than about 4 $M_\odot$ have convective interiors but radiative outer layers (Figure 18-12*a*). By contrast, less massive main-sequence stars such as the Sun have radiative interiors and convective outer layers (Figure 18-12*b*).

The internal structure is also different for main-sequence stars of very low mass. When such a star forms from a protostar,



INTERACTIVE EXERCISE 18.1

Figure 18-12

Main-Sequence Stars of Different Masses Stellar models show that when a protostar evolves into a main-sequence star, its internal structure depends on its mass. *Note:* The three stars shown here are *not* drawn to scale. Compared with a 1-$M_\odot$ main-sequence star like that shown in (b), a 6-$M_\odot$ main-sequence star like that in (a) has more than 4 times the radius, and a 0.2-$M_\odot$ main-sequence star like that in (c) has only one-third the radius.

the interior temperature is never high enough to fully ionize the interior. The interior remains too opaque for radiation to flow efficiently, so energy is transported by convection throughout the volume of the star (Figure 18-12*c*).

Arriving on the Main Sequence

All of the protostar evolutionary tracks in Figure 18-10 end on the main sequence. The main sequence therefore represents stars in which thermonuclear reactions are converting hydrogen into helium. For most stars, this is a stable situation. For example, our Sun will remain on or very near the main sequence, quietly fusing hydrogen into helium at its core, for a total of some 10^{10} years. The point along the main sequence where each evolutionary track ends depends on the star's mass. The most massive stars are the most luminous and their evolutionary tracks end at the upper left of the main sequence, while the least massive stars are the least luminous and their evolutionary tracks end at the lower right of the main sequence. The connection between a main-sequence star's mass and luminosity should be familiar: this is just the mass-luminosity relation (recall the *Cosmic Connections* figure in Section 17-9).

The theory of how protostars evolve helps explain why the main sequence has both an upper mass limit and a lower mass limit. As we saw in Section 17-5, protostars less massive than about 0.08 $M_\odot$ can never develop the necessary pressure and temperature to start hydrogen fusion in their cores. Instead, such "failed stars" end up as *brown dwarfs,* which shine faintly by Kelvin-Helmholtz contraction (see Figure 17-13).

Protostars with masses greater than about 200 solar masses also do not become main-sequence stars. Such a protostar rapidly becomes very luminous, resulting in tremendous internal pressures. This pressure is so great that it overwhelms the effects of gravity, expelling the outer layers into space and disrupting the star. Main-sequence stars therefore have masses between about 0.08 and 200 $M_\odot$, although the high-mass stars are extremely rare.

CAUTION! Two words of caution are in order here. First, while the evolutionary tracks of protostars begin in the red giant region of the H-R diagram (the upper right), protostars are *not* red giants. As we will see in Chapter 19, red giant stars represent a stage in the evolution of stars that comes *after* being a main-sequence star. Second, it is worth remembering that stars live out most of their lives on the main sequence, after only a relatively brief period as protostars. A 15-$M_\odot$ protostar takes only 20,000 years to become a main-sequence star, and a 1-$M_\odot$ protostar takes about 2×10^7 years. By contrast, the Sun has been a main-sequence star for about 4.56×10^9 years. By astronomical standards, pre–main-sequence stars are quite transitory.

18-5 During the birth process, stars both gain and lose mass

After reading the previous section, you may think that a main-sequence star forms simply by collapsing inward. In fact, much of the material of a cold, dark nebula is ejected into space and never incorporated into stars. As it is ejected, this material may help sweep away the dust surrounding a young star, making the star observable at visible wavelengths.

Before settling down, protostars eject more than half of their initial mass into space

T Tauri Stars

Mass ejection into space is a hallmark of **T Tauri stars.** These are protostars with emission lines as well as absorption lines in their spectra and whose luminosity can change irregularly on time scales of a few days. The namesake of this class of stars, T Tauri, is a protostar in the constellation Taurus (the Bull).

T Tauri stars have masses less than about 3 $M_\odot$ and ages around 10^6 years, so on an H-R diagram such as Figure 18-10 they appear above the right-hand end of the main sequence. The emission lines show that these protostars are surrounded by a thin, hot gas. The Doppler shifts of these emission lines suggest that the protostars eject gas at speeds around 80 km/s (300,000 km/h, or 180,000 mi/h).

On average, T Tauri stars eject about 10^{-8} to 10^{-7} solar masses of material per year. This may seem like a small amount, but by comparison the present-day Sun loses only about 10^{-14} $M_\odot$ per year. The T Tauri phase of a protostar may last 10^7 years or so, during which time the protostar may eject roughly a solar mass of material. Thus, the mass of the final main-sequence star is quite a bit less than that of the cloud of gas and dust from which the star originated. (The stellar masses shown in Figure 18-10 are those of the final, main-sequence stars.)

Young stars that are more massive than about 3 $M_\odot$ do not vary in luminosity like T Tauri stars. They do lose mass, however, because the pressure of radiation at their surfaces is so strong that it blows gas into space. One place this can be seen is in the Omega Nebula (**Figure 18-13*a***), where new stars are being formed (Figure 18-13*b*). The most massive of these stars eject gases with such high temperatures that they emit X rays (Figure 18-13*c*).

Bipolar Outflows and Herbig-Haro Objects

ANIMATION 18.1

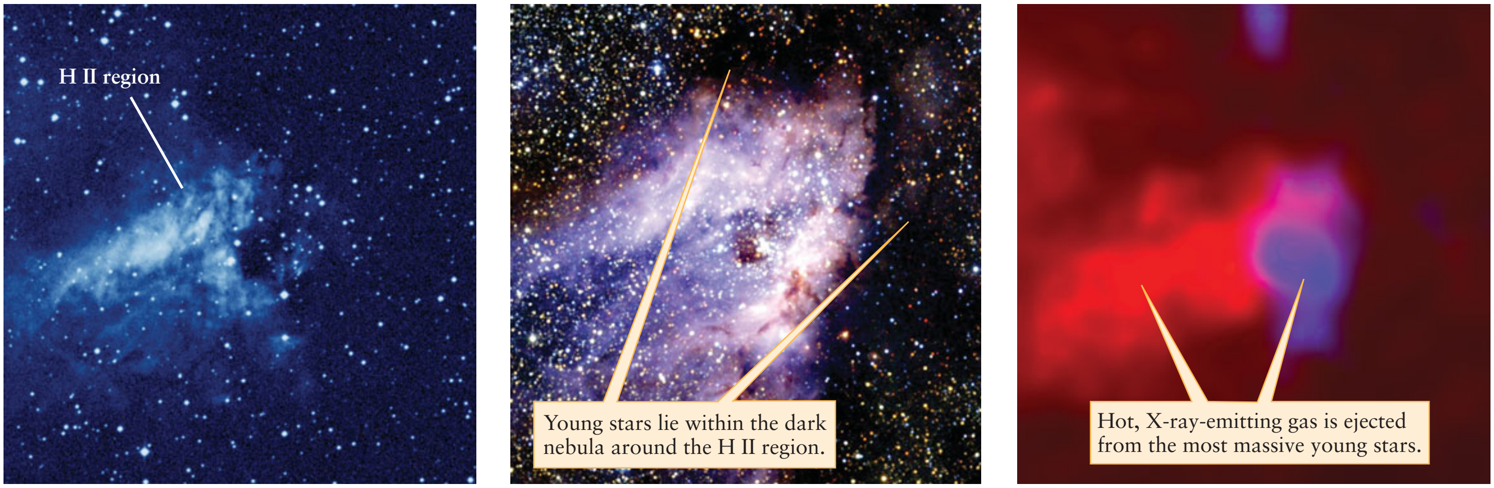
In the early 1980s, it was discovered that many young stars, including T Tauri stars, also lose mass by ejecting gas along two narrow, oppositely directed jets—a phenomenon called **bipolar outflow.** As this material is ejected into space at speeds of several hundred kilometers per second, it collides with the surrounding interstellar medium and produces knots of hot, ionized gas that glow with an emission-line spectrum. These glowing knots are called **Herbig-Haro objects** after the two astronomers, George Herbig in the United States and Guillermo Haro in Mexico, who discovered them independently.

Figure 18-14 is a Hubble Space Telescope image of the Herbig-Haro objects HH 1 and HH 2, which are produced by the two jets from a single young star in the constellation Orion. Herbig-Haro objects like these change noticeably in position, size, shape, and brightness from year to year, indicating the dynamic character of bipolar outflows.

Observations suggest that most protostars eject material in the form of jets at some point during their evolution. These bipolar outflows are very short-lived by astronomical standards, a mere 10^4 to 10^5 years, but they are so energetic that they typically eject into space more mass than ends up in the final protostar.

Accretion Disks

Protostars slowly add mass to themselves at the same time that they rapidly eject it into space. In fact, the two processes are related. As a protostar's nebula contracts, it spins faster and flattens into a disk with the protostar itself at the center. The same flattening took place in the solar nebula, from which the Sun and planets formed (see Section 8-4). Particles orbiting the protostar within this disk collide with each other, causing them to lose energy, spiral inward onto the protostar, and add to the protostar's



(a) Visible-light image R I **V** U X G (b) False-color infrared image R **I** V U X G (c) False-color X-ray image R I V U **X** G

Figure 18-13

Mass Loss from Young, Massive Stars **(a)** The Omega Nebula, also known as M17, is a region of star formation in the constellation Sagittarius about 1700 pc (5500 ly) from Earth. **(b)** This infrared image allows us to see through dust, revealing recently formed stars that cannot be seen in (a). **(c)** The most massive young stars eject copious amounts of hot gas. Red indicates X-ray emission from gas at a temperature of 1.5×10^6 K; blue indicates even hotter gas at a temperature of 7×10^6 K. Astronomers do not see such X-ray emission from the Orion Nebula (Figure 18-1), which has many young stars but very few massive ones. (a: Palomar Observatory DSS; b: 2MASS/UMass/IPAC-Caltech/NASA/NSF; c: NASA/CXC/PSU/L. Townsley et al.)

mass. This is a process of **accretion,** and the disk of material being added to the protostar in this way is called a **circumstellar accretion disk.** Figure 18-15 is an edge-on view of a circumstellar accretion disk, showing two oppositely directed jets emanating from a point at or near the center of the disk (where the protostar is located).

What causes some of the material in the disk to be blasted outward in a pair of jets? One model involves the magnetic field of the dark nebula in which the star forms (Figure 18-16). As material in the circumstellar accretion disk falls inward, it drags the magnetic field lines along with it. (We saw in Section 16-9 how a similar mechanism may explain the Sun's 22-year cycle.) Parts of the disk at different distances from the central protostar orbit at different speeds, and this can twist the magnetic field lines into two helix shapes, one on each side of the disk. The helices then act as channels that guide infalling material away from the protostar, forming two opposing jets.

Many astronomers suspect that interactions among the protostar, the accretion disk, and the jets help to slow the protostar's rotation. If so, this would explain why main-sequence stars generally spin much more slowly than protostars of the same final mass.

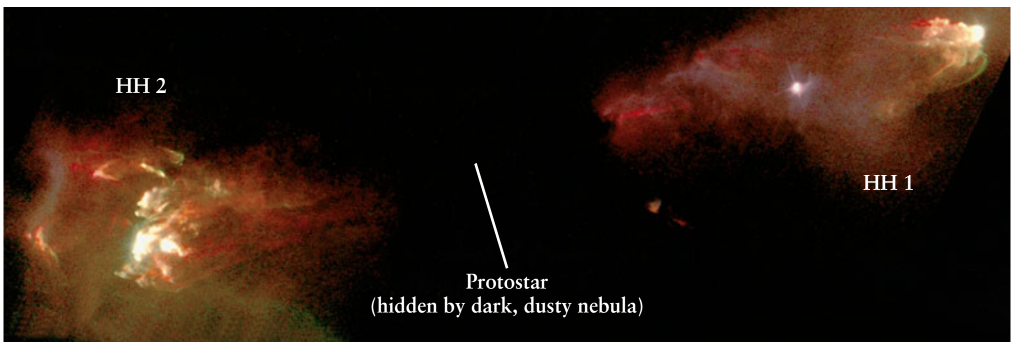
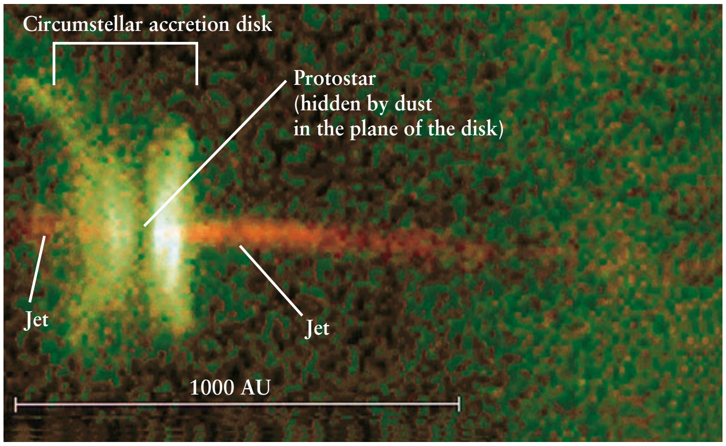


Figure 18-14 R I **V** U X G

Bipolar Outflow and Herbig-Haro Objects The two bright knots of glowing, ionized gas called HH 1 and HH 2 are Herbig-Haro objects. They are created when fast-moving gas ejected from a protostar slams into the surrounding interstellar medium, heating the gas to high temperature. HH 1 and HH 2 are 0.34 parsec (1.1 light-year) apart and lie 470 pc (1500 ly) from Earth in the constellation Orion. (J. Hester, the WFPC-2 Investigation Definition Team, and NASA)



WEB LINK 18.10

Figure 18-15 R I V U X G

A Circumstellar Accretion Disk and Jets This false-color image shows a star surrounded by an accretion disk, which we see nearly edge-on. Red denotes emission from ionized gas, while green denotes starlight scattered from dust particles in the disk. The midplane of the accretion disk is so dusty and opaque that it appears dark. Two oppositely directed jets flow away from the star, perpendicular to the disk and along the disk's rotation axis. This star lies 140 pc (460 ly) from Earth. (C. Burrows, the WFPC-2 Investigation Definition Team, and NASA)

ANIMATION 18.2

In the 1990s, astronomers using the Hubble Space Telescope discovered many examples of disks around newly formed stars in the Orion Nebula (see Figure 18-1), one of the most prominent star-forming regions in the northern sky. Figure 8-8 shows a number of these **protoplanetary disks,** or **proplyds,** that surround young stars within the nebula. As the name suggests, protoplanetary disks are thought to contain the material from which planets form around stars. They are what remains of a circumstellar accretion disk after much of the material has either fallen onto the star or been ejected by bipolar outflows.

Not all stars are thought to form protoplanetary disks; the exceptions probably include stars with masses in excess of about 3 $M_\odot$, as well as many stars in binary systems. But surveys of the Orion Nebula show that disks are found around most young, low-mass stars. Thus, disk formation may be a natural stage in the birth of many stars.

18-6 Young star clusters give insight into star formation and evolution

Dark nebulae contain tens or hundreds of solar masses of gas and dust, enough to form many stars. As a consequence, these nebulae tend to form groups or **clusters** of young stars. One such cluster is M16, shown in Figure 18-17; another is NGC 6520, depicted in Figure 18-4.

Star Clusters as Evolutionary Laboratories

In addition to being objects of great natural beauty, star clusters give us a unique way to compare the evolution of different stars. That's because clusters typically include stars with a range of different masses, all of which began to form out of the parent nebula at roughly the same time.

ANALOGY A foot race is a useful way to compare the performance of sprinters because all the competitors start the race simultaneously. A young star cluster gives us the same kind of opportunity to compare the evolution of stars of different masses that all began to form roughly simultaneously. Unlike a foot

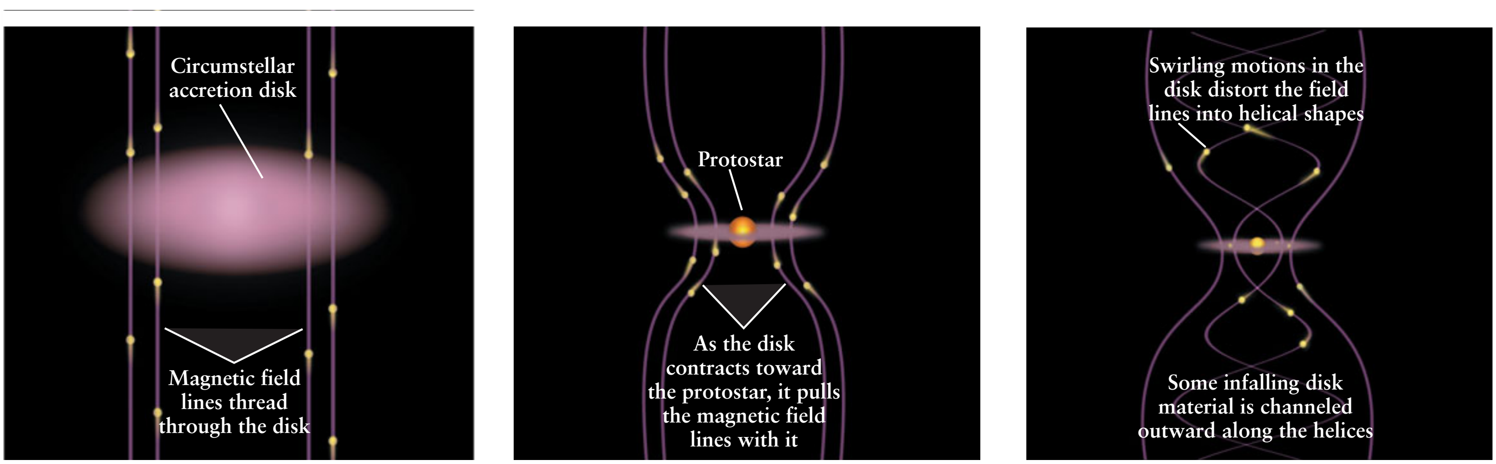
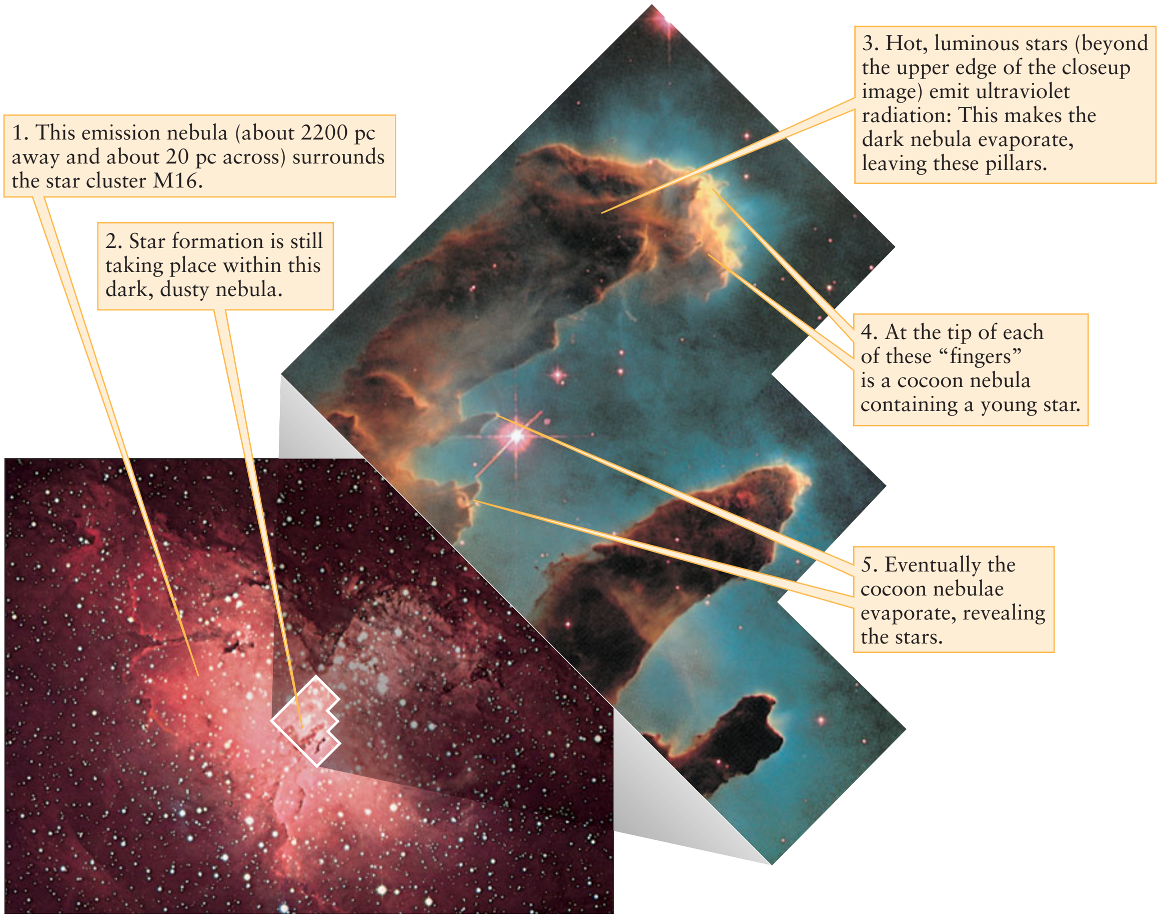


Figure 18-16

A Magnetic Model for Bipolar Outflow **(a)** Observations suggest that circumstellar accretion disks are threaded by magnetic field lines, as shown here. **(b), (c)** The contraction and rotation of the disk make the magnetic field lines distort and twist into helices. These helices steer some of the disk material into jets that stream perpendicular to the plane of the disk, as in Figure 18-15. (Adapted from Alfred T. Kamajian/Thomas P. Ray, "Fountain of Youth: Early Days in the Life of a Star," *Scientific American*, August 2000)



WEB LINK 18.11 ANIMATION 18.3

Figure 18-17 R I V U X G

A Star Cluster with an H II Region The star cluster M16 is thought to be no more than 800,000 years old, and star formation is still taking place within adjacent dark, dusty globules. The inset shows three dense, cold pillars of gas and dust silhouetted against the glowing background of the red emission nebula (called the Eagle Nebula for its shape). The pillar at the upper left extends about 0.3 parsec (1 light-year) from base to tip, and each of its “fingers” is somewhat broader than our entire solar system. (Anglo-Australian Observatory; J. Hester and P. Scowen, Arizona State University; NASA)

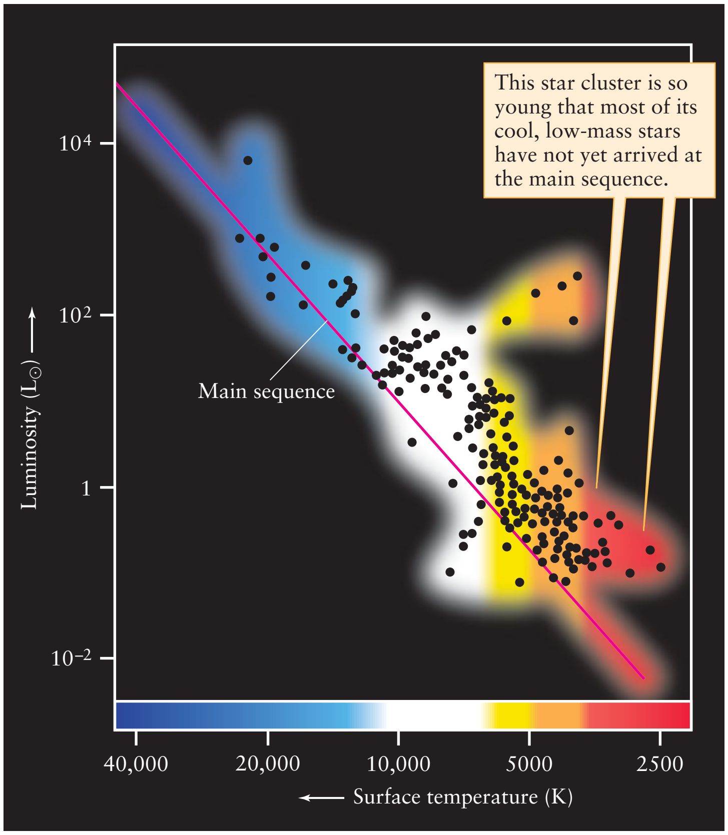
race, however, the entire “race” of evolution in a single cluster happens too slowly for us to observe; as Figure 18-10 shows, protostars take many thousands or millions of years to evolve significantly. Instead, we must compare different star clusters at various stages in their evolution to piece together the history of star formation in a cluster.

All the stars in a cluster may begin to form nearly simultaneously, but they do not all become main-sequence stars at the same time. As you can see from their evolutionary tracks (see Figure 18-10), high-mass stars evolve more rapidly than low-mass stars. The more massive the protostar, the sooner it develops the central pressures and temperatures needed for steady hydrogen fusion to begin.

Upon reaching the main sequence, *high-mass* protostars become hot, ultraluminous stars of spectral types O and B. As we saw in Section 18-2, these are the sorts of stars whose ultraviolet radiation ionizes the surrounding interstellar medium to produce an H II region. Figure 18-17 shows such an H II region, called the Eagle Nebula, surrounding the young star cluster M16. A few hundred thousand years ago, this region of space would have had a far less dramatic appearance. It was then a dark nebula, with protostars just beginning to form. Over the intervening millennia, mass ejection from these evolving protostars swept away the obscuring dust. The exposed young, hot stars heated the relatively thin remnants of the original dark nebula, creating the H II region that we see today.

When the most massive protostars to form out of a dark nebula have reached the main sequence, other *low-mass* protostars are still evolving nearby within their dusty cocoons. The evolution of these low-mass stars can be disturbed by their more massive neighbors. As an example, the inset in Figure 18-17 is a close-up of part of the Eagle Nebula. Within these opaque pillars of cold gas and dust, protostars are still forming. At the same time, however, the pillars are being eroded by intense ultraviolet light from hot, massive stars that have already shed their cocoons.

(a) The star cluster NGC 2264 R I V U X G



(b) An H-R diagram of the stars in NGC 2264

Figure 18-18

A Young Star Cluster and Its H-R Diagram (a) This photograph shows an H II region and the young star cluster NGC 2264 in the constellation Monoceros (the Unicorn). It lies about 800 pc (2600 ly) from Earth. (b) Each dot plotted on this H-R diagram represents a star in NGC 2264 whose luminosity and surface temperature have been determined. This star cluster probably started forming only 2 million years ago. (Anglo-Australian Observatory)

As each pillar evaporates, the embryonic stars within have their surrounding material stripped away prematurely, limiting the total mass that these stars can accrete.

Analyzing Young Clusters Using H-R Diagrams

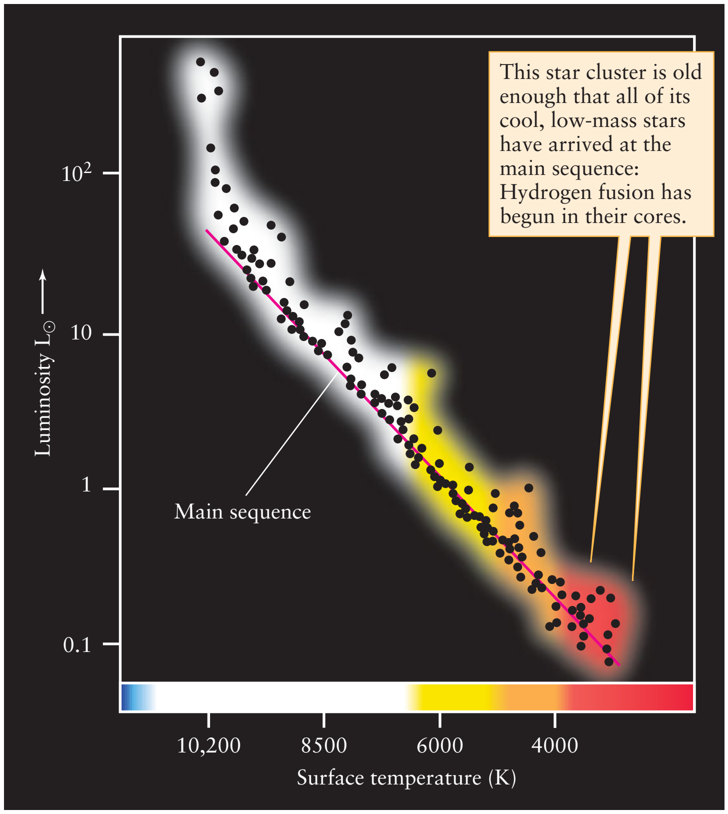
Star clusters tell us still more about how high-mass and low-mass stars evolve. Figure 18-18*a* shows the young star cluster NGC 2264 and its associated emission nebula. Astronomers have measured each star's apparent brightness and color ratio. Knowing the distance to the cluster, they have deduced the luminosities and surface temperatures of the stars (see Section 17-2 and Section 17-4). Figure 18-18*b* shows all these stars on an H-R diagram. Note that the hottest and most massive stars, with surface temperatures around 20,000 K, are on the main sequence. Stars cooler than about 10,000 K, however, have not yet quite arrived at the main sequence. These are less massive stars in the final stages of pre–main-sequence contraction and are just now beginning to ignite thermonuclear reactions at their centers. To find the ages of these stars, we can compare Figure 18-18*b* with the theoretical calculations of protostar evolution in Figure 18-10. It turns out that this particular cluster is probably about 2 million years old.

The H-R diagram of a young cluster reveals how much time has elapsed since its stars began to form

Figure 18-19*a* shows another young star cluster called the Pleiades. The photograph shows gas that must once have formed an H II region around this cluster and has dissipated into interstellar space, leaving only traces of dusty material that forms reflection nebulae around the cluster's stars. This implies that the Pleiades must be older than NGC 2264, the cluster in Figure 18-18*a*, which is still surrounded by an H II region. The H-R diagram for the Pleiades in Figure 18-19*b* bears out this idea. In contrast to the H-R diagram for NGC 2264, nearly all the stars in the Pleiades are on the main sequence. The cluster's age is about 50 million years, which is how long it takes for the least massive stars to finally begin hydrogen fusion in their cores.

CAUTION! Note that the data points for the most massive stars in the Pleiades (at the upper left of the H-R diagram in Figure

(a) The Pleiades star cluster



(b) An H-R diagram of the stars in the Pleiades

Figure 18-19

The Pleiades and Its H-R Diagram **(a)** The Pleiades star cluster is 117 pc (380 ly) from Earth in the constellation Taurus, and can be seen with the naked eye. **(b)** Each dot plotted on this H-R diagram represents a star in the Pleiades whose luminosity and surface temperature have been measured. (*Note:* The scales on this H-R diagram are different from those in Figure 18-18*b*.) The Pleiades is about 50 million (5×10^7) years old. (Anglo-Australian Observatory)

18-19*b*) lie above the main sequence. This is *not* because these stars have yet to arrive at the main sequence. Rather, these stars were the first members of the cluster to arrive at the main sequence some time ago and are now the first members to leave it. They have used up the hydrogen in their cores, so the steady process of core hydrogen fusion that characterizes main-sequence stars cannot continue. In Chapter 19 we will see why massive stars spend a rather short time as main-sequence stars and will study what happens to stars after the main-sequence phase of their lives.

WEB LINK 18.12

A loose collection of stars such as NGC 2264 or the Pleiades is referred to as an **open cluster** (or *galactic cluster*, since such clusters are usually found in the plane of the Milky Way Galaxy). Open clusters possess barely enough mass to hold themselves together by gravitation. Occasionally, a star moving faster than average will escape, or "evaporate," from an open cluster. Indeed, by the time the stars are a few billion years old, they may be so widely separated that a cluster no longer exists.

If a group of stars is gravitationally unbound from the very beginning—that is, if the stars are moving away from one another so rapidly that gravitational forces cannot keep them together—then the group is called a **stellar association.** Because young stellar associations are typically dominated by luminous O and B main-sequence stars, they are also called **OB associations.** The image that opens this chapter shows part of an OB association in the southern constellation Ara (the Altar).

18-7 Star birth can begin in giant molecular clouds

We have seen that star formation takes place within dark nebulae. But where within our Galaxy are these dark nebulae found? Does star formation take place everywhere within the Milky Way, or only in certain special locations? The answers to such questions can enhance our understanding of star formation and of the nature of our home Galaxy.

Exploring the Interstellar Medium at Millimeter Wavelengths

Dark nebulae are a challenge to locate simply because they *are* dark—they do not emit visible light. Nearby dark nebulae can be seen silhouetted against background stars or H II regions (see Figure 18-2), but sufficiently distant dark nebulae are impossible to see with visible light because of interstellar extinction. They can, however, be detected using longer-wavelength radiation that can pass unaffected through interstellar dust. In fact, dark nebulae actually emit radiation at millimeter wavelengths.

Such emission takes place because in the cold depths of interstellar space, atoms combine to form molecules. The laws of quantum mechanics predict that just as electrons within atoms can occupy only certain specific energy levels (see Section 5-8), molecules can vibrate and rotate only at certain specific rates. When a molecule goes from one vibrational state or rotational state to another, it either emits or absorbs a photon. (In the same way, an atom emits or absorbs a photon as an electron jumps from one energy level to another.) Most molecules are strong emitters of radiation with wavelengths of around 1 to 10 millimeters (mm). Consequently, observations with radio telescopes tuned to millimeter wavelengths make it possible to detect interstellar molecules of different types. More than 100 different kinds of molecules have so far been discovered in interstellar space, and the list is constantly growing.

Observing the Galaxy at millimeter wavelengths reveals the cold gas that spawns new stars

Hydrogen is by far the most abundant element in the universe. Unfortunately, in cold nebulae much of it is in a molecular form (H_2) that is difficult to detect. The reason is that the hydrogen molecule is symmetric, with two atoms of equal mass joined together, and such molecules do not emit many photons at radio frequencies. In contrast, asymmetric molecules that consist of two atoms of unequal mass joined together, such as carbon monoxide (CO), are easily detectable at radio frequencies. When a carbon monoxide molecule makes a transition from one rate of rotation to another, it emits a photon at a wavelength of 2.6 mm or shorter.

The ratio of carbon monoxide to hydrogen in interstellar space is reasonably constant: For every CO molecule, there are about 10,000 H_2 molecules. As a result, carbon monoxide is an excellent "tracer" for molecular hydrogen gas. Wherever astronomers detect strong emission from CO, they know molecular hydrogen gas must be abundant.

Giant Molecular Clouds

The first systematic surveys of our Galaxy looking for 2.6-mm CO radiation were undertaken in 1974 by the American astronomers Philip Solomon and Nicholas Scoville. In mapping the locations of CO emission, they discovered huge clouds, now called **giant molecular clouds,** that must contain enormous amounts of hydrogen. These clouds have masses in the range of 10^5 to 2×10^6 solar masses and diameters that range from about 15 to 100 pc (50 to 300 ly). Inside one of these clouds, there are about 200 hydrogen molecules per cubic centimeter. This is several thousand times greater than the average density of matter in the disk of our Galaxy, yet only 10^{-17} as dense as the air we breathe. Astronomers now estimate that our Galaxy contains about 5000 of these enormous clouds.

Figure 18-20 is a map of radio emissions from carbon monoxide in the constellations Orion and Monoceros. Note the extensive areas of the sky covered by giant molecular clouds. This part of the sky is of particular interest because it includes several star-forming regions. By comparing the radio map with the star chart overlay, you can see that the areas where CO emission is strongest, and, thus, where giant molecular clouds are densest, are sites of star formation. Therefore, giant molecular clouds are associated with the formation of stars. Particularly dense regions within these clouds form dark nebulae, and within these stars are born.

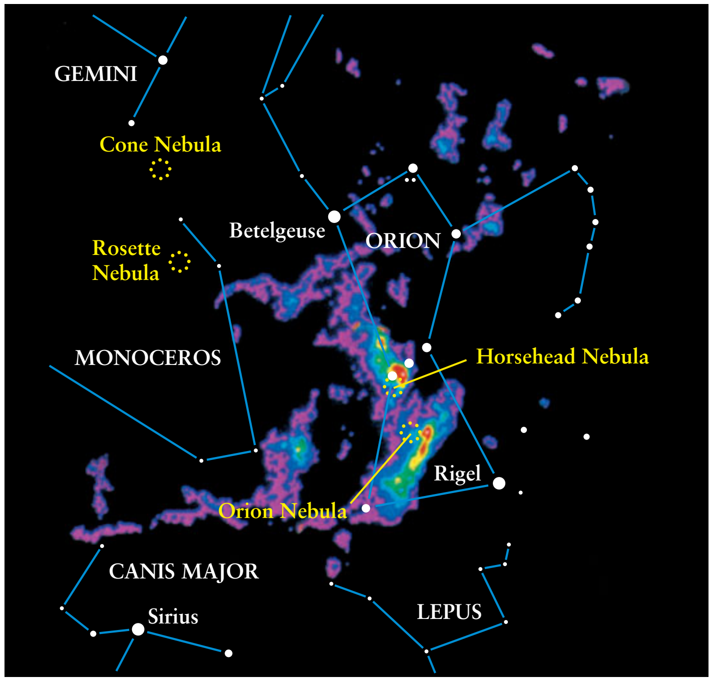


Figure 18-20 R I V U X G

Mapping Molecular Clouds A radio telescope was tuned to a wavelength of 2.6 mm to detect emissions from carbon monoxide (CO) molecules in the constellations Orion and Monoceros. The result was this false-color map, which shows a 35° × 40° section of the sky. The Orion and Horsehead star-forming nebulae are located at sites of intense CO emission (shown in red and yellow), indicating the presence of a particularly dense molecular cloud at these sites of star formation. The molecular cloud is much thinner at the positions of the Cone and Rosette nebulae, where star formation is less intense. (Courtesy of R. Maddalena, M. Morris, J. Moscowitz, and P. Thaddeus)

By using CO emissions to map out giant molecular clouds, astronomers can find the locations in our Galaxy where star formation occurs. These investigations reveal that molecular clouds clearly outline our Galaxy's spiral arms, as Figure 18-21 shows. These clouds lie roughly 1000 pc (3000 ly) apart and are strung along the spiral arms like beads on a string. This arrangement resembles the spacing of H II regions along the arms of other spiral galaxies, such as the galaxy shown in Figure 18-8*a*. The presence of both molecular clouds and H II regions shows that spiral arms are sites of ongoing star formation.

Star Formation in Spiral Arms

In Chapter 23 we will learn that spiral arms are locations where matter "piles up" temporarily as it orbits the center of the Galaxy.

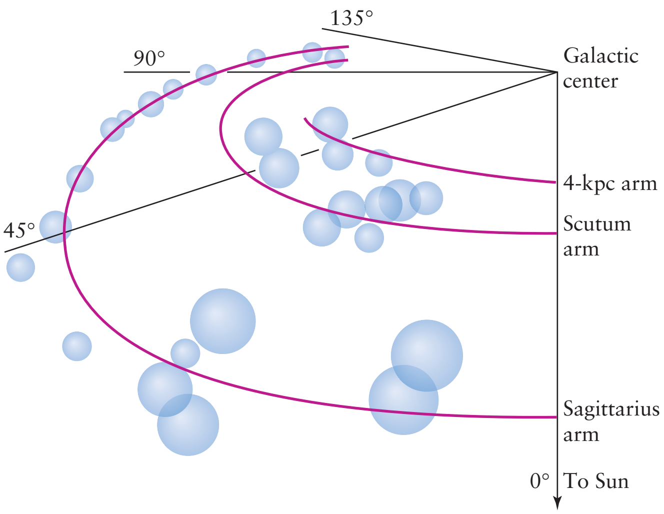


Figure 18-21

Giant Molecular Clouds in the Milky Way This perspective drawing shows the locations of giant molecular clouds in an inner part of our Galaxy as seen from a vantage point above the Sun. These clouds lie primarily along the Galaxy's spiral arms, shown by red arcs. The distance from the Sun to the galactic center is about 8000 pc (26,000 ly). (Adapted from T. M. Dame and colleagues)

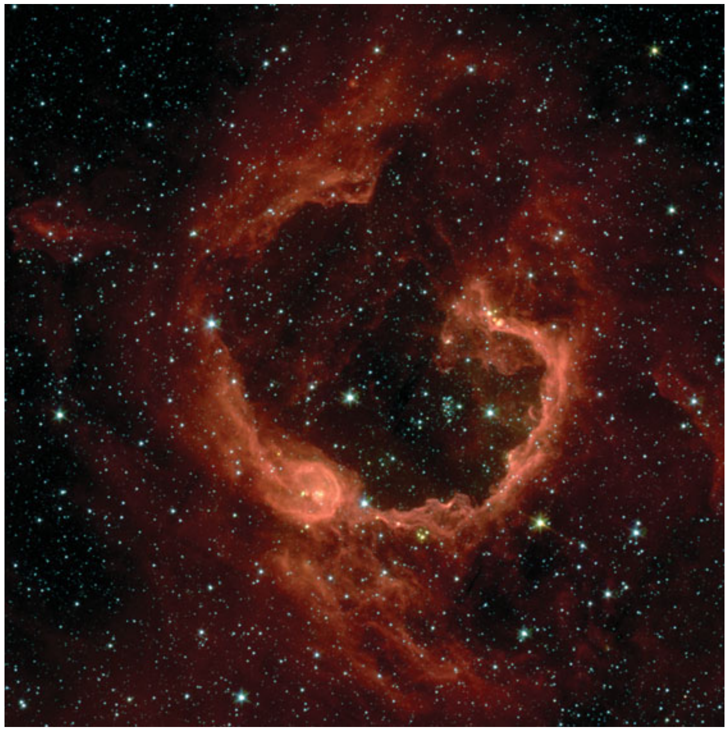


Figure 18-22 R I V U X G

A Star-Forming Bubble Radiation and winds from the hot, young O and B stars at the center of this Spitzer Space Telescope image have carved out a bubble about 20 pc (70 ly) in diameter in the surrounding gas and dust. The material around the surface of the bubble has been compressed and heated, making the dust glow at the infrared wavelengths used to record this image. The compressed material is so dense that new stars have formed within that material. This glowing cloud, called RCW 79, lies about 5300 pc (17,200 ly) from Earth in the constellation Centaurus. (NASA; JPL-Caltech; and E. Churchwell, University of Wisconsin-Madison)

You can think of matter in a spiral arm as analogous to a freeway traffic jam. Just as cars are squeezed close together when they enter a traffic jam, a giant molecular cloud is compressed when it passes through a spiral arm. When this happens, vigorous star formation begins in the cloud's densest regions.

As soon as massive O and B stars form, they emit ultraviolet light that ionizes the surrounding hydrogen, and an H II region is born. An H II region is thus a small, bright "hot spot" in a giant molecular cloud. An example is the Orion Nebula, shown in Figure 20-1*b*. Four hot, luminous O and B stars at the heart of the nebula produce the ionizing radiation that makes the surrounding gases glow. The Orion Nebula is embedded on the edge of a giant molecular cloud whose mass is estimated at 500,000 $M_\odot$. The H II regions in Figure 18-2 are located at a different point on the edge of the same molecular cloud, some 25 pc (80 ly) from the Orion Nebula.

Once star formation has begun and an H II region has formed, the massive O and B stars at the core of the H II region induce star formation in the rest of the giant molecular cloud. Ultraviolet radiation and vigorous stellar winds from the O and B stars carve out a cavity in the cloud, and the H II region, heated by the stars, expands into it. These winds travel faster than the speed of sound in the gas—that is, they are **supersonic.** Just as an airplane creates a shock wave (a sonic boom) if it flies faster than sound waves in our atmosphere, a shock wave forms where the expanding H II region pushes at supersonic speed into the rest of the giant molecular cloud. This shock wave compresses the gas through which it passes, stimulating more star birth (Figure 18-22).

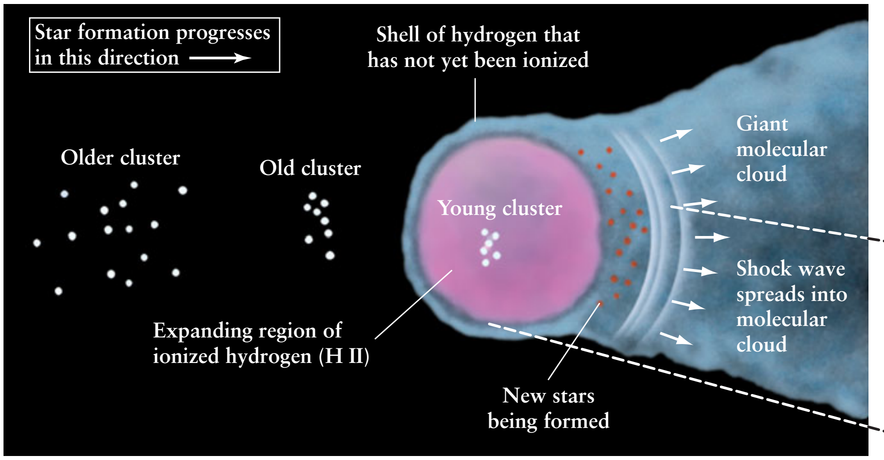
Newborn O and B stars further expand the H II region into the giant molecular cloud. Meanwhile, the older O and B stars, which were left behind, begin to disperse (Figure 18-23). In this way, an OB association "eats into" a giant molecular cloud, leaving stars in its wake.

18-8 Supernovae compress the interstellar medium and can trigger star birth

Spiral arms are not the only mechanism for triggering the birth of stars. Presumably, anything that compresses interstellar clouds will do the job. The most dramatic is a *supernova,* caused by the violent death of a massive star after it has left the main sequence. As we will see in Chapter 20, the core of the doomed star collapses suddenly, releasing vast quantities of particles and energy that blow the star apart. The star's outer layers are blasted into space at speeds of several thousand kilometers per second.

Supernova Remnants and Star Formation

Astronomers have found many nebulae across the sky that are the shredded funeral shrouds of these dead stars. Such nebulae, like the one shown in Figure 18-24, are known as **supernova remnants.**



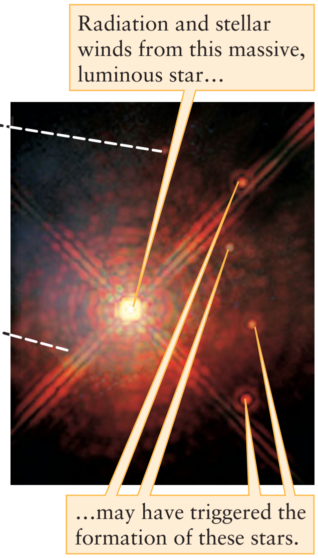
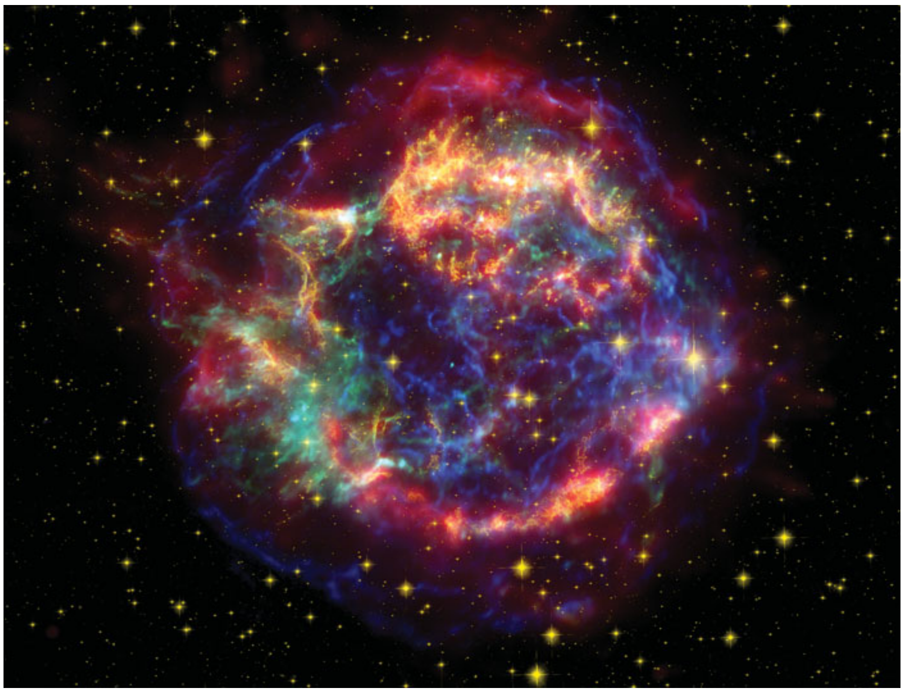


Figure 18-23 R I V U X G

How O and B Stars Trigger Star Formation Stellar winds and ultraviolet radiation from young O and B stars produce a shock wave that compresses gas farther into the giant molecular cloud. This stimulates star formation, producing more O and B stars, which stimulate still more star formation, and so on. Meanwhile, older stars are left behind. The inset shows a massive star that has spawned other, smaller stars in this way. These stars are about 770 pc (2500 ly) from Earth in the Cone Nebula, a star-forming region in the constellation Monoceros. The younger stars are just 0.04 to 0.08 ly (2500 to 5000 AU) from the central star. (Adapted from C. Lada, L. Blitz, and B. Elmegreen; inset: R. Thompson, M. Rieke, G, Schneider, and NASA)



ANIMATION 18.4

Figure 18-24 R I V U X G

A Supernova Remnant This composite image shows Cassiopeia A, the remnant of a supernova that occurred about 3000 pc (10,000 ly) from Earth. In the roughly 300 years since the supernova explosion, a shock wave has expanded about 3 pc (10 ly) outward in all directions from the explosion site. The shock wave has warmed interstellar dust to a temperature of about 300 K (Spitzer Space Telescope infrared image in red), and has heated interstellar gases to temperatures that range from 10^4 K (Hubble Space Telescope visible-light image in yellow) to 10^7 K (Chandra X-ray Observatory X-ray image in green and blue). (NASA; JPL-Caltech; and O. Krause, Steward Observatory)

Many supernova remnants have a distinctly circular or arched appearance, as would be expected for an expanding shell of gas. This wall of gas is typically moving away from the dead star faster than sound waves can travel through the interstellar medium. As we saw in Section 18-7, such supersonic motion produces a shock wave that abruptly compresses the medium through which it passes. When a gas is compressed rapidly, its temperature rises, and this temperature rise causes the gas to glow as shown in Figure 18-24.

When the expanding shell of a supernova remnant slams into an interstellar cloud, it squeezes the cloud, stimulating star birth. This kind of star birth is taking place in the stellar association seen in **Figure 18-25**. This stellar nursery is located along a luminous arc of gas about 30 pc (100 ly) in length that is presumably the remnant of an ancient supernova explosion. In fact, this arc is part of an almost complete ring of glowing gas with a diameter of about 60 pc (200 ly). Spectroscopic observations of stars along this arc reveal substantial T Tauri activity. This activity results from newborn stars undergoing mass loss in their final stages of contraction before they become main-sequence stars.

ANIMATION 18.5

Figure 18-25 R I V U X G

The Canis Major R1 Association This luminous arc of gas, about 30 pc (100 ly) long, is studded with numerous young stars. Both the luminous arc and the young stars can be traced to the same source, a supernova explosion. The shock wave from the supernova explosion is exciting the gas and making it glow; the same shock wave also compresses the interstellar medium through which it passes, triggering star formation. (Courtesy of H. Vehrenberg)

Supernovae produce a variety of atomic nuclei, including some that are not produced in any other way. These nuclei are dispersed into space by the explosion. Some of these telltale nuclei have been discovered in meteorites that have fallen to Earth. Since meteorites formed very early in the history of our solar system (see Section 8-3), this suggests that a supernova occurred nearby when our solar system was very young. Some astronomers have used this idea to propose that the Sun was once a member of a loose stellar association created by a supernova. Individual stellar motions soon carry the stars of such an association in various directions away from their birthplaces. About 4.56 billion years have passed since the birth of our star, so if the Sun was once part of a stellar association, its brothers and sisters are now widely scattered across the Galaxy.

Our Sun may have formed in association with a number of other stars

Many other processes can also trigger star formation. For example, a collision between two interstellar clouds can create new stars. Compression occurs at the interface between the two colliding clouds and vigorous star formation follows. Similarly, stellar winds from a group of O and B stars may exert strong enough pressure on interstellar clouds to cause compression, followed by star formation. (This process is similar to the one depicted in Figure 18-23.)

Star Birth in Perspective

Our understanding of star birth has improved dramatically in recent years, primarily through infrared- and millimeter-wavelength observations. The *Cosmic Connections* figure on the next page summarizes our present state of knowledge about the formation of stars.

Nevertheless, many puzzles and mysteries remain. One problem is that different modes of star birth tend to produce different percentages of different kinds of stars. For example, the passage of a spiral arm through a giant molecular cloud tends to produce an abundance of massive O and B stars. In contrast, the shock wave from a supernova seems to produce fewer O and B stars but many more of the less massive A, F, G, and K stars. We do not yet know why this is so.

Despite these unanswered questions, it is now clear that star birth involves mechanisms on a colossal scale, from the deaths of massive stars to the rotation of an entire galaxy. In many respects, we have just begun to appreciate these cosmic processes. The study of cold, dark stellar nurseries will be an active and exciting area of astronomical research for many years to come.

In the chapters that follow, we will learn that the interstellar medium is both the birthplace of new stars and a dumping ground for dying stars. At the end of its life, a star can shed most of its mass in an outburst that enriches interstellar space with new chemical elements. The interstellar medium is therefore both nursery and graveyard. Because of this intimate relationship with stars, the interstellar medium evolves as successive generations of stars live out their lives. Understanding the details of this cosmic symbiosis is one of the challenges of modern astronomy.

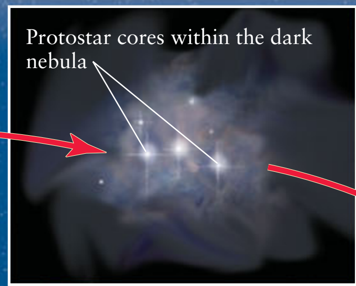
COSMIC CONNECTIONS

How Stars Are Born

If a clump of interstellar matter is cold and dense enough, it will begin to collapse thanks to the mutual gravitational attraction of its parts. If the clump is massive enough, it will evolve into a main-sequence star through the sequence of events shown here.

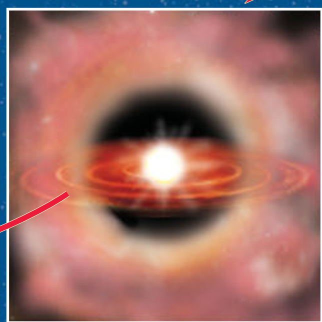
In this cold, dark nebula, gas atoms and dust particles move so slowly that gravity can draw them together.

Gas and dust begin to condense into clumps, forming the cores of protostars.



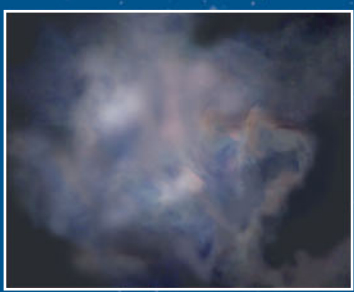
As the cores condense, their density and temperature both increase.

As the protostars continue to heat up and accrete matter from the nebula, they begin to glow due to their increasing temperature.



Once the temperature at the center of a protostar becomes sufficiently high, thermonuclear fusion of hydrogen into helium begins. The mass that is continuing to fall onto the star forms an accretion disk.

In the T Tauri stage, the young star ejects mass into space in a bipolar outflow. A stellar wind blows away the remaining parts of the nebula that surround the star, exposing the star to space.



The ejected mass can induce a shock wave in the surrounding interstellar material, triggering the formation of additional stars.

Processes that cause the star to lose or gain mass come to an end, and the star stabilizes as a main-sequence star in hydrostatic equilibrium. The remnants of the accretion disk may remain as a protoplanetary disk, from which a system of planets may form around the star.

Key Words

The term preceded by an asterisk (*) is discussed in Box 18-1.

Key Ideas

Stellar Evolution: Because stars shine by thermonuclear reactions, they have a finite life span. The theory of stellar evolution describes how stars form and change during that life span.

The Interstellar Medium: Interstellar gas and dust, which make up the interstellar medium, are concentrated in the disk of the Galaxy. Clouds within the interstellar medium are called nebulae.

• Dark nebulae are so dense that they are opaque. They appear as dark blots against a background of distant stars.

• Emission nebulae, or H II regions, are glowing, ionized clouds of gas. Emission nebulae are powered by ultraviolet light that they absorb from nearby hot stars.

• Reflection nebulae are produced when starlight is reflected from dust grains in the interstellar medium, producing a characteristic bluish glow.

Protostars: Star formation begins in dense, cold nebulae, where gravitational attraction causes a clump of material to condense into a protostar.

• As a protostar grows by the gravitational accretion of gases, Kelvin-Helmholtz contraction causes it to heat and begin glowing. Its relatively low temperature and high luminosity place it in the upper right region on an H-R diagram.

• Further evolution of a protostar causes it to move toward the main sequence on the H-R diagram. When its core temperatures become high enough to ignite steady hydrogen burning, it becomes a main-sequence star.

• The more massive the protostar, the more rapidly it evolves.

Mass Loss by Protostars: In the final stages of pre–main-sequence contraction, when thermonuclear reactions are about to begin in its core, a protostar may eject large amounts of gas into space.

• Low-mass stars that vigorously eject gas are called T Tauri stars.

• A circumstellar accretion disk provides material that a young star ejects as jets. Clumps of glowing gas called Herbig-Haro objects are sometimes found along these jets and at their ends.

Star Clusters: Newborn stars may form an open or galactic cluster. Stars are held together in such a cluster by gravity. Occasionally a star moving more rapidly than average will escape, or "evaporate," from such a cluster.

• A stellar association is a group of newborn stars that are moving apart so rapidly that their gravitational attraction for one another cannot pull them into orbit about one another.

O and B Stars and Their Relation to H II Regions: The most massive protostars to form out of a dark nebula rapidly become main sequence O and B stars. They emit strong ultraviolet radiation that ionizes hydrogen in the surrounding cloud, thus creating the reddish emission nebulae called H II regions.

• Ultraviolet radiation and stellar winds from the O and B stars at the core of an H II region create shock waves that move outward through the gas cloud, compressing the gas and triggering the formation of more protostars.

Giant Molecular Clouds: The spiral arms of our Galaxy are laced with giant molecular clouds, immense nebulae so cold that their constituent atoms can form into molecules.

• Star-forming regions appear when a giant molecular cloud is compressed. This can be caused by the cloud's passage through one of the spiral arms of our Galaxy, by a supernova explosion, or by other mechanisms.

Questions

Review Questions

1. If no one has ever seen a star go through the complete formation process, how are we able to understand how stars form?
2. Why is it more difficult to observe the life cycles of stars than the life cycles of planets or animals?
3. If an interstellar medium fills the space between the stars, how is that we are able to see the stars at all?
4. Summarize the evidence that interstellar space contains (**a**) gas and (**b**) dust.
5. What are H II regions? Near what kinds of stars are they found? Why do only these stars give rise to H II regions?
6. What are stationary absorption lines? In what sort of spectra are they seen? How do they give evidence for the existence of the interstellar medium?
7. In Figure 18-2, what makes the Horsehead Nebula dark? What makes IC 434 glow?
8. Why is the daytime sky blue? Why are distant mountains purple? Why is the Sun red when seen near the horizon at sunrise or sunset? In what ways are your answers analogous to the explanations for the bluish color of reflection nebulae and the process of interstellar reddening?

9. To see the constellation Coma Berenices (Berenice's Hair) you must look perpendicular to the plane of the Milky Way. By contrast, the Milky Way passes through the constellation Cassiopeia (named for a mythical queen). Would you expect H II regions to be more abundant in Coma Berenices or in Cassiopeia? Explain your reasoning.
10. The interior of a dark nebula is billions of times less dense than the air that you breathe. How, then, are dark nebulae able to block out starlight?
11. Why are low temperatures necessary in order for protostars to form inside dark nebulae?
12. Compare and contrast Barnard objects and Bok globules. How many Sun-sized stars could you make out of a Barnard object? Out of a Bok globule?
13. Describe the energy source that causes a protostar to shine. How does this source differ from the energy source inside a main-sequence star?
14. What is an evolutionary track? How can evolutionary tracks help us interpret the H-R diagram?
15. What happens inside a protostar to slow and eventually halt its gravitational contraction?
16. Why are the evolutionary tracks of high-mass stars different from those of low-mass stars? For which kind of star is the evolution more rapid? Why?
17. Why are protostars more easily seen with an infrared telescope than with a visible-light telescope?
18. In what ways is the internal structure of a 1-$M_\odot$ main-sequence star different from that of a 5-$M_\odot$ main-sequence star? From that of a 0.5-$M_\odot$ main-sequence star? What features are common to all these stars?
19. What sets the limits on the maximum and minimum masses of a main-sequence star?
20. What are T Tauri stars? How do we know that they eject matter at high speed? How does their rate of mass loss compare to that of the Sun?
21. What are Herbig-Haro objects? Why are they often found in pairs?
22. Why do disks form around contracting protostars? What is the connection between disks and bipolar outflows?
23. Young open clusters like those shown in Figures 18-18 and 18-19 are found only in the plane of the Galaxy. Explain why this should be.
24. Why are observations at millimeter wavelengths so much more useful in exploring interstellar clouds than observations at visible wavelengths?
25. What are giant molecular clouds? What role do these clouds play in the birth of stars?
26. Giant molecular clouds are among the largest objects in our Galaxy. Why, then, were they discovered only relatively recently?
27. Consider the following stages in the evolution of a young star cluster: (i) H II region; (ii) dark nebula; (iii) formation of O and B stars; (iv) giant molecular cloud. Put these stages in the correct chronological order and discuss how they are related.
28. Briefly describe four mechanisms that compress the interstellar medium and trigger star formation.

Advanced Questions

Problem-solving tips and tools

You may find it helpful to review Box 17-4, which describes the relationship among a star's luminosity, radius, and surface temperature. The small-angle formula is described in Box 1-1. Orbital periods are described by Kepler's third law, which we discussed in Boxes 4-2 and 4-4. Remember that the Stefan-Boltzmann law (Box 5-2) relates the temperature of a blackbody to its energy flux. Remember, too, that the volume of a sphere of radius r is $4\pi r^3/3$.

29. If you looked at the spectrum of a reflection nebula, would you see absorption lines, emission lines, or no lines? Explain your answer. As part of your explanation, describe how the spectrum demonstrates that the light was reflected from nearby stars.
30. In the direction of a particular star cluster, interstellar extinction allows only 15% of a star's light to pass through each kiloparsec (1000 pc) of the interstellar medium. If the star cluster is 3.0 kiloparsecs away, what percentage of its photons survive the trip to the Earth?
31. The visible-light photograph below shows the Trifid Nebula in the constellation Sagittarius. Label the following features on this photograph: (**a**) reflection nebulae (and the star or stars whose light is being reflected); (**b**) dark nebulae; (**c**) H II regions; (**d**) regions where star formation may be occurring. Explain how you identified each feature.

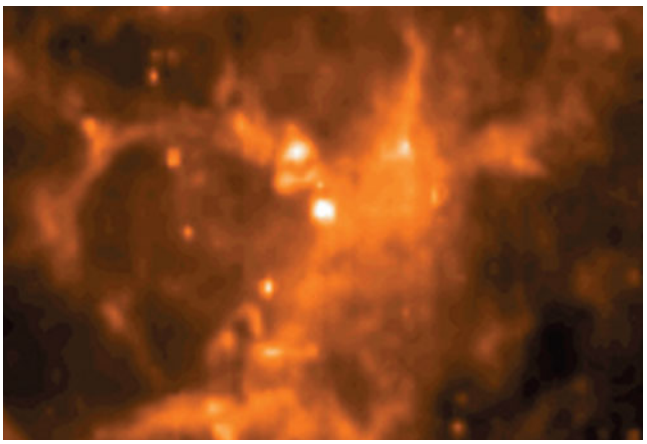
R I V U X G

(Anglo-Australian Observatory)

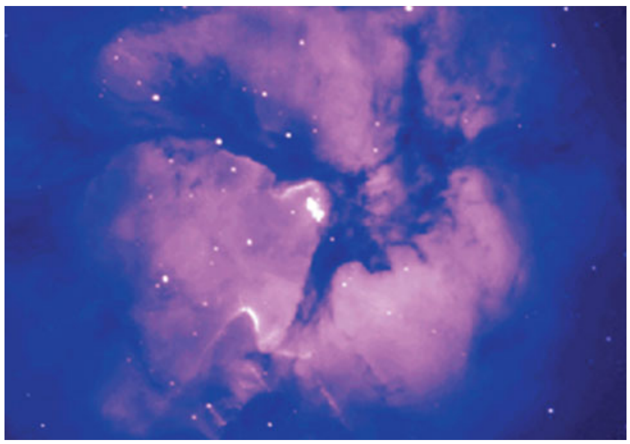
32. Find the density (in atoms per cubic centimeter) of a Bok globule having a radius of 1 light-year and a mass of 100 $M_\odot$. How does your result compare with the density of a typical H II region, between 80 and 600 atoms per cm^3? (Assume that the globule is made purely of hydrogen atoms.)
33. The *Becklin-Neugebauer object* is a newly formed star within the Orion Nebula. It is substantially more luminous than the other newly formed stars in that nebula. Assuming that all these stars began the process of formation of the same time, what can you conclude about the mass of the Becklin-Neugebauer object compared with those of the other newly formed stars? Does your conclusion depend on whether or not the stars have reached the main sequence? Explain your reasoning.
34. The two false-color images below show a portion of the Trifid Nebula (see Question 31). The reddish-orange view is a false-color infrared image, while the bluish picture (shown to the same scale) was made with visible light. Explain why the dark streaks in the visible-light image appear bright in the infrared image.
35. At one stage during its birth, the protosun had a luminosity of 1000 $L_\odot$ and a surface temperature of about 1000 K. At this time, what was its radius? Express your answer in three ways: as a multiple of the Sun's present-day radius, in kilometers, and in astronomical units.
36. A newly formed protostar and a red giant are both located in the same region on the H-R diagram. Explain how you could distinguish between these two.
37. (**a**) Determine the radius of the circumstellar accretion disk in Figure 18-15. (You will need to measure this image with a ruler. Note the scale bar in this figure.) Give your answer in astronomical units and in kilometers. (**b**) Assume that the young star at the center of this disk has a mass of 1 $M_\odot$. What is the orbital period (in years) of a particle at the outer edge of the disk? (**c**) Using your ruler again, determine the length of the jet that extends to the right of the circumstellar disk in Figure 18-15. At a speed of 200 km/s, how long does it take gas to traverse the entire visible length of the jet?
38. The star cluster NGC 2264 (Figure 18-18) contains numerous T Tauri stars, while the Pleiades (Figure 18-19) contains none. Explain why there is a difference.
39. The concentration or abundance of ethyl alcohol in a typical molecular cloud is about 1 molecule per 10^8 cubic meters. What volume of such a cloud would contain enough alcohol to make a martini (about 10 grams of alcohol)? A molecule of ethyl alcohol has 46 times the mass of a hydrogen atom (that is, ethyl alcohol has a molecular weight of 46).
40. From the information given in the caption to Figure 18-24, calculate the angular diameter in arcminutes of Cassiopeia A as seen from Earth.
41. From the information given in the caption for Figure 18-24, calculate the average speed at which the shock wave has spread away from the site of the supernova explosion. Give your answer in kilometers per second and as a fraction of the speed of light. (*Hint:* There are 3.16×10^7 seconds in a year and the speed of light is 3.00×10^5 km/s.)

Discussion Questions

42. Some science-fiction movies show stars suddenly becoming dramatically brighter when they are "born" (that is, when thermonuclear fusion reactions begin in their cores). Discuss whether this is a reasonable depiction.
43. Suppose that the electrons in hydrogen atoms were not as strongly attracted to the nuclei of those atoms, so that these atoms were easier to ionize. What consequences might this have for the internal structure of main-sequence stars? Explain your reasoning.
44. What do you think would happen if our solar system were to pass through a giant molecular cloud? Do you think the Earth has ever passed through such clouds?
45. Many of the molecules found in giant molecular clouds are *organic* molecules (that is, they contain carbon). Speculate about the possibility of life-forms and biological processes occurring in giant molecular clouds. In what ways might the conditions existing in giant molecular clouds favor or hinder biological evolution?



R **I** V U X G



R I **V** U X G

(ESA/ISO, ISOCAM, and J. Cernicharo et al.; IAC, Observatorio del Teide, Tenerife)

46. Speculate on why a shock wave from a supernova seems to produce relatively few high-mass O and B stars, compared to the lower-mass A, F, G, and K stars.

Web/eBook Questions

47. In recent years astronomers have been able to learn about the character of the interstellar medium in the vicinity of the Sun. Search the World Wide Web for information about aspects of the nearby interstellar medium, including features called the Local Interstellar Cloud and the Local Bubble. How do astronomers study the nearby interstellar medium? What makes these studies difficult? Is the interstellar medium relatively uniform in our neighborhood, or is it clumpy? If the latter, is our solar system in a relatively thin or thick part of the interstellar medium? How is our solar system moving through the interstellar medium?

48. Search the World Wide Web for recent discoveries about how brown dwarfs form. Do they tend to form in the same locations as "real" stars? Do they form in relatively small or relatively large numbers compared to "real" stars? What techniques are used to make these discoveries?

49. Measuring a Stellar Jet. Access the animation "A Stellar Jet in the Trifid Nebula" in Chapter 18 of the *Universe* Web site or eBook. (**a**) The Trifid Nebula as a whole has an angular diameter of 28 arcmin. By stepping through the animation, estimate the angular size of the stellar jet shown at the end of the animation. (**b**) The Trifid Nebula is about 2800 pc (9000 ly) from Earth. Estimate the length of the jet in light-years. (**c**) If gas in the jet travels at 200 km/s, how long does it take to traverse the length of the jet? Give your answer in years. (*Hint:* There are 3.16×10^7 seconds in a year.)

Activities

Observing Projects

Observing tips and tools

After looking at the beautiful color photographs of nebulae in this chapter, you may find the view through a telescope a bit disappointing at first, but fear not. You can see a great deal with even a small telescope. To get the best view of a dim nebula using a telescope, use the same "averted vision" trick we described in "Observing tips and tools" in Chapter 17: If you direct your vision a little to one side of the object that you are looking at, the light from that object will go onto a more sensitive part of the retina.

50. Use a telescope to observe at least two of the H II regions listed in the following table. In each case, can you guess which stars are probably responsible for the ionizing radiation that causes the nebula to glow? Can you see any obscuration or silhouetted features that suggest the presence of interstellar dust? Draw a picture of what you see through the telescope and compare it with a photograph of the object. Take note of which portions of the nebula were not visible through your telescope.

Nebula	Right ascension	Declination
M42 (Orion)	$5^h\ 35.4^m$	$-5°\ 27'$
M43	$5\ 35.6^m$	$-5\ 16$
M20 (Trifid)	$18\ 02.6^m$	$-23\ 02$
M8 (Lagoon)	$18\ 03.8^m$	$-24\ 23$
M17 (Omega)	$18\ 20.8^m$	$-16\ 11$

Note: The right ascensions and declinations are given for epoch 2000.

51. On an exceptionally clear, moonless night, use a telescope to observe at least one of the dark nebulae listed in the following table. These nebulae are very difficult to find, because they are recognizable only by the absence of stars in an otherwise starry part of the sky. Are you confident that you actually saw the dark nebula? Does the pattern of background stars suggest a particular shape to the nebula?

Nebula	Right ascension	Declination
Barnard 72 (the Snake)	$17^h\ 23.5^m$	$-23°\ 38'$
Barnard 86	18 02.7	$-27\ 50$
Barnard 133	19 06.1	$-6\ 50$
Barnard 142 and 143	19 40.7	$-10\ 57$

Note: The right ascensions and declinations are given for epoch 2000.

52. A few fine objects cover such large regions of the sky that they are best seen with binoculars. If you have access to a high-quality pair of binoculars, observe the North American Nebula in Cygnus and the Pipe Nebula in Ophiuchus. Both nebulae are quite faint, so you should attempt to observe them only on an exceptionally dark, clear, moonless night. The North America Nebula is a cloud of glowing hydrogen gas located about 3° east of Deneb, the brightest star in Cygnus. While searching for the North America Nebula, you may glimpse another diffuse H II region, the Pelican Nebula, located about 2° southeast of Deneb. The Pipe Nebula is a 7°-long, meandering, dark nebula to the south and to the east of the star θ (theta) Ophiuchi, which is in a section of Ophiuchus that extends southward between the constellations of Scorpius and Sagittarius. Located about 12° east of the bright red star Antares, you can locate θ Ophiuchi using the *Starry Night Enthusiast™* software on the CD-ROM that accompanies certain printed copies of this book.

53. Use the *Starry Night Enthusiast™* program to investigate a star-forming region. Use the **Find . . .** command in the **Edit** menu to find and center on M20 (the Trifid Nebula, shown in the figure that accompanies Question 31) as seen from your location. Zoom out as far as possible using the **Zoom** controls at the right-hand end of the toolbar. Set the **Time** appropriately and adjust the **Month** and **Day** in the **Date** to answer the following questions. (*Hint:* You may want to remove daylight and display the local meridian to provide precise answers.) (**a**) On what day is M20 highest in the sky at noon? Explain how you determined this. (**b**) On what day is M20 highest in the sky at midnight, so that it is best placed for observing with a telescope? Explain how you determined this.

54. Use the *Starry Night Enthusiast*™ program to examine the Milky Way Galaxy. Open the **Favourites** pane and click on **Stars > Sun in Milky Way** to display our Galaxy from a position 0.150 million light-years above the galactic plane. (You can remove the astronaut's feet from this view if desired by clicking on **View > Feet.**) You can zoom in or out on the Galaxy using the + and − buttons at the upper right end of the toolbar. You can move the Galaxy by holding down the mouse button while moving the mouse. You can also rotate the Galaxy by putting the mouse cursor over the image and holding down the **Shift** key while holding down the mouse button and moving the mouse. (**a**) You can identify H II regions by their characteristic magenta color. Describe where in the Galaxy you find these. Are most found in the inner part of the Galaxy or in its outer regions? (**b**) Where do you find dark lanes of dust—in the inner part of the Galaxy or in its outer regions? Do you see any connection between the locations of dust and of H II regions? If there is a connection, what do you think causes it? If there is not a connection, why is this the case? You can examine the location of our Galaxy in relation to neighboring galaxies by turning the Milky Way edge-on and by increasing the distance from the Earth using the **up** key below the Viewing Location on the toolbar.

Collaborative Exercises

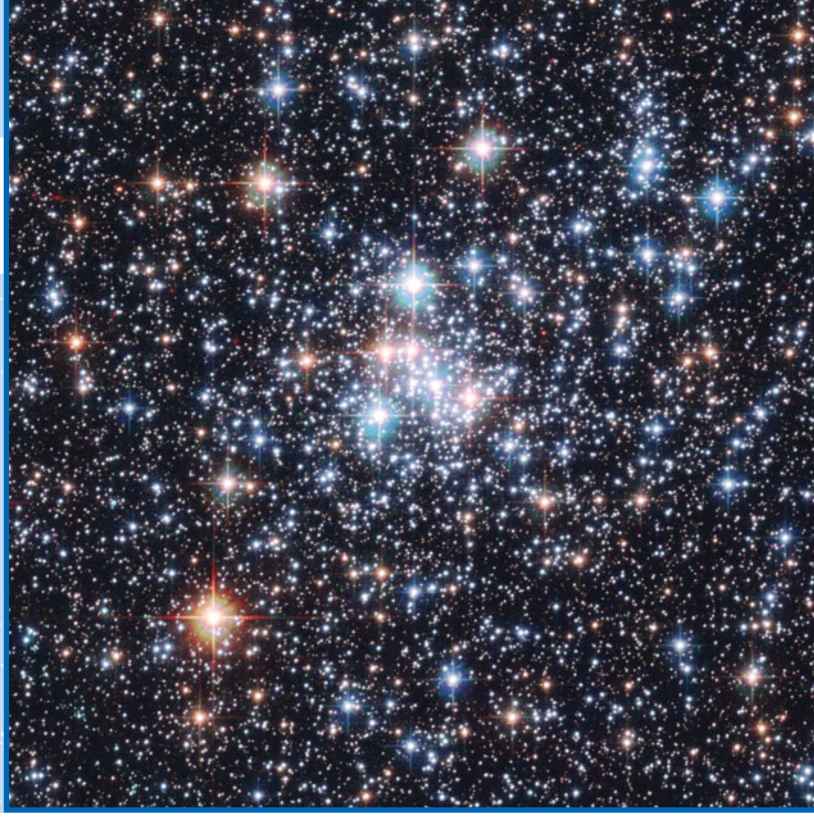
55. Imagine that your group walks into a store that specializes in selling antique clothing. Prepare a list of observable characteristics that you would look for to distinguish which items were from the early, middle, and late twentieth century. Also, write a paragraph that specifically describes how this task is similar to how astronomers understand the evolution of stars.

56. Consider advertisement signs visible at night in your community and provide specific examples of ones that are examples of the three different types of nebulae that astronomers observe and study. If an example doesn't exist in your community, creatively design an advertisement sign that could serve as an example.

57. The pre–main-sequence evolutionary tracks shown in Figure 18-10 describe the tracks of seven protostars of different masses. Imagine a new sort of H-R diagram that plots a human male's increasing age versus decreasing hair density on the head instead of increasing luminosity versus decreasing temperature. Create and carefully label a sketch of this imaginary HaiR diagram showing both the majority of the U.S. male population and a few oddities. Finally, draw a line that clearly labels your sketch to show how a typical male undergoing male-pattern baldness might slowly change position on the HaiR diagram over the course of a human life span.

19

Stellar Evolution: On and After the Main Sequence



WEB LINK 19.1

R I V U X G

The red stars in this image of open cluster NGC 290 are red giants, a late stage in stellar evolution. (ESA/NASA/Edward W. Olszewski, U. of Arizona)

Imagine a world like Earth, but orbiting a star more than 100 times larger and 2000 times more luminous than our Sun. Bathed in the star's intense light, the surface of this world is utterly dry, airless, and hot enough to melt iron. If you could somehow survive on the daytime side of this world, you would see the star filling almost the entire sky.

This bizarre planet is not a creation of science fiction—it is our own Earth some 7.6 billion years from now. The bloated star is our own Sun, which in that remote era will have become a red-giant star.

A main-sequence star evolves into a red giant when all the hydrogen in its core is consumed. The star's core contracts and heats up, but its outer layers expand and cool. In the hot, compressed core, helium fusion becomes a new energy source. The more massive a star, the more rapidly it consumes its core's hydrogen and the sooner it evolves into a giant.

The interiors of stars are hidden from our direct view, so much of the story in this chapter is based on theory. We back up those calculations with observations of star clusters, which contain stars of different masses but roughly the same age. (An example is the cluster shown here, many of whose stars have evolved into luminous red giants.) Other observations show that some red-giant stars actually pulsate, and that stars can evolve along very different paths if they are part of a binary star system.

19-1 During a star's main-sequence lifetime, it expands and becomes more luminous

At the core, main-sequence stars are all fundamentally alike. As we saw in Section 18-4, it is in their cores that all such stars convert hydrogen into helium by thermonuclear reactions. This process is called **core hydrogen fusion.** The total time that a star will spend fusing hydrogen into helium in its core, and thus the total time that it will spend as a main-sequence star, is called its **main-sequence lifetime.** For our Sun, the main-sequence lifetime is about 12 billion (1.2×10^{10}) years. Hydrogen

Over the past 4.56 billion years, thermonuclear reactions have caused an accumulation of helium in our Sun's core

Learning Goals

By reading the sections of this chapter, you will learn

19-1 How a main-sequence star changes as it converts hydrogen to helium

19-2 What happens to a star when it runs out of hydrogen fuel

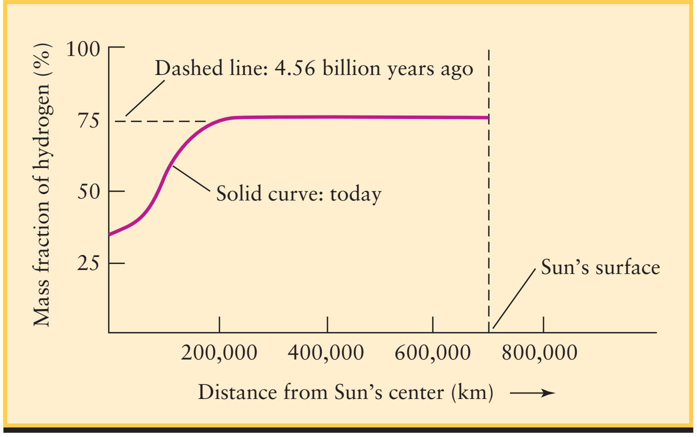
19-3 How aging stars can initiate a second stage of thermonuclear fusion

19-4 How H-R diagrams for star clusters reveal the later stages in the evolution of stars

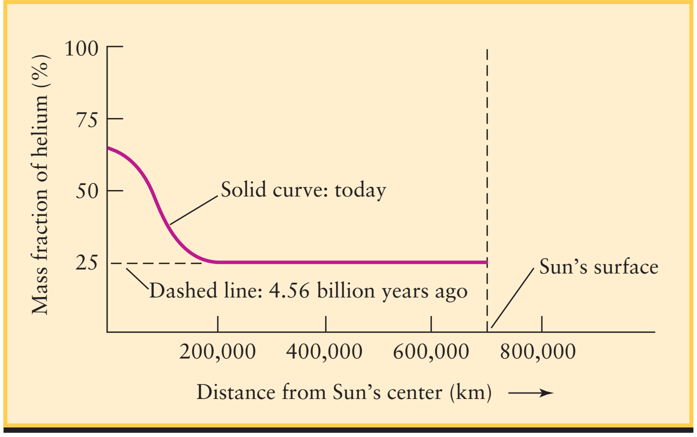
19-5 The two kinds of stellar populations and their significance

19-6 Why some aging stars pulsate and vary in luminosity

19-7 How stars in a binary system can evolve very differently from single, isolated stars



(a) Hydrogen in the Sun's interior



(b) Helium in the Sun's interior

Figure 19-1

Changes in the Sun's Chemical Composition These graphs show the percentage by mass of **(a)** hydrogen and **(b)** helium at different points within the Sun's interior. The dashed horizontal lines show that these percentages were the same throughout the Sun's volume when it first formed. As the solid curves show, over the past 4.56×10^9 years, thermonuclear reactions at the core have depleted hydrogen in the core and increased the amount of helium in the core.

fusion has been going on in the Sun's core for the past 4.56 billion (4.56×10^9) years, so our Sun is less than halfway through its main-sequence lifetime.

What happens to a star like the Sun after the core hydrogen has been used up, so that it is no longer a main-sequence star? As we will see, it expands dramatically to become a red giant. To understand why this happens, it is useful to first look at how a star evolves *during* its main-sequence lifetime. The nature of that evolution depends on whether its mass is less than or greater than about 0.4 $M_\odot$.

Main-Sequence Stars of 0.4 $M_\odot$ or Greater: Consuming Core Hydrogen

A protostar becomes a main-sequence star when steady hydrogen fusion begins in its core and it achieves *hydrostatic equilibrium*—a balance between the inward force of gravity and the outward pressure produced by hydrogen fusion (see Section 16-2 and Section 18-4). Such a freshly formed main-sequence star is called a **zero-age main-sequence star.**

We make the distinction between "main sequence" and "zero-age main sequence" because a star undergoes noticeable changes in luminosity, surface temperature, and radius during its main-sequence lifetime. These changes are a result of core hydrogen fusion, which alters the chemical composition of the core. As an example, when our Sun first formed, its composition was the same at all points throughout its volume: by mass, about 74% hydrogen, 25% helium, and 1% heavy elements. But as Figure 19-1 shows, the Sun's core now contains a greater mass of helium than of hydrogen. (There is still enough hydrogen in the Sun's core for another 7 billion years or so of core hydrogen fusion.)

CAUTION! Although the outer layers of the Sun are also predominantly hydrogen, there are two reasons why this hydrogen cannot undergo fusion. The first reason is that while the temperature and pressure in the core are high enough for thermonuclear reactions to take place, the temperatures and pressure in the outer layers are not. The second reason is that there is no flow of material between the Sun's core and outer layers, so the hydrogen in the outer layers cannot move into the hot, high-pressure core to undergo fusion. The same is true for main-sequence stars with masses of about 0.4 $M_\odot$ or greater. (We will see below that the outer layers *can* undergo fusion in main-sequence stars with a mass less than about 0.4 $M_\odot$.)

Thanks to core hydrogen fusion, the total number of atomic nuclei in a star's core decreases with time: In each reaction, four hydrogen nuclei are converted to a single helium nucleus (see the *Cosmic Connections* figure in Section 16-1, as well as Box 16-1). With fewer particles bouncing around to provide the core's internal pressure, the core contracts slightly under the weight of the star's outer layers. Compression makes the core denser and increases its temperature. (Box 19-1 gives some everyday examples of how the temperature of a gas changes when it compresses or expands.) As a result of these changes in density and temperature, the pressure in the compressed core is actually higher than before.

As the star's core shrinks, its outer layers expand and shine more brightly. Here's why: As the core's density and temperature increase, hydrogen nuclei in the core collide with one another more frequently, and the rate of core hydrogen fusion increases. Hence, the star's luminosity increases. The radius of the star as a whole also increases slightly, because increased core pressure pushes outward on the star's outer layers. The star's surface temperature changes as well, because it is related to the luminosity and radius (see Section 17-6 and Box 17-4). As an example, theoretical calculations indicate that over the past 4.56×10^9 years, our Sun has become 40% more luminous, grown in radius by 6%, and increased in surface temperature by 300 K (Figure 19-2).

As a main-sequence star ages and evolves, the increase in energy outflow from its core also heats the material immediately surrounding the core. As a result, hydrogen fusion can begin in

BOX 19-1 The Heavens on the Earth

Compressing and Expanding Gases

As a star evolves, various parts of the star either contract or expand. When this happens, the gases behave in much the same way as gases here on the Earth when they are forced to compress or allowed to expand.

When a gas is compressed, its temperature rises. You know this by personal experience if you have ever had to inflate a bicycle tire with a hand pump. As you pump, the compressed air gets warm and makes the pump warm to the touch. The same effect happens on a larger scale in southern California during Santa Ana winds or downwind from the Rocky Mountains when there are Chinook winds. Both of these strong winds blow from the mountains down to the lowlands. Even though the mountain air is cold, the winds that reach low elevations can be very hot. (Chinook winds have been known to raise the temperature by as much as 27°C, or 49°F, in only 2 minutes!) The explanation is compression. Air blown downhill by the winds is compressed by the greater air pressure at lower altitudes, and this compression raises the temperature of the air.

Expanding gases tend to drop in temperature. When you open a bottle of carbonated beverage, the gases trapped in the bottle expand and cool down. The cooling can be so great that a little cloud forms within the neck of the bottle. Clouds form in the atmosphere in the same way. Rising air cools as it goes to higher altitudes, where the pressure is lower, and the cooling makes water in the air condense into droplets.

Here's an experiment you can do to feel the cooling of expanding gases. Your breath is actually quite warm, as you can feel if you open your mouth wide, hold the back of your hand next to your mouth, and exhale. But if you bring your lips together to form an "o" and again blow on your hand, your breath feels cool. In the second case, your exhaled breath has to expand as it passes between your lips to the outside air. That makes its temperature drop.

this surrounding material. By tapping this fresh supply of hydrogen, a star manages to eke out a few million more years of main-sequence existence.

Main-Sequence Stars of Less than 0.4 $M_\odot$: Consuming All Their Hydrogen

The story is somewhat different for the least massive main-sequence stars, with masses between 0.08 $M_\odot$ (the minimum mass for sustained thermonuclear reactions to take place in a star's core) and about 0.4 $M_\odot$. These stars, of spectral class M, are called **red dwarfs** because they are small in size and have a red color due to their low surface temperature. They are also very numerous; about 85% of all stars in the Milky Way Galaxy are red dwarfs.

In a red dwarf, helium does *not* accumulate in the core to the same extent as in the Sun's core. The reason is that in a red dwarf there are convection cells of rising and falling gas that extend throughout the star's volume and penetrate into the core (see Figure 18-12*c*). These convection cells drag helium outward from the core and replace it with hydrogen from the outer layers (Figure 19-3). The fresh hydrogen can undergo thermonuclear fusion that releases energy and makes additional helium. This helium is then dragged out of the core by convection and replaced by even more hydrogen from the red dwarf's outer layers.

As a consequence, over a red dwarf's main-sequence lifetime essentially all of the star's hydrogen can be consumed and converted to helium. The core temperature and pressure in a red dwarf is less than in the Sun, so thermonuclear reactions happen more slowly than in our Sun. Calculations indicate that it takes hundreds of billions of years for a red dwarf to convert all of its hydrogen completely to helium. The present age of the universe is only 13.7 billion years, so there has not yet been time for any red dwarfs to become pure helium.

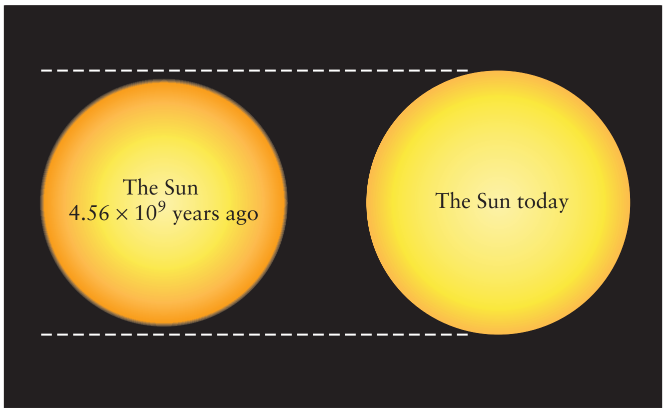


Figure 19-2

The Zero-Age Sun and Today's Sun Over the past 4.56×10^9 years, much of the hydrogen in the Sun's core has been converted into helium, the core has contracted a bit, and the Sun's luminosity has gone up by about 40%. These changes in the core have made the Sun's outer layers expand in radius by 6% and increased the surface temperature from 5500 K to 5800 K.

A Star's Mass Determines Its Main-Sequence Lifetime

The main-sequence lifetime of a star depends critically on its mass. As Table 19-1 shows, massive stars have short main-sequence lifetimes because they are also very luminous (see Section 17-9, and particularly the *Cosmic Connections* figure for Chapter 17). To emit energy so rapidly, these stars must be

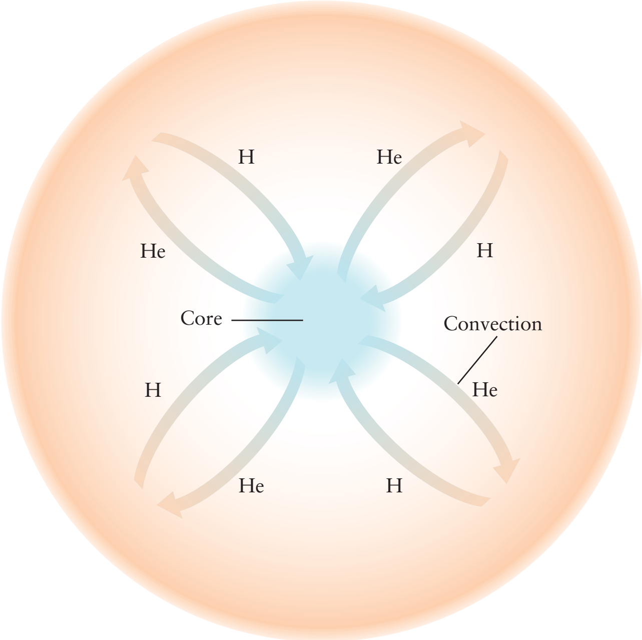


Figure 19-3

A Fully Convective Red Dwarf In a red dwarf—a main-sequence star with less than about 0.4 solar masses—helium (He) created in the core by thermonuclear reactions is carried to the star's outer layers by convection. Convection also brings fresh hydrogen (H) from the outer layers into the core. This process continues until the entire star is helium.

depleting the hydrogen in their cores at a prodigious rate. Hence, even though a massive O or B main-sequence star contains much more hydrogen fuel in its core than is in the entire volume of a red dwarf of spectral class M, the O or B star exhausts its hydrogen much sooner. High-mass O and B stars gobble up the available hydrogen fuel in only a few million years, while red dwarf stars of very low mass take hundreds of billions of years to use up their hydrogen. Thus, a main-sequence star's mass determines not only its luminosity and spectral type, but also how long it can remain a main-sequence star (see Box 19-2 for details).

We saw in Section 18-4 how more-massive stars evolve more quickly through the protostar phase to become main-sequence stars (see Figure 18-10). In general, the more massive the star, the more rapidly it goes through *all* the phases of its life. Still, most of the stars we are able to detect are in their main-sequence phase, because this phase lasts so much longer than other luminous phases.

In the remainder of this chapter we will look at the luminous phases that can take place after the end of a star's main-sequence lifetime. (In Chapters 20, 21, and 22 we will explore the final phases of a star's existence, when it ceases to have an appreciable luminosity.)

19-2 When core hydrogen fusion ceases, a main-sequence star like the Sun becomes a red giant

Like so many properties of stars, what happens at the end of a star's main-sequence lifetime depends on its mass. If the star is a red dwarf of less than about 0.4 $M_\odot$, after hundreds of billions of years the star has converted all of its hydrogen to helium. It is possible for helium to undergo thermonuclear fusion, but this requires temperatures and pressures far higher than those found within a red dwarf. Thus, a red dwarf will end its life as an inert ball of helium that still glows due to its internal heat. As it radiates energy into space, it slowly cools and shrinks. This slow, quiet demise is the ultimate fate of the 85% of stars in the Milky Way that are red dwarfs. (As we have seen, there has not yet been time in the history of the universe for any red dwarf to reach this final stage in its evolution.)

What is the fate of stars more massive than about 0.4 $M_\odot$, including the Sun? As we will see, the late stages of their evolution are far more dramatic. Studying these stages will give us insight into the fate of our solar system and of life on Earth.

Stars of 0.4 $M_\odot$ or Greater: From Main-Sequence Star to Red Giant

When a star of at least 0.4 solar masses reaches the end of its main-sequence lifetime, all of the hydrogen in its core has been used up and hydrogen fusion ceases there. In this new stage, hydrogen fusion continues only in the hydrogen-rich material just outside the core, a situation called **shell hydrogen fusion.** At first, this process occurs only in the hottest region just outside the core,

Table 19-1 Approximate Main-Sequence Lifetimes

Mass ($M_\odot$)	Surface temperature (K)	Spectral class	Luminosity ($L_\odot$)	Main-sequence lifetime (10^6 years)
25	35,000	O	80,000	4
15	30,000	B	10,000	15
3	11,000	A	60	800
1.5	7000	F	5	4500
1.0	6000	G	1	12,000
0.75	5000	K	0.5	25,000
0.50	4000	M	0.03	700,000

The main-sequence lifetimes were estimated using the relationship $t \propto 1/M^{2.5}$ (see Box 19-2).

BOX 19-2 Tools of the Astronomer's Trade

Main-Sequence Lifetimes

Hydrogen fusion converts a portion of a star's mass into energy. We can use Einstein's famous equation relating mass and energy to calculate how long a star will remain on the main sequence.

Suppose that M is the mass of a star and f is the fraction of the star's mass that is converted into energy by hydrogen fusion. The total energy E supplied by the hydrogen fusion can be expressed as

$$E = fMc^2$$

In this equation c is the speed of light.

This energy from hydrogen fusion is released gradually over millions or billions of years. If L is the star's luminosity (energy released per unit time) and t is the star's main-sequence lifetime (the total time over which the hydrogen fusion occurs), then we can write

$$L = \frac{E}{t}$$

(Actually, this equation is only an approximation. A star's luminosity is not quite constant over its entire main-sequence lifetime. But the variations are not important for our purposes.) We can rewrite this equation as

$$E = Lt$$

From this equation and $E = fMc^2$, we see that

$$Lt = fMc^2$$

We can rearrange this equation as

$$t = \frac{fMc^2}{L}$$

Thus, a star's lifetime on the main sequence is proportional to its mass (M) divided by its luminosity (L). Using the symbol $\propto$ to denote "is proportional to," we write

$$t \propto \frac{M}{L}$$

We can carry this analysis further by recalling that main-sequence stars obey the mass-luminosity relation (see Section 17-9, especially the *Cosmic Connections* figure). The distribution of data on the graph in the *Cosmic Connections* figure in Section 17-9 tells us that a star's luminosity is roughly proportional to the 3.5 power of its mass:

$$L \propto M^{3.5}$$

Substituting this relationship into the previous proportionality, we find that

$$t \propto \frac{M}{M^{3.5}} = \frac{1}{M^{2.5}} = \frac{1}{M^2\sqrt{M}}$$

This approximate relationship can be used to obtain rough estimates of how long a star will remain on the main sequence. It is often convenient to relate these estimates to the Sun (a typical 1-$M_\odot$ star), which will spend 1.2×10^{10} years on the main sequence.

EXAMPLE: How long will a star whose mass is 4 $M_\odot$ remain on the main sequence?

Situation: Given the mass of a star, we are asked to determine its main-sequence lifetime.

Tools: We use the relationship $t \propto 1/M^{2.5}$.

Answer: The star has 4 times the mass of the Sun, so it will be on the main sequence for approximately

$$\frac{1}{4^{2.5}} = \frac{1}{4^2\sqrt{4}} = \frac{1}{32} \text{ times the Sun's main-sequence lifetime}$$

Thus, a 4-$M_\odot$ main-sequence star will fuse hydrogen in its core for about $(1/32) \times 1.2 \times 10^{10}$ years, or about 4×10^8 (400 million) years.

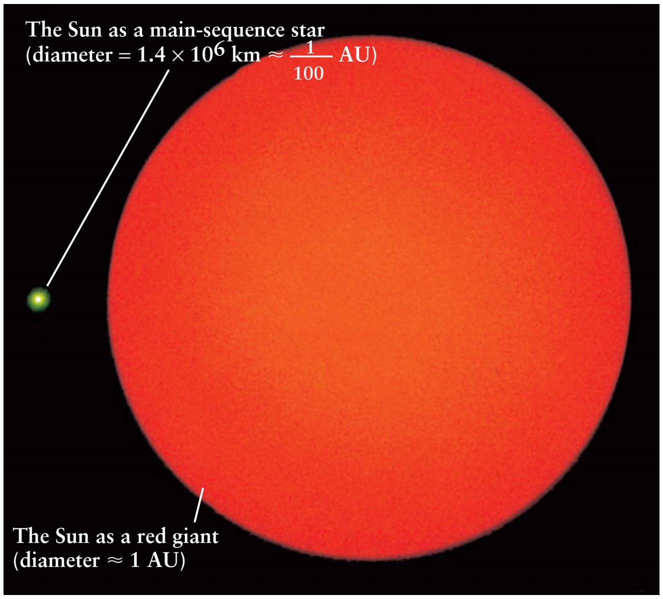
Review: Our result makes sense: A star more massive than the Sun must have a shorter main-sequence lifetime.

where the hydrogen fuel has not yet been exhausted. Outside this region, no fusion reactions take place.

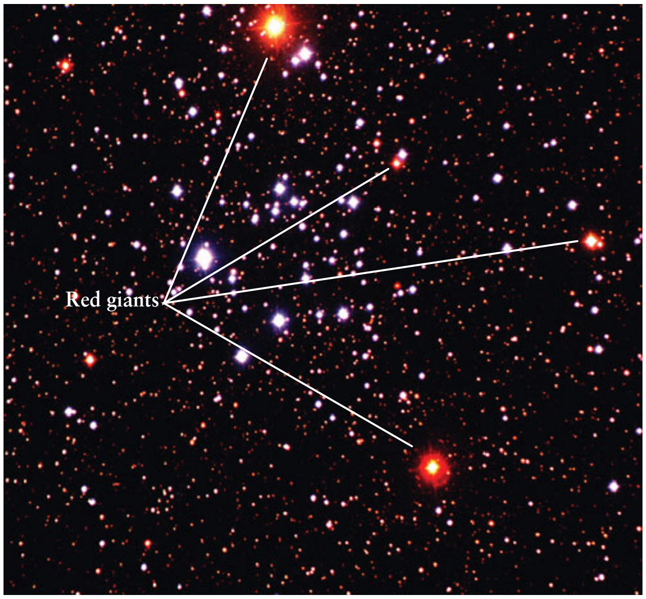
Strangely enough, the end of core hydrogen fusion *increases* the core's temperature. When thermonuclear reactions first cease in the core, nothing remains to generate heat there. Hence, the core starts to cool and the pressure in the core starts to decrease. This pressure decrease allows the star's core to again compress under the weight of the outer layers. As the core contracts, its temperature again increases, and heat begins to flow outward from the core even though no nuclear reactions are taking place there. (Technically, gravitational energy is converted into thermal energy, as in Kelvin-Helmholtz contraction; see Section 8-4 and Section 16-1).

In the transition from main-sequence star to red giant, the star's core contracts while its outer layers expand

This new flow of heat warms the gases around the core, increasing the rate of shell hydrogen fusion and making the shell eat further outward into the surrounding matter. Helium produced by reactions in the shell falls down onto the core, which continues



(a) The Sun today and as a red giant



(b) Red giant stars in the star cluster M50 R I V U X G

WEB LINK 19.2

Figure 19-4

Red Giants **(a)** The present-day Sun produces energy in a hydrogen-fusing core about 100,000 km in diameter. Some 7.6 billion years from now, when the Sun becomes a red giant, its energy source will be a shell only about 30,000 km in diameter within which hydrogen fusion will take place at a furious rate. The Sun's luminosity will be about 2000 times greater than today, and the increased luminosity will make the Sun's outer layers expand to approximately 100 times their present size. **(b)** This composite of visible and infrared images shows bright red giant stars in the open cluster M50 in the constellation Monoceros (the Unicorn). (T. Credner and S. Kohle, Calar Alto Observatory)

to contract and heat up as it gains mass. Over the course of hundreds of millions of years, the core of a 1-$M_\odot$ star compresses to about one-third of its original radius, while its central temperature increases from about 15 million (1.5×10^7) K to about 100 million (10^8) K.

During this post–main-sequence phase, the star's outer layers expand just as dramatically as the core contracts. As the hydrogen-fusing shell works its way outward, egged on by heat from the contracting core, the star's luminosity increases substantially. This increases the star's internal pressure and makes the star's outer layers expand to many times their original radius. This tremendous expansion causes those layers to cool down, and the star's surface temperature drops (see Box 19-1). Once the temperature of the star's bloated surface falls to about 3500 K, the gases glow with a reddish hue, in accordance with Wien's law (see Figure 17-7*a*). The star is then appropriately called a **red giant** (Figure 19-4). Thus, we see that red-giant stars are former main-sequence stars that have evolved into a new stage of existence. We can summarize these observations as a general rule:

Stars join the main sequence when they begin hydrogen fusion in their cores. They leave the main sequence and become giant stars when the core hydrogen is depleted.

Red-giant stars undergo substantial **mass loss** because of their large diameters and correspondingly weak surface gravity. This makes it relatively easy for gases to escape from the red giant into space. Mass loss can be detected in a star's spectrum, because gas escaping from a red giant toward a telescope on the Earth produces narrow absorption lines that are slightly blueshifted by the Doppler effect (review Figure 5-26). Typical observed blueshifts correspond to a speed of about 10 km/s. A typical red giant loses roughly 10^{-7} $M_\odot$ of matter per year. For comparison, the Sun's present-day mass loss rate is only 10^{-14} $M_\odot$ per year. Hence, an evolving star loses a substantial amount of mass as it becomes a red giant. Figure 19-5 shows a star losing mass in this way.

The Distant Future of Our Solar System

We can use these ideas to peer into the future of our planet and our solar system. The Sun's luminosity will continue to increase as it goes through its main-sequence lifetime, and the temperature of the Earth will increase with it. One and a half billion years from now the Earth's average surface temperature will be 50°C (122°F). By 3½ billion years from now the surface temperature of the Earth will exceed the boiling temperature of water. All the oceans will boil away, and the Earth will become utterly incapable of supporting life. These increasingly hostile conditions will pose the ultimate challenge to whatever intelligent beings might inhabit the Earth in the distant future.

About 7 billion years from now, our Sun will finish converting hydrogen into helium at its core. As the Sun's core contracts, its atmosphere will expand to envelop Mercury and perhaps reach to the orbit of Venus. Roughly 700 million years after leaving the main sequence, our red-giant Sun will have swollen to a diameter

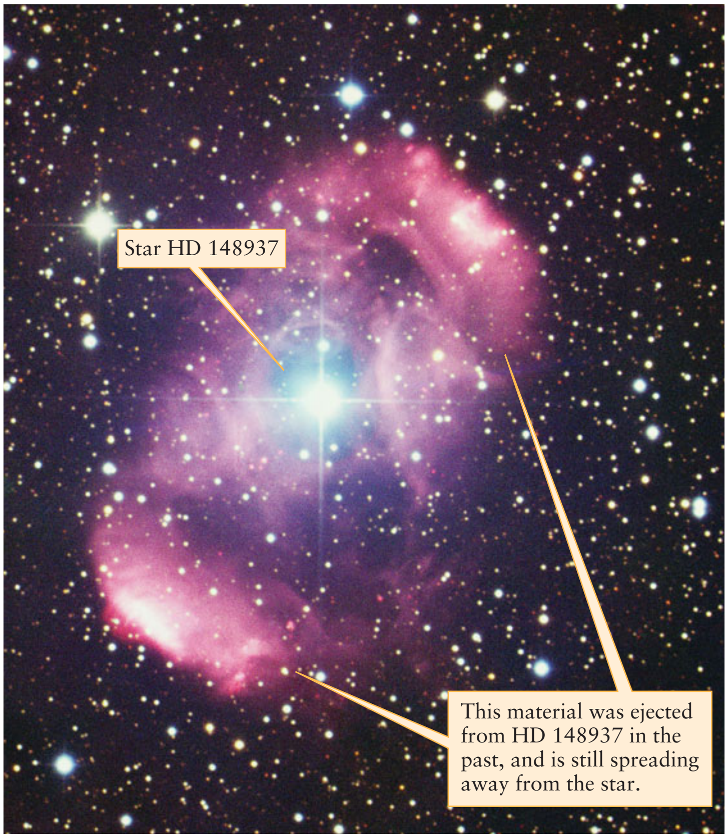


Figure 19-5 R I **V** U X G

A Mass-Loss Star As stars age and become giant stars, they expand tremendously and shed matter into space. This star, HD 148937, is losing matter at a high rate. Other strong outbursts in the past ejected the clouds that surround HD 148937. These clouds absorb ultraviolet radiation from the star, which excites the atoms in the clouds and causes them to glow. The characteristic red color of the clouds reveals the presence of hydrogen (see Section 5-6) that was ejected from the star's outer layers. (David Malin, Anglo-Australian Observatory)

of about 1 AU—roughly 100 times larger than its present size—and its surface temperature will have dropped to about 3500 K. Shell hydrogen fusion will proceed at such a furious rate that our star will shine with the brightness of 2000 present-day Suns. Some of the inner planets will be vaporized, and the thick atmospheres of the outer planets will evaporate away to reveal tiny, rocky cores. Thus, in its later years, the aging Sun may destroy the planets that have accompanied it since its birth.

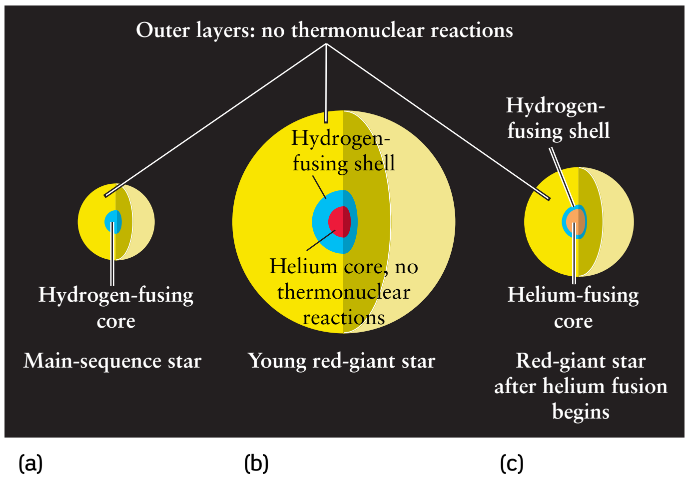
19-3 Fusion of helium into carbon and oxygen begins at the center of a red giant

When a star with a mass greater than 0.4 $M_\odot$ first changes from a main-sequence star (Figure 19-6*a*) to a red giant (Figure 19-6*b*), its hydrogen-fusing shell surrounds a small, compact core of almost pure helium. In a red giant of moderately low mass, which the Sun will become 7 billion years from now, the dense helium core is about twice the diameter of the Earth. Most of this core helium was *produced* by thermonuclear reactions during the star's main-sequence lifetime; during the red-giant era, this helium will *undergo* thermonuclear reactions.

Core Helium Fusion

Helium, the "ash" of hydrogen fusion, is a potential nuclear fuel: **Helium fusion,** the thermonuclear fusion of helium nuclei to make heavier nuclei, releases energy. But this reaction cannot take place within the core of our present-day Sun because the temperature there is too low. Each helium nucleus contains two protons, so it has twice the positive electric charge of a hydrogen nucleus, and there is a much stronger electric repulsion between two helium nuclei than between two hydrogen nuclei. For helium nuclei to overcome this repulsion and get close enough to fuse together, they must be moving at very high speeds, which means that the temperature of the helium gas must be very high. (For more on the relationship between the temperature of a gas and the speed of atoms in the gas, see Box 7-2.)

When a star first becomes a red giant, the temperature of the contracted helium core is still too low for helium nuclei to fuse. But as the hydrogen-fusing shell adds mass to the helium core, the core contracts even more, further increasing the star's central temperature. When the central temperature finally reaches 100 million (10^8) K, **core helium fusion**—that is, thermonuclear fusion of helium in the core—begins. As a result, the aging star again has a central energy source for the first time since leaving the main sequence (Figure 19-6*c*).



INTERACTIVE EXERCISE 19.1

Figure 19-6

Stages in the Evolution of a Star with More than 0.4 Solar Masses **(a)** During the star's main-sequence lifetime, hydrogen is converted into helium in the star's core. **(b)** When the core hydrogen is exhausted, hydrogen fusion continues in a shell, and the star expands to become a red giant. **(c)** When the temperature in the red giant's core becomes high enough because of contraction, core helium fusion begins (right). (These three pictures are *not* drawn to scale. The star is about 100 times larger in its red-giant phase than in its main-sequence phase, then shrinks somewhat when core helium fusion begins.)

Helium fusion occurs in two steps. First, two helium nuclei combine to form a beryllium nucleus:

$$^{4}\mathrm{He} + {}^{4}\mathrm{He} \rightarrow {}^{8}\mathrm{Be}$$

This particular beryllium isotope, which has four protons and four neutrons, is very unstable and breaks into two helium nuclei soon after it forms. However, in the star's dense core a third helium nucleus may strike the ^{8}Be nucleus before it has a chance to fall apart. Such a collision creates a stable isotope of carbon and releases energy in the form of a gamma-ray photon (γ):

$$^{8}\mathrm{Be} + {}^{4}\mathrm{He} \rightarrow {}^{12}\mathrm{C} + \gamma$$

This process of fusing three helium nuclei to form a carbon nucleus is called the **triple alpha process,** because helium nuclei (^{4}He) are also called **alpha particles** by nuclear physicists. Some of the carbon nuclei created in this process can fuse with an additional helium nucleus to produce a stable isotope of oxygen and release more energy:

$$^{12}\mathrm{C} + {}^{4}\mathrm{He} \rightarrow {}^{16}\mathrm{O} + \gamma$$

Thus, both carbon and oxygen make up the "ash" of helium fusion. The *Cosmic Connections* figure summarizes the reactions involved in helium fusion.

It is interesting to note that ^{12}C and ^{16}O are the most common isotopes of carbon and oxygen, respectively; the vast majority of the carbon atoms in your body are ^{12}C, and almost all the oxygen you breathe is ^{16}O. We will discuss the significance of this in Section 19-5.

The second step in the triple alpha process and the process of oxygen formation both release energy. The onset of these reactions reestablishes thermal equilibrium and prevents any further gravitational contraction of the star's core. A mature red giant fuses helium in its core for a much shorter time than it spent fusing hydrogen in its core as a main-sequence star. For example, in the distant future the Sun will sustain helium core fusion for only about 100 million years. (While this is going on, hydrogen fusion is still continuing in a shell around the core.)

The Helium Flash and Electron Degeneracy

How helium fusion begins at a red giant's center depends on the mass of the star. In high-mass red giants (greater than about 2 to 3 $M_{\odot}$), helium fusion begins gradually as temperatures in the star's core approach 10^{8} K. In red giants with a mass less than about 2 to 3 $M_{\odot}$, helium fusion begins explosively and suddenly, in what is called the **helium flash.** Table 19-2 summarizes these differences.

The helium flash occurs because of unusual conditions that develop in the core of a moderately low-mass star as it becomes a red giant. To appreciate these conditions we must first understand how an ordinary gas behaves. Then we can explore how the densely packed electrons at the star's center alter this behavior.

When a gas is compressed, it usually becomes denser and warmer. To describe this process, scientists use the convenient concept of an **ideal gas,** which has a simple relationship between pressure, temperature, and density. Specifically, the pressure exerted by an ideal gas is directly proportional to both the density and the temperature of the gas. Many real gases actually behave like an ideal gas over a wide range of temperatures and densities.

Table 19-2 How Helium Core Fusion Begins in Different Red Giants

Mass of star	Onset of helium fusion in core
More than about 0.4 but less than 2–3 solar masses	Explosive (helium flash)
More than 2–3 solar masses	Gradual

Stars with less than about 0.4 solar masses do not become red giants (see Section 19-2).

Under most circumstances, the gases inside a star act like an ideal gas. If the gas expands, it cools down, and if it is compressed, it heats up (see Box 19-1). This behavior serves as a safety valve, ensuring that the star remains in thermal equilibrium (see Section 16-2). For example, if the rate of thermonuclear reactions in the star's core should increase, the additional energy releases heat and expands the core. This cools the core's gases and slows the rate of thermonuclear reactions back to the original value. Conversely, if the rate of thermonuclear reactions should decrease, the core will cool down and compress under the pressure of the overlying layers. The compression of the core will make its temperature increase, thus speeding up the thermonuclear reactions and returning them to their original rate.

In a red giant with a mass between about 0.4 $M_{\odot}$ and 2-3 $M_{\odot}$, however, the core behaves very differently from an ideal gas. The core must be compressed tremendously in order to become hot enough for helium fusion to begin. At these extreme pressures and temperatures, the atoms are completely ionized, and most of the core consists of nuclei and detached electrons. Eventually, the free electrons become so closely crowded that a limit to further compression is reached, as predicted by a remarkable law of quantum mechanics called the **Pauli exclusion principle.** Formulated in 1925 by the Austrian physicist Wolfgang Pauli, this principle states that two electrons cannot simultaneously occupy the same quantum state. A quantum state is a particular set of circumstances concerning locations and speeds that are available to a particle. In the submicroscopic world of electrons, the Pauli exclusion principle is analogous to saying that you can't have two things in the same place at the same time.

Before helium fusion begins in a red giant's core, the helium gas there behaves like a metal

Just before the onset of helium fusion, the electrons in the core of a low-mass star are so closely crowded together that any further compression would violate the Pauli exclusion principle. Because the electrons cannot be squeezed any closer together, they produce a powerful pressure that resists further core contraction.

This phenomenon, in which closely packed particles resist compression as a consequence of the Pauli exclusion principle, is called **degeneracy.** Astronomers say that the electrons in the helium-rich core of a low-mass red giant are "degenerate," and

COSMIC CONNECTIONS

Helium Fusion In A Red Giant

A star becomes a red giant after the fusion of hydrogen into helium in its core has come to an end. As the red giant's core shrinks and heats up, a new cycle of reactions can occur that create the even heavier elements carbon and oxygen.

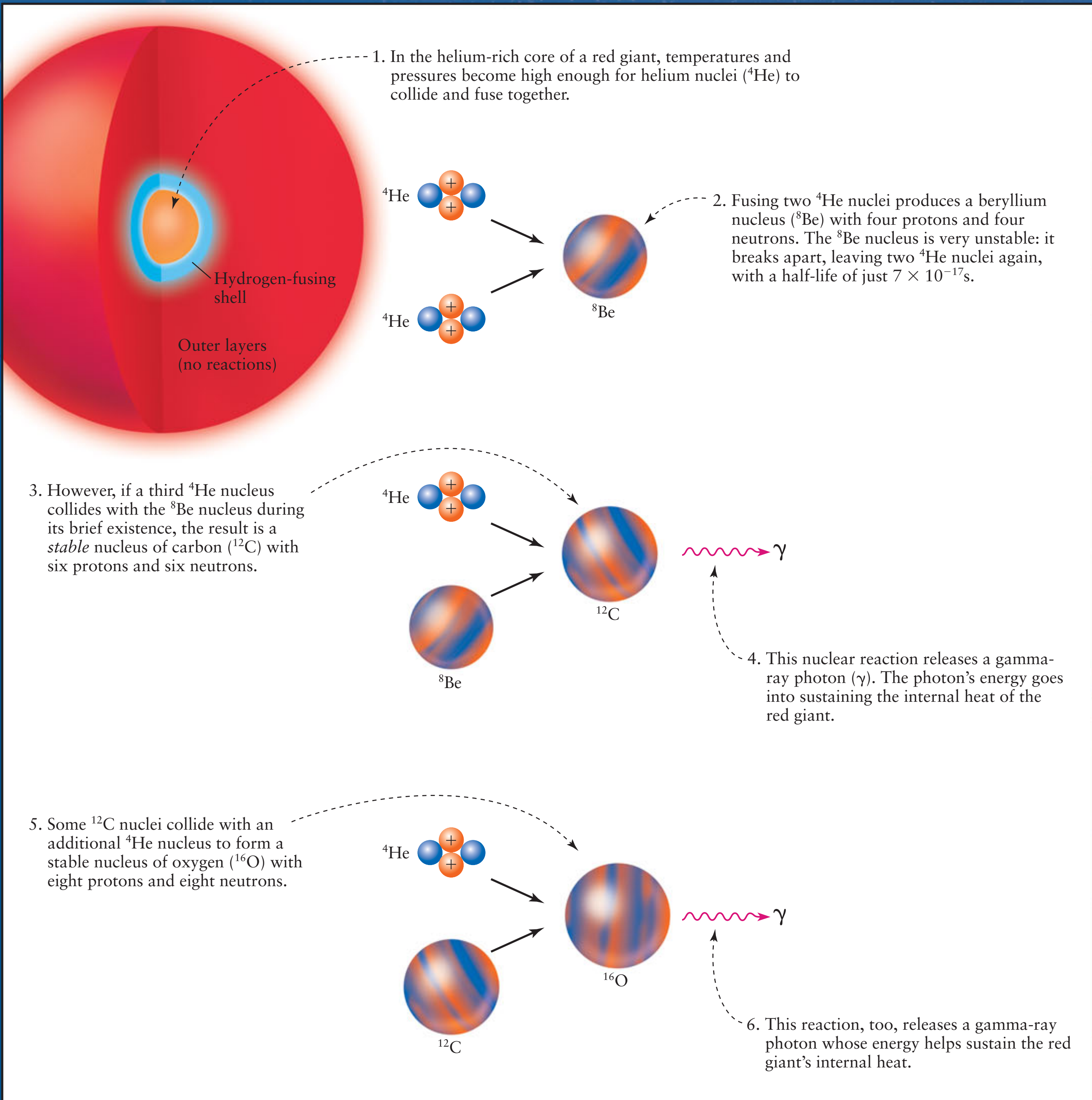


Figure 19-7 R I V U X G

Degenerate Electrons The electrons in an ordinary piece of metal, like the chrome grille on this classic car, are so close together that they are affected by the Pauli exclusion principle. The resulting degenerate electron pressure helps make metals strong and difficult to compress. A more powerful version of this same effect happens inside the cores of low-mass red-giant stars. (Santokh Kochar/PhotoDisc)

that the core is supported by **degenerate-electron pressure.** This degenerate pressure, unlike the pressure of an ideal gas, does not depend on temperature. Remarkably, you can find degenerate electrons on Earth in an ordinary piece of metal (Figure 19-7).

When the temperature in the core of a low-mass red giant reaches the high level required for the triple alpha process, energy begins to be released. The helium heats up, which makes the triple alpha process happen even faster. However, the pressure provided by the degenerate electrons is independent of the temperature, so the pressure does not change. Without the "safety valve" of increasing pressure, the star's core cannot expand and cool. The rising temperature causes the helium to fuse at an ever-increasing rate, producing the helium flash.

Eventually, the temperature becomes so high that the electrons in the core are no longer degenerate. The electrons then behave like an ideal gas and the star's core expands, terminating the helium flash. These events occur so rapidly that the helium flash is over in seconds, after which the star's core settles down to a steady rate of helium fusion.

CAUTION! The term "helium flash" might give you the impression that a star emits a sudden flash of light when the helium flash occurs. If this were true, it would be an incredible sight. During the brief time interval when the helium flash occurs, the helium-fusing core is 10^{11} times more luminous than the present-day Sun, comparable to the total luminosity of all the stars in the Milky Way Galaxy! But, in fact, the helium flash has no immediately visible consequences—for two reasons. First, much of the energy released during the helium flash goes into heating the core and terminating the degenerate state of the electrons. Second, the energy that does escape the core is largely absorbed by the star's outer layers, which are quite opaque (just like the Sun's present-day interior; see Section 16-2). Therefore, the explosive drama of the helium flash takes place where it cannot be seen directly.

The Continuing Evolution of a Red Giant

Whether a helium flash occurs or not, the onset of core helium fusion actually causes a *decrease* in the luminosity of the star. This is the opposite of what you might expect—after all, turning on a new energy source should make the luminosity greater, not less. What happens is that after the onset of core helium fusion, a star's superheated core expands like an ideal gas. (If the star is of sufficiently low mass to have had a degenerate core, the increased temperature after the helium flash makes the core too hot to remain degenerate. Hence, these stars also end up with cores that behave like ideal gases.) Temperatures drop around the expanding core, so the hydrogen-fusing shell reduces its energy output and the star's luminosity decreases. This allows the star's outer layers to contract and heat up. Consequently, a post–helium-flash star is less luminous, hotter at the surface, and smaller than a red giant.

Core helium fusion lasts for only a relatively short time. Calculations suggest that a 1-$M_\odot$ star like the Sun sustains core

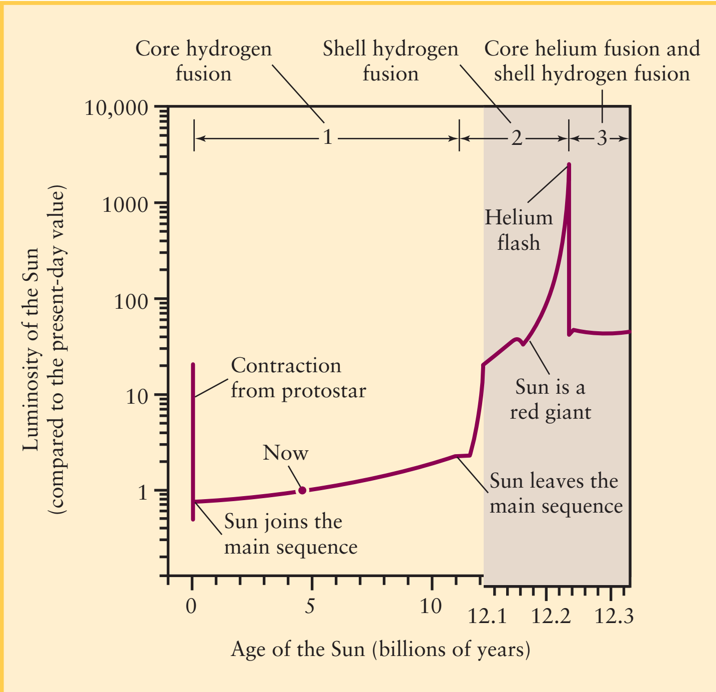


Figure 19-8

Stages in the Evolution of the Sun This diagram shows how the luminosity of the Sun (a 1-$M_\odot$ star) changes over time. The Sun began as a protostar whose luminosity decreased rapidly as the protostar contracted. Once established as a main-sequence star with core hydrogen fusion, the Sun's luminosity increases slowly over billions of years. The post-main-sequence evolution is much more rapid, so a different time scale is used in the right-hand portion of the graph. (Adapted from Mark A. Garlick, based on calculations by I.-Juliana Sackmann and Kathleen E. Kramer) hydrogen fusion for about 12 billion (1.2×10^{10}) years, followed by about 250 million (2.5×10^8) years of shell hydrogen fusion leading up to the helium flash. After the helium flash, such a star can fuse helium in its core (while simultaneously fusing hydrogen in a shell around the core) for only 100 million (10^8) years, a mere 1% of its main-sequence lifetime. Figure 19-8 summarizes these evolutionary stages in the life of a 1-$M_\odot$ star. In Chapter 20 we will take up the story of what happens after a star has consumed all the helium in its core.

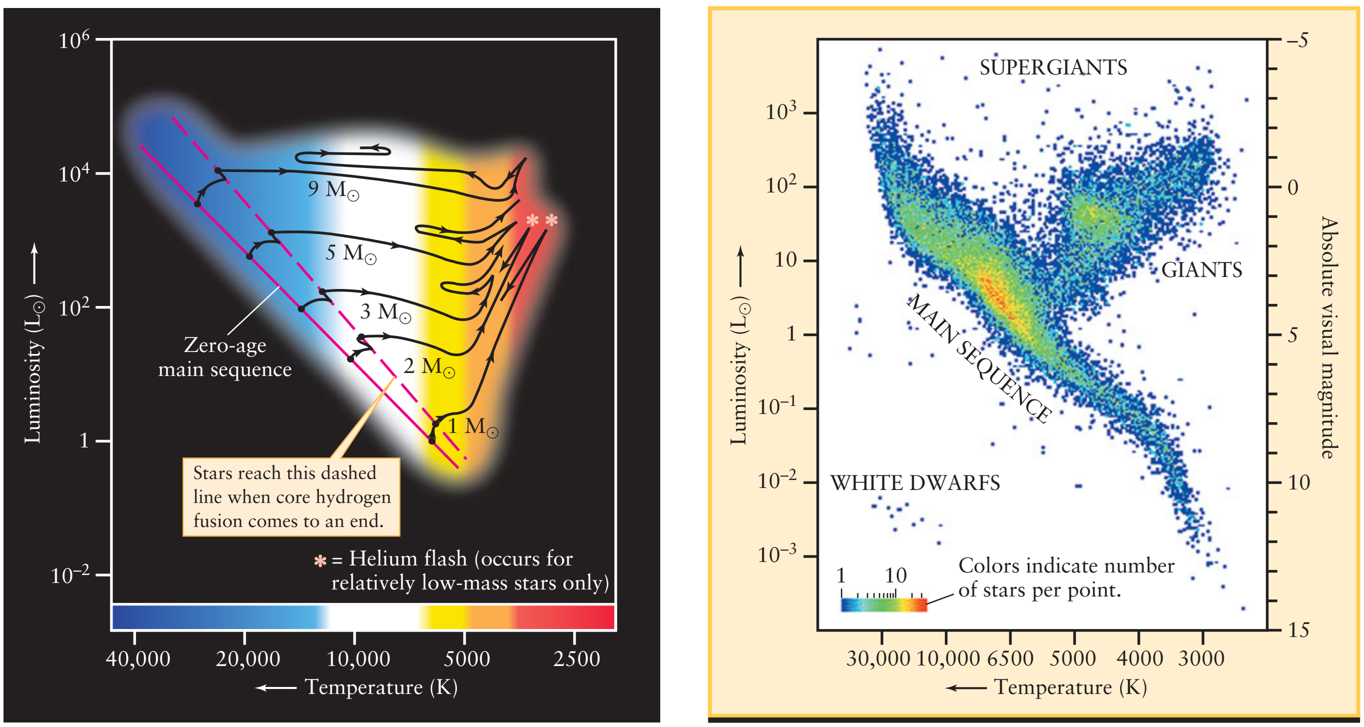
Here is the story of post–main-sequence evolution in its briefest form: Before the beginning of core helium fusion, the star's core compresses and the outer layers expand, and just after core helium fusion begins, the core expands and the outer layers compress. We will see in Chapter 20 that this behavior, in which the inner and outer regions of the star change in opposite ways, occurs again and again in the final stages of a star's evolution.

19-4 H-R diagrams and observations of star clusters reveal how red giants evolve

To see how stars evolve during and after their main-sequence lifetimes, it is helpful to follow them on a Hertzsprung-Russell (H-R) diagram. On such a diagram, zero-age main-sequence stars lie along a line called the **zero-age main sequence**, or **ZAMS** (Figure 19-9*a*). These stars have just emerged from their protostar stage, are steadily fusing hydrogen into helium in their cores, and have attained hydrostatic equilibrium. With the passage of time, hydrogen in a main-sequence star's core is converted to helium, the luminosity slowly increases, the star slowly expands, and the star's position on the H-R diagram inches away from the ZAMS. As a result, the main sequence on an H-R diagram is a fairly broad band rather than a narrow line (Figure 19-9*b*).

Post-Main-Sequence Evolution on an H-R Diagram

The dashed line in Figure 19-9*a* denotes stars whose cores have been exhausted of hydrogen and in which core hydrogen fusion has ceased. These stars have reached the ends of their main-sequence lifetimes. From there, the points representing high-mass stars (3 $M_\odot$, 5 $M_\odot$, and 9 $M_\odot$) move rapidly from left to right across the H-R diagram. This means that, although the star's surface temperature is decreasing, its surface area is increasing at a rate that keeps its overall luminosity roughly constant. During this transition, the star's core contracts and its outer layers expand as energy flows outward from the hydrogen-fusing shell.



(a) Post-main-sequence evolutionary tracks of five stars with different mass

(b) H-R diagram of 20,853 stars—note the width of the main sequence

WEB LINK 19.3

Figure 19-9

H-R Diagrams of Stellar Evolution on and off the Main Sequence **(a)** The two lowest-mass stars shown here (1 $M_\odot$ and 2 $M_\odot$) undergo a helium flash at their centers, as shown by the asterisks. In the high-mass stars, core helium fusion ignites more gradually where the evolutionary tracks make a sharp downward turn in the red-giant region on the right hand side of the H-R diagram. **(b)** Data from the *Hipparcos* satellite (see Section 17-1) was used to create this H-R diagram. The thickness of the main sequence is due in large part to stars evolving during their main-sequence lifetimes. (a: Adapted from I. Iben; b: Adapted from M. A. C. Perryman)

Just before core helium fusion begins, the evolutionary tracks of high-mass stars turn upward in the red-giant region of the H-R diagram (to the upper right of the main sequence). After core helium fusion begins, however, the cores of these stars expand, the outer layers contract, and the evolutionary tracks back away from these temporary peak luminosities. The tracks then wander back and forth in the red-giant region while the stars readjust to their new energy sources.

Figure 19-9*a* also shows the evolutionary tracks of two stars of moderately low mass (1 $M_\odot$ and 2 $M_\odot$). The onset of core helium fusion in these stars occurs with a helium flash, indicated by the red asterisks in the figure. As we saw in the previous section, after the helium flash, these stars shrink and become less luminous. The decrease in size is proportionately greater than the decrease in luminosity, and so the surface temperatures increase. Hence, after the helium flash, the evolutionary tracks for the 1-$M_\odot$ and 2-$M_\odot$ stars move down and to the left.

A Simulated Star Cluster: Tracking 4½ Billion Years of Stellar Evolution

We can summarize our understanding of stellar evolution from birth through the onset of helium fusion by following the evolution of a hypothetical cluster of stars. We saw in Section 18-6 that the stars that make up a cluster all begin to form at essentially the same time but have different initial masses. Hence, studying star clusters allows us to compare how stars of different masses evolve.

The eight H-R diagrams in Figure 19-10 are from a computer simulation of the evolution of 100 stars that all form at the same moment and differ only in initial mass. All 100 stars begin as cool protostars on the right side of the H-R diagram (see Figure 19-10*a*). The protostars are spread out on the diagram according to their masses, and the greater the mass, the greater the protostar's initial luminosity. As we saw in Section 18-3, the source of a protostar's luminosity is its gravitational energy. As the protostar contracts, this is converted to thermal energy and radiated into space.

The most massive protostars contract and heat up very rapidly. After only 5000 years, they have already moved across the H-R diagram toward the main sequence (see Figure 19-10*b*). After 100,000 years, these massive stars have ignited hydrogen fusion in their cores and have settled down on the main sequence as O stars (see Figure 19-10*c*). After 3 million years, stars of moderate mass have also ignited core hydrogen fusion and become main-sequence stars of spectral classes B and A (see Figure 19-10*d*). Meanwhile, low-mass protostars continue to inch their way toward the main sequence as they leisurely contract and heat up.

After 30 million years (see Figure 19-10*e*), the most massive stars have depleted the hydrogen in their cores and become red giants. These stars have moved from the upper left end of the main sequence to the upper right corner of the H-R diagram. (This simulation follows stars only to the red-giant stage, after which they are simply deleted from the diagram.) Intermediate-mass stars lie on the main sequence, while the lowest-mass stars are still in the protostar stage and lie above the main sequence.

After 66 million years (see Figure 19-10*f*), even the lowest-mass protostars have finally ignited core hydrogen fusion and have settled down on the main sequence as cool, dim, M stars. These lowest-mass stars can continue to fuse hydrogen in their cores for hundreds of billions of years.

In the final two H-R diagrams (Figures 19-10*g* and 19-10*h*), the main sequence gets "eaten away" from the upper left to the lower right as stars exhaust their core supplies of hydrogen and evolve into red giants. The stars that leave the main sequence between Figure 19-10*g* and Figure 19-10*h* have masses between about 1 $M_\odot$ and 3 $M_\odot$ and undergo the helium flash in their cores.

For all stars in this simulation, the giant stage lasts only a brief time compared to the star's main-sequence lifetime. Compared to a 1-$M_\odot$ star (see Figure 19-8), a more massive star has a shorter main-sequence lifetime *and* spends a shorter time as a giant star. Thus, at any given time, only a small fraction of the stellar population is passing through the giant stage. Hence, most of the stars we can see through telescopes are main-sequence stars. As an example, of the stars within 4.00 pc (13.05 ly) of the Sun listed in Appendix 4, only one—Procyon A—is presently evolving from a main-sequence star into a giant. Two other nearby stars are white dwarfs, an even later stage in stellar evolution that we will discuss in Chapter 20. (By contrast, most of the *brightest* stars listed in Appendix 5 are giants and supergiants. Although they make up only a small fraction of the stellar population, these stars stand out due to their extreme luminosity.)

Real Star Clusters: Cluster Ages and Turnoff Points

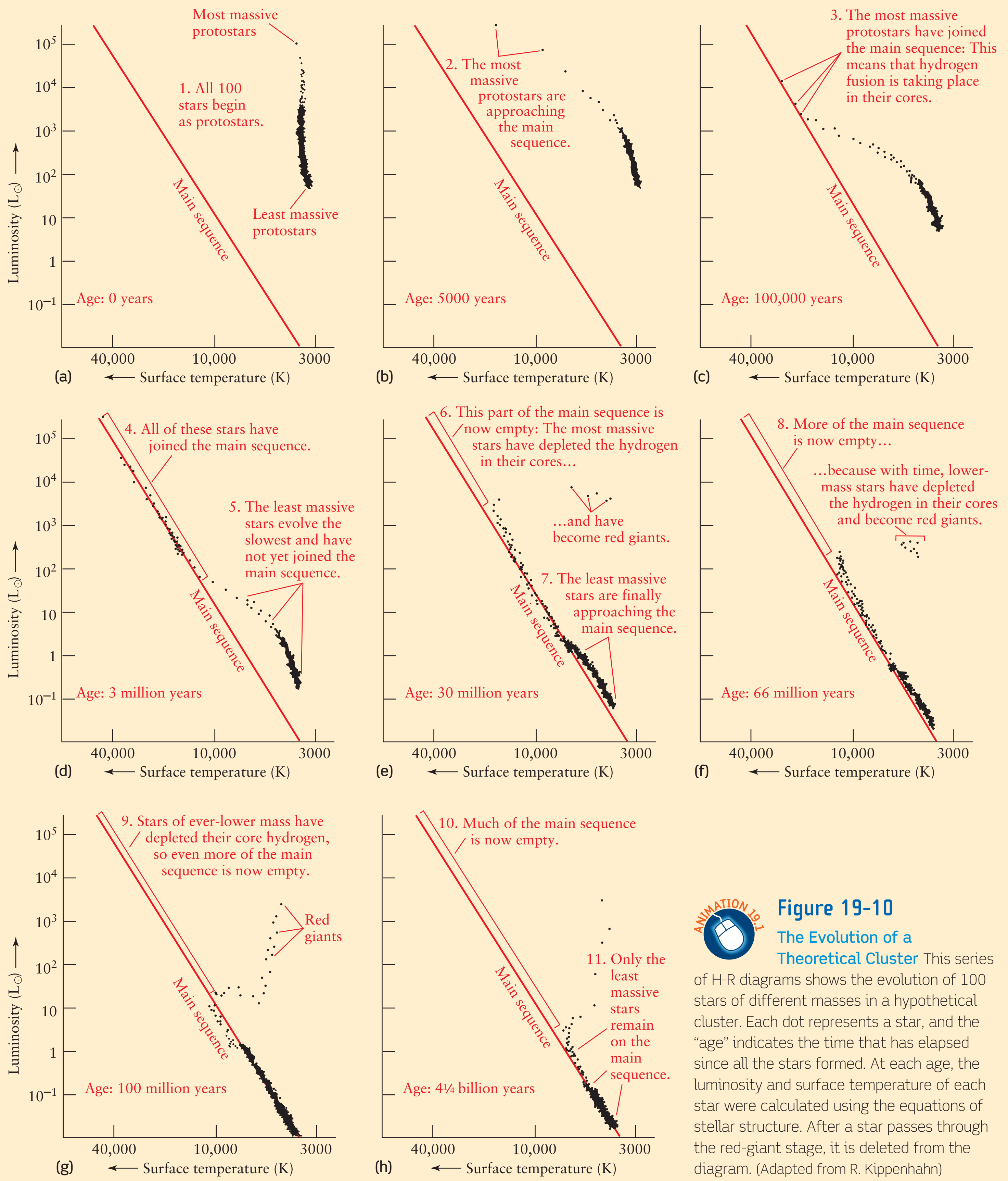
WEB LINK 19.4

The evolution of the hypothetical cluster displayed in Figure 19-10 helps us interpret what we see in actual star clusters. We can observe the early stages of stellar evolution in **open clusters,** which typically contain a few hundred to a few thousand stars. Many open clusters are just a few million years old, so their H-R diagrams resemble Figure 19-10*d*, 19-10*e*, or 19-10*f* (see Section 18-6, especially Figure 18-18 and Figure 18-19).

Figure 19-11 shows two open clusters of different ages. The nearer cluster, called M35, must be relatively young because it contains several dozen luminous, blue, high-mass main-sequence stars. These stars lie in the upper part of the main sequence on an H-R diagram. They have main-sequence lifetimes of only a few hundred million years, so M35 can be no older than that. Some of the most luminous stars in M35 are red or yellow in color; these are stars that ended their main-sequence lifetimes some time ago and have evolved into red giants. The H-R diagram for this cluster resembles Figure 19-10*g*.

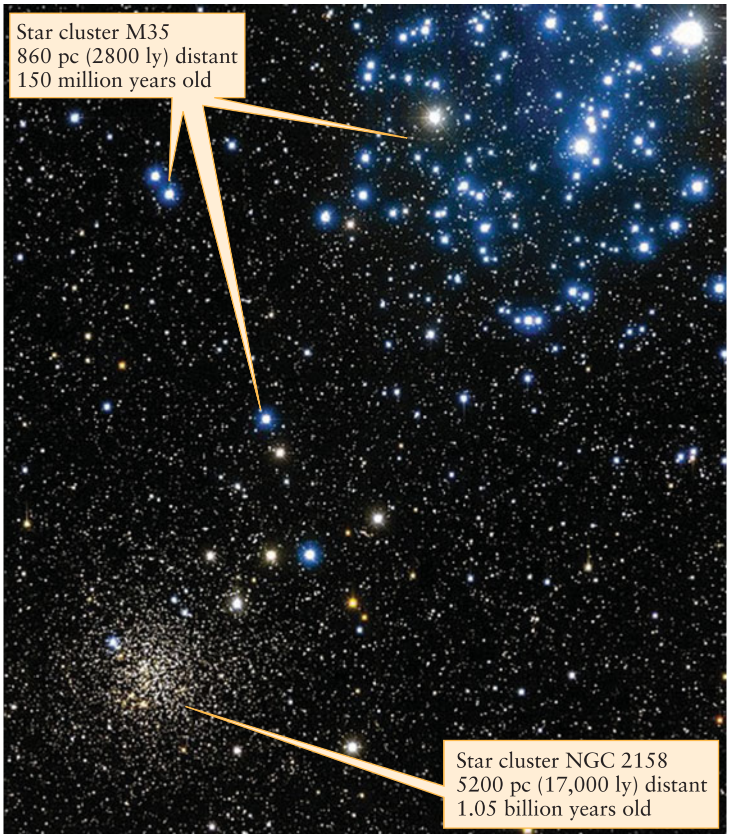
As a cluster ages and its stars end their main-sequence lifetimes, the cluster's color changes from blue to red

There are no high-mass blue main-sequence stars at all in NGC 2158, the more distant cluster shown in Figure 19-11. Any such stars that were once in NGC 2158 have long since come to the end of their main-sequence lifetimes. As a result, the main sequence in this cluster has been "eaten away" more than that of M35, leaving only stars that are yellow or red in color. This tells us that NGC 2158 must be older than M35 (compare Figures



ANIMATION 19.1

Figure 19-10 **The Evolution of a Theoretical Cluster** This series of H-R diagrams shows the evolution of 100 stars of different masses in a hypothetical cluster. Each dot represents a star, and the "age" indicates the time that has elapsed since all the stars formed. At each age, the luminosity and surface temperature of each star were calculated using the equations of stellar structure. After a star passes through the red-giant stage, it is deleted from the diagram. (Adapted from R. Kippenhahn)



WEB LINK 19.5

Figure 19-11 R I V U X G

Two Open Clusters The two clusters in this image, M35 and NGC 2158, lie in almost the same direction in the constellation Gemini. The nearer cluster, M35, has a number of luminous blue main-sequence stars with surface temperatures around 10,000 K as well as a few red giants. Hence its H-R diagram resembles that shown in Figure 19-10*g*, and its age is around 100 million years (more accurately, around 150 million). The more distant cluster, NGC 2158, has *no* blue main-sequence stars; long ago, all of these massive stars came to the end of their main-sequence lifetimes and became giants. The H-R diagram for NGC 2158 is intermediate between Figure 19-10*g* and Figure 19-10*h*, and its age (1.05 billion years) is as well. (Canada-France-Hawaii Telescope; J.-C. Cuillandre, CFHT; and Coelum)

19-10*g* and 19-10*h*). This example shows that as a cluster ages, it generally becomes redder in its average color.

WEB LINK 19.6

We can see even later stages in stellar evolution by studying **globular clusters,** so called because of their spherical shape. A typical globular cluster contains up to 1 million stars in a volume less than 100 parsecs across (Figure 19-12). Among these are many highly evolved post–main-sequence stars.

Globular clusters must be old, because they contain no high-mass main-sequence stars. To show this, you would measure the apparent magnitude (a measure of apparent brightness, which we introduced in Section 17-3) and color ratio of many stars in a globular cluster, then plot the data as shown in Figure 19-13. Such a **color-magnitude diagram** for a cluster is equivalent to an H-R diagram. The color ratio of a star tells you its surface temperature (as described in Section 17-4), and because all the stars in the cluster are at essentially the same distance from us, their relative brightnesses indicate their relative luminosities. What you would discover is that a globular cluster's main sequence has been "eaten away" even more extensively than the open cluster NGC 2158 in Figure 19-11. Hence, globular clusters must be even older than NGC 2158. In a typical globular cluster, all the main-sequence stars with masses more than about $1M_\odot$ or $2\ M_\odot$ evolved long ago into red giants. Only low-mass, slowly evolving stars still have core hydrogen fusion. (Compare Figure 19-13 with Figure 19-10*h*.)

CAUTION! The inset in Figure 19-12 shows something surprising: There are luminous *blue* stars in the ancient globular cluster M10. This seems to contradict our earlier statements that blue main-sequence stars evolve into red giants after just a few hundred million years. The explanation is that these are *not* main-sequence stars, but rather **horizontal-branch stars.** These stars get their name because in the color-magnitude diagram of a globular cluster, they form a horizontal grouping in the left-of-center portion of the diagram (see Figure 19-13). Horizontal-branch stars are relatively low-mass stars that have already become red giants and undergone a helium flash, so there is both core helium fusion and shell hydrogen fusion taking place in their interiors. After the helium flash their luminosity decreased to about 50 $L_\odot$ (compared to about 1000 $L_\odot$ before the flash) and their outer layers contracted and heated, giving these stars their blue color. In years to come, these stars will move back toward the red-giant region as their fuel is devoured. Our own Sun will go through a horizontal-branch phase in the distant future; this is the phase labeled by the number 3 at the far right of Figure 19-8.

Measuring the Ages of Star Clusters

The idea that a cluster's main sequence is progressively "eaten away" is the key to determining the age of a cluster. In the H-R diagram for a very young cluster, all the stars are on or near the main sequence. (An example is the open cluster NGC 2264, shown in Figure 18-18.) As a cluster gets older, however, stars begin to leave the main sequence. The high-mass, high-luminosity stars are the first to consume their core hydrogen and become red giants. As time passes the main sequence gets shorter and shorter, like a candle burning down (see parts *d* through *h* of Figure 19-10).

The age of a cluster can be found from the **turnoff point,** which is the top of the surviving portion of the main sequence on the cluster's H-R diagram (see Figure 19-13). The stars at the turnoff point are just now exhausting the hydrogen in their cores, so their main-sequence lifetime is equal to the age of the cluster. For example, in the case of the globular cluster M55 plotted in Figure 19-13, 0.8-$M_\odot$ stars have just left the main sequence, indicating that the cluster's age is more than 12 billion (1.2×10^{10}) years (see Table 19-1).

Figure 19-14 shows data for several star clusters plotted on a single H-R diagram. This graph also shows turnoff-point times from which the ages of the clusters can be estimated.

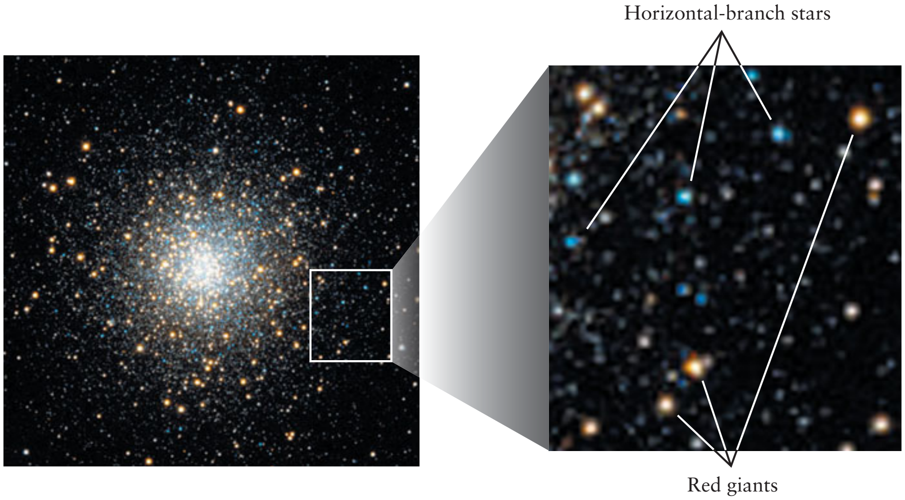


Figure 19-12 R I V U X G

A Globular Cluster This cluster, called M10, contains a few hundred thousand stars within a diameter of only 20 pc (70 ly). It lies approximately 5000 pc (16,000 ly) from the Earth in the constellation Ophiuchus (the Serpent Holder). Most of the stars in this image are either red giants or blue, horizontal-branch stars with both core helium fusion and shell hydrogen fusion. (T. Credner and S. Kohle, Astronomical Institutes of the University of Bonn)

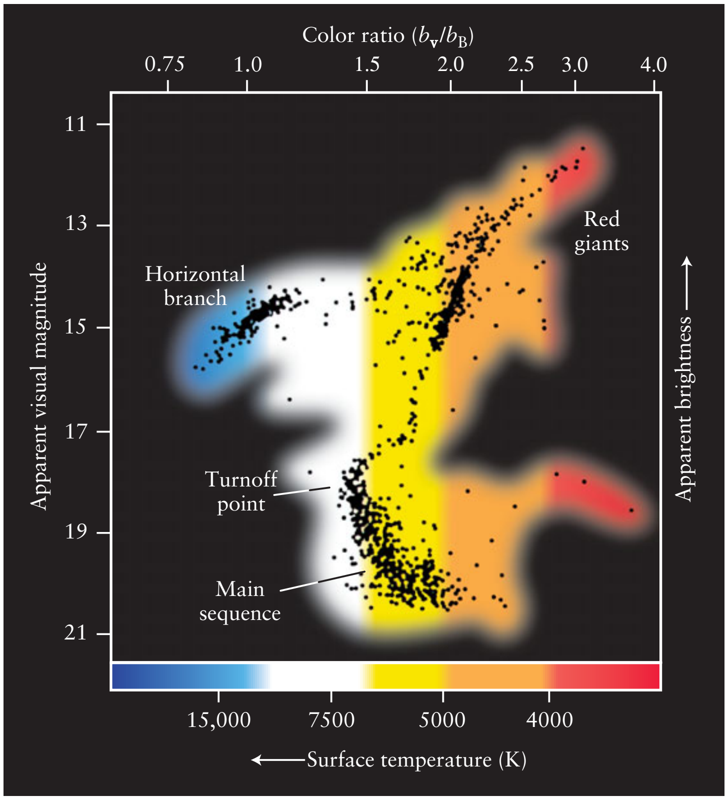


Figure 19-13

A Color-Magnitude Diagram of a Globular Cluster Each dot in this diagram represents the apparent visual magnitude (a measure of the brightness as seen through a V filter) and surface temperature (as measured by the color ratio b_V/b_B) of a star in the globular cluster M55 in Sagittarius. Because all the stars in M55 are at essentially the same distance from the Earth (about 6000 pc or 20,000 ly), their apparent visual magnitudes are a direct measure of their luminosities. Note that the upper half of the main sequence is missing. (Adapted from D. Schade, D. VandenBerg, and F. Hartwick)

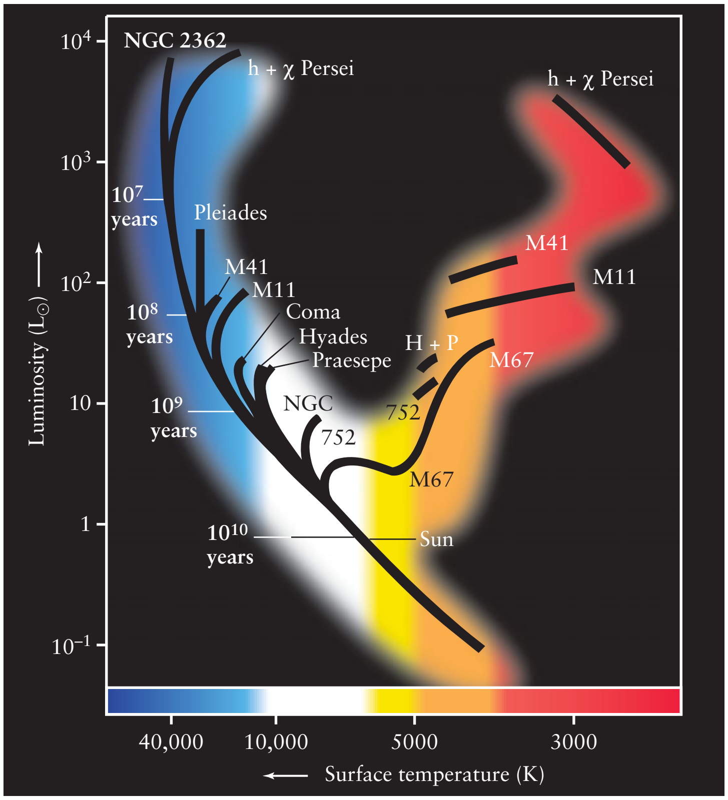


Figure 19-14

An H-R Diagram for Open Star Clusters The black bands indicate where stars from various open clusters fall on the H-R diagram. The age of a cluster can be estimated from the location of the cluster's turnoff point, where the cluster's most massive stars are just now leaving the main sequence. The times for these turnoff points are listed alongside the main sequence. For example, the Pleiades cluster turnoff point is near the 10^8-year point, so this cluster is about 10^8 years old. (Adapted from A. Sandage)

19-5 Stellar evolution has produced two distinct populations of stars

Studies of star clusters reveal a curious difference between the youngest and oldest stars in our Galaxy. Stars in the youngest clusters (those with most of their main sequences still intact) are said to be **metal rich,** because their spectra contain many prominent spectral lines of heavy elements. (Recall from Section 17-5 that astronomers use the term "metal" to denote any element other than hydrogen and helium, which are the two lightest elements.) Such stars are also called **Population I stars.** The Sun is a relatively young, metal-rich, Population I star.

By contrast, the spectra of stars in the oldest clusters show only weak lines of heavy elements. These ancient stars are thus said to be **metal poor,** because heavy elements are only about 3% as abundant in these stars as in the Sun. They are also called **Population II stars.** The stars in globular clusters are metal poor, Population II stars. Figure 19-15 shows the difference in spectra between a metal-poor, Population II star and the Sun (a metal-rich, Population I star).

CAUTION! Note that "metal rich" and "metal poor" are relative terms. In even the most metal-rich star known, metals make up just a few percent of the total mass of the star.

Stellar Populations and the Origin of Heavy Elements

To explain why there are two distinct populations of stars, we must go back to the Big Bang, the explosive origin of the universe that took place some 13.7 billion years ago. As we will discuss in Chapter 27, the early universe consisted almost exclusively of hydrogen and helium, with almost no heavy elements (metals). The first stars to form were likewise metal poor. The least massive of these have survived to the present day and are now the ancient stars of Population II.

The more massive of the original stars evolved more rapidly and no longer shine. But as these stars evolved, helium fusion in their cores produced metals—carbon and oxygen. In the most massive stars, as we will learn in Chapter 21, further thermonuclear reactions produced even heavier elements. As these massive original stars aged and died, they expelled their metal-enriched gases into space. (The star shown in Figure 19-5 is going through such a mass-loss phase late in its life.) This expelled material joined the interstellar medium and was eventually incorporated into a second generation of stars that have a higher concentration of heavy elements. These metal-rich members of the second stellar generation are the Population I stars, of which our Sun is an example.

Stars like the Sun contain material that was processed through an earlier generation of stars

CAUTION! Be careful not to let the designations of the two stellar populations confuse you. Population *I* stars are members of a *second* stellar generation, while Population *II* stars belong to an older *first* generation.

The relatively high concentration of heavy elements in the Sun means that the solar nebula, from which both the Sun and planets formed (see Section 8-4), must likewise have been metal rich. The Earth is composed almost entirely of heavy elements, as are our bodies. Thus, our very existence is intimately linked to the Sun's being a Population I star. A planet like the Earth probably could not have formed from the metal-poor gases that went into making Population II stars.

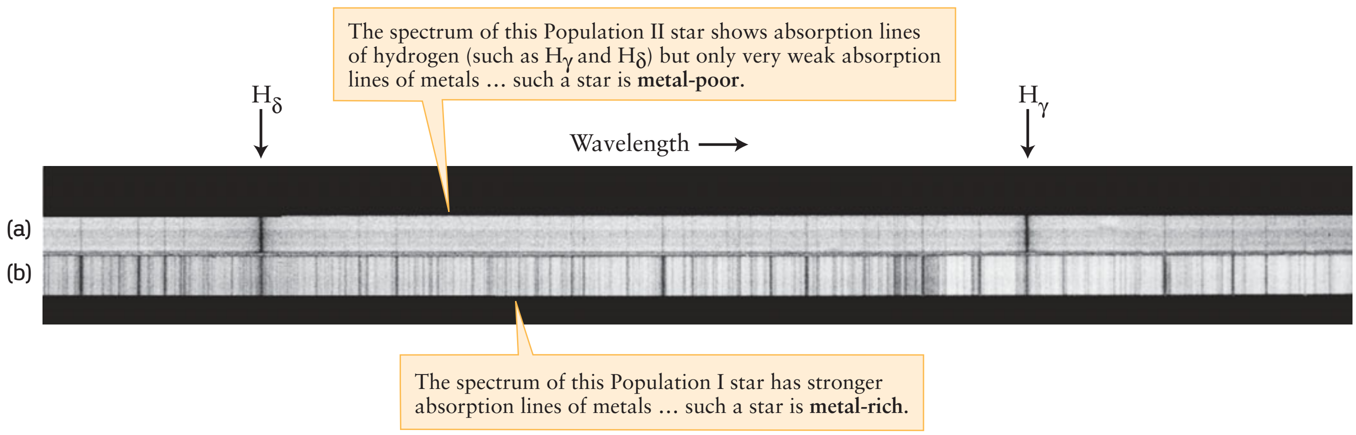


Figure 19-15 R I **V** U X G

Spectra of a Metal-Poor Star and a Metal-Rich Star The abundance of metals (elements heavier than hydrogen and helium) in a star can be inferred from its spectrum. These spectra compare **(a)** a metal-poor, Population II star and **(b)** a metal-rich, Population I star (the Sun) of the same surface temperature. We described the hydrogen absorption lines H_γ (wavelength 434 nm) and H_δ (wavelength 410 nm) in Section 5-8. (Lick Observatory)

The concept of two stellar populations provides insight into our own origins. Recall from Section 19-3 that helium fusion in red-giant stars produces the same isotopes of carbon (^{12}C) and oxygen (^{16}O) that are found most commonly on the Earth. The reason is that the Earth's carbon and oxygen atoms, including all of those in your body, actually *were* produced by helium fusion. These reactions occurred billions of years ago within an earlier generation of stars that died and gave up their atoms to the interstellar medium—the same atoms that later became part of our solar system, our planet, and our bodies. We are literally children of the stars.

19-6 Many mature stars pulsate

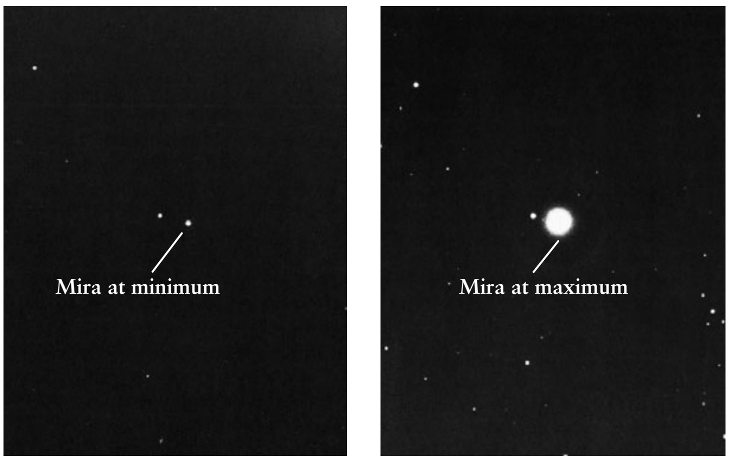
We saw in Section 16-3 that the surface of our Sun vibrates in and out, although by only a small amount. But other stars undergo substantial changes in size, alternately swelling and shrinking. As these stars pulsate, they also vary dramatically in brightness. We now understand that these **pulsating variable stars** are actually evolved, post–main-sequence stars.

Long-Period Variables

Pulsating variable stars were first discovered in 1595 by David Fabricius, a Dutch minister and amateur astronomer. He noticed that the star o (omicron) Ceti is sometimes bright enough to be easily seen with the naked eye but at other times fades to invisibility (Figure 19-16). By 1660, astronomers realized that these brightness variations repeated with a period of 332 days. Seventeenth-century astronomers were so enthralled by this variable star that they renamed it Mira ("wonderful").

Mira is an example of a class of pulsating stars called **long-period variables.** These stars are cool red giants that vary in brightness by a factor of 100 or more over a period of months or years. With surface temperatures of about 3500 K and average luminosities that range from about 10 to 10,000 $L_\odot$, they occupy the upper right side of the H-R diagram (Figure 19-17). Some, like Mira, are periodic, but others are irregular. Many eject large amounts of gas and dust into space.

Astronomers do not fully understand why some cool red giants become long-period variables. It is difficult to calculate accurate stellar models to describe such huge stars with extended, tenuous atmospheres.



(a) (b)

Figure 19-16 R I V U X G

Mira—A Long-Period Variable Star Mira, or o (omicron) Ceti, is a variable star whose luminosity varies with a 332-day period. At its dimmest, as in **(a)** (photographed in December 1961), Mira is less than 1% as bright as when it is at maximum, as in **(b)** (January 1965). These brightness variations occur because Mira pulsates. (Lowell Observatory)

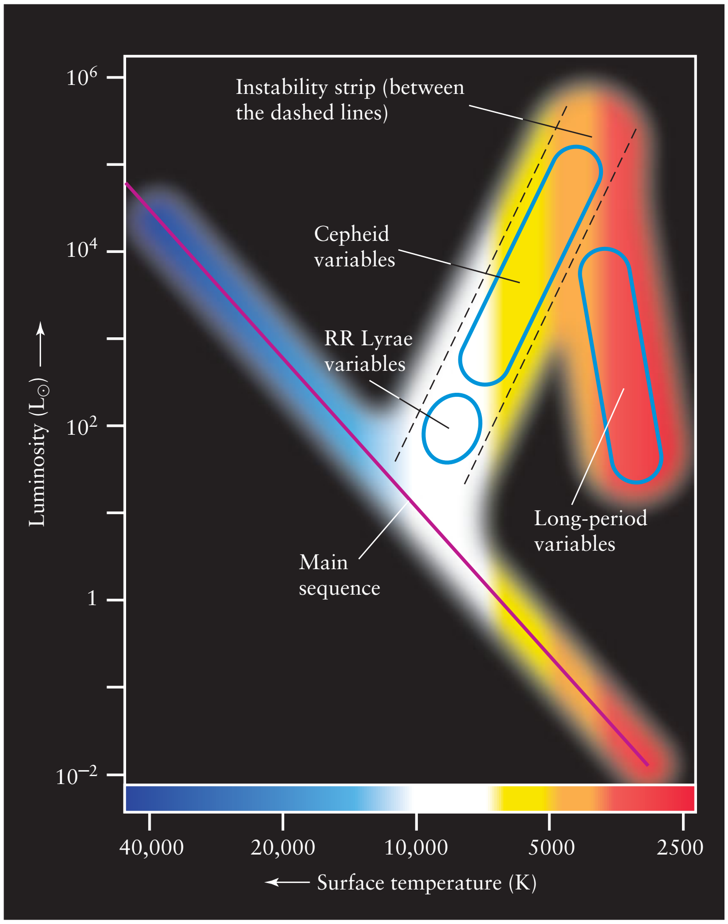


Figure 19-17

Variable Stars on the H-R Diagram Pulsating variable stars are found in the upper right of the H-R diagram. Long-period variables like Mira are cool red giant stars that pulsate slowly, changing their brightness in a semi-regular fashion over months or years. Cepheid variables and RR Lyrae variables are located in the instability strip, which lies between the main sequence and the red-giant region. A star passing through this strip along its evolutionary track becomes unstable and pulsates.

Cepheid Variables

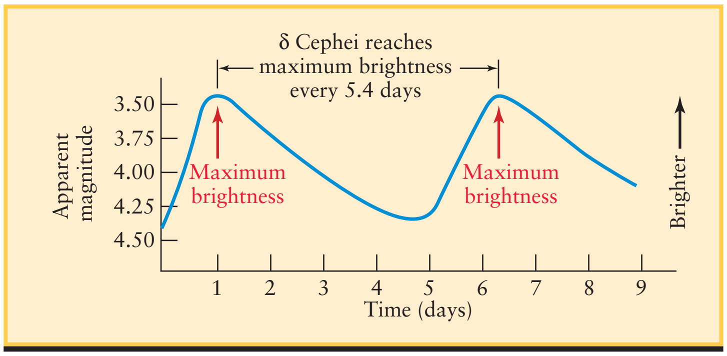
Astronomers have a much better understanding of other pulsating stars, called **Cepheid variables,** or simply Cepheids. A Cepheid variable is recognized by the characteristic way in which its light output varies—rapid brightening followed by gradual dimming. They are named for δ (delta) Cephei, an example of this type of star discovered in 1784 by John Goodricke, a deaf, mute, 19-year-old English amateur astronomer. He found that at its most brilliant, δ Cephei is 2.3 times as bright as at its dimmest. The cycle of brightness variations repeats every 5.4 days. (Sadly, Goodricke paid for his discoveries with his life; he caught pneumonia while making his nightly observations and died before his twenty-second birthday.) The surface temperatures and luminosities of the Cepheid variables place them in the upper middle of the H-R diagram (see Figure 19-17).

By studying variable stars, astronomers gain insight into late stages of stellar evolution

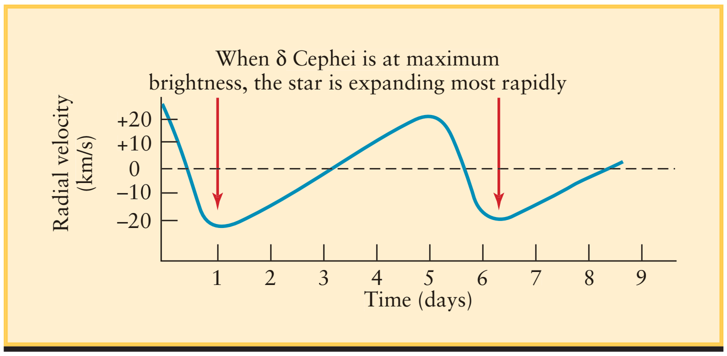
After core helium fusion begins, mature stars move across the middle of the H-R diagram. Figure 19-9*a* shows the evolutionary tracks of high-mass stars crisscrossing the H-R diagram. Post–helium-flash stars of moderate mass also cross the middle of the H-R diagram between the red-giant region and the horizontal branch.

During these transitions across the H-R diagram, a star can become unstable and pulsate. In fact, there is a region on the H-R diagram between the upper main sequence and the red-giant branch called the **instability strip** (see Figure 19-17). When an evolving star passes through this region, the star pulsates and its brightness varies periodically. Figure 19-18*a* shows the brightness variations of δ Cephei, which lies within the instability strip.

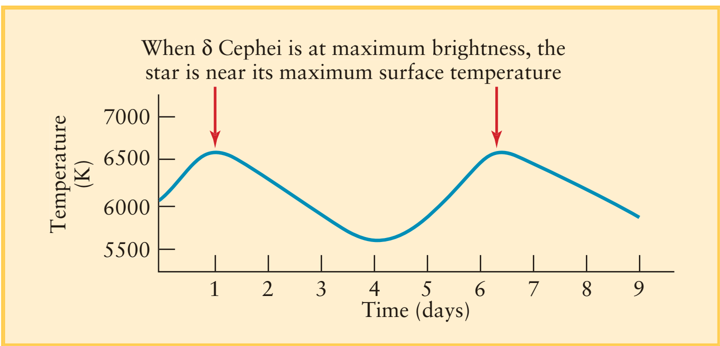
A Cepheid variable brightens and fades because the star's outer envelope cyclically expands and contracts. The first to observe this was the Russian astronomer Aristarkh Belopol'skii, who noticed in 1894 that spectral lines in the spectrum of δ Cephei shift back and forth with the same 5.4-day period as that of the magnitude variations. From the Doppler effect, we can



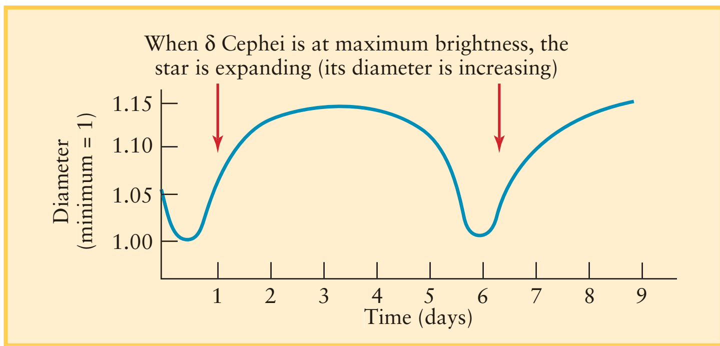
(a) The light curve of δ Cephei (a graph of brightness versus time)



(b) Radial velocity versus time for δ Cephei (positive: star is contracting; negative: star is expanding)



(c) Surface temperature versus time for δ Cephei



(d) Diameter versus time for δ Cephei

WEB LINK 19.10

Figure 19-18

δ Cephei—A Pulsating Star **(a)** As δ Cephei pulsates, it brightens quickly (the light curve moves upward sharply) but fades more slowly (the curve declines more gently). The increases and decreases in brightness are nearly in step with variations in **(b),** the star's radial velocity (positive when the star contracts and the surface moves away from us, negative when the star expands and the surface approaches us), as well as in **(c),** the star's surface temperature. **(d)** The star is still expanding when it is at its brightest and hottest (compare with parts a and b).

translate these wavelength shifts into radial velocities and draw a velocity curve (Figure 19-18*b*). Negative velocities mean that the star's surface is expanding toward us; positive velocities mean that the star's surface is receding. Note that the light curve and velocity curve are mirror images of each other. The star is brighter than average while it is expanding and dimmer than average while contracting.

When a Cepheid variable pulsates, the star's surface oscillates up and down like a spring. During these cyclical expansions and contractions, the star's gases alternately heat up and cool down. Figure 19-18*c* shows the resulting changes in the star's surface temperature. Figure 19-18*d* graphs the periodic changes in the star's diameter.

Just as a bouncing ball eventually comes to rest, a pulsating star would soon stop pulsating without something to keep its oscillations going. In 1914, the British astronomer Arthur Eddington suggested that a Cepheid pulsates because the star is more opaque when compressed than when expanded. When the star is compressed, trapped heat increases the internal pressure, which pushes the star's surface outward. When the star expands, the heat escapes, the internal pressure drops, and the star's surface falls inward.

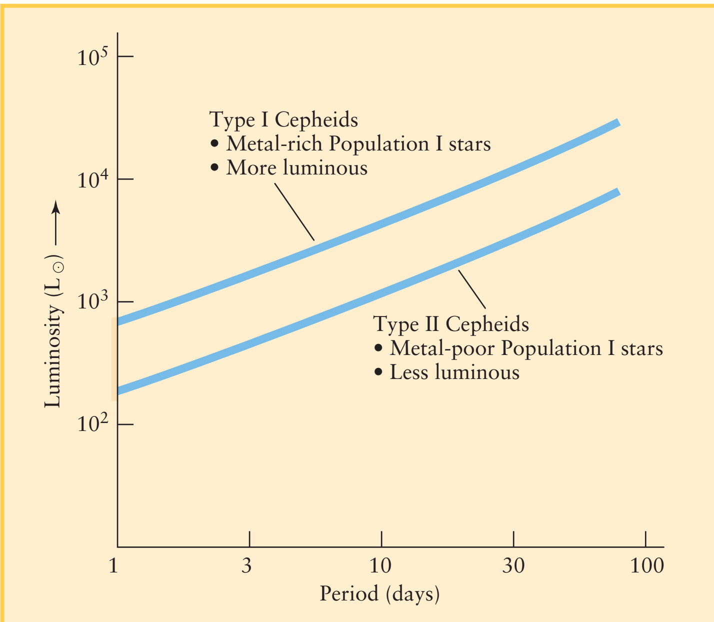
In the 1960s, the American astronomer John Cox followed up on Eddington's idea and proved that helium is what keeps Cepheids pulsating. Normally, when a star's helium is compressed, the gas increases in temperature and becomes more transparent. But in certain layers near the star's surface, compression may ionize helium (remove one of its electrons) instead of raising its temperature. Ionized helium gas is quite opaque, so these layers effectively trap heat and make the star expand, as Eddington suggested. This expansion cools the outer layers and makes the helium ions recombine with electrons, which makes the gas more transparent and releases the trapped energy. The star's surface then falls inward, recompressing the helium, and the cycle begins all over again.

CAUTION! In our discussion of the behavior of gases (see Box 19-1, Section 19-1, and Section 19-3) we saw that a gas cools when it expands and heats up when it is compressed. Hence, you would expect that the gases in a pulsating star like δ Cephei would reach their maximum temperature when the star is at its smallest diameter, so that the gases are most compressed. The hotter the gas, the more brightly it glows, so δ Cephei should also have its maximum brightness when its diameter is smallest. But Figure 19-18 shows that the star's brightness and the temperature of the gases at the surface reach their maximum values when the star is expanding, some time *after* the star has contracted to its smallest diameter. How can this be? The explanation is again related to how opaque the gases are inside the star. The rate at which energy is emitted from the central regions of the star is indeed greatest when the star is at its minimum diameter, but the opaque gases in the star's outer layers impede the flow of energy to the surface. Hence, δ Cephei reaches its maximum brightness and maximum surface temperature about half a day after the star is at its smallest size.

Cepheid variables are important because they have two properties that allow astronomers to determine the distances to very remote objects. First, Cepheids can be seen even at distances of millions of parsecs. This is because they are very luminous, ranging from a few hundred times solar luminosity to more than 10^4 $L_\odot$. Second, there is a direct relationship between a Cepheid's period and its average luminosity: The dimmest Cepheid variables pulsate rapidly, with periods of 1 to 2 days, while the most luminous Cepheids pulsate with much slower periods of about 100 days.

Figure 19-19 shows this **period-luminosity relation.** By measuring the period of a distant Cepheid's brightness variations and using a graph like Figure 19-19, an astronomer can determine the star's luminosity. By also measuring the star's apparent brightness, the distance to the Cepheid can be found by using the inverse-square law (see Section 17-2). By applying the period-luminosity relation in this way to Cepheids in other galaxies, astronomers have been able to calculate the distances to those galaxies with great accuracy. (Box 17-2 gives an example of such a calculation.) As we will see in Chapters 24 and 26, such measurements play an important role in determining the overall size and structure of the universe.

How a Cepheid pulsates depends on the amount of heavy elements in the star's outer layers, because even trace amounts of these elements can have a large effect on how opaque the stellar gases are. Hence, Cepheids are classified according to their metal content. If the star is a metal-rich, Population I star, it is called a **Type I Cepheid;** if it is a metal-poor, Population II star, it is called



WEB LINK 19.11

Figure 19-19

Period-Luminosity Relations for Cepheids The greater the average luminosity of a Cepheid variable, the longer its period and the slower its pulsations. Note that there are actually two distinct period-luminosity relations—one for Type I Cepheids and one for the less luminous Type II Cepheids. (Adapted from H. C. Arp) a **Type II Cepheid.** As Figure 19-19 shows, these two types of Cepheids exhibit different period-luminosity relations. In order to know which period-luminosity relation to apply to a given Cepheid, an astronomer must determine the star's metal content from its spectrum (see Figure 19-15).

The evolutionary tracks of mature, high-mass stars pass back and forth through the upper end of the instability strip on the H-R diagram. These stars become Cepheids when helium ionization occurs at just the right depth to drive the pulsations. For stars on the high-temperature (left) side of the instability strip, helium ionization occurs too close to the surface and involves only an insignificant fraction of the star's mass. For stars on the cool (right) side of the instability strip, convection in the star's outer layers prevents the storage of the energy needed to drive the pulsations. Thus, Cepheids exist only in a narrow temperature range on the H-R diagram.

RR Lyrae Variables

Stars of lower mass do not become Cepheids. Instead, after leaving the main sequence, becoming red giants, and undergoing the helium flash, their evolutionary tracks pass through the lower end of the instability strip as they move along the horizontal branch. Some of these stars become **RR Lyrae variables,** named for their prototype in the constellation Lyra (the Harp). RR Lyrae variables all have periods shorter than one day and roughly the same average luminosity as horizontal-branch stars, about 100 $L_{\odot}$. In fact, the RR Lyrae region of the instability strip (see Figure 19-17) is actually a segment of the horizontal branch. RR Lyrae stars are all metal-poor, Population II stars. Many have been found in globular clusters, and they have been used to determine the distances to those clusters in the same way that Cepheids are used to find the distances to other galaxies. In Chapter 23 we will see how RR Lyrae stars helped astronomers determine the size of the Milky Way Galaxy.

In some cases the expansion speed of a pulsating star exceeds the star's escape speed. When this happens, the star's outer layers are ejected completely. We will see in Chapter 20 that dying stars eject significant amounts of mass in this way, renewing and enriching the interstellar medium for future generations of stars.

19-7 Mass transfer can affect the evolution of stars in a close binary system

We have outlined what happens when a main-sequence star evolves into a red giant. What we have ignored is that more than half of all stars are members of multiple-star systems, including binaries. If the stars in such a system are widely separated, the individual stars follow the same course of evolution as if they were isolated. In a **close binary,** however, when one star expands to become a red giant, its outer layers can be gravitationally captured by the nearby companion star. In other words, a bloated red giant in a close binary system can dump gas onto its companion, a process called **mass transfer.**

Roche Lobes and Lagrangian Points

Our modern understanding of mass transfer in close binaries is based on the work of the French mathematician Edouard Roche. In the mid-1800s, Roche studied how rotation and mutual tidal interaction affect the stars in a binary system. Tidal forces cause the two stars in a close binary to keep the same sides facing each other, just as our Moon keeps its same side facing the Earth (see Section 4-8). But because stars are gaseous, not solid, rotation and tidal forces can have significant effects on their shapes.

If the stars in a binary system are sufficiently close, tidal forces can pull gases off one star and onto the other

In widely separated binaries, the stars are so far apart that tidal effects are small, and, therefore, the stars are nearly perfect spheres. In close binaries, where the separation between the stars is not much greater than their sizes, tidal effects are strong, causing the stars to be somewhat egg-shaped.

Roche discovered a mathematical surface that marks the gravitational domain of each star in a close binary. (This is not a real physical surface, like the surface of a balloon, but a mathematical construct.) Figure 19-20*a* shows the outline of this surface as a dashed line. The two halves of this surface, each of which encloses one of the stars, are known as **Roche lobes.** The more massive star is always located inside the larger Roche lobe. If gas from a star leaks over its Roche lobe, it is no longer bound by gravity to that star. This escaped gas is free either to fall onto the companion star or to escape from the binary system.

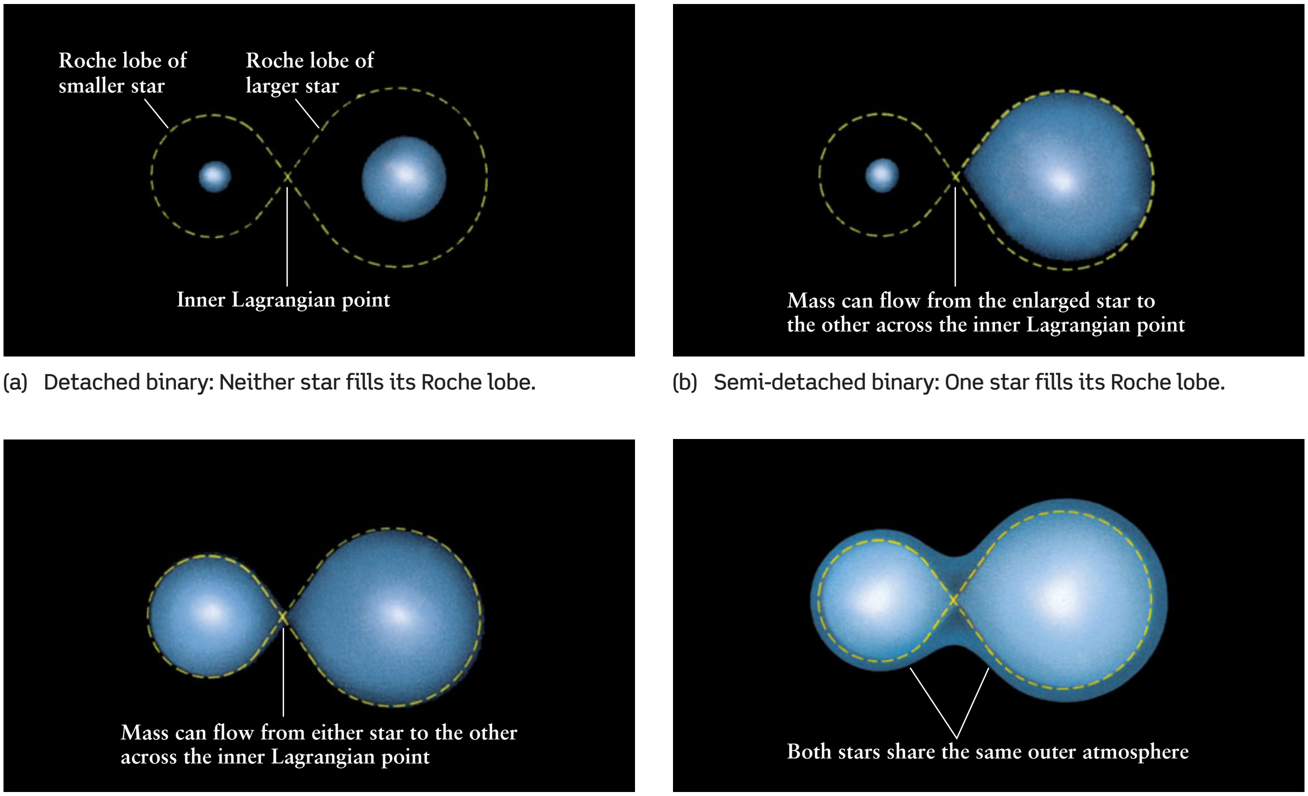
The point where the two Roche lobes touch, called the **inner Lagrangian point,** is a kind of balance point between the two stars in a binary. It is here that the effects of gravity and rotation cancel each other. When mass transfer occurs in a close binary, gases flow through the inner Lagrangian point from one star to the other.

In many binaries, the stars are so far apart that even during their red-giant stages the surfaces of the stars remain well inside their Roche lobes. As a result, little mass transfer can occur and each star lives out its life as if it were single and isolated. A binary system of this kind is called a **detached binary** (Figure 19-20*a*).

However, if the two stars are close enough, when one star expands to become a red giant, it may fill or overflow its Roche lobe. Such a system is called a **semidetached binary** (Figure 19-20*b*). If both stars fill their Roche lobes, the two stars actually touch and the system is called a **contact binary** (Figure 19-20*c*). It is quite unlikely, however, that both stars exactly fill their Roche lobes at the same time. (This would only be the case if the two stars had identical masses, so that they both evolved at exactly the same rate.) It is more likely that they overflow their Roche lobes, giving rise to a common envelope of gas. Such a system is called an **overcontact binary** (Figure 19-20*d*).

Observations of Mass Transfer

The binary star system Algol (from an Arabic term for "demon") provided the first clear evidence of mass transfer in close binaries. Also called β (beta) Persei, Algol can easily be seen with the naked



(a) Detached binary: Neither star fills its Roche lobe.

(b) Semi-detached binary: One star fills its Roche lobe.

(c) Contact binary: Both stars fill their Roche lobes.

(d) Overcontact binary: Both stars overfill their Roche lobes.

Figure 19-20

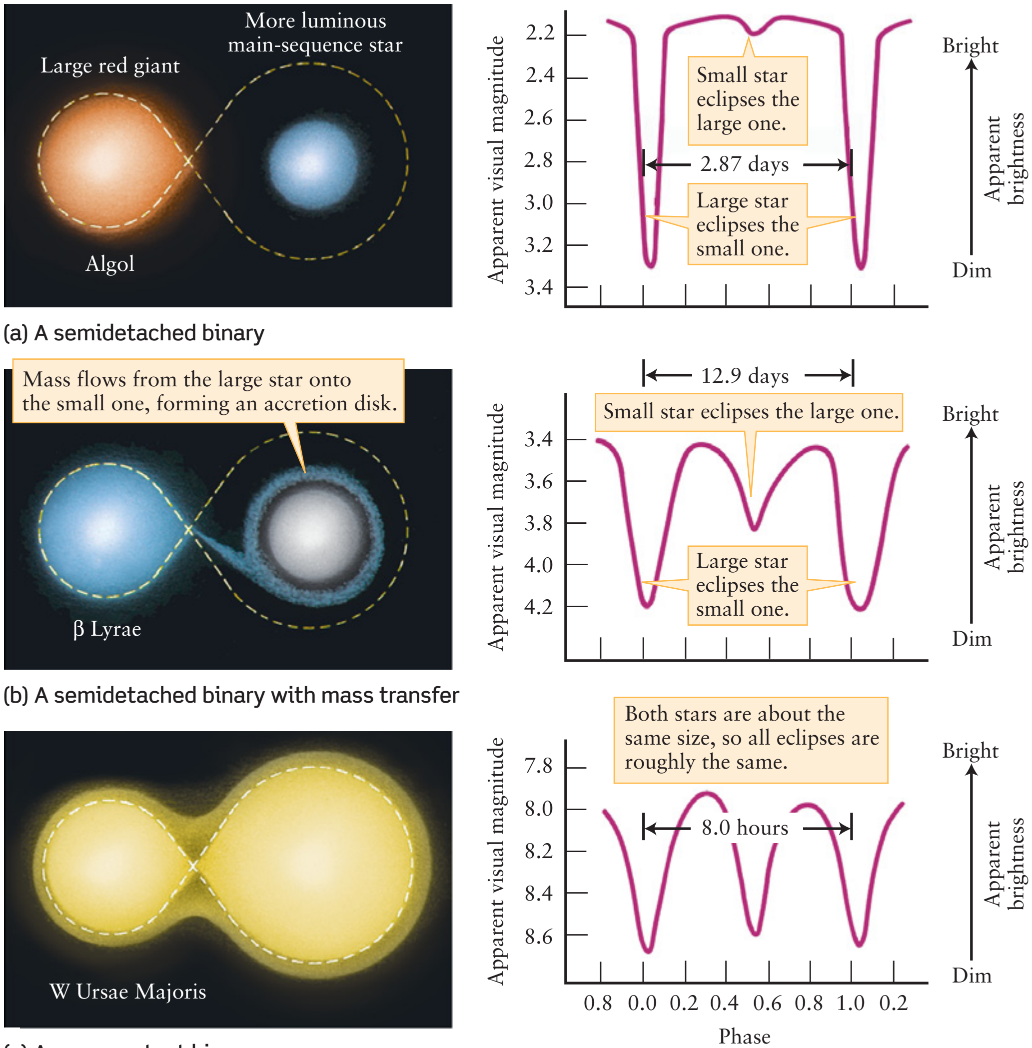
Close Binary Star Systems The gravitational domain of a star in a close binary system is called its Roche lobe. The two Roche lobes meet at the inner Lagrangian point. The sizes of the stars relative to their Roche lobes determine whether the system is **(a)** a detached binary, **(b)** a semidetached binary, **(c)** a contact binary, or **(d)** an overcontact binary.

eye. Ancient astronomers knew that Algol varies periodically in brightness by a factor of more than 2. In 1782, John Goodricke (the discoverer of δ Cephei's variability) first suggested that these brightness variations take place because Algol is an *eclipsing* binary. (We discussed this type of binary in Section 17-11.) The orbital plane of the two stars that make up the binary system happens to be nearly edge-on to our line of sight, so one star periodically eclipses the other. Algol's light curve (Figure 19-21*a*) and spectrum show that Goodricke's brilliant hypothesis is correct, and that Algol is a semidetached binary. The detached star (on the right in Figure 19-21*a*) is a luminous blue main-sequence star, while its less massive companion is a dimmer red giant that fills its Roche lobe.

CAUTION! According to stellar evolution theory, the more massive a star, the more rapidly it should evolve. Since the two stars in a binary system form simultaneously and thus are the same age, the more massive star should become a red giant before the less massive one. But in Algol and similar binaries, the *more* massive star (on the right in Figure 19-21*a*) is still on the main sequence, whereas the *less* massive star (on the left in Figure 19-21*a*) has evolved to become a red giant. How can we explain this apparent contradiction? The answer is that the red giant in Algol-type binaries was *originally* the more massive star. As it left the main sequence to become a red giant, this star expanded until it overflowed its Roche lobe and dumped gas onto its originally less massive companion. Because of the resulting mass transfer, that companion (which is still on the main sequence) became the more massive star.

Mass transfer is also important in another class of semidetached binaries, called β (beta) Lyrae variables, after their prototype in the constellation Lyra. As with Algol, the less massive star in β Lyrae (on the left in Figure 19-21*b*) fills its Roche lobe. Unlike Algol, however, the more massive detached star (on the right in Figure 19-21*b*) is the dimmer of the two stars. Apparently, this detached star is enveloped in a rotating *accretion disk* of gas captured from its bloated companion. This disk partially blocks the light coming from the detached star, making it appear dimmer.

What is the fate of an Algol or β Lyrae system? If the detached star is massive enough, it will evolve rapidly, expanding to also fill its Roche lobe. The result will be an overcontact binary in which the two stars share the gases of their outer layers. Such binaries are sometimes called W Ursae Majoris stars, after the prototype of this class (Figure 19-21*c*).



(a) A semidetached binary

(b) A semidetached binary with mass transfer

(c) An overcontact binary

Figure 19-21

WEB LINK 19.12

Three Eclipsing Binaries Compare the light curves of these three eclipsing binaries to those in Figure 17-24. **(a)** Algol is a semidetached binary. The deep eclipse occurs when the large red-giant star blocks the light from the smaller but more luminous main-sequence star. (Compare Figure 17-24*b*.) **(b)** β Lyrae's light curve is also at its lowest when the larger star completely eclipses the smaller one. Half an orbital period later, the smaller star partially eclipses the larger one, making a shallower dip in the light curve. **(c)** W Ursae Majoris is an overcontact binary in which both stars overfill their Roche lobes. The extremely short period of this binary indicates that the two stars are very close to each other.

Mass transfer can also continue even after nuclear reactions cease in one of the stars in a close binary. In Chapters 20 and 21 we will see how mass transfer onto different types of dead stars called *white dwarfs*, *neutron stars*, and *black holes* can produce some of the most unusual and dramatic objects in the sky.

Key Words

alpha particle, p. 504
Cepheid variable, p. 514
close binary, p. 516
color-magnitude diagram, p. 510
contact binary, p. 516
core helium fusion, p. 503
core hydrogen fusion, p. 497
degeneracy, p. 504
degenerate-electron pressure, p. 506
detached binary, p. 516
globular cluster, p. 510
helium flash, p. 504
helium fusion, p. 503
horizontal-branch star, p. 510
ideal gas, p. 504
inner Lagrangian point, p. 516
instability strip, p. 514
long-period variable, p. 513
main-sequence lifetime, p. 497
mass loss, p. 502
mass transfer, p. 516
metal-poor star, p. 512
metal-rich star, p. 512
open cluster, p. 508
overcontact binary, p. 516
Pauli exclusion principle, p. 504
period-luminosity relation, p. 515
Population I and Population II stars, p. 512
pulsating variable star, p. 513
red dwarf, p. 499
red giant, p. 502

Roche lobe, p. 516
RR Lyrae variable, p. 516
semidetached binary, p. 516
shell hydrogen fusion, p. 500
triple alpha process, p. 504
turnoff point, p. 510
Type I and Type II Cepheids, p. 515–516
zero-age main sequence (ZAMS), p. 507
zero-age main-sequence star, p. 498

Key Ideas

The Main-Sequence Lifetime: The duration of a star's main-sequence lifetime depends on the amount of hydrogen available to be consumed in the star's core and the rate at which this hydrogen is consumed.

• The more massive a star, the shorter is its main-sequence lifetime. The Sun has been a main-sequence star for about 4.56 billion years and should remain one for about another 7 billion years.

• During a star's main-sequence lifetime, the star expands somewhat and undergoes a modest increase in luminosity.

• If a star's mass is greater than about 0.4 $M_\odot$, only the hydrogen present in the core can undergo thermonuclear fusion during the star's main-sequence lifetime. If the star is a red dwarf with a mass less than about 0.4 $M_\odot$, over time convection brings all of the star's hydrogen to the core where it can undergo fusion.

Becoming a Red Giant: Core hydrogen fusion ceases when the hydrogen has been exhausted in the core of a main-sequence star with mass greater than about 0.4 $M_\odot$. This leaves a core of nearly pure helium surrounded by a shell through which hydrogen fusion works its way outward in the star. The core shrinks and becomes hotter, while the star's outer layers expand and cool. The result is a red giant star.

• As a star becomes a red giant, its evolutionary track moves rapidly from the main sequence to the red-giant region of the H-R diagram. The more massive the star, the more rapidly this evolution takes place.

Helium Fusion: When the central temperature of a red giant reaches about 100 million K, helium fusion begins in the core. This process, also called the triple alpha process, converts helium to carbon and oxygen.

• In a more massive red giant, helium fusion begins gradually; in a less massive red giant, it begins suddenly, in a process called the helium flash.

• After the helium flash, a low-mass star moves quickly from the red-giant region of the H-R diagram to the horizontal branch.

Star Clusters and Stellar Populations: The age of a star cluster can be estimated by plotting its stars on an H-R diagram.

• The cluster's age is equal to the age of the main-sequence stars at the turnoff point (the upper end of the remaining main sequence).

• As a cluster ages, the main sequence is "eaten away" from the upper left as stars of progressively smaller mass evolve into red giants.

• Relatively young Population I stars are metal rich; ancient Population II stars are metal poor. The metals (heavy elements) in Population I stars were manufactured by thermonuclear reactions in an earlier generation of Population II stars, then ejected into space and incorporated into a later stellar generation.

Pulsating Variable Stars: When a star's evolutionary track carries it through a region in the H-R diagram called the instability strip, the star becomes unstable and begins to pulsate.

• Cepheid variables are high-mass pulsating variables. There is a direct relationship between their periods of pulsation and their luminosities.

• RR Lyrae variables are low-mass, metal-poor pulsating variables with short periods.

• Long-period variable stars also pulsate but in a fashion that is less well understood.

Close Binary Systems: Mass transfer in a close binary system occurs when one star in a close binary overflows its Roche lobe. Gas flowing from one star to the other passes across the inner Lagrangian point. This mass transfer can affect the evolutionary history of the stars that make up the binary system.

Questions

Review Questions

1. How does the chemical composition of the present-day Sun's core compare to the core's composition when the Sun formed? What caused the change?
2. Which regions of the Sun are denser today than they were a billion years ago? Which regions are less dense? What has caused these changes?
3. What is a red dwarf? How are thermonuclear reactions in the core of a red dwarf able to consume hydrogen from the star's outer layers?
4. Why do high-mass main-sequence stars have shorter lifetimes than those of lower mass?
5. On what grounds are astronomers able to say that the Sun has about 7×10^9 years remaining in its main-sequence stage?
6. What will happen inside the Sun 7 billion years from now, when it begins to mature into a red giant?
7. Explain why the Earth is expected to become inhospitable to life long before the Sun becomes a red giant.
8. Explain how it is possible for the core of a red giant to contract at the same time that its outer layers expand.
9. Why does helium fusion require much higher temperatures than hydrogen fusion?
10. How is a degenerate gas different from ordinary gases?
11. What is the helium flash? Why does it happen in some stars but not in others?
12. Why does a star's luminosity decrease after helium fusion begins in its core?
13. What does it mean when an astronomer says that a star "moves" from one place to another on an H-R diagram?
14. Explain why the majority of the stars visible through telescopes are main-sequence stars.
15. On an H-R diagram, main-sequence stars do not lie along a single narrow line but are spread out over a band (see Figure 19-9*b*). On the basis of how stars evolve during their main-sequence lifetimes, explain why this should be so.

16. Explain how and why the turnoff point on the H-R diagram of a cluster is related to the cluster's age.
17. There is a good deal of evidence that our universe is about 13.7 billion years old (see Chapter 24). Explain why no main-sequence stars of spectral class M have yet evolved into red-giant stars.
18. How do astronomers know that globular clusters are made of old stars?
19. Red-giant stars appear more pronounced in composites of infrared images and visible-light images, like those in Figure 19-4*b* and Figure 19-12. Explain why.
20. The horizontal-branch stars in Figure 19-12 appear blue. (**a**) Explain why this is consistent with the color-magnitude diagram shown in Figure 19-13. (**b**) All horizontal-branch stars were once red giants. Explain what happened to these stars to change their color.
21. What is the difference between Population I and Population II stars? In what sense can the stars of one population be regarded as the "children" of the other population?
22. Both diamonds and graphite (the material used in pencils to make marks on paper) are crystalline forms of carbon. Most of the carbon atoms in these substances have nuclei with 6 protons and 6 neutrons (^{12}C). Where did these nuclei come from?
23. Why do astronomers attribute the observed Doppler shifts of a Cepheid variable to pulsation, rather than to some other cause, such as orbital motion?
24. Why do Cepheid stars pulsate? Why are these stars important to astronomers who study galaxies beyond the Milky Way?
25. What is a Roche lobe? What is the inner Lagrangian point? Why are Roche lobes important in close binary star systems?
26. What is the difference between a detached binary, a semidetached binary, a contact binary, and an overcontact binary?
27. Massive main-sequence stars turn into red giants before less massive stars. Why, then, is the more massive star in Algol a main-sequence star and the less massive star a red giant?

Advanced Questions

Questions preceded by an asterisk (*) involve topics discussed in Box 7-2, the Boxes in Chapter 17, or the Boxes in this chapter.

Problem-solving tips and tools

Recall from Section 16-1 that 6×10^{11} kg of hydrogen is converted into helium each second at the Sun's center. Recall also that you must use absolute (Kelvin) temperatures when using the Stefan-Boltzmann law. You may find it helpful to review the discussion of apparent magnitude, absolute magnitude, and luminosity in Box 17-3. Section 17-4 discusses the connection between the surface temperatures and colors of stars. Newton's form of Kepler's third law (see Section 17-9) describes the orbits of stars in binary systems. Box 7-2 gives the formula for escape speed and for the average speeds of gas molecules.

28. The radius of the Sun has increased over the past several billion years. Over the same time period, the size of the Moon's orbit around the Earth has also increased. A few billion years ago, were annular eclipses of the Sun (see Figure 3-12) more or less common than they are today? Explain.
29. The Sun has increased in radius by 6% over the past 4.56 billion years. Its present-day radius is 696,000 km. What was its radius 4.56 billion years ago? (*Hint:* The answer is *not* 654,000 km.)

*30. Calculate the escape speed from (**a**) the surface of the present-day Sun and (**b**) the surface of the Sun when it becomes a red giant, with essentially the same mass as today but with a radius that is 100 times larger. (**c**) Explain how your results show that a red-giant star can lose mass more easily than a main-sequence star.

*31. Calculate the average speed of a hydrogen atom (mass 1.67×10^{-27} kg) (**a**) in the atmosphere of the present-day Sun, with temperature 5800 K, and (**b**) in the atmosphere of a 1-$M_\odot$ red giant, with temperature 3500 K. (**c**) Compare your results with the escape speeds that you calculated in Question 27. Use this comparison to discuss how well the present-day Sun and a 1-$M_\odot$ red giant can retain hydrogen in their atmospheres.

32. Use the value of the Sun's luminosity (3.90×10^{26} watts, or 3.90×10^{26} joules per second) to calculate what mass of hydrogen the Sun will convert into helium during its entire main-sequence lifetime of 1.2×10^{10} years. (Assume that the Sun's luminosity remains nearly constant during the entire 1.2×10^{10} years.) What fraction does this represent of the total mass of hydrogen that was originally in the Sun?
33. (**a**) The main-sequence stars Sirius (spectral type A1), Vega (A0), Spica (B1), Fomalhaut (A3), and Regulus (B7) are among the 20 brightest stars in the sky. Explain how you can tell that all these stars are younger than the Sun. (**b**) The third-brightest star in the sky, although it can be seen only south of 29° north latitude, is α (alpha) Centauri A. It is a main-sequence star of spectral type G2, the same as the Sun. Can you tell from this whether α Centauri A is younger than the Sun, the same age, or older? Explain your reasoning.
34. Using the same horizontal and vertical scales as in Figure 19-9*a*, make points on an H-R diagram for each of the stars listed in Table 19-1. Label each point with the star's mass and its main-sequence lifetime. Which of these stars will remain on the main sequence after 10^9 years? After 10^{11} years?

*35. Explain why the quantity *f* in Box 19-2 has a different value for stars with masses less than 0.4 $M_\odot$ than for stars with masses greater than 0.4 $M_\odot$. In which case does *f* have a greater value?

*36. Calculate the main-sequence lifetimes of (**a**) a 9-$M_\odot$ star and (**b**) a 0.25-$M_\odot$ star. Compare these lifetimes with that of the Sun.

*37. The earliest fossil records indicate that life appeared on the Earth about a billion years after the formation of the solar system. What is the most mass that a star could have in order that its lifetime on the main sequence is long enough to permit life to form on one or more of its planets? Assume that the evolutionary processes would be similar to those that occurred on the Earth.

38. As a red giant, the Sun's luminosity will be about 2000 times greater than it is now, so the amount of solar energy falling on the Earth will increase to 2000 times its present-day value. Hence, to maintain thermal equilibrium, each square meter of the Earth's surface will have to radiate 2000 times as much energy into space as it does now. Use the Stefan-Boltzmann law to determine what the Earth's surface temperature will be under these conditions. (*Hint:* The present-day Earth has an average surface temperature of 14°C.)

39. When the Sun becomes a red giant, its luminosity will be about 2000 times greater than it is today. Assuming that this luminosity is caused *only* by fusion of the Sun's remaining hydrogen, calculate how long our star will be a red giant. (In fact, only a fraction of the remaining hydrogen will be consumed, and the luminosity will vary over time as shown in Figure 19-8.)

40. What observations would you make of a star to determine whether its primary source of energy is hydrogen fusion or helium fusion?

41. The star whose spectrum is shown in Figure 19-15*a* has a lower percentage of heavy elements than the Sun, whose spectrum is shown in Figure 19-15*b*. Hence, the star in Figure 19-15*a* has a higher percentage of hydrogen. Why, then, isn't the H_δ absorption line of hydrogen noticeably darker for the star in Figure 19-15*a*?

42. Would you expect the color of a Cepheid variable star (see Figure 19-18) to change during the star's oscillation period? If not, why not? If so, describe why the color should change, and describe the color changes you would expect to see during an oscillation period.

43. The brightness of a certain Cepheid variable star increases and decreases with a period of 10 days. (**a**) What must this star's luminosity be if its spectrum has strong absorption lines of hydrogen and helium, but no strong absorption lines of heavy elements? (**b**) Repeat part (a) for the case in which the star's spectrum also has strong absorption lines of heavy elements.

44. The star X Arietis is an RR Lyrae variable. Its apparent brightness varies between 2.0×10^{-15} and 4.9×10^{-15} that of the Sun with a period of 0.65 day. Interstellar extinction dims the star by 37%. Approximately how far away is the star?

45. The apparent brightness of δ Cephei (a Type I Cepheid variable) varies with a period of 5.4 days. Its average apparent brightness is 5.1×10^{-13} that of the Sun. Approximately how far away is δ Cephei? (Ignore interstellar extinction.)

46. Suppose you find a binary star system in which the more massive star is a red giant and the less massive star is a main-sequence star. Would you expect that mass transfer between the stars has played an important role in the evolution of these stars? Explain your reasoning.

47. The larger star in the Algol binary system (see Figure 19-21*a*) is of spectral class K, while the smaller star is of spectral class B. Discuss how the color of Algol changes as seen through a small telescope (through which Algol appears as a single star). What is the color during a deep eclipse, when the large star eclipses the small one? What is the color when the small star eclipses the large one?

48. Suppose the detached star in β Lyrae (Figure 19-21*b*) did not have an accretion disk. Would the deeper dips in the light curve be deeper, shallower, or about the same? What about the shallower dip? Explain your answers.

49. The two stars that make up the overcontact binary W Ursae Majoris (Figure 19-21*c*) have estimated masses of 0.99 $M_\odot$ and 0.62 $M_\odot$. (**a**) Find the average separation between the two stars. Give your answer in kilometers. (**b**) The radii of the two stars are estimated to be 1.14 $R_\odot$ and 0.83 $R_\odot$. Show that these values and your result in part (a) are consistent with the statement that this is an overcontact binary.

50. The stars that make up the binary system W Ursae Majoris (see Figure 19-21*c*) have particularly strong magnetic fields. Explain how astronomers could have discovered this. (*Hint:* See Section 16-9.)

51. Consult recent issues of *Sky & Telescope* and *Astronomy* to find out when Mira will next reach maximum brightness. Look up the star's location in the sky using the *Starry Night Enthusiast*™ program on the CD-ROM that comes with selected printed copies of this book. (Use the **Find . . .** command in the **Edit** menu to search for Omicron Ceti.) Why is it unlikely that you will be able to observe Mira at maximum brightness?

Discussion Questions

52. Eventually the Sun's luminosity will increase to the point where the Earth can no longer sustain life. Discuss what measures a future civilization might take to preserve itself from such a calamity.

53. The half-life of the ^{8}Be nucleus, 2.6×10^{-16} second, is the average time that elapses before this unstable nucleus decays into two alpha particles. How would the universe be different if instead the ^{8}Be half-life were zero? How would the universe be different if the ^{8}Be nucleus were stable and did not decay?

54. Discuss how H-R diagrams of star clusters could be used to set limits on the age of the universe. Could they be used to set lower limits on the age? Could they be used to set upper limits? Explain your reasoning.

Web/eBook Questions

55. Suppose that an oxygen nucleus (^{16}O) were fused with a helium nucleus (^{4}He). What element would be formed? Look up the relative abundance of this element in, for example, the *Handbook of Chemistry and Physics* or on the World Wide Web. Based on the abundance, comment on whether such a process is likely. (*Hint:* See Figure 8-4.)

56. Although Polaris, the North Star, is a Cepheid variable, it pulsates in a somewhat different way than other Cepheids. Search the World Wide Web for information about this star's pulsations and how they have been measured by astronomers at the U.S. Naval Observatory. How does Polaris pulsate? How does this differ from other Cepheids?

57. **Observing Stellar Evolution.** Step through the animation "The Hertzsprung-Russell Diagram and Stellar Evolution" in Chapter 19 of the *Universe* Web site or eBook. Use this animation to answer

the following questions. (**a**) How does a 1-$M_\odot$ star move on the H-R diagram during its first 4.56 billion (4560 million) years of existence? Compare this with the discussion in Section 19-1 of how the Sun has evolved over the past 4.56 billion years. (**b**) What is the zero-age spectral class of a 2-$M_\odot$ star? At what age does such a star evolve into a red giant of spectral class K? (**c**) What is the approximate zero-age luminosity of a 1.3-$M_\odot$ star? What is its approximate luminosity when it becomes a red giant? (**d**) Suppose a star cluster has no main-sequence stars of spectral classes O or B. What is the approximate age of the cluster? (**e**) Approximately how long do the most massive stars of spectral class B live before leaving the main sequence? What about the most massive stars of spectral class F?

Activities

Observing Projects

Observing tips and tools

An excellent resource for learning how to observe variable stars is the Web site of the American Association of Variable Star Observers. A wealth of data about specific variable stars can be found on the *Sky & Telescope* Web site and in the three volumes of *Burnham's Celestial Handbook: An Observer's Guide to the Universe Beyond the Solar System* (Dover, 1978). This book also provides useful information about observing star clusters.

58. Observe several of the red giants and supergiants listed below with the naked eye and through a telescope. (Note that γ Andromedae, α Tauri, and ε Pegasi are all multiple-star systems. The spectral type and luminosity class refer to the brightest star in these systems.) You can locate these stars in the sky using the *Starry Night Enthusiast™* program on the CD-ROM that comes with selected printed copies of this book. Is the reddish color of these stars apparent when they are compared with neighboring stars?

Star	Spectral type	R.A.	Decl.
Almach (γ Andromedae)	K3 II	2^{h} 03.9^{m}	+42° 20′
Aldebaran (α Tauri)	K5 III	4 35.9	+16 31
Betelgeuse (α Orionis)	M2 I	5 55.2	+07 24
Arcturus (α Boötis)	K2 III	14 15.7	+19 11
Antares (α Scorpii)	M1 I	16 29.5	−26 26
Eltanin (γ Draconis)	K5 III	17 56.7	+51 29
Enif (ε Pegasi)	K2 I	21 44.2	+09 52

Note: The right ascensions and declinations are given for epoch 2000.

59. Several of the open clusters referred to in Figure 19-14 can be seen with a good pair of binoculars. Observe as many of these clusters as you can, using both a telescope and binoculars. (Some are actually so large that they will not fit in the field of view of many telescopes.) You can locate these clusters in the sky using the *Starry Night Enthusiast™* program on the CD-ROM that comes with selected printed copies of this book. Note that in the following table, the M prefix refers to the Messier Catalog, NGC refers to the New General Catalogue, and Mel refers to the Melotte Catalog. In making your own observations, note the overall distribution of stars in each cluster. Which clusters are seen better through binoculars than through a telescope? Which clusters can you see with the naked eye?

Star cluster	Constellation	R.A.	Decl.
h Persei (NGC 869)	Perseus	2^{h} 19.0^{m}	+57° 09′
χ Persei (NGC 884)	Perseus	2 22.4	+57 07
Pleiades (M45)	Taurus	3 47.0	+24 07
Hyades (Mel 25)	Taurus	4 27	+16 00
Praesepe (M44)	Cancer	8 40.1	+19 59
Coma (Mel 111)	Coma Berenices	12 25	+26 00
Wild Duck (M11)	Scutum	18 51.1	−06 16

Note: The right ascensions and declinations are given for epoch 2000.

60. There are many beautiful globular clusters scattered around the sky that can be easily seen with a small telescope. Several of the brightest and nearest globulars are listed below. You can locate these clusters in the sky using the *Starry Night Enthusiast™* program on the CD-ROM that comes with selected printed copies of this book. Observe as many of these globular clusters as you can. Can you distinguish individual stars toward the center of each cluster? Do you notice any differences in the overall distribution of stars between clusters?

Star cluster	Constellation	R.A.	Decl.
M3 (NGC 5272)	Canes Venatici	13^{h} 42.2^{m}	+28° 23′
M5 (NGC 5904)	Serpens	15 18.6	+2 05
M4 (NGC 6121)	Scorpius	16 23.6	−26 32
M13 (NGC 6205)	Hercules	16 41.7	+36 28
M12 (NGC 6218)	Ophiuchus	16 47.2	−1 57
M28 (NGC 6626)	Sagittarius	18 24.5	−24 52
M22 (NGC 6656)	Sagittarius	18 36.4	−23 54
M55 (NGC 6809)	Sagittarius	19 40.0	−30 58
M15 (NGC 7078)	Pegasus	21 30.0	+12 10

Note: The right ascensions and declinations are given for epoch 2000.

61. Use the *Starry Night Enthusiast™* program to view some of the objects described in this chapter. Click the **Home** button in the toolbar. **Stop time** flow and change the **Time** in the toolbar to

12:00:00 A.M. Use the **Find** pane to locate the giant star Aldebaran, the open cluster M44, and the globular cluster M12. (**a**) Which of these objects are visible from your home location at midnight tonight? If the object is visible from your home location at midnight, click the menu button next to the object's name in the Find pane and select the **Show Info** option. (**b**) For each object that is visible, use the Zoom controls at the upper right of the toolbar to get the best view. Describe the appearance of the object. (**c**) For each object that is visible, in which direction of the compass would you have to look at midnight to see it (that is, what is its *azimuth*)? How far above the horizon would you have to look (that is, what is its *altitude*)? (Look under the **Position in Sky** layer of the **Info** pane for this information.)

62. Use the *Starry Night Enthusiast*™ program to look for signs of stellar evolution in our Milky Way Galaxy. In the menu, select **Favourites > Stars > Sun in Milky Way.** You can zoom in or out using the buttons at the right of the toolbar. You can move your viewpoint around the Galaxy by holding down the **Shift** key and the mouse button as you drag the mouse. (**a**) What is the color of the central part of the Galaxy? Based on the color of this region, do you expect that there are many massive main-sequence stars there? Would you expect to find many young stars there? Explain your reasoning. (**b**) Repeat part (a) for the outer regions of the Galaxy.

Collaborative Exercises

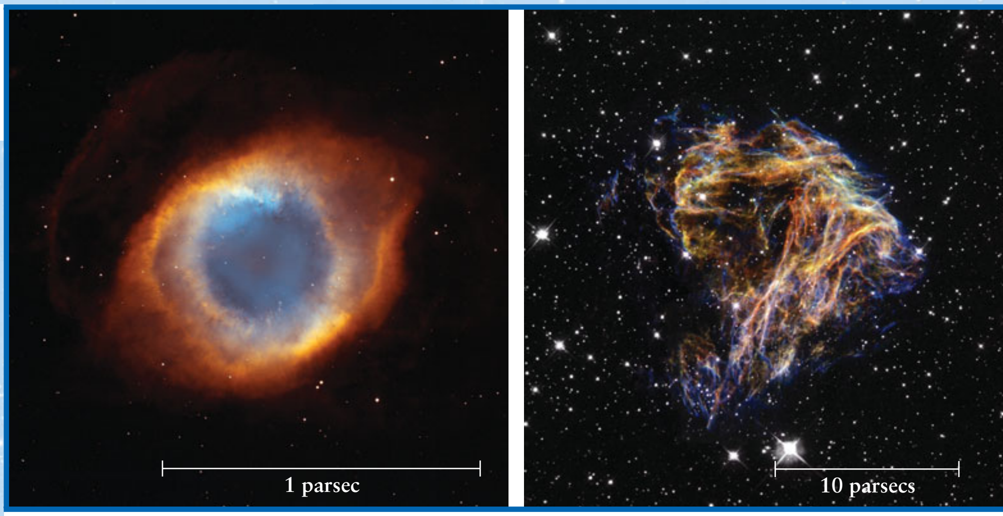
63. The inverse relationship between a star's mass and its main-sequence lifetime is sometimes likened to automobiles in that the more massive vehicles, such as commercial semi–tractor-trailer trucks, need to consume significantly more fuel to travel at highway speeds than more lightweight and economical vehicles. As a group, create a table called "Maximum Vehicle Driving Distances," much like Table 19-1, "Main-Sequence Lifetimes," by making estimates for any five vehicles of your groups' choosing. The table's column headings should be: (1) vehicle make and model; (2) estimated gas tank size; (3) cost to fill tank; (4) estimated mileage (in miles per gallon); and (5) number of miles driven on a single fill-up.

64. Consider Figure 19-19, showing the period-luminosity relation for Cepheids. If the length of time it has been since someone in your group last purchased milk or juice at the grocery store was identical to the pulsation period of a Type I Cepheid, how much longer a pulsation period would a Type II Cepheid need in order to reach the same luminosity as the Type I Cepheid in this time frame?

65. Figure 19-18*a* shows a light curve of apparent magnitude versus time in days for δ Cephei—a pulsating star that reaches maximum brightness every 5.4 days. Create a new sketch of apparent magnitude versus time in days showing three different stars: (1) δ Cephei; (2) a slightly smaller pulsating star; and (3) a slightly larger pulsating star, all of which have about the same total change in apparent magnitude.

20

Stellar Evolution: The Deaths of Stars



(a) A planetary nebula (b) A supernova remnant

WEB LINK 20.1

R I V U X G

Left: The planetary nebula NGC 7293. (the Helix Nebula) Right: The supernova remnant LMC N49. (NASA, NOAO, ESA, the Hubble Helix Nebula Team, M. Meixner/STScI, and T. A. Rector/NRAO; NASA and the Hubble Heritage Team, STScI/AURA)

When a star of 0.4 solar mass or more reaches the end of its main-sequence lifetime and becomes a red giant, it comes to have a compressed core and a bloated atmosphere. Finally, it devours its remaining nuclear fuels and begins to die. The character of its death depends crucially on the value of the star's mass.

A star of relatively low mass—such as our own Sun—ends its evolution by gently expelling its outer layers into space. These ejected gases form a glowing cloud called a *planetary nebula* such as the one shown here in the left-hand image. The burned-out core that remains is called a *white dwarf*.

In contrast, a high-mass star ends its life in almost inconceivable violence. At the end of its short life, the core of such a star collapses suddenly. This triggers a powerful *supernova* explosion that can be as luminous as an entire galaxy of stars. A white dwarf, too, can become a supernova if it accretes gas from a companion star in a close binary system.

Thermonuclear reactions in supernovae produce a wide variety of heavy elements, which are ejected into the interstellar medium. (The supernova remnant shown here in the right-hand image is rich in these elements.) Such heavy elements are essential building blocks for terrestrial worlds like our Earth. Thus, the deaths of massive stars can provide the seeds for planets orbiting succeeding generations of stars.

20-1 Stars of between 0.4 and 4 solar masses go through two distinct red-giant stages

All main-sequence stars convert hydrogen to helium in their cores in a series of energy-releasing thermonuclear reactions. As we saw in Section 19-1, convection within a low-mass main-sequence star—a so-called *red dwarf* with a mass between 0.08 and 0.4 $M_\odot$—will eventually bring all of the star's hydrogen into the core. Over hundreds of billions of years a red dwarf evolves into an inert ball of helium. Convection is less important in main-sequence stars with masses greater than 0.4 $M_\odot$, so these stars are able to consume only the hydrogen that is present within the core. These stars of greater mass then leave the main sequence. Let's examine what happens next to a star of moderately low mass, between 0.4 and 4 $M_\odot$. One example of such a star is our own Sun, with a mass of 1 $M_\odot$. We'll begin by reviewing what we learned

Learning Goals

By reading the sections of this chapter, you will learn

20-1 What kinds of thermonuclear reactions occur inside a star of moderately low mass as it ages

20-2 How evolving stars disperse carbon into the interstellar medium

20-3 How stars of moderately low mass eventually die

20-4 The nature of white dwarfs and how they are formed

20-5 What kinds of reactions occur inside a high-mass star as it ages

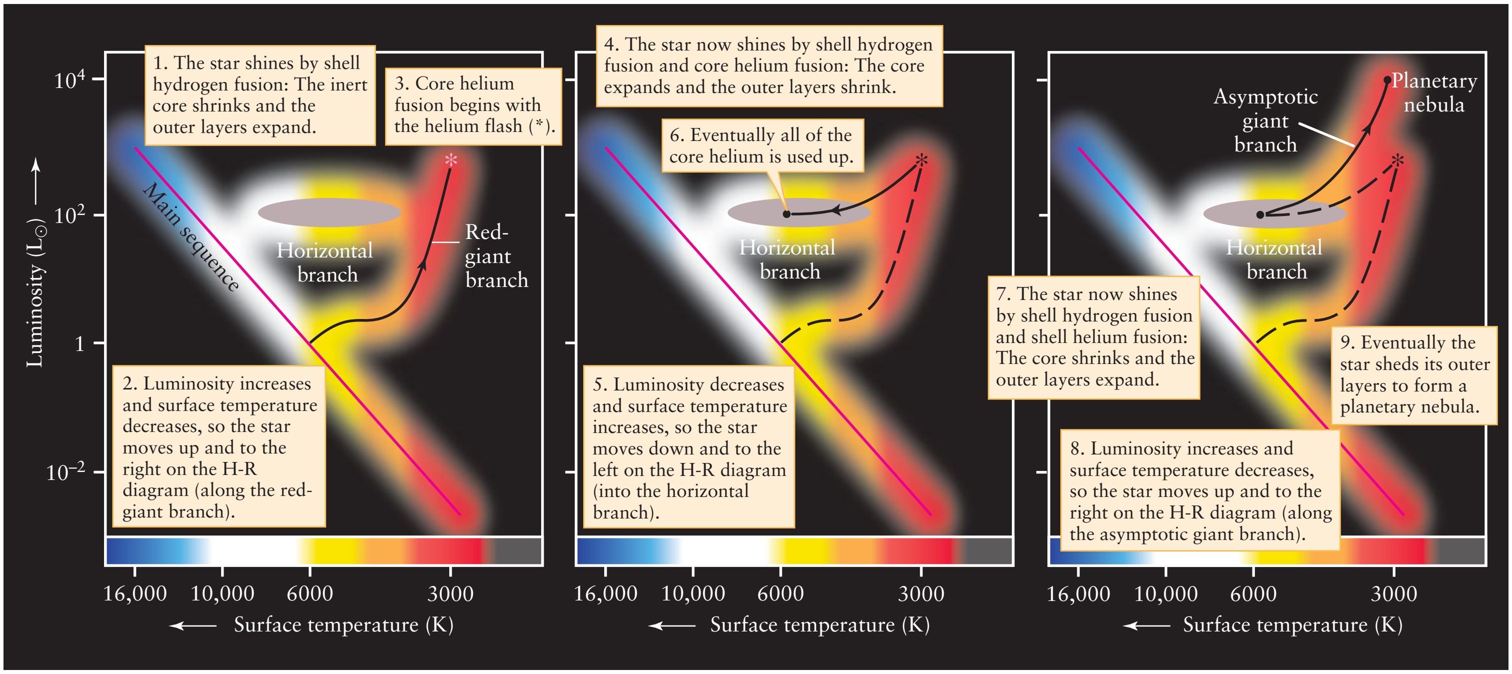
20-6 How high-mass stars explode and die

20-7 Why supernova SN 1987A was both important and unusual

20-8 What role neutrinos play in the death of a massive star

20-9 How white dwarfs in close binary systems can explode

20-10 What remains after a supernova explosion



(a) Before the helium flash: A red-giant star (b) After the helium flash: A horizontal-branch star (c) After core helium fusion ends: An AGB star

ANIMATION 20.1

Figure 20-1

The Post-Main-Sequence Evolution of a 1-$M_{\odot}$ Star These H-R diagrams show the evolutionary track of a star like the Sun as it goes through the stages of being **(a)** a red-giant star, **(b)** a horizontal-branch star, and **(c)** an asymptotic giant branch (AGB) star. The star eventually evolves into a planetary nebula (described in Section 20-3).

in Chapter 19 about the first stages of post–main-sequence evolution for such a star. (Later in this chapter we'll study the evolution of more massive stars.)

The Red-Giant and Horizontal-Branch Stages: A Review

We can describe a star's post–main-sequence evolution using an evolutionary track on a H-R diagram. Figure 20-1 shows the track for a 1-$M_{\odot}$ star like the Sun. Once core hydrogen fusion ceases, the core shrinks, heating the surrounding hydrogen and triggering shell hydrogen fusion. The new outpouring of energy causes the star's outer layers to expand and cool, and the star becomes a red giant. As the luminosity increases and the surface temperature drops, the post–main-sequence star moves up and to the right along the **red-giant branch** on an H-R diagram (Figure 20-1*a*).

Next, the helium-rich core of the star shrinks and heats until eventually **core helium fusion** begins. This second post–main-sequence stage begins gradually in stars more massive than about 2–3 $M_{\odot}$, but for less massive stars it comes suddenly—in a *helium flash*. During core helium fusion, the surrounding hydrogen-fusing shell still provides most of the red giant's luminosity.

As we learned in Section 19-3, the core expands when core helium fusion begins, which makes the core cool down a bit. (We saw in Box 19-1 that letting a gas expand tends to lower its temperature, while compressing a gas tends to increase its temperature.) The cooling of the core also cools the surrounding hydrogen-fusing shell, so that the shell releases energy more slowly. Hence, the luminosity goes down a bit after core helium fusion begins.

The slower rate of energy release also lets the star's outer layers contract. As they contract, they heat up, so the star's surface temperature increases and its evolutionary track moves to the left on the H-R diagram in Figure 20-1*b*. The luminosity changes relatively little during this stage, so the evolutionary track moves almost horizontally, along a path called the **horizontal branch.** Horizontal-branch stars have helium-fusing cores surrounded by hydrogen-fusing shells. Figure 19-12 shows horizontal-branch stars in a globular cluster, and Figure 19-8 shows the evolution of the luminosity of a 1-$M_{\odot}$ star up to this point in its history.

AGB Stars: The Second Red-Giant Stage

Helium fusion produces nuclei of carbon and oxygen. After about a hundred million (10^8) years of core helium fusion, essentially all the helium in the core of a 1-$M_{\odot}$ star has been converted into carbon and oxygen, the fusion of helium in the core ceases. (This corresponds to the right-hand end of the graph in Figure 19-8.) Without thermonuclear reactions to maintain the core's internal pressure, the core again contracts, until it is stopped by degenerate-electron pressure (described in Section 19-3). This contraction releases heat into the surrounding helium-rich gases, and

Stars like the Sun go through a second red-giant phase, during which helium fusion takes place in a shell around an inert core

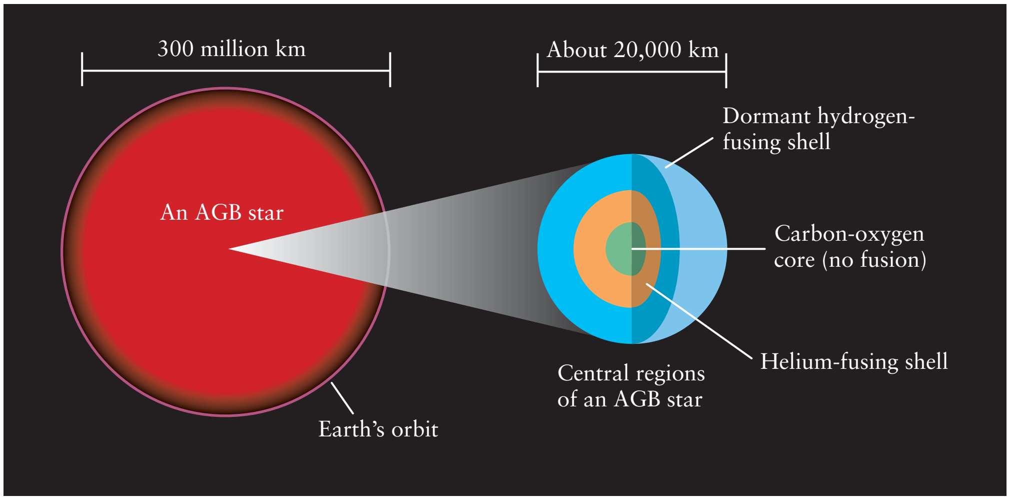


Figure 20-2

The Structure of an Old, Moderately Low-Mass AGB Star Near the end of its life, a star like the Sun becomes an immense, red, asymptotic giant branch (AGB) star. The star's inert core, active helium-fusing shell, and dormant hydrogen-fusing shell are all contained within a volume roughly the size of the Earth. Thermonuclear reactions in the helium-fusing shell are so rapid that the star's luminosity is thousands of times that of the present-day Sun. (The relative sizes of the shells in the star's interior are not shown to scale.)

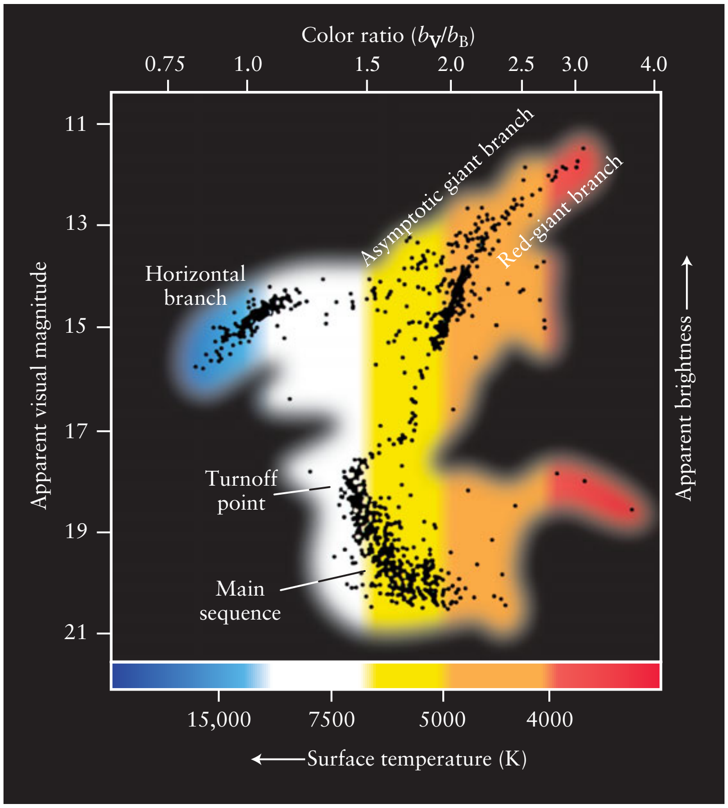
a new stage of helium fusion begins in a thin shell around the core. This process is called **shell helium fusion.**

History now repeats itself—the star enters a *second* red-giant phase. A star first becomes a red giant at the end of its main-sequence lifetime, when the outpouring of energy from shell hydrogen fusion makes the star's outer layers expand and cool. In the same way, the outpouring of energy from shell *helium* fusion causes the outer layers to expand again. The low-mass star ascends into the red-giant region of the H-R diagram for a second time (Figure 20-1*c*), but now with even greater luminosity than during its first red-giant phase.

Stars in this second red-giant phase are commonly called **asymptotic giant branch stars,** or **AGB stars,** and their evolutionary tracks follow what is called the **asymptotic giant branch.** (*Asymptotic* means "approaching"; the name means that a star on the asymptotic giant branch approaches the red-giant branch from the left on an H-R diagram.)

When a low-mass star first becomes an AGB star, it consists of an inert, degenerate carbon-oxygen core and a helium-fusing shell, both inside a hydrogen-fusing shell, all within a volume not much larger than the Earth. This small, dense central region is surrounded by an enormous hydrogen-rich envelope about as big as Earth's orbit around the Sun. After a while, the expansion of the star's outer layers causes the hydrogen-fusing shell to also expand and cool, and thermonuclear reactions in this shell temporarily cease. This leaves the aging star's structure as shown in Figure 20-2.

We saw in Section 19-1 that the more massive a star, the shorter the amount of time it remains on the main sequence. Similarly, the greater the star's mass, the more rapidly it goes through the stages of post–main-sequence evolution. Hence, we can see all of these stages by studying star clusters, which contain stars that are all the same age but that have a range of masses (see Section 19-4). Figure 20-3 shows a color-magnitude diagram for the globular cluster M55, which is at least 13 billion years old. The least massive stars in this cluster are still on the main sequence. Progressively more massive stars have evolved to the red giant branch, the horizontal branch, and the asymptotic giant branch.



WEB LINK 20.2

Figure 20-3

Stellar Evolution in a Globular Cluster In the old globular cluster M55, stars with masses less than about 0.8 $M_\odot$ are still on the main sequence, converting hydrogen into helium in their cores. Slightly more massive stars have consumed their core hydrogen and are ascending the red-giant branch; even more massive stars have begun helium core fusion and are found on the horizontal branch. The most massive stars (which still have less than 4 $M_\odot$) have consumed all the helium in their cores and are ascending the asymptotic giant branch. (Compare with Figure 21-11.) (Adapted from D. Schade, D. VandenBerg, and F. Hartwick)

A 1-$M_\odot$ AGB star can reach a maximum luminosity of nearly 10^4 $L_\odot$, as compared with approximately 10^3 $L_\odot$ when it reached the helium flash and a relatively paltry 1 $L_\odot$ during its main-sequence lifetime. When the Sun becomes an AGB star some 12.3 billion years from now, this tremendous increase in luminosity will cause Mars and the Jovian planets to largely evaporate away. The Sun's bloated outer layers will reach to the Earth's orbit. Mercury and perhaps Venus will simply be swallowed whole.

20-2 Dredge-ups bring the products of nuclear fusion to a giant star's surface

As we saw in Section 16-2, energy is transported outward from a star's core by one of two processes—radiative diffusion or convection. The first is the passage of energy in the form of electromagnetic radiation, and it dominates only when a star's gases are relatively transparent. The second involves up-and-down movement of the star's gases. Convection plays a very important role in giant stars, because it helps supply the cosmos with the elements essential to life.

Convection, Dredge-ups, and Carbon Stars

ANIMATION 20.2

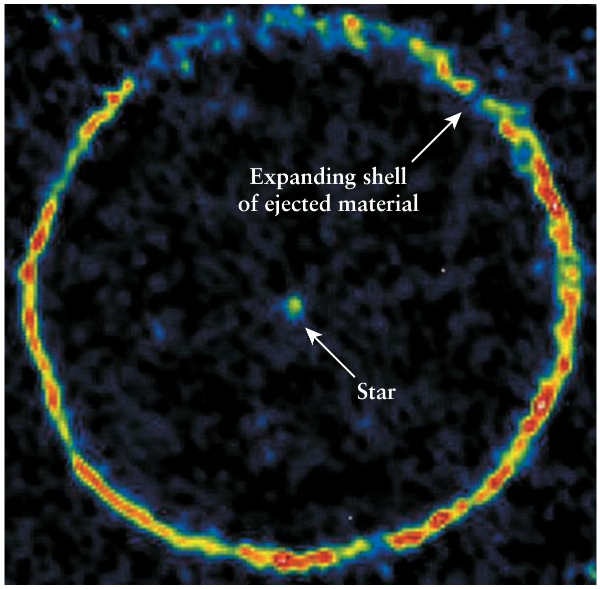
In the Sun, convection dominates only the outer layers, from around 0.71 solar radius (measured from the center of the Sun) up to the photosphere (recall Figure 16-4). During the final stages of a star's life, however, the convective zone can become so broad that it extends down to the star's core. At these times, convection can "dredge up" the heavy elements produced in and around the core by thermonuclear fusion, transporting them all the way to the star's surface.

LOOKING DEEPER 20.1

The *first* **dredge-up** takes place after core hydrogen fusion stops, when the star becomes a red giant for the first time. Convection dips so deeply into the star that material processed by the CNO cycle of hydrogen fusion (see Section 16-1) is carried up to the star's surface, changing the relative abundances of carbon, nitrogen, and oxygen. A *second* dredge-up occurs after core helium fusion ceases, further altering the abundances of carbon, nitrogen, and oxygen. Still later, during the AGB stage, a *third* dredge-up can occur if the star has a mass greater than about 2 $M_\odot$. This third dredge-up transports large amounts of freshly synthesized carbon to the star's surface, and the star's spectrum thus exhibits prominent absorption bands of carbon-rich molecules like C_2, CH, and CN. For this reason, an AGB star that has undergone a third dredge-up is called a **carbon star.**

All AGB stars have very strong stellar winds that cause them to lose mass at very high rates, up to 10^{-4} $M_\odot$ per year (a thousand times greater than that of a red giant, and 10^{10} times greater than the rate at which our present-day Sun loses mass). The surface temperature of AGB stars is relatively low, around 3000 K, so any ejected carbon-rich molecules can condense to form tiny grains of soot. Indeed, carbon stars are commonly found to be obscured in sooty cocoons of ejected matter (Figure 20-4).

The carbon that forms the basis of all life on Earth was ejected billions of years ago from giant stars



WEB LINK 20.3

Figure 20-4 R I V U X G

A Carbon Star TT Cygni is an AGB star in the constellation Cygnus that ejects some of its carbon-rich outer layers into space. Some of the ejected carbon combines with oxygen to form molecules of carbon monoxide (CO), whose emissions can be detected with a radio telescope. This radio image shows the CO emissions from a shell of material that TT Cygni ejected some 7000 years ago. Over that time, the shell has expanded to a diameter of about ½ light-year. (H. Olofsson, Stockholm Observatory, et al./NASA)

Carbon stars are important because they enrich the interstellar medium with carbon and some nitrogen and oxygen. The triple alpha process that occurs in helium fusion is the *only* way that carbon can be made, and carbon stars are the primary avenue by which this element is dispersed into interstellar space. Indeed, most of the carbon in your body was produced long ago inside a star by the triple alpha process (see Section 19-3). This carbon was later dredged up to the star's surface and ejected into space. Some 4.56 billion years ago a clump of the interstellar medium which contained this carbon coalesced into the solar nebula from which our Earth—and all of the life on it—eventually formed. In this sense you can think of your body as containing "recycled" material—substances that were once in the heart of a star that formed and evolved long before our solar system existed.

20-3 Stars of moderately low mass die by gently ejecting their outer layers, creating planetary nebulae

For a star that began with a moderately low mass (between about 0.4 and 4 $M_\odot$), the AGB stage in its evolution is a dramatic turning point. Before this stage, a star loses mass only gradually through steady stellar winds. But as it evolves during its AGB stage, a star divests itself completely of its outer layers. The aging star undergoes a series of bursts in luminosity, and in each burst it ejects a shell of material into space. (The shell around the AGB star TT Cygni, shown in Figure 20-4, was probably created in this way.) Eventually, all that remains of a low-mass star is a fiercely hot, exposed core, surrounded by glowing shells of ejected gas. This late stage in the life of a star is called a **planetary nebula.** The left-hand image on the opening page of this chapter shows one such planetary nebula, called the Ring Nebula for its shape.

Like a human going through a midlife crisis, an aging AGB star casts off much of the mass that it possesses and makes a cosmic spectacle of itself

Making a Planetary Nebula

To understand how an AGB star can eject its outer layers in shells, consider the internal structure of such a star as shown in Figure 20-2. As the helium in the helium-fusing shell is used up, the pressure that holds up the dormant hydrogen-fusing shell decreases. Hence, the dormant hydrogen shell contracts and heats up, and hydrogen fusion begins anew. This revitalized hydrogen fusion creates helium, which rains downward onto the temporarily dormant helium-fusing shell. As the helium shell gains mass, it shrinks and heats up. When the temperature of the helium shell reaches a certain critical value, it reignites in a **helium shell flash** that is similar to (but less intense than) the helium flash that occurred earlier in the evolution of a low-mass star (see Section 19-3). The released energy pushes the hydrogen-fusing shell outward, making it cool off, so that hydrogen fusion ceases and this shell again becomes dormant. The process then starts over again.

When a helium shell flash occurs, the luminosity of an AGB star increases substantially in a relatively short-lived burst called a **thermal pulse.** Figure 20-5, which is based on a theoretical calculation of the evolution of a 1-$M_\odot$ star, shows that thermal pulses begin when the star is about 12.365 billion years old. The calculations predict that thermal pulses occur at ever-shorter intervals of about 100,000 years.

During these thermal pulses, the dying star's outer layers can separate completely from its carbon-oxygen core. As the ejected material expands into space, dust grains condense out of the cooling gases. Radiation pressure from the star's hot, burned-out core acts on the specks of dust, propelling them further outward, and the star sheds its outer layers altogether. In this way an aging 1-$M_\odot$ star loses as much as 40% of its mass. More massive stars eject even greater fractions of their original mass.

As a dying star ejects its outer layers, the star's hot core becomes exposed. With a surface temperature of about 100,000 K, this exposed core emits ultraviolet radiation intense enough to ionize and excite the expanding shell of ejected gases. These gases therefore glow and emit visible light through the process of fluorescence (see Box 18-1), producing a planetary nebula like those shown in Figure 20-6.

CAUTION! Despite their name, planetary nebulae have nothing to do with planets. This misleading term was introduced in the nineteenth century because these glowing objects looked like distant Jovian planets when viewed through the small telescopes then available. The difference between planets and planetary nebulae became obvious with the advent of spectroscopy: Planets have *absorption* line spectra (see Section 7-3), but the excited gases of planetary nebulae have *emission* line spectra.

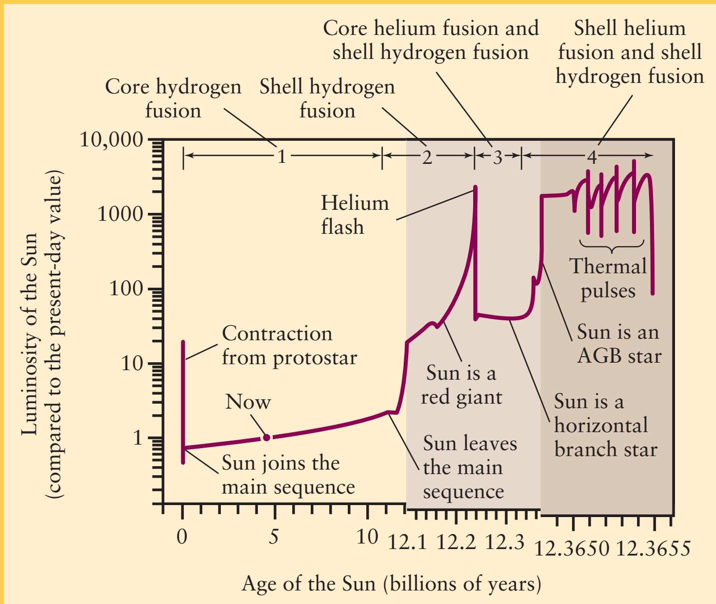


Figure 20-5

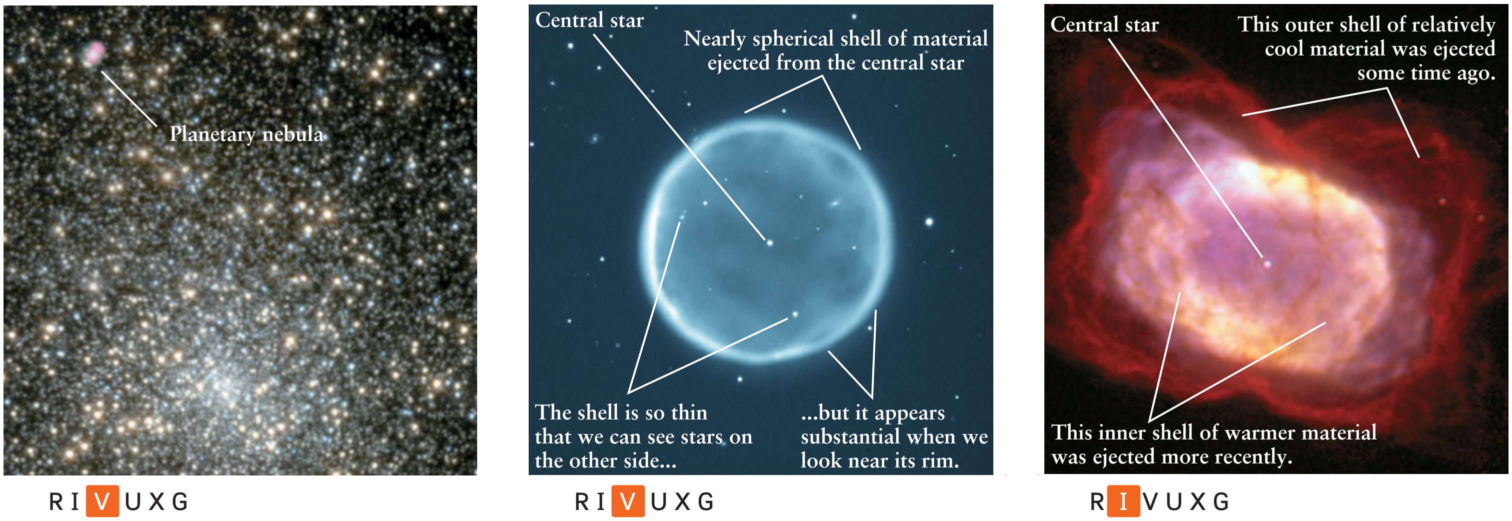
Further Stages in the Evolution of the Sun This diagram, which shows how the luminosity of the Sun (a 1-$M_\odot$ star) changes over time, is an extension of Figure 19-8. We use different scales for the final stages because the evolution is so rapid. During the AGB stage there are brief periods of runaway helium fusion, causing spikes in luminosity called thermal pulses. (Adapted from Mark A. Garlick, based on calculations by I.-Juliana Sackmann and Kathleen E. Kramer)

The Properties of Planetary Nebulae

Planetary nebulae are quite common. Astronomers estimate that there are 20,000 to 50,000 planetary nebulae in our Galaxy alone. Many planetary nebulae, such as those in Figure 20-6, are more or less spherical in shape. This is a result of the symmetrical way in which the gases were ejected. But if the rate of expansion is not the same in all directions, the resulting nebula takes on an hourglass or dumbbell appearance (Figure 20-7).

Spectroscopic observations of planetary nebulae show emission lines of ionized hydrogen, oxygen, and nitrogen. From the Doppler shifts of these lines, astronomers have concluded that the expanding shell of gas moves outward from a dying star at speeds from 10 to 30 km/s. For a shell expanding at such speeds to have attained the typical diameter of a planetary nebula, about 1 light-year, it must have begun expanding about 10,000 years ago. Thus, by astronomical standards, the planetary nebulae we see today were created only very recently.

We do not observe planetary nebulae that are more than about 50,000 years old. After this length of time, the shell has spread out so far from the cooling central star that its gases cease to glow and simply fade from view. The nebula's gases then mix with the surrounding interstellar medium.

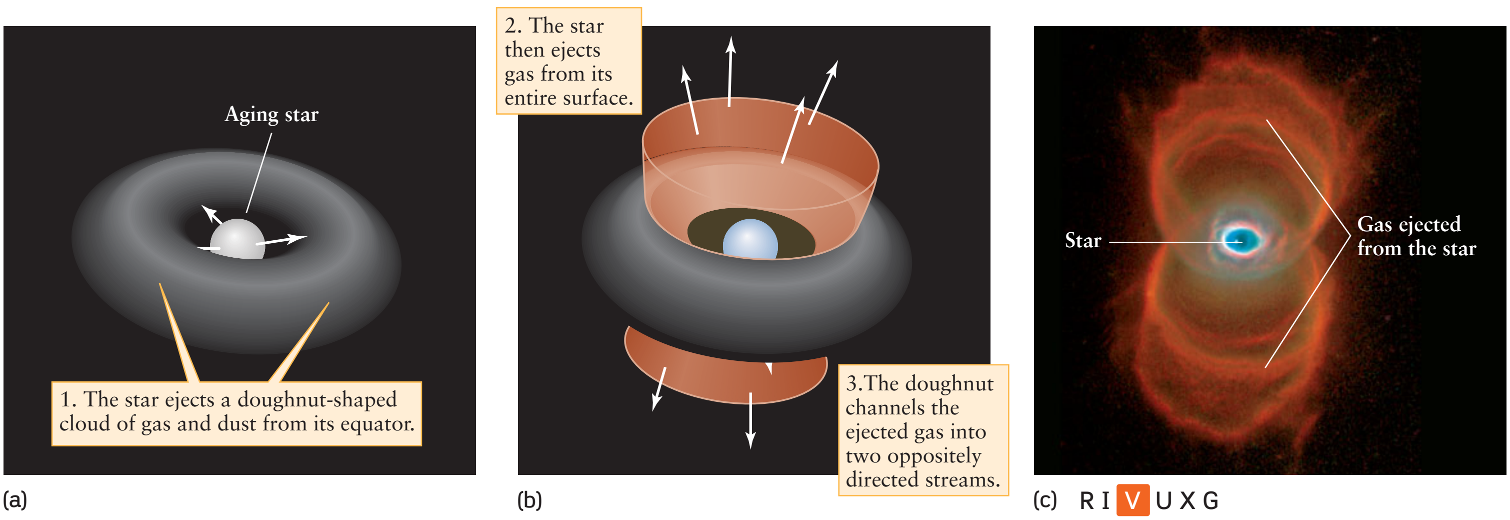


WEB LINK 20.4

Figure 20-6

Planetary Nebulae (a) The pinkish blob is a planetary nebula surrounding a star in the globular cluster M15, about 10,000 pc (33,000 ly) from Earth in the constellation Pegasus. **(b)** The planetary nebula Abell 39 lies about 2200 pc (7000 ly) from Earth in the constellation Hercules. The almost perfectly spherical shell that comprises the nebula is about 1.5 pc (5 ly) in diameter; the thickness of the shell is only about 0.1 pc (0.3 ly). **(c)** This infrared image of the planetary nebula NGC 7027 suggests a more complex evolutionary history than that of Abell 39. NGC 7027 is about 900 pc (3000 ly) from Earth in the constellation Cygnus and is roughly 14,000 AU across. (a: NASA/Hubble Heritage Team, STScI/AURA; b: WIYN/NOAO/NSF; c: William B. Latter, SIRTF Science Center/Caltech, and NASA)

Astronomers estimate that all the planetary nebulae in the Galaxy return a total of about 5 $M_\odot$ to the interstellar medium each year. This amounts to 15% of all the matter expelled by all the various sorts of stars in the Galaxy each year. Because this contribution is so significant, and because the ejected material includes heavier elements (metals) manufactured within a nebula's central star, planetary nebulae play an important role in the chemical evolution of the Galaxy as a whole.



WEB LINK 20.5

Figure 20-7

Making an Elongated Planetary Nebula (a), (b) These illustrations show one proposed explanation for why many planetary nebulae have an elongated shape. **(c)** The planetary nebula MyCn18, shown here in false color, may have acquired its elongated shape in this way. It lies some 2500 pc (8000 ly) from Earth in the constellation Musca (the Fly). (R. Sahai and J. Trauger, Jet Propulsion Laboratory; the WFPC-2 Science Team; and NASA)

20-4 The burned-out core of a moderately low-mass star cools and contracts until it becomes a white dwarf

We have seen that after a moderately low-mass star (from about 0.4 to about 4 solar masses) consumes all the hydrogen in its core, it is able to ignite thermonuclear reactions that convert helium to carbon and oxygen. Given sufficiently high temperature and pressure, carbon and oxygen can also undergo fusion reactions that release energy. But for such a moderately low-mass star, the core temperature and pressure never reach the extremely high values needed for these reactions to take place. Instead, as we have seen, the process of mass ejection just strips away the star's outer layers and leaves behind the hot carbon-oxygen core. With no thermonuclear reactions taking place, the core simply cools down like a dying ember. Such a burnt-out relic of a star's former glory is called a **white dwarf.** Such white dwarfs prove to have exotic physical properties that are wholly unlike any object found on Earth.

CAUTION! Unfortunately, the word *dwarf* is used in astronomy for several very different kinds of small objects. Here's a review of the three kinds that we have encountered so far in this book. A *white* dwarf is the relic that remains at the very end of the evolution of a star of initial mass between about 0.4 $M_\odot$ and 4 $M_\odot$. Thermonuclear reactions are no longer taking place in its interior; it emits light simply because it is still hot. A *red* dwarf, discussed in Section 19-1, is a cool main-sequence star with a mass between about 0.08 $M_\odot$ and 0.4 $M_\odot$. The energy emitted by a red dwarf in the form of light comes from its core, where fusion reactions convert hydrogen into helium. Finally, a *brown* dwarf (see Section 8-6 and Section 17-5) is an object like a main-sequence star but with a mass less than about 0.08 $M_\odot$. Because its mass is so small, its internal pressure and temperature are too low to sustain thermonuclear reactions. Instead, a brown dwarf emits light because it is slowly contracting, a process that releases energy (see Section 16-1). White dwarfs are comparable in size to the Earth (see Section 17-7); by contrast, brown dwarfs are larger than the planet Jupiter, and red dwarfs are even larger.

Properties of White Dwarfs

You might think that without thermonuclear reactions to provide internal heat and pressure, a white dwarf should keep on shrinking under the influence of its own gravity as it cools. Actually, however, a cooling white dwarf maintains its size, because the burnt-out stellar core is so dense that most of its electrons are degenerate (see Section 17-3). Thus, degenerate-electron pressure supports the star against further collapse. This pressure does not depend on temperature, so it continues to hold up the star even as the white dwarf cools and its temperature drops.

A white dwarf is kept from collapsing by the pressure of its degenerate electrons

Many white dwarfs are found in the solar neighborhood, but all are too faint to be seen with the naked eye. One of the first white dwarfs to be discovered is a companion to Sirius, the brightest star in the night sky. In 1844 the German astronomer Friedrich Bessel noticed that Sirius was moving back and forth slightly, as if it was being orbited by an unseen object. This companion, designated Sirius B (Figure 20-8), was first glimpsed in 1862 by the American astronomer Alvan Clark. Recent Hubble Space Telescope observations at ultraviolet wavelengths, where hot white dwarfs emit most of their light, show that the surface temperature of Sirius B is 25,200 K. (By contrast, the main-sequence star Sirius A has a surface temperature of 10,500 K, while the Sun's surface temperature is a relatively frosty 5800 K.)

Observations of white dwarfs in binary systems like Sirius allow astronomers to determine the mass, radius, and density of these stars (see Sections 17-9, 17-10, and 17-11). Such observations show that the density of the degenerate matter in a white dwarf is typically 10^9 kg/m^3 (a million times denser than water). A teaspoonful of white dwarf matter brought to Earth would weigh nearly 5.5 tons—as much as an elephant!

The Mass-Radius Relation for White Dwarfs

As we learned in Section 17-3, degenerate matter has a very different relationship between its pressure, density, and temperature

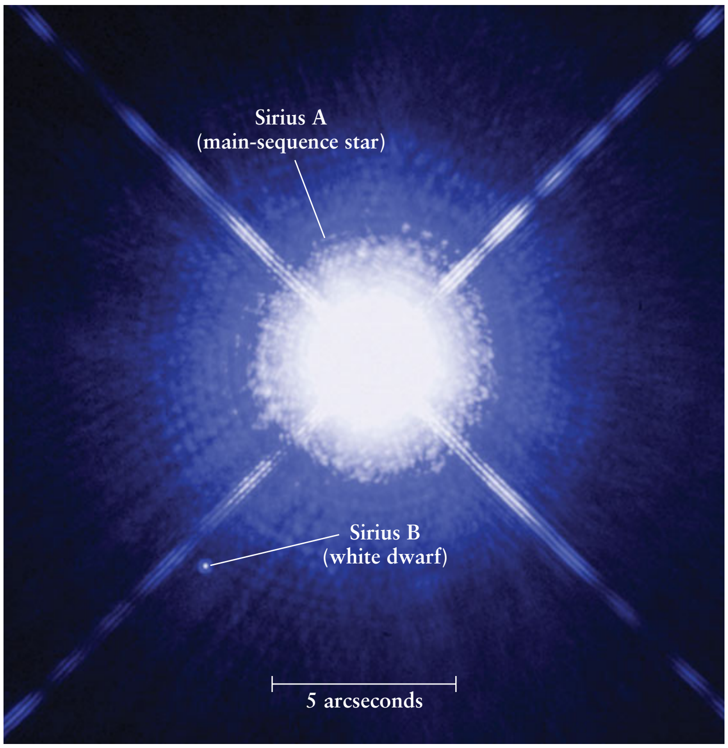


Figure 20-8 R I V U X G
Sirius A and Its White Dwarf Companion Sirius, the brightest-appearing star in the sky, is actually a binary star: The secondary star, called Sirius B, is a white dwarf. In this Hubble Space Telescope image, Sirius B is almost obscured by the glare of the overexposed primary star, Sirius A, which is about 10^4 times more luminous than Sirius B. The halo and rays around Sirius A are the result of optical effects within the telescope. (NASA; H. E. Bond and E. Nelan, STScI; M. Barstow and M. Burleigh, U. of Leicester; and J. B. Holberg, U. of Arizona)

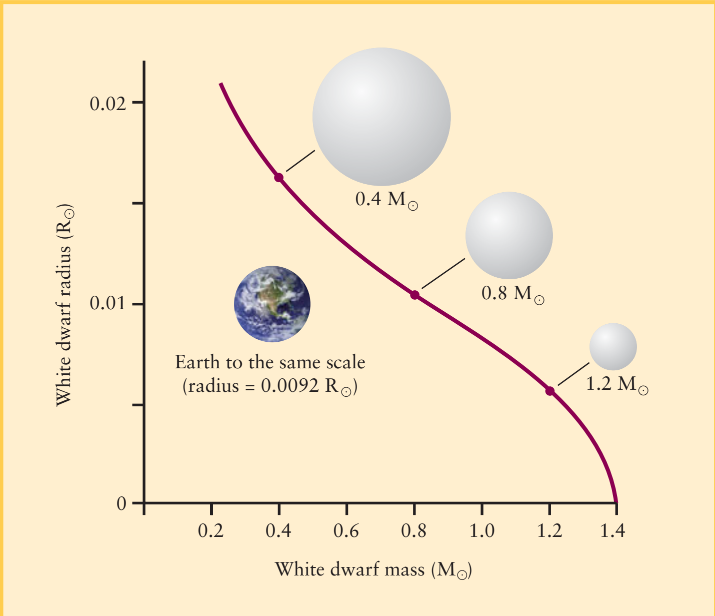


Figure 20-9

The Mass-Radius Relationship for White Dwarfs The more massive a white dwarf is, the smaller its radius. (The drawings of white dwarfs of different mass are drawn to the same scale as the image of the Earth.) This unusual relationship is a result of the degenerate-electron pressure that supports the star. The maximum mass of a white dwarf, called the Chandrasekhar limit, is 1.4 $M_\odot$.

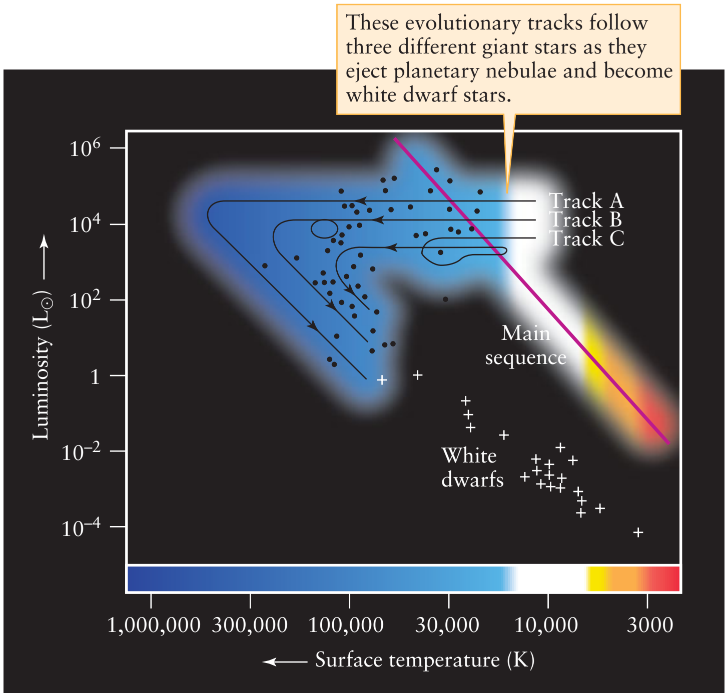
than that of ordinary gases. Consequently, white dwarf stars have an unusual **mass-radius relation:** The more massive a white dwarf star, the *smaller* it is.

Figure 20-9 displays the mass-radius relation for white dwarfs. Note that the more degenerate matter you pile onto a white dwarf, the smaller it becomes. However, there is a limit to how much pressure degenerate electrons can produce. As a result, there is an upper limit to the mass that a white dwarf can have. This maximum mass is called the **Chandrasekhar limit,** after the Indian-American scientist Subrahmanyan Chandrasekhar, who pioneered theoretical studies of white dwarfs in the 1930s. (The orbiting Chandra X-ray Observatory, described in Section 6-7, is named in his honor.) The Chandrasekhar limit is equal to 1.4 $M_\odot$, meaning that all white dwarfs must have masses less than 1.4 $M_\odot$.

The material inside a white dwarf consists mostly of ionized carbon and oxygen atoms floating in a sea of degenerate electrons. As the dead star cools, the carbon and oxygen ions slow down, and electric forces between the ions begin to prevail over the random thermal motions. About 5×10^9 years after the star first becomes a white dwarf, when its luminosity has dropped to about 10^{-4} $L_\odot$ and its surface temperature is a mere 4000 K, the ions no longer move freely. Instead, they arrange themselves in orderly rows, like an immense crystal lattice. From this time on, you could say that the star is "solid." The degenerate electrons move around freely in this crystal material, just as electrons move freely through an electrically conducting metal like copper or silver. A diamond is also crystallized carbon, so a cool carbon-oxygen white dwarf resembles an immense spherical diamond!

From Red Giant to Planetary Nebula to White Dwarf

Figure 20-10 shows the evolutionary tracks followed by three burned-out stellar cores as they pass through the planetary nebula stage and become white dwarfs. When these three stars were red giants, they had masses of 0.8, 1.5, and 3.0 $M_\odot$. Mass ejection strips these dying stars of up to 60% of their matter. During their final spasms, the luminosity and surface temperature of these stars change quite rapidly. The points representing these stars on an H-R diagram race along their evolutionary tracks, sometimes executing loops corresponding to thermal pulses (see



Evolutionary track	Mass ($M_\odot$): Giant star	Ejected nebula	White dwarf
A	3.0	1.8	1.2
B	1.5	0.7	0.8
C	0.8	0.2	0.6

Figure 20-10

Evolution from Giants to White Dwarfs This H-R diagram shows the evolutionary tracks of three low-mass giant stars as they eject planetary nebulae. The table gives the extent of mass loss in each case. The dots represent the central stars of planetary nebulae whose surface temperatures and luminosities have been determined; the crosses represent white dwarfs of known temperature and luminosity. (Adapted from B. Paczynski)

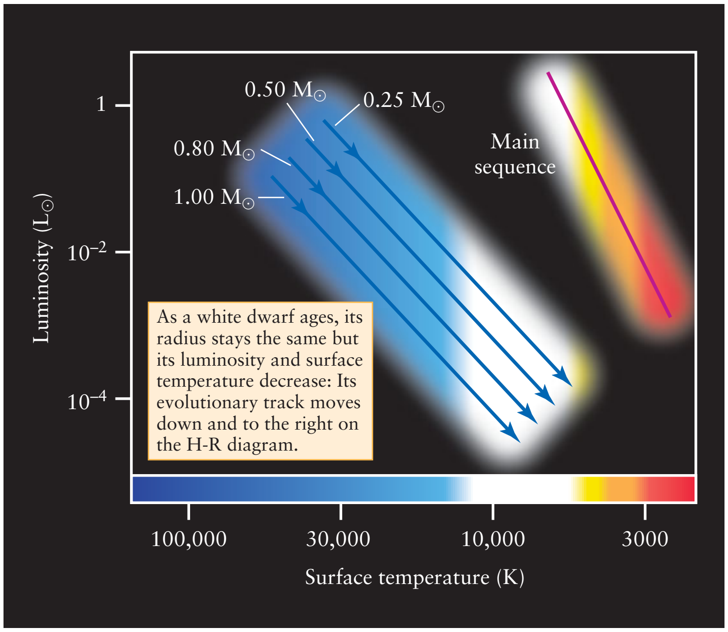


Figure 20-11

White Dwarf "Cooling Curves" As white dwarf stars radiate their internal energy into space, they become dimmer and cooler. The blue lines show the evolutionary tracks of four white dwarfs of different mass: The more massive a white dwarf, the smaller and hence fainter it is. Compare these "cooling curves" with the lines of constant radius in Figure 17-15*b*.

Track B and Track C in Figure 20-10). Finally, as the ejected nebulae fade and the stellar cores cool, the evolutionary tracks of these dying stars take a sharp turn toward the white dwarf region of the H-R diagram. As the table accompanying Figure 20-10 shows, the final white dwarf has only a fraction of the mass of the giant star from which it evolved.

Although a white dwarf maintains the same size as it cools, its luminosity and surface temperature both decrease with time. Consequently, the evolutionary tracks of aging white dwarfs point toward the lower right corner of the H-R diagram. You can see this in Figure 20-10; Figure 20-11 shows it in more detail. The energy that the white dwarf radiates into space comes only from the star's internal heat, which is a relic from the white dwarf's past existence as a stellar core. Over billions of years, white dwarfs grow dimmer and dimmer as their surface temperatures drop toward absolute zero.

After ejecting much of its mass into space, our own Sun will eventually evolve into a white dwarf star about the size of the Earth and with perhaps one-tenth of its present luminosity. It will become even dimmer as it cools. After 5 billion years as a white dwarf, the Sun will radiate with no more than one ten-thousandth of its present brilliance. With the passage of eons, our Sun will simply fade into obscurity. The *Cosmic Connections* figure summarizes the full evolutionary cycle of a 1-$M_\odot$ star like the Sun, from its birth as a main-sequence star to its demise as a white dwarf.

20-5 High-mass stars create heavy elements in their cores

During the entire lifetime of a low-mass red dwarf star (with an initial mass less than about 0.4 $M_\odot$), the only thermonuclear reaction that takes place is the fusion of hydrogen nuclei to form helium nuclei. In stars with initial masses from about 0.4 $M_\odot$ to about 4 $M_\odot$, a second kind of thermonuclear reaction takes place—helium fusion. The heaviest elements manufactured by helium fusion are carbon and oxygen.

The life story of a *high-mass* star (with an initial, zero-age mass greater than about 4 $M_\odot$) begins with these same reactions. But theoretical calculations show that high-mass stars can also go through several additional stages of thermonuclear reactions involving the fusion of carbon, oxygen, and other heavy nuclei. As a result, high-mass stars end their lives quite differently from low-mass stars.

Heavy-Element Fusion in Massive Stars

Why is fusion of heavy nuclei possible only in a high-mass star? The reason is that heavy nuclei have large electric charges: For example, a nucleus of carbon has 6 positively charged protons and hence 6 times the charge of a hydrogen nucleus (which has a single proton). This means that there are strong electric forces that tend to keep these nuclei apart. Only at the great speeds associated with extremely high temperatures can the nuclei travel fast enough to overcome their mutual electric repulsion and fuse together. To produce these very high temperatures at a star's center, the pressure must also be very high. Hence, the star must have a very large mass, because only such a star has strong enough gravity trying to pull it together and thus strong enough pressure at its center.

As we discussed in Section 19-2, when a main-sequence star with a mass greater than about 0.4 $M_\odot$ uses up its core hydrogen, it begins shell hydrogen fusion and enters a red-giant phase. Such a star then begins core helium fusion when the core temperature becomes high enough. The differences between moderately low-mass (from about 0.4 $M_\odot$ to about 4 $M_\odot$) and high-mass stars (more than about 4 $M_\odot$) become pronounced after helium core fusion ends, when the core is composed primarily of carbon and oxygen.

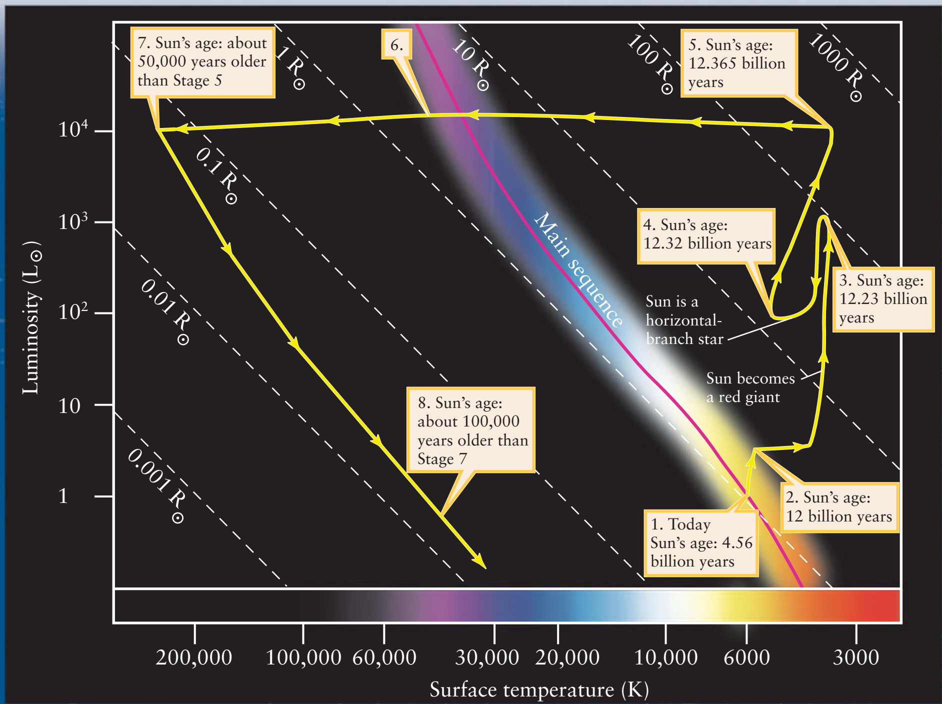
Let us consider how the late stages in the evolution of a high-mass star differ from those of a low-mass star. In low-mass stars, as we saw in Section 20-4, the carbon-oxygen core eventually becomes exposed and becomes a white dwarf. But in stars whose overall mass is more than about 4 $M_\odot$, the carbon-oxygen core is more massive than the Chandrasekhar limit of 1.4 $M_\odot$, so degenerate-electron pressure cannot prevent the core from contracting and heating. Hence, a high-mass star is able to enter a new round of core thermonuclear reactions. When the central temperature of such a high-mass star reaches 600 million kelvins (6×10^8 K), the first of the new thermonuclear reactions, **carbon fusion**, begins. Carbon fusion consumes carbon nuclei (^{12}C, with 6 protons in each nucleus) and produces oxygen (^{16}O, 8 protons), neon (^{20}Ne, 10 protons), sodium (^{23}Na, 11 protons), and magnesium (^{23}Mg and ^{24}Mg, each with 12 protons).

If a star has an even larger main-sequence mass of about 8 $M_\odot$ or so (before mass ejection), even more thermonuclear

COSMIC CONNECTIONS

Our Sun: The Next Eight Billion Years

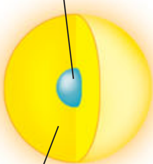
The Sun is presently less than halfway through its lifetime as a main-sequence star. The H-R diagram and cross-sections on this page summarize the dramatic changes that will take place when the Sun's main-sequence lifetime comes to an end.



NOTE: The illustrations below do *not* show the dramatic changes in the Sun's radius as it evolves. The sizes of the various layers are not shown to scale.

1. On the main sequence

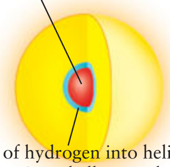
The present-day Sun is a main-sequence star – in its core, hydrogen fuses to produce helium.



Fusion does not occur in the outer layers (which contain predominantly hydrogen and helium).

2. Becoming a red giant

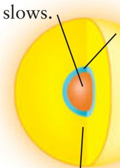
At the end of the Sun's main-sequence lifetime, fusion stops in the core (which has been converted to helium).



Fusion of hydrogen into helium continues in a shell around the core. The core shrinks, accelerating the fusion reactions in the shell and making the outer layers expand and cool.

3. The helium flash

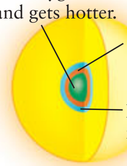
As the core contracts and heats, the core helium begins to fuse to make carbon and oxygen. The core expands and the rate of energy release slows.



The outer layers (where there are still no fusion reactions) contract and get hotter due to the slower rate of energy release.

4. Beginning the second red giant phase

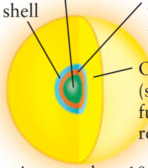
Once the core helium is consumed, what remains is an inert core of carbon and oxygen. The core again shrinks and gets hotter.



The shrinkage of the core again accelerates fusion reactions in the shells, making the inert outer layers expand and cool.

5. The Sun reaches its maximum size

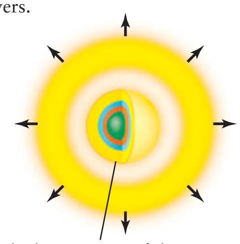
Inert carbon-oxygen core



The Sun is more than 100 times larger in radius than when it was a main-sequence star. Part of the outer layers escapes into space in a stellar wind.

6. A planetary nebula

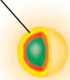
Thermal pulses cause spikes in luminosity that eject the star's outer layers.



As the hot interior of the star is exposed, we observe an increase in the star's surface temperature.

7. The end of nuclear reactions

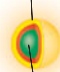
With the outer layers gone, the pressure on the shells around the core is too little to sustain nuclear reactions.



The star still glows intensely because of its high temperature. As energy is lost in the form of electromagnetic radiation, the star slowly cools.

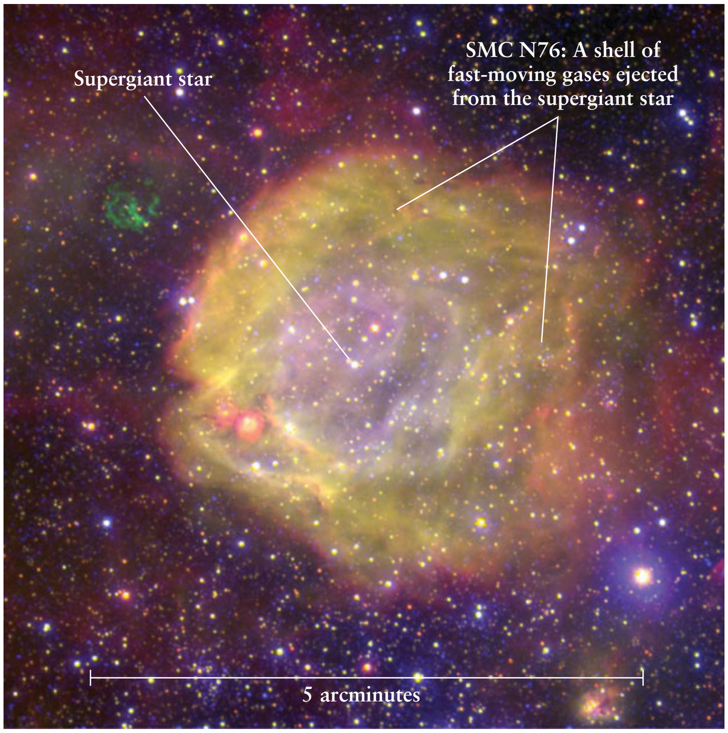
8. A white dwarf

The core is now a white dwarf star, and the former shells around the core become its thin atmosphere.



The carbon-oxygen interior of the white dwarf is degenerate, so it does not contract as it cools. Hence the white dwarf's radius no longer changes.

Hydrogen and helium, no fusion | Hydrogen fusion producing helium | Helium, no fusion | Helium fusion producing carbon and oxygen | Carbon and oxygen, no fusion



WEB LINK 20.7

Figure 20-12 R I V U X G

Mass Loss from a Supergiant Star At the heart of this nebulosity, called SMC N76, lies a supergiant star with a mass of at least 18 $M_\odot$. This star is losing mass at a rapid rate in a strong stellar wind. As this wind collides with the surrounding interstellar gas and dust, it creates the "bubble" shown here. SMC N76, which has an angular diameter of 130 arcsec, lies within the Small Magellanic Cloud, a small galaxy that orbits our Milky Way. It is about 60,600 pc (198,000 ly) distant. (Y. Nazé, G. Rauw, J. Manfroid, and J.-M. Vreux, Liège Institute; Y.-H. Chu, U. of Illinois; and ESO)

reactions can take place. After the cessation of carbon fusion, the core will again contract, and the star's central temperature can rise to 1 billion kelvins (10^9 K). At this temperature **neon fusion** begins. This uses up the neon accumulated from carbon fusion and further increases the concentrations of oxygen and magnesium in the star's core.

After neon fusion ends, the core will again contract, and **oxygen fusion** will begin when the central temperature of the star reaches about 1.5 billion kelvins (1.5×10^9 K). The principal product of oxygen fusion is silicon (^{28}Si, 14 protons). Once oxygen fusion is over, the core will contract yet again. If the central temperature reaches about 2.7 billion kelvins (2.7×10^9 K), **silicon fusion** begins, producing a variety of nuclei from sulfur (^{32}S, 16 protons) to iron (^{56}Fe, 26 protons) and nickel (^{56}Ni, 28 protons). While all of this is going on in the star's interior, at the surface the star is losing mass at a rapid rate (Figure 20-12).

As a high-mass star consumes increasingly heavier nuclei, the thermonuclear reactions produce a wider variety of products. For example, oxygen fusion produces not only silicon but also magnesium (^{24}Mg, with 12 protons), phosphorus (^{31}P, with 15 protons), and sulfur (^{31}S and ^{32}S, each with 16 protons). Some thermonuclear reactions that create heavy elements also release neutrons. A neutron is like a proton except that it carries no electric charge. Therefore, neutrons are not repelled by positively charged nuclei, and so can easily collide and combine with them. This absorption of neutrons by nuclei, called **neutron capture,** creates many elements and isotopes that are not produced directly in fusion reactions.

Each stage of thermonuclear reactions in a high-mass star helps to trigger the succeeding stage. In each stage, when the star exhausts a given variety of nuclear fuel in its core, gravitational contraction takes the core to ever-higher densities and temperatures, thereby igniting the "ash" of the previous fusion stage—and possibly the outlying shell of unburned fuel as well.

Supergiant Stars and Their Evolution

The increasing density and temperature of the core make each successive thermonuclear reaction more rapid than the one that preceded it. As an example, Table 20-1 shows a theoretical calculation of the evolutionary stages for a star with a zero-age mass of 25 $M_\odot$. This calculation indicates that carbon fusion in such a star lasts for 600 years, neon fusion for 1 year, and oxygen fusion for only 6 months. The last, and briefest, stage of nuclear reactions is silicon fusion. The entire core supply of silicon in a 25-$M_\odot$ star is used up in only one day!

Table 20-1 Evolutionary Stages of a 25-$M_\odot$ Star

Stage	Core temperature (K)	Core density (kg/m^3)	Duration of stage
Hydrogen fusion	4×10^7	5×10^3	7×10^6 years
Helium fusion	2×10^8	7×10^5	7×10^5 years
Carbon fusion	6×10^8	2×10^8	600 years
Neon fusion	1.2×10^9	4×10^9	1 year
Oxygen fusion	1.5×10^9	10^{10}	6 months
Silicon fusion	2.7×10^9	3×10^{10}	1 day
Core collapse	5.4×10^9	3×10^{12}	¼ second
Core bounce	2.3×10^{10}	4×10^{15}	milliseconds
Explosive (supernova)	about 10^9	varies	10 seconds

Based on calculations by Stanford Woosley (University of California, Santa Cruz) and Thomas Weaver (Lawrence Livermore National Laboratory).

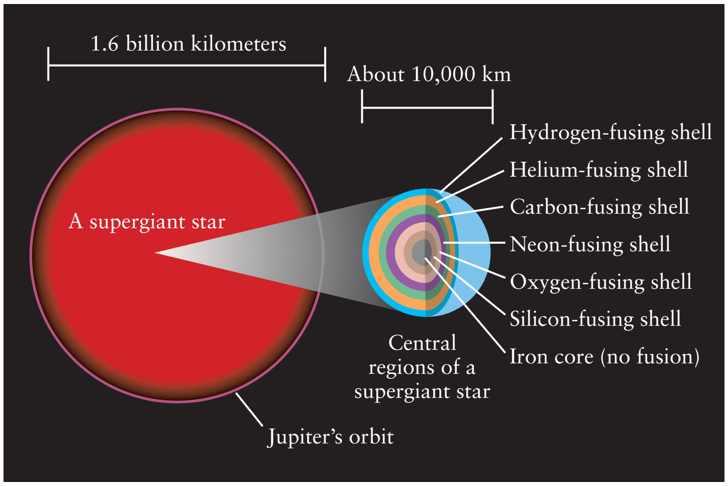


Figure 20-13

The Structure of an Old High-Mass Star Near the end of its life, a star with an initial mass greater than about 8 $M_\odot$ becomes a red supergiant. The star's overall size can be as large as Jupiter's orbit around the Sun. The stars energy comes from a series of concentric fusing shells, all combined within a volume roughly the same size as the Earth. Thermonuclear reactions do not occur within the iron core, because fusion reactions that involve iron absorb energy rather than release it.

Each stage of core fusion in a high-mass star generates a new shell of material around the core. After several such stages, the internal structure of a truly massive star—say, 25 to 30 $M_\odot$ or greater—resembles that of an onion (see Figure 20-13). Because thermonuclear reactions can take place simultaneously in several shells, energy is released at such a rapid rate that the star's outer layers expand tremendously. The result is a **supergiant** star, whose luminosity and radius are much larger than those of a giant (see Section 17-7).

A star of 8 or more solar masses evolves into a supergiant 100 times (or more) larger than the Sun

Several of the brightest stars in the sky are supergiants, including Betelgeuse and Rigel in the constellation Orion and Antares in the constellation Scorpius. (Figure 17-15 shows the locations of these stars on an H-R diagram.) They appear bright not because they are particularly close, but because they are extraordinarily luminous.

A supergiant star cannot keep adding shells to its "onion" structure forever, because the sequence of thermonuclear reactions cannot go on indefinitely. In order for an element to serve as a thermonuclear fuel, energy must be given off when its nuclei collide and fuse. This released energy is a result of the strong nuclear force of attraction that draws nucleons (neutrons and protons) together. However, protons also repel one another by the weaker electric force. As a result of this electric repulsion, adding extra protons to nuclei larger than iron, which has 26 protons, requires an *input* of energy rather than causing energy to be released. Nuclei of this size or larger cannot act as fuel for thermonuclear reactions. Hence, the sequence of fusion stages ends with silicon fusion. One of the products of silicon fusion is iron, and the result is a star with an iron-rich core in which no thermonuclear reactions take place (see Figure 20-13).

Shell fusion in the layers surrounding the iron-rich core consumes the star's remaining reserves of fuel. At this stage the entire energy-producing region of the star is contained in a volume no bigger than the Earth, some 10^6 times smaller in radius than the overall size of the star. This state of affairs will soon come to an end, because the buildup of an inert, iron-rich core signals the impending violent death of a massive supergiant star.

20-6 High-mass stars violently blow apart in core-collapse supernova explosions

Our present understanding is that all stars of about 8 $M_\odot$ or less divest most of their mass in the form of planetary nebulae. The burned-out core that remains settles down to become a white dwarf star. But the truly massive stars—stellar heavyweights that begin their lives with more than 8 solar masses of material—do not pass through a planetary nebula phase. Instead, they die in spectacular *core-collapse supernova explosions*.

The Violent End of a High-Mass Star

To understand what happens in a core-collapse supernova explosion, we must look deep inside a massive star at the end of its life. Of course, we cannot do this in actuality, because the interiors of stars are opaque. But astronomers have developed theoretical models based on what we know about the behavior of gases and atomic nuclei. The story that follows, while largely theoretical, describes our observations of supernovae fairly well. And, as we will see in Section 20-8, a special kind of "telescope" has allowed us to glimpse the interior of at least one relatively nearby supernova.

The core of an aging, massive star gets progressively hotter as it contracts to ignite successive stages of thermonuclear fusion (see Stage 1 in Figure 20-14). Wien's law (Section 5-4) and Planck's law (Section 5-5) together tell us that as the temperature of an object like a star increases, so does the energy of the photons it emits. When the temperature in the core of a massive star reaches a few hundred million kelvins, the photons are energetic enough to initiate a host of nuclear reactions that create neutrinos. These neutrinos, which carry off energy, escape from the star's core, just as solar neutrinos flow freely out of the Sun (see Section 16-4).

To compensate for the energy drained by the neutrinos, the star must provide energy either by consuming more thermonuclear fuel, by contracting, or both. But when the star's core is converted into iron, no more energy-producing thermonuclear reactions are possible, and the only source of energy is contraction and rapid heating (see Stage 2 in Figure 20-14).

By forming a dense core of iron, a massive star sows the seeds of its own destruction

Once a star with an original mass of about 8 $M_\odot$ or more develops an iron-rich core, the core contracts very rapidly, so that the core temperature skyrockets to 5×10^9 K within a tenth of a second. The gamma-ray photons emitted by the intensely hot core have so much energy that when they collide with iron nuclei, they begin to break the iron nuclei down into much smaller helium nuclei (^{4}He). This process is called

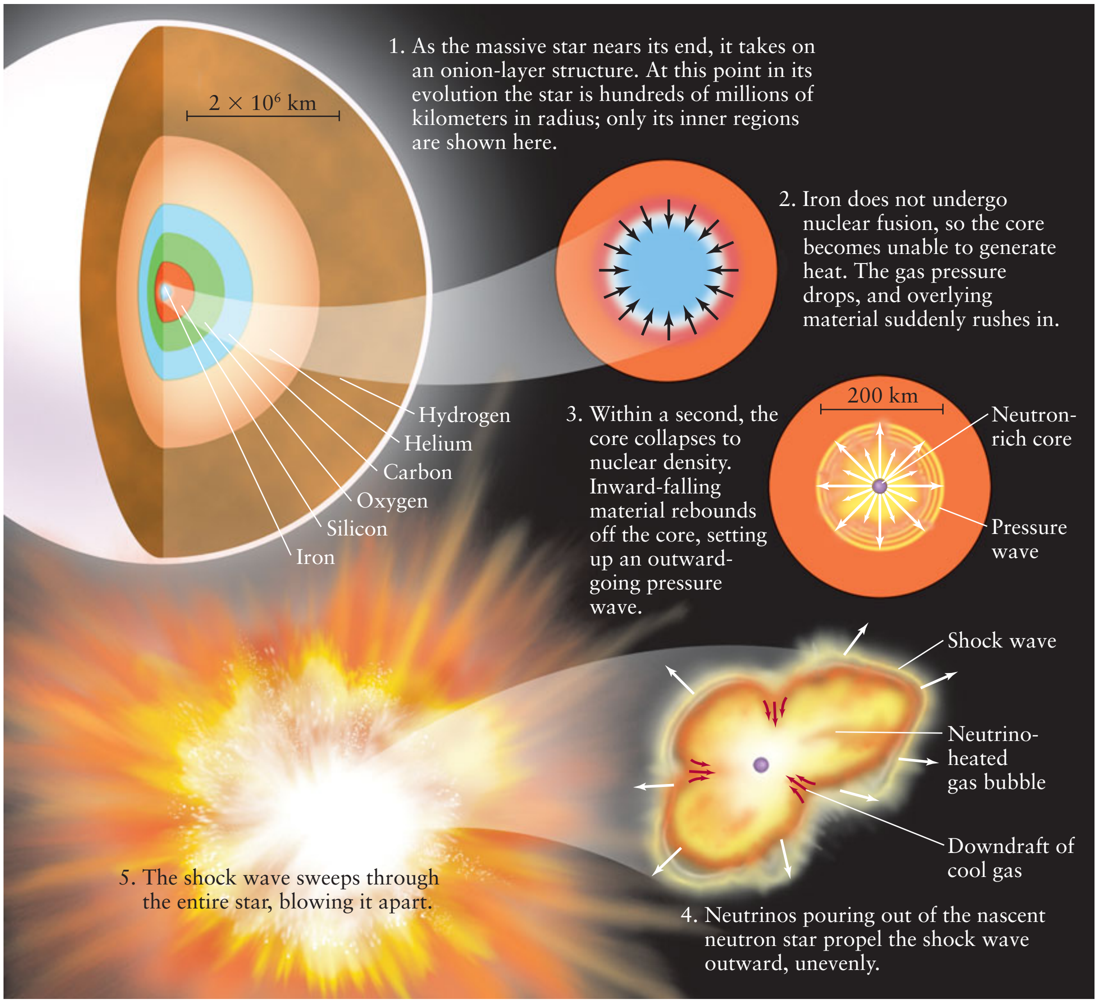


Figure 20-14

A Core-Collapse Supernova This series of illustrations depicts our understanding of the last day in the life of a star of more than about 8 $M_\odot$. (Illustration by Don Dixon, adapted from Wolfgang Hillebrandt, Hans-Thomas Janka, and Ewald Müller, "How to Blow Up a Star," *Scientific American*, October 2006)

photodisintegration. As Table 20-1 shows, it takes a high-mass star millions of years and several stages of thermonuclear reactions to build up an iron core; within a fraction of a second, photodisintegration undoes the result of those millions of years of reactions.

Within another tenth of a second, the core becomes so dense that the negatively charged electrons within it are forced to combine with the positively charged protons to produce electrically neutral neutrons. This process also releases a flood of neutrinos, denoted by the Greek letter ν (nu):

$$e^- + p \rightarrow n + \nu$$

Although neutrinos interact only very weakly with matter (see Section 16-4), the core is now so dense that even neutrinos cannot escape from it immediately. But because these neutrinos carry away a substantial amount of energy as they escape from the core, the core cools down and condenses even further.

At about 0.25 second after its rapid contraction begins, the core is less than 20 km in diameter and its density is in excess of 4×10^{17} kg/m^3. This is **nuclear density,** the density with which neutrons and protons are packed together inside nuclei. (If the Earth were compressed to this density, it would be only 300 meters, or 1000 feet, in diameter.)

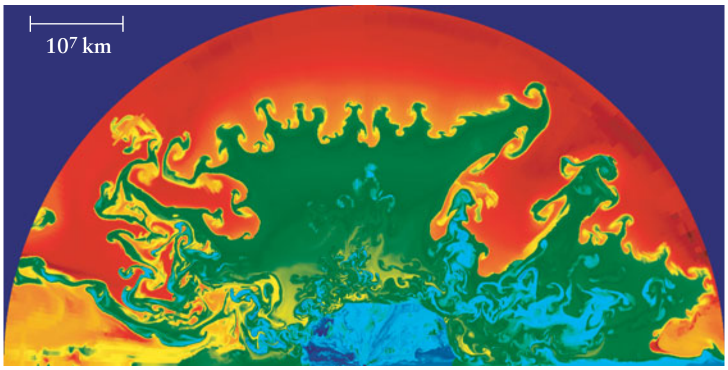
Matter at nuclear density or higher is extraordinarily difficult to compress. Thus, when the density of the neutron-rich core begins to exceed nuclear density, the core suddenly becomes very stiff and rigid. The core's contraction comes to a sudden halt, and the innermost part of the core actually bounces back and expands somewhat. This *core bounce* sends a powerful wave of pressure, like an unimaginably intense sound wave, outward into the outer core (see Stage 3 in Figure 20-14).

During this critical stage, the cooling of the core has caused the pressure to decrease profoundly in the regions surrounding the core. Without pressure to hold it up against gravity, the material from these regions plunges inward at speeds up to 15% of the speed of light. When this inward-moving material crashes down onto the rigid core, it encounters the outward-moving pressure wave. In just a fraction of a second, the material that fell onto the core begins to move back out toward the star's surface, propelled in part by the flood of neutrinos trying to escape from the star's core.

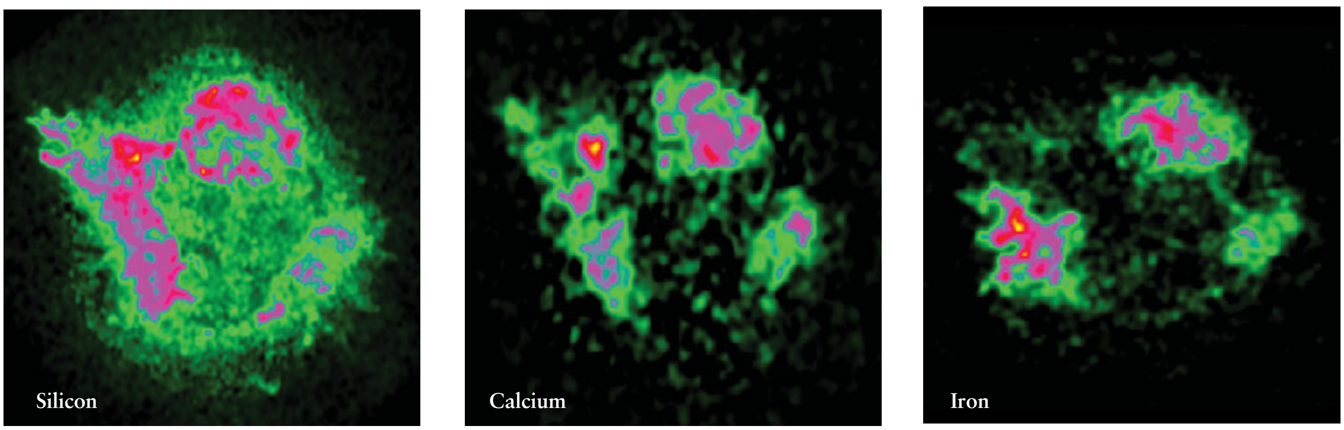
Supercomputer simulations of this complex process show that if the pressure wave were to spread outward at precisely the same speed in all directions, its energy would be absorbed by the gas around the core and the wave would fizzle out. But when the simulations allow for the presence of convection and turbulence in the dying star's gases, the result is quite different: The material surrounding the core behaves more like water boiling furiously in a heated pot. Rising bubbles of superheated gases deliver extra energy to the pressure wave, sustaining it and making it accelerate as it plows outward through the doomed star's outer layers. The wave soon reaches a speed greater than the speed of sound waves in the star's outer layers. When this happens, the wave becomes a *shock wave,* like the sonic boom produced by a supersonic airplane (see Stage 4 in Figure 20-14).

After a few hours, the shock wave reaches the star's surface, by which time the star's outer layers have begun to lift away from the core. When the star's outer layers thin out sufficiently, a portion of this energy escapes in a torrent of light (see Stage 5 in Figure 20-14). The star has become a **supernova** (plural **supernovae**). Specifically, what we have described is the formation of a **core-collapse supernova.** We use this term because, as we will see in Section 20-9, it is possible for supernova explosions to occur that do not involve the collapse of the core of a massive star.

CAUTION! The energy released in a core-collapse supernova is an incomprehensibly large 10^{46} joules—a hundred times more energy than the Sun has emitted due to thermonuclear reactions over its entire 4.56-billion year history. However, it is important to recognize that the source of the supernova's energy release is *not* thermonuclear reactions. Rather, it is the *gravitational* energy released by the collapse of the core and by the inward fall of the star's outer layers. (You release gravitational energy when you fall off a diving board, and this released energy goes into making a big splash in the swimming pool.) The energy released by the collapse of the core reappears in the form of neutrinos;



(a) A simulated supernova $5\frac{1}{2}$ hours after the core "bounce" (red = hydrogen, green = helium, turquoise and blue = carbon, oxygen, silicon, and iron)



(b) Material was ejected in "blobs" from the supernova that produced the Cassiopeia A supernova remnant R I V U X G

ANIMATION 20.3 WEB LINK 20.8

Figure 20-15

Turbulence in a Core-Collapse Supernova

(a) This image from a supercomputer simulation show a cross section of a massive star several hours into the supernova explosion. The colors show the turbulent mixing of material from the star's inner regions (turquoise and blue) with hydrogen and helium from the outer layers (green and red). **(b)** Turbulence causes material to be ejected from the supernova in irregular "blobs," as shown by these images of the supernova remnant Cassiopeia A. Each image was made using an X-ray wavelength emitted by a particular element. (Figure 18-24 shows a false-color image of Cassiopeia A made using visible, infrared, and X-ray wavelengths.) (a: Konstantinos Kifonidis, Max-Planck-Institut für Astrophysik; b: U. Hwang et al., NASA/GSFC) the fall of the outer layers provides the energy to power the nuclear reactions that generate the supernova's electromagnetic radiation. The amount of energy release from the supernova is so great because the star is so massive; hence, the amount of material that falls inward is immense, it falls a great distance, and is acted on by a strong gravitational pull as it falls.

The Wreckage of a Core-Collapse Supernova

Supercomputer simulations provide many insights into the violent, complex, and rapidly changing conditions deep inside a star as it is torn apart by a supernova explosion (Figure 20-15). For example, Figure 20-15*a* shows a snapshot of the interior of a high-mass star $5\frac{1}{2}$ hours after the stiffening of the core. The simulation predicts turbulent swirls and eddies that grow behind the shock wave as it moves outward from the star's core. Evidence in favor of such turbulence comes from images of the remnants of long-ago supernovae (Figure 20-15*b*). Such images show that material is ejected from the supernova not in uniform shells but in irregular clumps. This is just what would be expected from a turbulent explosion. (We discuss supernova remnants further in Section 20-10.)

Detailed computer calculations suggest that a 25-$M_\odot$ star ejects about 96% of its material to the interstellar medium for use in producing future generations of stars. Less massive stars eject a smaller percentage of their mass into space when they become core-collapse supernovae.

Before this material is ejected into space, it is compressed so much by the passage of the shock wave through the star's outer layers that a new wave of thermonuclear reactions sets in. These reactions can produce many more chemical elements, including elements heavier than iron. Reactions of this kind require a tremendous input of energy, and thus cannot take place during the star's pre-supernova lifetime.

The energy-rich environment of a supernova shock wave is almost the only place in the universe where such heavy elements as zinc, silver, tin, gold, mercury, lead, and uranium can be produced. (In Chapter 21 we will see another, even more exotic mechanism for producing these heaviest elements.) Remarkably, all of these elements are found on the Earth. Hence, some of the material that makes up our solar system, our Earth, and our bodies must long ago have been part of a star that lived, evolved, and died as a supernova.

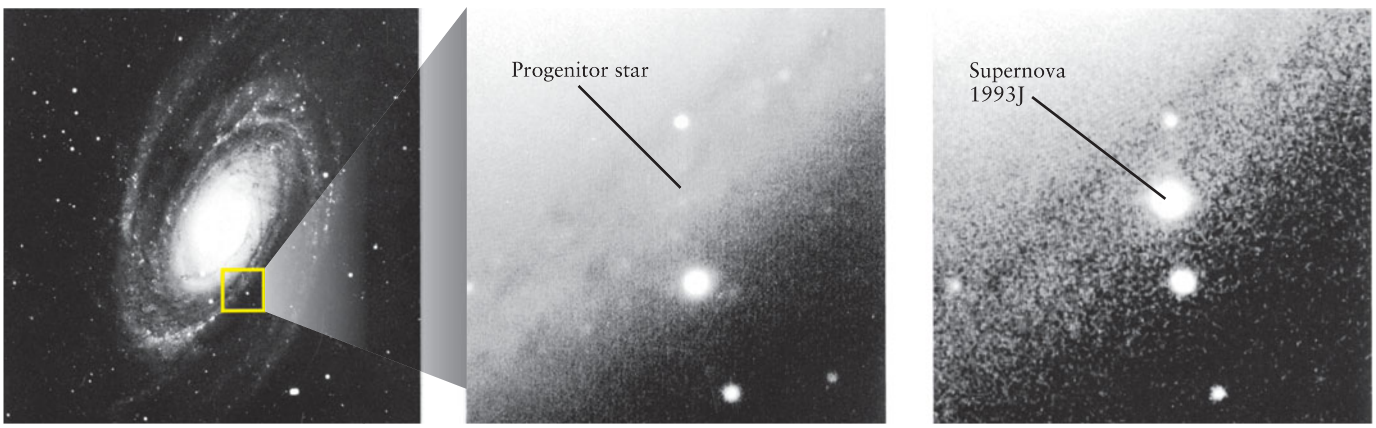
20-7 In 1987 a nearby supernova gave us a close-up look at the death of a massive star

Supernovae have peak luminosities as great as 10^9 $L_\odot$, rivaling the light output of an entire galaxy. This makes it possible to see supernovae in galaxies far beyond our own Milky Way Galaxy, and indeed hundreds of these distant supernovae are observed each year (Figure 20-16). In a handful of cases, images made before the explosion have allowed astronomers to identify the star that subsequently exploded into a supernova, called the **progenitor star.** For example, the progenitor star shown in Figure 20-16*b* was a red supergiant star whose internal structure probably resembled that shown in cross section in Figure 20-13.

One frustrating aspect of these distant supernovae is simply that they *are* distant, and so cannot be studied in as much detail as astronomers would like. But one recent and unusually close supernova has provided astronomers with a unique opportunity to check the theoretical ideas presented in Section 20-6.

A Supernova in the Galaxy Next Door

On February 23, 1987, a supernova was discovered in the Large Magellanic Cloud (LMC), a companion galaxy to our Milky Way some 51,500 pc (168,000 ly) from Earth. The supernova, designated SN 1987A because it was the first discovered that year, occurred near an enormous H II region in the LMC called the



(a) Spiral galaxy M81 (b) Before the explosion (c) After the explosion

WEB LINK 20.9

Figure 20-16 R I V U X G

A Supernova in a Distant Galaxy On the night of March 28, 1993, Francisco Garcia Diaz, a Spanish amateur astronomer, discovered supernova SN 1993J in the galaxy M81 in Ursa Major. **(a)** M81 lies some 3.6 million pc (12 million ly) from Earth. Its angular size is about half that of the full moon. **(b)** The progenitor star that later exploded into SN 1993J was a K0 red supergiant. **(c)** This image shows the same part of the sky as (b). Like SN 1987A, SN 1993J resulted from the core collapse and subsequent explosion of a massive star. (a: Palomar Observatory; b, c: D. Jones and E. Telles, Isaac Newton Telescope)

Tarantula Nebula (Figure 20-17). The supernova was so bright that observers in the southern hemisphere could see it without a telescope.

Supernova 1987A was the first nearby supernova to be seen since the invention of the telescope

Such a bright supernova is a rare event. In the last thousand years, only five other supernovae—in 1006, 1054, 1181, 1572, and 1604—have been bright enough to be seen with the naked eye, and all of these occurred in our own Milky Way Galaxy. (As we described in Section 4-3, the supernova of 1572 had a major influence on Tycho Brahe's ideas about the heavens.) Although outside our Galaxy, SN 1987A occurred relatively close to us in a part of the heavens that is obscured only slightly by the Milky Way's interstellar dust. Furthermore, not long after SN 1987A appeared, several new orbiting telescopes were placed into service. These enabled astronomers to study the supernova's evolution with unprecedented resolution and in wavelength ranges not accessible from the Earth's surface (see Section 6-7). As a result, SN 1987A has given astronomers an unprecedented view of the violent death of a massive star.

The light from a supernova such as SN 1987A does not all come in a single brief flash; the outer layers continue to glow as they expand into space. For the first 20 days after the detonation of SN 1987A, its glow was powered primarily by the tremendous heat that the shock wave deposited in the star's outer layers. As the expanding gases cooled, the light energy began to be provided by a different source—the decay of radioactive isotopes of cobalt, nickel, and titanium produced in the supernova explosion.

Astronomers have been able to pinpoint the specific isotopes involved because different radioactive nuclei emit gamma rays of different wavelengths when they decay. These emissions have been detected by orbiting gamma-ray telescopes (see Figure 6-31). Thanks to these radioactive decays, the brightness of SN 1987A actually *increased* for the first 85 days after the detonation, then settled into a slow decline as the radioactive isotopes were used up. The supernova remained visible to the naked eye for several months after the detonation.

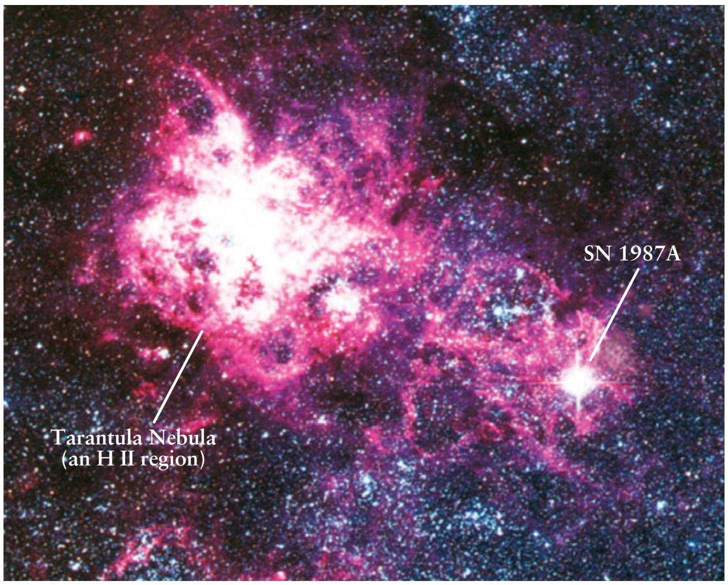


Figure 20-17 R I V U X G

Supernova 1987A This photograph, taken soon after the discovery of SN 1987A, shows a portion of the Large Magellanic Cloud that includes the supernova and a huge H II region called the Tarantula Nebula. Although it was 51,000 pc from Earth, SN 1987A was bright enough to be seen without a telescope. (European Southern Observatory)

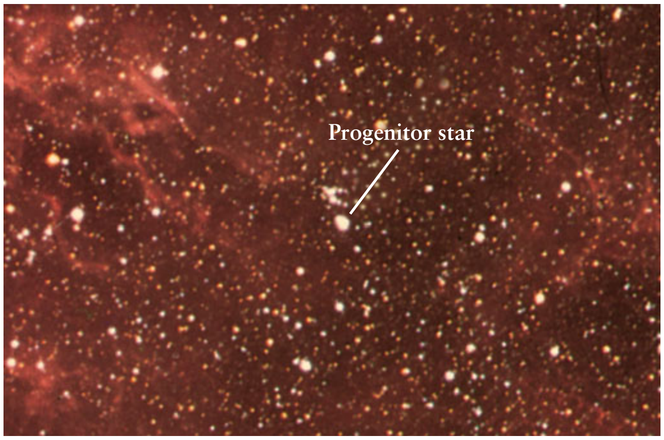
Why SN 1987A Was Unusual

Ideally, SN 1987A would have confirmed the theories of astronomers about typical supernovae. But SN 1987A was *not* typical. Its luminosity peaked at roughly 10^8 $L_\odot$, only a tenth of the maximum luminosity observed for other, more distant supernovae. Fortunately, the doomed star had been observed prior to becoming a supernova, and these observations helped explain why SN 1987A was an exceptional case.

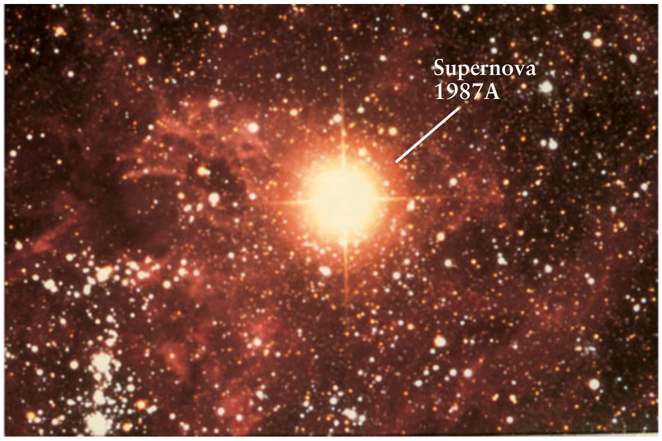
Figure 20-13 suggests that a massive star should be in a red supergiant stage when the iron core collapses and the star becomes a supernova. As we have mentioned, this was indeed the case for the progenitor star shown in Figure 20-16*b*, and is also the case for other supernovae seen in distant galaxies. The progenitor star of SN 1987A was indeed a high-mass star: Its estimated main sequence mass was about 20 $M_\odot$, although by the time it exploded—some 10^7 years after it first formed—it probably had shed a few solar masses. However, this progenitor star was identified not as a red supergiant, but as a *blue* B3 I supergiant (Figure 20-18).

The explanation of this seeming contradiction is that stars in the Large Magellanic Cloud, including the progenitor of SN 1987A, are Population II stars with a very low percentage of metals—that is, elements heavier than hydrogen and helium (see Section 19-5). A small difference in the amount of metals present can affect whether a star's interior is relatively transparent or opaque, just as a small amount of dirt can make a window difficult to see through. As a result, a high-mass Population II star follows a somewhat different evolutionary track than does a Population I star (with a greater percentage of metals) of the same mass. On an H-R diagram, an aging high-mass star of Population I goes directly from the main sequence to the red supergiant region at the upper right of the diagram. By contrast, the track for an equally massive Population II star wanders from left to right and back again across the top of the H-R diagram as the star alternates between being a hot, blue supergiant (on the left in an H-R diagram) and a cool red supergiant (on the right in an H-R diagram). Apparently the progenitor star of SN 1987A developed an iron core and became a supernova when it was in the blue supergiant stage.

A massive Population II star like the SN 1987A progenitor changes dramatically in size as it alternates between being a red supergiant and a blue supergiant. When in a blue supergiant phase, its radius may be less than 1/10 as large as when it is in a red supergiant phase. Hence, the progenitor star was relatively small when its core collapsed (though its radius was still more than 10 times that of our Sun). This means that the star's outer layers were close to the core and thus held more strongly by the core's gravitational attraction. When the detonation occurred, a relatively large fraction of the shock wave's energy had to be used against this gravitational attraction to push the outer layers into



(a) Before the star exploded



(b) After the star exploded

WEB LINK 20.10

Figure 20-18 R I V U X G

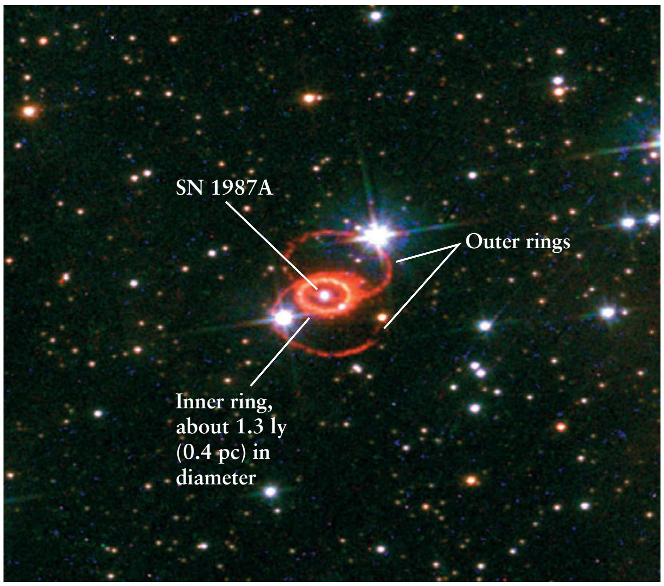
SN 1987A—Before and After **(a)** This photograph shows a small section of the Large Magellanic Cloud as it appeared before the explosion of SN 1987A. The supernova's progenitor star was a B3 blue supergiant. **(b)** This image shows a somewhat larger region of the sky a few days after the supernova exploded into brightness. (Anglo-Australian Observatory)

space. Hence, the amount of shock wave energy available to be converted into light was smaller than for most supernovae. This explains why SN 1987A was only a tenth as bright as an exploding red supergiant would have been.

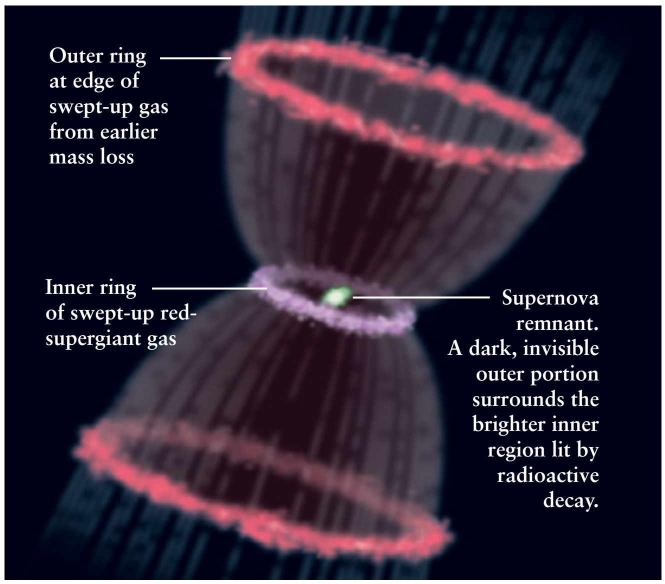
The Aftermath of SN 1987A

Three and a half years after SN 1987A exploded, astronomers used the newly launched Hubble Space Telescope to obtain a picture of the supernova. To their surprise, the image showed a ring of glowing gas around the exploded star. After the optics of the Hubble Space Telescope were repaired in 1994 (see Section 6-7), SN 1987A was observed again and a set of *three* glowing rings was revealed (Figure 20-19*a*).

These rings are relics of a hydrogen-rich outer atmosphere that was ejected by gentle stellar winds from the progenitor star when it was a red supergiant, about 20,000 years ago. This



(a) Supernova 1987A seen in 1996 R I V U X G



(b) An explanation of the rings

ANIMATION 20.4

Figure 20-19

SN 1987A and Its "Three-Ring Circus" **(a)** This true-color view from the Hubble Space Telescope shows three bright rings around SN 1987A. **(b)** This drawing shows the probable origin of the rings. A wind from the progenitor star, shown in Figure 20-18*a*, formed an hourglass-shaped shell surrounding the star. (Compare with Figure 20-7.) Ultraviolet light from the supernova explosion ionized ring-shaped regions in the shell, causing them to glow. (Robert Kirshner and Peter Challis, Harvard-Smithsonian Center for Astrophysics; STScI) diffuse gas expanded in a hourglass shape (Figure 20-19*b*), because it was blocked from expanding around the star's equator either by a preexisting ring of gas or by the orbit of an as yet unseen companion star. (Figure 20-7 shows a similar model used to explain the shapes of certain planetary nebulae.) The outer rings in Figure 20-19*a* are parts of the hourglass that were ionized by the initial flash of ultraviolet radiation from the supernova; as electrons recombine with the ions, the rings emit visible-light photons. (We described this process of *recombination* in Section 18-2.)

WEB LINK 20.11 ANIMATION 20.5

By the early years of the twenty-first century, the shock wave from the supernova was beginning to collide with the "waist" of the hourglass shown in Figure 20-19*b*. This collision is making the hourglass glow more brightly in visible wavelengths—though not enough, unfortunately, to make the supernova again visible to the naked eye—and emit copious radiation at X-ray and ultraviolet wavelengths.

By studying this cosmic collision, astronomers hope to learn more about the shock wave and the supernova explosion that spawned it. They hope as well to learn about the matter in the inner ring, which will give us insight into the stellar winds that blew from the progenitor star thousands of years ago. Because SN 1987A provides a unique laboratory for studying the evolution of a supernova, astronomers will monitor it carefully for decades to come.

20-8 Neutrinos emanate from supernovae like SN 1987A

In addition to electromagnetic radiation, supernovae also emit a brief but intense burst of neutrinos from their collapsing cores. In fact, theory suggests that *most* of the energy released by the exploding star is in the form of neutrinos. If it were possible to detect the flood of neutrinos from an exploding star, astronomers would have direct evidence of the nuclear processes that occur within the star during its final seconds before becoming a supernova.

Observations of SN 1987A confirm that a core-collapse supernova emits most of its energy in the form of neutrinos

Unfortunately, detecting neutrinos is difficult because under most conditions matter is transparent to neutrinos (see Section 16-4). Consequently, when supernova neutrinos encounter the Earth, almost all of them pass completely through the planet as if it were not there. The challenge to scientists is to detect the tiny fraction of neutrinos that *do* interact with the matter through which they pass.

Neutrino Telescopes

During the 1980s, two "neutrino telescopes" uniquely suited for detecting supernova neutrinos went into operation—the Kamiokande detector in Japan (a joint project of the University of Tokyo and the University of Pennsylvania), and the IMB detector (a collaboration of the University of California, Irvine, the University of Michigan, and Brookhaven National Laboratory). Both detectors consisted of large tanks containing thousands of tons of water. On the rare occasion when a neutrino collided with one of the water molecules, it produced a brief flash of light. Any light flash was recorded by photomultiplier tubes lining the walls of the tank. Because other types of subatomic particles besides neutrinos could also produce similar light flashes, the detectors were placed deep underground so that hundreds of meters of earth would screen out almost all particles except neutrinos. (The more recent Sudbury Neutrino Observatory, shown in Figure 16-6, has a similar design.)

What causes the light flashes? And how do we know whether a neutrino comes from a supernova? The key is that a single supernova neutrino carries a relatively large amount of energy, typically 20 MeV or more. (We introduced the unit of energy called the electron volt, or eV, in Section 5-5. One MeV is equal to 10^6 electron volts. Only the most energetic nuclear reactions produce particles with energies of more than 1 MeV.) If such a high-energy neutrino hits a proton in the water-filled tank of a neutrino telescope, the collision produces a positron (see Box 16-1). The positron then recoils at a speed greater than the speed of light in water, which is 2.3×10^5 km/s. (Such motion does not violate the ultimate speed limit in the universe, 3×10^5 km/s, which is the speed of light in a *vacuum*.)

Just as an airplane that flies faster than sound produces a shock wave (a sonic boom), a positron that moves through a substance such as water faster than the speed of light in that substance produces a shock wave of light. This shock wave is called **Cerenkov radiation,** after the Russian physicist Pavel A. Cerenkov, who first observed it in 1934. It is this radiation that is detected by the photomultiplier tubes that line the detector walls.

By measuring the properties of the Cerenkov radiation from a recoiling positron, scientists can determine the positron's energy and, therefore, the energy of the neutrino that created the positron. This allows them to tell the difference between the high-energy neutrinos from supernovae and neutrinos from the Sun, which typically have energies of 1 MeV or less.

Neutrinos from SN 1987A

Both the Kamiokande and IMB detectors were operational on February 23, 1987, when SN 1987A was first observed in the Large Magellanic Cloud. Soon afterward, the physicists working with these detectors excitedly reported that they had detected Cerenkov flashes from a 12-second burst of neutrinos that reached the Earth 3 hours before astronomers saw the light from the exploding star. Only a few neutrinos were seen: The Kamiokande detector saw flashes from 12 neutrinos at about the same time that 8 were recorded by the IMB detector. But when the physicists factored in the sensitivity of their detectors, they calculated that Kamiokande and IMB had actually been exposed to a torrent of more than 10^{16} neutrinos.

Given the flux of neutrinos measured by the detectors and the distance of 168,000 light-years from SN 1987A to the Earth, physicists used the inverse-square law to determine the total number of neutrinos that had been emitted from the supernova. (This law applies to neutrinos just as it does to electromagnetic radiation; see Section 17-2.) They found that over a 10-second period, SN 1987A emitted 10^{58} neutrinos with a total energy of 10^{46} joules. This is more than 100 times as much energy as the Sun has emitted in its entire history and more than 100 times the amount of energy that the supernova emitted in the form of electromagnetic radiation. Indeed, for a few seconds the supernova's neutrino luminosity—that is, the *rate* at which it emitted energy in the form of neutrinos—was 10 times greater than the total luminosity in electromagnetic radiation of all of the stars in the observable universe! Such comparisons give a hint of the incomprehensible violence with which a supernova explodes.

Why did the neutrinos from SN 1987A arrive 3 hours *before* the first light was seen? As we saw in Section 20-6, neutrinos are produced when thermonuclear reactions cease in the core of a massive star and the core collapses. These neutrinos encounter little delay as they pass through the volume of the star. The tremendous increase in the star's light output, by contrast, occurs only when the shock wave reaches the star's outermost layers (which are thin enough to allow light to pass through them). It took 3 hours for this shock wave to travel outward from the star's core to its surface, by which time the neutrino burst was already billions of kilometers beyond the dying star. For the next 168,000 years, the neutrinos that would eventually produce light flashes in Kamiokande and IMB remained in front of the photons emitted from the star's surface, and so the neutrinos were detected before the supernova's light. Thus, the neutrino data from SN 1987A gave astronomers direct confirmation of theoretical ideas about how supernova explosions take place.

Kamiokande and IMB have both been replaced by a new generation of neutrino telescopes (see Section 16-4). While these new detectors are intended primarily to observe neutrinos from the Sun, they are also fully capable of measuring neutrino bursts from nearby supernovae. Astronomers have identified a number of supergiant stars in our Galaxy that are likely to explode into supernovae, among them the bright red supergiant Betelgeuse in the constellation Orion (see Figure 2-2 and Figure 6-27).

Unfortunately, astronomers do not yet know how to predict precisely when such stars will explode into supernovae. It may be many thousands of years before Betelgeuse explodes. Then again, it could happen tomorrow. If it does, neutrino telescopes will be ready to record the collapse of its massive core.

20-9 White dwarfs in close binary systems can also become supernovae

Astronomers discover dozens of supernovae in distant galaxies every year, but not all of these are the result of massive stars dying violently. A totally different type of supernova occurs when a white dwarf star in a binary system blows itself completely apart. The first clue that two entirely distinct chains of events could produce supernovae was rather subtle: Some supernovae have prominent hydrogen emission lines in their spectra but others do not.

In just a few seconds, a thermonuclear supernova completely destroys an entire white dwarf star

Types of Supernovae

Supernovae with hydrogen emission lines, called **Type II supernovae,** are core-collapse supernovae of the sort we described in Section 20-6. They are caused by the death of highly evolved massive stars that still have ample hydrogen in their atmospheres when they explode. When the star explodes, the hydrogen atoms are excited and glow prominently, producing hydrogen emission lines. SN 1987A (the topic of Sections 20-7 and 20-8) and SN 1993J (shown in Figure 20-16*c*) were both Type II supernovae.

Hydrogen lines are missing in the spectrum of a **Type I supernova,** which tells us that little or no hydrogen is left in the debris from the explosion. Type I supernovae are further divided into three important subclasses. **Type Ia supernovae** have spectra that include a strong absorption line of ionized silicon. **Type Ib** and **Type Ic supernovae** both lack the ionized silicon line. The difference between them is that the spectra of Type Ib supernovae have a strong helium absorption line, while those of Type Ic supernovae do not. Figure 20-20 shows these different supernova spectra.

Astronomers suspect that Type Ib and Ic supernovae are caused by core collapse in dying massive stars, just like Type II supernovae. The difference is that the progenitor stars of Type Ib and Ic supernovae have been stripped of their outer layers before they explode. A star can lose its outer layers to a strong stellar wind (see Figure 20-12) or, if it is part of a close binary system, by transferring mass to its companion star (see Figure 19-21*b*). If enough mass remains for the star's core to collapse, the star dies as a Type Ib supernova. Because the outer layers of hydrogen are absent, the supernova's spectrum exhibits no hydrogen lines but many helium lines (Figure 20-20*b*). Type Ic supernovae have apparently undergone even more mass loss prior to their explosion; their spectra show that they have lost much of their helium as well as their hydrogen (Figure 20-20*c*).

A key piece of evidence that Type II, Type Ib, and Type Ic supernovae all begin as massive stars is that all three types are found only near sites of recent star formation. The life span of a massive star from its formation to its explosive death as a supernova is only about 10^7 years, less time than it took our own Sun to condense from a protostar to a main-sequence star (see Figure 18-10) and a mere blink of an eye on the time scale of stellar evolution. Because massive stars live such a short time, it makes sense that they should meet their demise very close to where they were formed.

Type Ia Supernovae: Detonating a White Dwarf

Type Ia supernovae, by contrast, are found even in galaxies where there is no ongoing star formation. Hence, they are probably *not* the death throes of massive supergiant stars. Instead, Type Ia supernovae are thought to result from the thermonuclear explosion of a white dwarf star. This may sound contradictory, because we saw in Section 20-4 that white dwarf stars have no thermonuclear reactions going on in their interiors. But these reactions *can* occur if a carbon-oxygen–rich white dwarf is in a close, semidetached binary system with a red giant star (see Figure 19-20*b* and Figure 20-20*a*).

ANIMATION 20.6 ANIMATION 20.7

Figure 20-21 shows the likely series of events that lead to a Type Ia supernova. Stage 1 in this figure shows a close binary system in which both stars have less than 4 $M_\odot$. The more massive star on the left evolves more rapidly than its less massive companion and eventually becomes a white dwarf. As the companion evolves and its outer layers expand, it overflows its Roche lobe

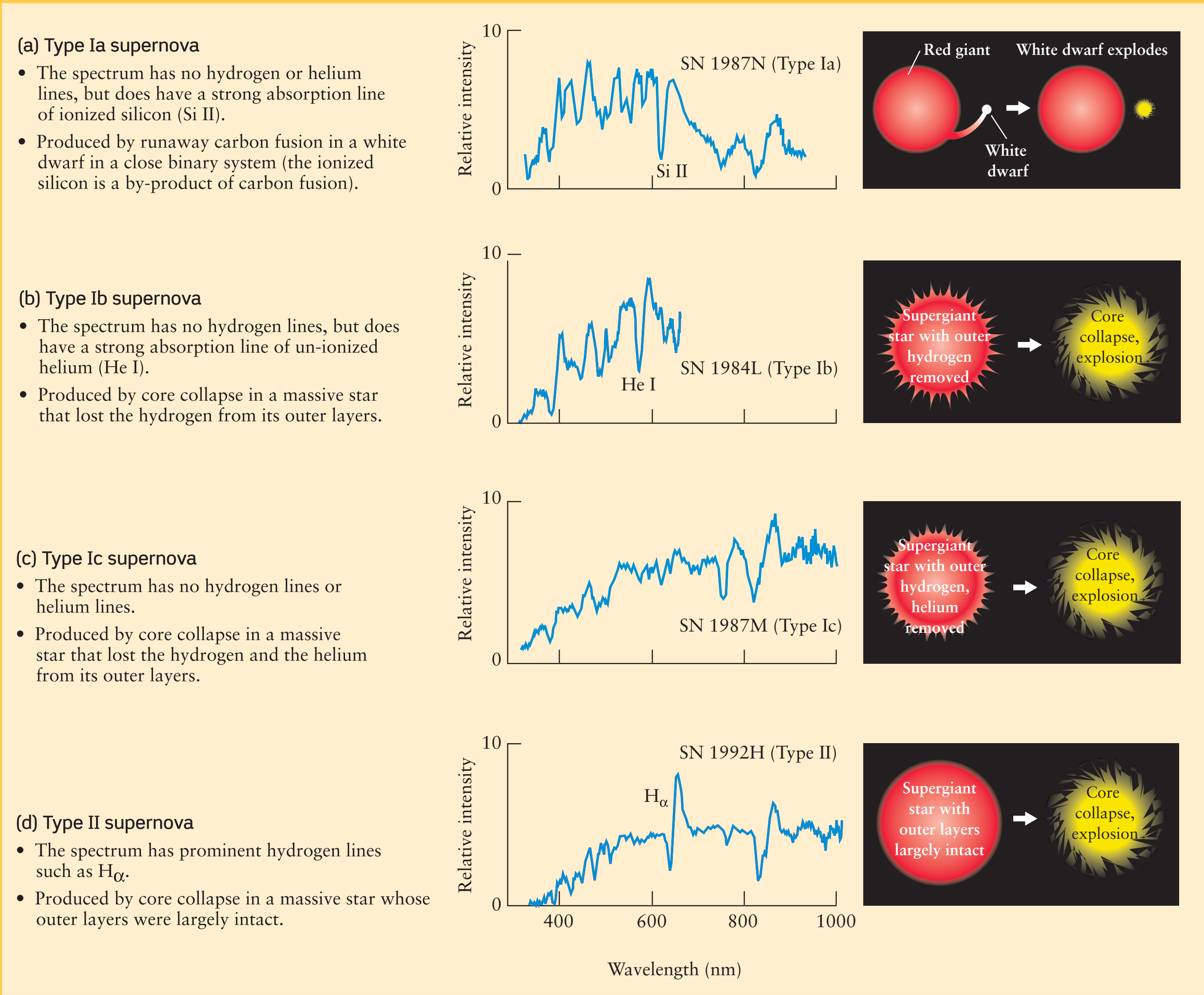


Figure 20-20

Supernova Types These illustrations show the characteristic spectra and the probable origins of supernovae of **(a)** Type Ia, **(b)** Type Ib, **(c)** Type Ic, and **(d)** Type II. (Spectra courtesy of Alexei V. Filippenko, University of California, Berkeley)

and dumps gas from its outer layers onto the white dwarf (see Stage 2 in Figure 20-21). When the total mass of the white dwarf approaches the Chandrasekhar limit, the increased pressure applied to the white dwarf's interior causes carbon fusion to begin there (Stage 3 in Figure 20-21). Hence, the interior temperature of the white dwarf increases.

If the white dwarf were made of ordinary matter, the temperature increase would cause a further increase in pressure, the white dwarf would expand and cool, and the carbon-fusing reactions would abate. But because the white dwarf is composed of degenerate matter, this "safety valve" between temperature and pressure does not operate. Instead, the increased temperature just makes the reactions proceed at an ever-increasing rate, in a catastrophic runaway process reminiscent of the helium flash in low-mass stars (Stage 4 in Figure 20-21). The reaction spreads rapidly outward from the white dwarf's center, with its leading edge (called the *flame front*) being propelled by convection and turbulence in a manner analogous to what happens to the shock wave in a core-collapse supernova. Within seconds the white dwarf blows apart, dispersing 100% of its mass into space (Stage 5 in Figure 20-21).

Before exploding, the white dwarf contained primarily carbon and oxygen and almost no hydrogen or helium, which explains the absence of hydrogen and helium lines in the spectrum

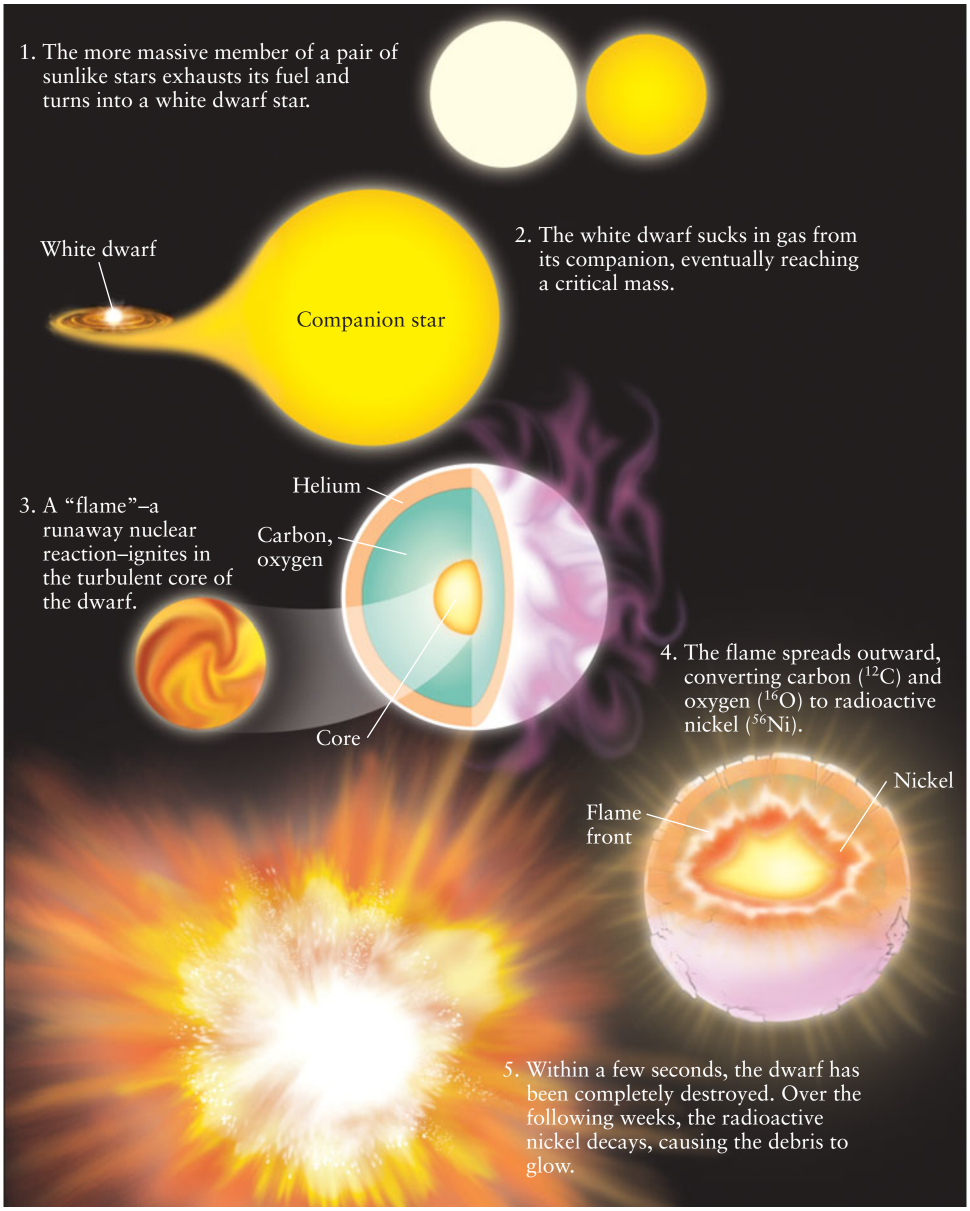


Figure 20-21

A Type Ia Supernova This series of illustrations depicts our understanding of how a white dwarf in a close binary system can undergo a sudden nuclear detonation that destroys it completely. Such a cataclysmic event is called a Type Ia supernova or thermonuclear supernova. (Illustration by Don Dixon, adapted from Wolfgang Hillebrandt, Hans-Thomas Janka, and Ewald Müller, "How to Blow Up a Star," *Scientific American*, October 2006)

of the resulting supernova. Silicon is a by-product of the carbon-fusing reaction and gives rise to the silicon absorption line characteristic of Type Ia supernovae.

CAUTION! Different types of supernovae have fundamentally different energy sources. We saw in Section 20-6 that core-collapse supernovae (which we have now sorted into Types II, Ib, and Ic) are powered by *gravitational* energy released as the star's iron-rich core and outer layers fall inward. Type Ia supernovae, by contrast, are powered by *nuclear* energy released in the explosive thermonuclear fusion of a white dwarf star. For this reason we also use the term **thermonuclear supernova** to refer to a Type Ia supernova. While Type Ia supernovae typically emit more energy in the form of electromagnetic radiation than supernovae of other types, they do not emit copious numbers of neutrinos because there is no core collapse. If we include the energy emitted in the form of neutrinos, the most luminous supernovae by far are those of Type II.

The Decay of a Supernova: Light Curves

In addition to the differences in their spectra, different types of supernovae can be distinguished by their light curves (Figure 20-22). All supernovae begin with a sudden rise in brightness that occurs in less than a day. After reaching peak luminosity, Type Ia, Ib, and Ic supernovae settle into a steady, gradual decline in luminosity. (An example is the supernova of 1006, which is thought to have been of Type Ia. This supernova, which at its peak was more than 200 times brighter than any other star in the

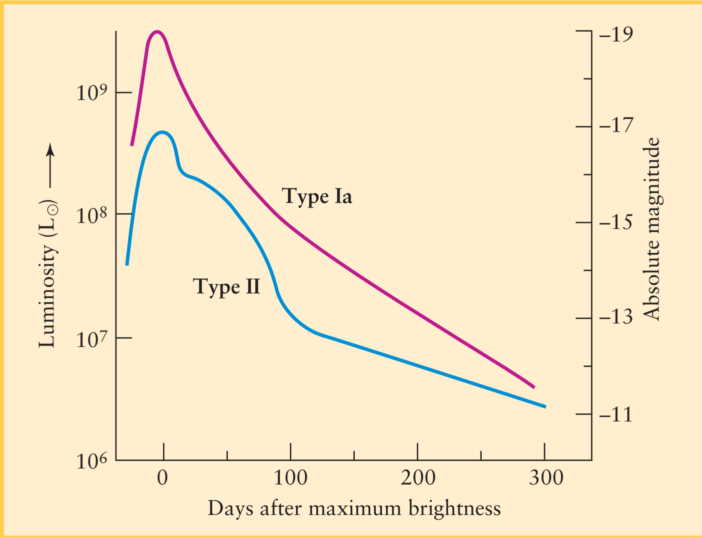


Figure 20-22

Supernova Light Curves A Type Ia supernova reaches maximum brightness in about a day, followed by a gradual decline in brightness. A Type II supernova reaches a maximum brightness only about one-fourth that of a Type Ia supernova and usually has alternating intervals of steep and gradual declines.

sky, took three years to fade into invisibility.) By contrast, the Type II light curve has a steplike appearance caused by alternating periods of steep and gradual declines in brightness.

For all supernova types, the energy source during the period of declining brightness is the decay of radioactive isotopes produced during the supernova explosion. Because a different set of thermonuclear reactions occurs for each type of supernova, each type produces a unique set of isotopes that decay at different rates. This helps explains the distinctive light curves for different supernova types.

For the same reason, each type of supernova ejects a somewhat different mix of elements into the interstellar medium. As an example, Type Ia supernovae are primarily responsible for the elements near iron in the periodic table, because they generate these elements in more copious quantities than Type II supernovae.

A number of astronomers are now measuring the distances to remote galaxies by looking for Type Ia supernovae in those galaxies. This is possible because there is a simple relationship between the rate at which a Type Ia supernova fades away and its peak luminosity: The slower it fades, the greater its luminosity. Hence, by observing how rapidly a distant Type Ia supernova fades, astronomers can determine its peak luminosity. A measurement of the supernova's peak apparent brightness then tells us (through the inverse-square law) the distance to the supernova, and, therefore, the distance to the supernova's host galaxy. The tremendous luminosity of Type Ia supernovae allows this method to be used for galaxies more than 10^9 light-years distant. In Chapter 26 we will learn what such studies tell us about the size and evolution of the universe as a whole.

20-10 A supernova remnant can be detected at many wavelengths for centuries after the explosion

WEB LINK 20.12

Astronomers find the debris of supernova explosions, called **supernova remnants,** scattered across the sky. A beautiful example of a supernova remnant is the Veil Nebula, shown in Figure 20-23. The doomed star's outer layers were blasted into space with such violence that they are still traveling through the interstellar medium at supersonic speeds 15,000 years later. As this expanding shell of gas plows through space, it collides with atoms in the interstellar medium, exciting the gas and making it glow. We saw in Section 18-8 that the passage of a supernova remnant through the interstellar medium can trigger the formation of new stars, so the death of a single massive star (in a core-collapse supernova) or white dwarf (in a thermonuclear supernova) can cause a host of new stars to be born.

Many supernovae in our Galaxy are hidden from us by the obscuring interstellar medium

A few nearby supernova remnants cover sizable areas of the sky. The largest is the Gum Nebula, named after the astronomer Colin Gum, who first noticed its faint glowing wisps on photographs of the southern sky (Figure 20-24). Its 40° angular diameter is centered on the constellation Vela (the Ship's Sail).

Figure 20-23 R I V U X G

The Veil Nebula—A Supernova Remnant This nebulosity is a portion of the Cygnus Loop, which is the roughly spherical remnant of a supernova that exploded about 15,000 years ago. The distance to the nebula is about 800 pc (2600 ly), and the overall diameter of the loop is about 35 pc (120 ly). (Palomar Observatory)

Figure 20-24 R I V U X G

The Gum Nebula—A Supernova Remnant This nebula has the largest angular size (40°) of any known supernova remnant. Only the central regions of the Gum Nebula are shown here. The supernova explosion occurred about 11,000 years ago, and the remnant now has a diameter of about 700 pc (2300 ly). (Royal Observatory, Edinburgh)

The Gum Nebula looks big simply because it is quite close to us. Its center is only about 460 pc (1300 ly) from Earth, and its near side is just 100 pc (330 ly) away. Studies of the nebula's expansion rate suggest that this supernova exploded around 9000 B.C. At maximum brilliance, the exploding star probably was as bright as the Moon at first quarter. Like the first quarter moon, it would have been visible in the daytime!

Many supernova remnants are virtually invisible at optical wavelengths. However, when the expanding gases collide with the interstellar medium, they radiate energy at a wide range of wavelengths, from X rays through radio waves. For example, Figure 20-25 shows a radio image of the supernova remnant Cassiopeia A. (Compare to Figure 18-24, which is a composite of observations of Cassiopeia A at X-ray, visible, and infrared wavelengths.) As a rule, radio searches for supernova remnants are more fruitful than optical searches. Only two dozen supernova remnants have been found in visible-light images, but more than 100 remnants have been discovered by radio astronomers.

From the expansion rate of Cassiopeia A, astronomers conclude that this supernova explosion occurred about 300 years ago. Although telescopes were in wide use by the late 1600s, no one saw the outburst (and no one today knows why). The last supernova seen in our Galaxy, which occurred in 1604, was observed by Johannes Kepler. In 1572, Tycho Brahe also recorded the sudden appearance of an exceptionally bright star in the sky. To find any other accounts of nearby bright supernovae, we must delve into astronomical records that are almost 1000 years old.

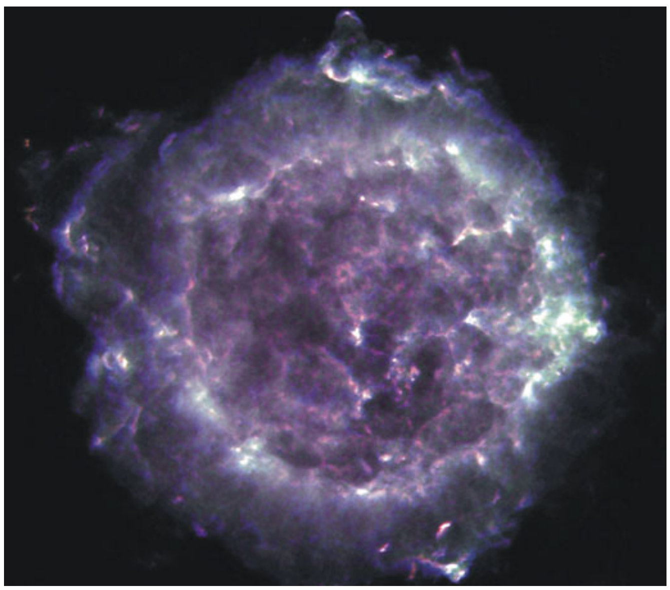
At first glance, this apparent lack of nearby supernovae may seem puzzling. From the frequency with which supernovae occur in distant galaxies, it is reasonable to suppose that a galaxy such as our own should have as many as five supernovae per century. Where have they been?

As we will learn when we study galaxies in Chapters 23 and 24, the plane of our Galaxy is where massive stars are born and supernovae explode. This region is so rich in interstellar dust, however, that we simply cannot see very far into space in the directions occupied by the Milky Way (see Section 20-2). In other words, supernovae probably do in fact erupt every few decades in remote parts of our Galaxy, but their detonations are hidden from our view by intervening interstellar matter.

Relics of the Fall: White Dwarfs, Neutron Stars, and Black Holes

A supernova remnant may be all that is left after some supernovae explode. But for core-collapse supernovae of Types II, Ib, and Ic, the core itself may also remain. If there is a relic of the core, it may be either a *neutron star* or a *black hole,* depending on the mass of the core and the conditions within it during the collapse. Neutron stars, as the name suggests, are made primarily of neutrons. Wholly unlike anything we have studied so far, these exotic objects are the subject of Chapter 22. We will study black holes, which are far stranger even than neutron stars, in Chapter 23.

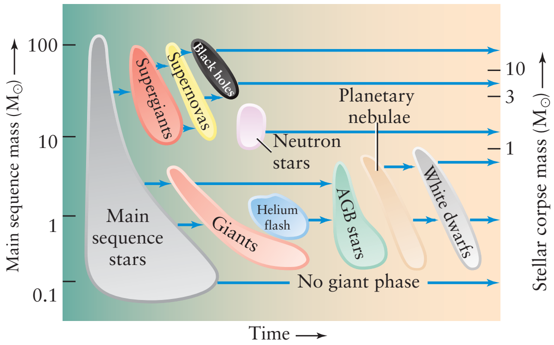
CAUTION! Although neutron stars and black holes can be part of the debris from a supernova explosion, they are *not* called "supernova remnants." That term is applied exclusively to the gas and dust that spreads away from the site of the supernova explosion.



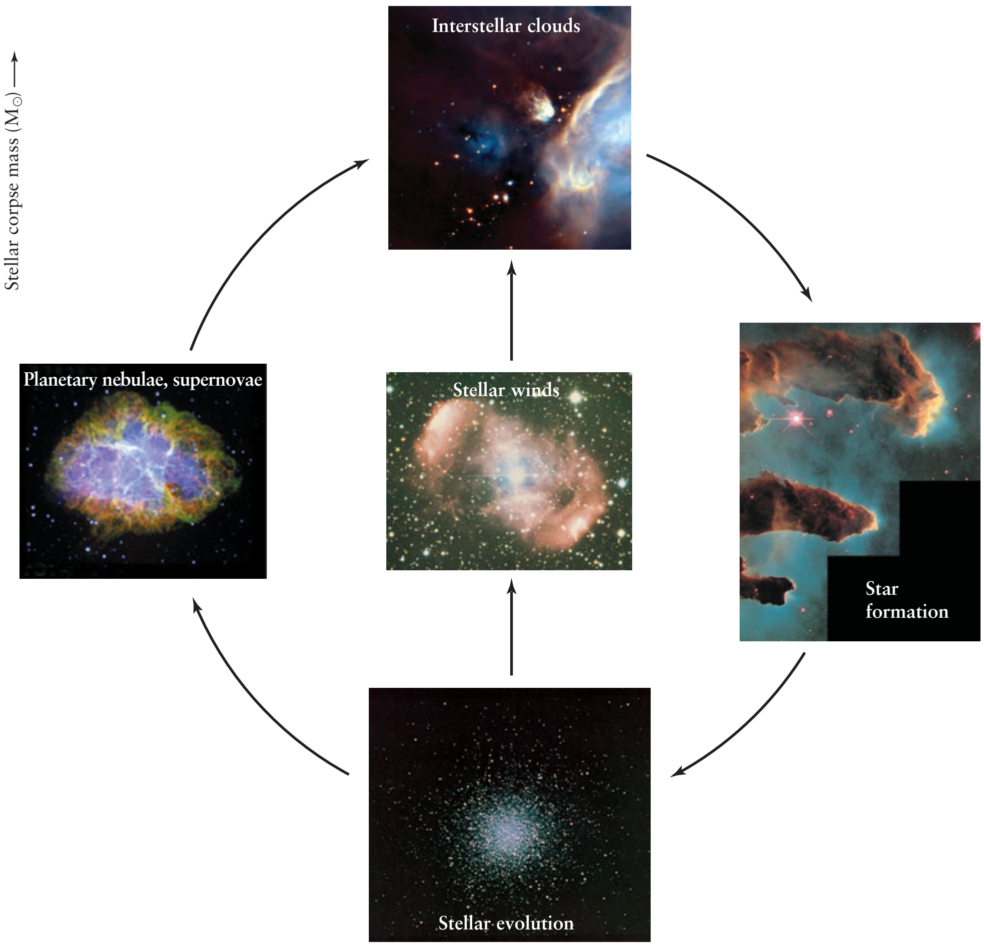
WEB LINK 20.13 ANIMATION 20.8

Figure 20-25 R I V U X G

Cassiopeia A—A Supernova Remnant This false-color radio image of Cassiopeia A was produced by the Very Large Array (see Figure 6-26). The impact of supernova material on the interstellar medium causes ionization, and the liberated electrons generate radio waves as they move. Cassiopeia A is roughly 3300 pc (11,000 ly) from Earth. (Image courtesy of NRAO/AUI)



(a)



(b) R I V U X G

Figure 20-26

A Summary of Stellar Evolution **(a)** The evolution of an isolated star (one that is not part of a close multiple-star system) depends on the star's mass. The more massive the star, the more rapid its evolution. The scale on the left gives the mass of the star when it is on the main sequence, and the scale at the right shows the mass of the resulting stellar corpse. If the star's initial mass is less than about 0.4 $M_\odot$, it evolves slowly over the eons into an inert ball of helium. If the initial mass is in the range from about 0.4 $M_\odot$ to about 8 $M_\odot$, it ejects enough mass over its lifetime so that what remains is a white dwarf with a mass less than the Chandrasekhar limit of 1.4 $M_\odot$. If the star's initial mass is more than about 8 $M_\odot$ it ends as a core-collapse supernova, leaving behind a neutron star or black hole **(b)** This figure summarizes the key stages in the cycle of stellar evolution. (b: top inset, Infrared Space Observatory, NASA; right inset, Anglo-Australian Observatory/J. Hester and P. Scowen, Arizona State University/NASA; bottom inset, NASA; left inset, NASA; middle inset, Anglo-Australian Observatory)

We have seen in this chapter that only the most massive stars end their lives as core-collapse supernovae. We learned earlier that mass plays a central role in determining the speed with which a star forms and joins the main sequence (see Figure 18-10), the star's luminosity and surface temperature while on the main sequence (see Figure 17-21 and the *Cosmic Connections* figure on page 459 of Chapter 17), and how long a star can remain on the main sequence (see Table 19-1). Now we see that a star's initial mass also determines its eventual fate (Figure 20-26*a*).

Perhaps the most remarkable aspect of this rich and varied story of stellar evolution is that stars are wholly natural phenomena. Beginning with a handful of simple ingredients—hydrogen, helium, and perhaps a dash of heavier elements—stars evolve into immense, glowing orbs that shine because one chemical element naturally converts into another deep in the star's interior. As they evolve and eventually cease to shine, many stars naturally return much of their material to interstellar space, where over time it may spawn new generations of stars (Figure 20-26*b*). For anyone who seeks an example of the magnificence and elegance of nature, a wise piece of advice would be to reflect on the processes of stellar evolution.

Key Words

asymptotic giant branch, p. 527
asymptotic giant branch star (AGB star), p. 527
carbon fusion, p. 533
carbon star, p. 528
Cerenkov radiation, p. 542
Chandrasekhar limit, p. 532
core-collapse supernova, p. 538
core helium fusion, p. 526
dredge-up, p. 528

Key Ideas

Late Evolution of Low-Mass Stars: A star of moderately low mass (about 0.4 $M_\odot$ to about 4 $M_\odot$) becomes a red giant when shell hydrogen fusion begins, a horizontal-branch star when core helium fusion begins, and an asymptotic giant branch (AGB) star when the helium in the core is exhausted and shell helium fusion begins.

• As a moderately low-mass star ages, convection occurs over a larger portion of its volume. This takes heavy elements formed in the star's interior and distributes them throughout the star.

Planetary Nebulae and White Dwarfs: Helium shell flashes in an old, moderately low-mass star produce thermal pulses during which more than half the star's mass may be ejected into space. This exposes the hot carbon-oxygen core of the star.

• Ultraviolet radiation from the exposed core ionizes and excites the ejected gases, producing a planetary nebula.

• No further nuclear reactions take place within the exposed core. Instead, it becomes a degenerate, dense sphere about the size of the Earth and is called a white dwarf. It glows from thermal radiation; as a white dwarf cools, it becomes dimmer.

Late Evolution of High-Mass Stars: Unlike a moderately low-mass star, a high-mass star (initial mass more than about 4 $M_\odot$) undergoes an extended sequence of thermonuclear reactions in its core and shells. These include carbon fusion, neon fusion, oxygen fusion, and silicon fusion.

• In the last stages of its life, a high-mass star has an iron-rich core surrounded by concentric shells hosting the various thermonuclear reactions. The sequence of thermonuclear reactions stops here, because the formation of elements heavier than iron requires an input of energy rather than causing energy to be released.

The Deaths of the Most Massive Stars: A star with an initial mass greater than 8 $M_\odot$ dies in a violent cataclysm in which its core collapses and most of its matter is ejected into space at high speeds. The luminosity of the star increases suddenly by a factor of around 10^8 during this explosion, producing a supernova.

• More than 99% of the energy from a core-collapse supernova is emitted in the form of neutrinos from the collapsing core.

• The matter ejected from the supernova, moving at supersonic speeds through interstellar gases and dust, glows as a nebula called a supernova remnant.

Other Types of Supernovae: An accreting white dwarf in a close binary system can also become a supernova when carbon fusion ignites explosively throughout such a degenerate star. This is called a thermonuclear supernova.

• A Type Ia supernova is produced by accreting white dwarfs in close binaries. A Type II supernova is the result of the collapse of the core of a massive star, as are supernovae of Type Ib and Type Ic; these latter types occur when the star has lost a substantial part of its outer layers before exploding.

• Most supernovae occurring in our Galaxy are hidden from our view by interstellar dust and gases.

Questions

Review Questions

1. What is the horizontal branch? Where is it located on an H-R diagram? How do stars on the horizontal branch differ from red giants or main-sequence stars?
2. Horizontal-branch stars are sometimes referred to as "helium main-sequence stars." In what sense is this true?
3. What is the asymptotic giant branch? Where is it located on an H-R diagram? How do asymptotic giant branch stars differ from red giants or main-sequence stars?
4. Is a carbon star a star that is made of carbon? Explain.
5. What is the connection between dredge-ups in old stars and life on Earth?
6. What are thermal pulses in AGB stars? What causes them? What effect do they have on the luminosity of the star?
7. How is a planetary nebula formed?
8. How can an astronomer tell the difference between a planetary nebula and a planet?
9. What is the evidence that typical planetary nebulae are only a few thousand years old?
10. Why do we not observe planetary nebulae that are more than about 50,000 years old?
11. What is a white dwarf? Does it produce light in the same way as a star like the Sun?
12. How does the radius of a white dwarf depend on its mass? How is this different from other types of stars?
13. What is the significance of the Chandrasekhar limit?
14. On an H-R diagram, sketch the evolutionary track that the Sun will follow from when it leaves the main sequence to when it becomes a white dwarf. Approximately how much mass will the Sun have when it becomes a white dwarf? Where will the rest of the mass go?
15. What prevents thermonuclear reactions from occurring at the center of a white dwarf? If no thermonuclear reactions are occurring in its core, why doesn't the star collapse?
16. A white dwarf has a greater mass than either a red dwarf or a brown dwarf. Yet a white dwarf has a smaller radius than either a red dwarf or a brown dwarf. Explain why, in terms of the types of pressure that keep the different kind of dwarfs from collapsing under their own gravity.

17. Why do you suppose that all the white dwarfs known to astronomers are relatively close to the Sun?
18. Why does the mass of a star play such an important role in determining the star's evolution?
19. Why is the temperature in a star's core so important in determining which nuclear reactions can occur there?
20. What is the difference between a red giant and a red supergiant?
21. Why does the evolutionary track of a high-mass star move from left to right and back again in the H-R diagram?
22. In what way does the structure of an aging supergiant resemble that of an onion?
23. What is nuclear density? Why is it significant when a star's core reaches this density?
24. Why is SN 1987A so interesting to astronomers? In what ways was it not a typical supernova?
25. Why are neutrinos emitted by core-collapse supernovae? How can these neutrinos be detected? How can they be distinguished from solar neutrinos?
26. What causes a thermonuclear supernova? How does a thermonuclear supernova compare with a core-collapse supernova?
27. What are the differences among Type Ia, Type Ib, Type Ic, and Type II supernovae? Which type is most unlike the other three, and why?
28. How can a supernova continue to shine for many years after it explodes?
29. How do supernova remnants produce radiation at nonvisible wavelengths?
30. There may have been recent supernovae in our Galaxy that have not been observable even though they are incredibly luminous. How is it this possible?
31. Is our own Sun likely to become a supernova? Why or why not?

Advanced Questions

Questions preceded by an asterisk (*) involve topics discussed in the Boxes in Chapter 1, Chapter 7, or Chapter 17.

Problem-solving tips and tools

The small-angle formula is given in Box 1-1. You may find it useful to review Box 17-4, which discusses stellar radii and their relationship to temperature and luminosity. Section 4-7 explains the formula for gravitational force, and Box 7-2 explains the concept of escape speed. Sections 5-2 and 5-4 describe some key properties of light, especially blackbody radiation. We discussed the relationship among luminosity, apparent brightness, and distance in Box 17-2. The relationship among absolute magnitude, apparent magnitude, and distance was the topic of Box 17-3. In our discussion of binary stars in Section 17-9 we saw how the masses of the stars, the orbital period, and the average distance between the two stars are all related.

32. Some blue main-sequence stars in our region of the Galaxy have the same luminosity and surface temperature as the horizontal-branch stars in the globular cluster M55 (see Figure 20-3). How do we know that the horizontal-branch stars in M55 are not main-sequence stars?
33. Stellar winds from an AGB star can cause it to lose mass at a rate of up to 10^{-4} $M_{\odot}$ per year. (**a**) Express this rate in metric tons per second. (One metric ton equals 1000 kilograms.) (**b**) At this rate, how long would it take an AGB star to eject an amount of mass equal to the mass of the Earth? Express your answer in days.
34. The globular cluster M15 depicted in Figure 20-6*a* contains 30,000 old stars, but only one of these stars is presently in the planetary nebula stage of its evolution. Explain why planetary nebulae are not more prevalent in M15.
35. Explain why astronomers had to use infrared light to detect the faint wisps in planetary nebula NGC 7027 (Figure 20-6*c*).
36. The central star in a newly formed planetary nebula has a luminosity of 1000 $L_{\odot}$ and a surface temperature of 100,000 K. What is the star's radius? Give your answer as a multiple of the Sun's radius.
37. You want to determine the age of a planetary nebula. What observations should you make, and how would you use the resulting data?
38. Figure 20-6*b* shows the planetary nebula Abell 39. Use the information given in the figure caption to calculate the angular diameter of the nebula as seen from Earth.

*39. The Ring Nebula is a planetary nebula in the constellation Lyra. It has an angular size of 1.4 arcmin × 1.0 arcmin and is expanding at the rate of about 20 km/s. Approximately how long ago did the central star shed its outer layers? Assume that the nebula is 2,700 ly from Earth.

40. The accompanying image shows the planetary nebula IC 418 in the constellation Lepus (the Hare). (**a**) The image shows a small shell of glowing gas (shown in blue) within a larger glowing gas shell (shown in orange). Discuss how IC 418 could have acquired this pair of gas shells. (**b**) Explain why the outer shell looks thicker around the edges than near the middle.

R I **V** U X G

(NASA and Hubble Heritage Team, STScI/AURA)

41. (**a**) Calculate the wavelength of maximum emission of the white dwarf Sirius B. In what part of the electromagnetic spectrum does this wavelength lie? (**b**) In a visible-light photograph such as Figure 20-8, Sirius B appears much fainter than its primary star. But in an image made with an X-ray telescope, Sirius B is the brighter star. Explain the difference.

***42.** Sirius is 2.63 pc from Earth. By making measurements on Figure 20-8, calculate the distance between the centers of Sirius A and Sirius B at the time that this image was made in October 2003. Give your answer in astronomical units (AU). (Note that your result is the true distance only if Sirius A and Sirius B were exactly the same distance from Earth at the time the image was made. If one of the stars was closer to us than the other, the actual distance between them is greater than what you calculate.)

43. (**a**) Find the average density of a 1-$M_{\odot}$ white dwarf having the same diameter as the Earth. (**b**) What speed is required to eject gas from the white dwarf's surface? (This is also the speed with which interstellar gas falling from a great distance would strike the star's surface.)

44. In the classic 1960s science-fiction comic book *The Atom,* a physicist discovers a basketball-sized meteorite (about 10 cm in radius) that is actually a fragment of a white dwarf star. With some difficulty, he manages to hand-carry the meteorite back to his laboratory. Estimate the mass of such a fragment, and discuss the plausibility of this scenario.

***45.** (**a**) Use the information in the caption to Figure 20-12 to calculate the diameter of the nebula SMC N76. Express your answer in parsecs. (**b**) How does your answer to part (a) compare to the diameters of the planetary nebulae depicted in Figure 20-6? Explain how this is consistent with the observation that gases ejected from a supergiant travel faster than gases ejected from an AGB star.

***46.** The supergiant star depicted in Figure 20-12 is actually one member of a binary star system. The masses of the two stars are 18 $M_{\odot}$ and 34 $M_{\odot}$, and the orbital period is 19.56 days. (**a**) What is the average separation between the two stars? Give your answer in AU. (**b**) Compare your answer in part (a) to the sizes of the orbits of Mercury, Venus, and Earth around the Sun.

47. (**a**) What kinds of stars would you monitor if you wished to observe a core-collapse supernova explosion from its very beginning? (**b**) Examine Appendices 4 and 5, which list the nearest and brightest stars, respectively. Which, if any, of these stars are possible supernova candidates? Explain.

***48.** Consider a high-mass star just prior to a supernova explosion, with a core of diameter 20 km and density 4×10^{17} kg/m^3. (**a**) Calculate the mass of the core. Give your answer in kilograms and in solar masses. (**b**) Calculate the force of gravity on a 1-kg object at the surface of the core. How many times larger is this than the gravitational force on such an object at the surface of the Earth, which is about 10 newtons? (**c**) Calculate the escape speed from the surface of the star's core. Give your answer in m/s and as a fraction of the speed of light. What does this tell you about how powerful a supernova explosion must be in order to blow material away from the star's core?

49. The shock wave that traveled through the progenitor star of SN 1987A took 3 hours to reach the star's surface. (**a**) Given the size of a blue supergiant star (see Section 20-7), estimate the speed with which the shock wave traveled through the star's outer layers. (The core of the progenitor star was very small, so you may consider the shock wave to have started at the very center of the star.) Give your answer in meters per second. (**b**) Compare your answer with the speed of sound waves in our atmosphere, about 340 m/s, and with the speed of light. (**c**) A shock wave traveling through a gas is a special case of a sound wave. In general, sound waves travel faster through denser, less easily compressed materials. Thus, sound travels faster through water (about 1500 m/s) than through our atmosphere, and faster still through steel (about 5900 m/s). Use this idea to compare the gases within the progenitor star of SN 1987A with the gases in our atmosphere in terms of their average density and how easily they are compressed.

50. The neutrinos from SN 1987A arrived 3 hours before the visible light. While they were en route to the Earth, what was the distance between the neutrinos and the first photons from SN 1987A? Assume that neutrinos are massless and thus travel at the speed of light. Give your answers in kilometers and in AU.

***51.** Compared to SN 1987A (see Figure 20-17), the supernova SN 1993J (see Figure 20-16) had a maximum apparent brightness only 9.1×10^{-4} as great. Using the distances from Earth to each of these supernovae, determine the ratio of the maximum luminosity of SN 1993J to that of SN 1987A. Which of the two supernovae had the greater maximum luminosity?

***52.** Suppose that the brightness of a star becoming a supernova increases by 20 magnitudes. Show that this corresponds to an increase of 10^8 in luminosity.

***53.** Suppose that the red-supergiant star Betelgeuse, which lies some 425 light-years from the Earth, becomes a Type II supernova. (**a**) At the height of the outburst, how bright would it appear in the sky? Give your answer as a fraction of the brightness of the Sun ($b_{\odot}$). (**b**) How would it compare with the brightness of Venus (about 10^{-9} $b_{\odot}$)?

***54.** In July 1997, a supernova named SN 1997cw exploded in the galaxy NGC 105 in the constellation Cetus (the Whale). It reached an apparent magnitude of +16.5 at maximum brilliance, and its spectrum showed an absorption line of ionized silicon. Use this information to find the distance to NGC 105. (*Hint:* Inspect the light curves in Figure 20-22 to find the *absolute* magnitudes of typical supernovae at peak brightness.)

55. Figure 20-23 shows a portion of the Veil Nebula in Cygnus. Use the information given in the caption to find the average speed at which material has been moving away from the site of the supernova explosion over the past 15,000 years. Express your answer in km/s and as a fraction of the speed of light.

56. The images that open this chapter show two kinds of glowing gas clouds: a planetary nebula and a supernova remnant. (**a**) Explain what makes the planetary nebula glow and what makes the supernova remnant glow. (*Hint:* The explanations are different for the two kinds of gas clouds.) (**b**) Which of these two kinds of gas clouds continues to glow for a longer time? Why?

57. The planetary nebula and supernova remnant shown in the images that open this chapter are both about the same age. Both objects consist of glowing gases that have expanded away from a central star. Based on these images, in which of these objects have the gases expanded more rapidly? Explain your reasoning.

Discussion Questions

58. Suppose that you discover a small, glowing disk of light while searching the sky with a telescope. How would you decide if this object is a planetary nebula? Could your object be something else? Explain.

59. Suppose the convective zone in AGB stars did *not* reach all the way down into their carbon-rich cores. How might this have affected the origin and evolution of life on Earth?

60. Imagine that our Sun was somehow replaced by a 1-$M_{\odot}$ white dwarf star, and that our Earth continued in an orbit of semimajor axis 1 AU around this star. Discuss what effects this would have on our planet. What would the white dwarf look like as seen from Earth? Could you look at it safely with the unaided eye? Would the Earth's surface temperature remain the same as it is now?

61. The similar names *white dwarf*, *red dwarf*, and *brown dwarf* describe three very different kinds of objects. Suggest better names for these three kinds of objects, and describe how your names more accurately describe the objects' properties.

62. The major final product of silicon fusion is ^{56}Fe, an isotope of iron with 26 protons and 30 neutrons. This is also the most common isotope of iron found on Earth. Discuss what this tells you about the origin of the solar system.

63. SN 1987A did not agree with the theoretical picture outlined in Section 20-6. Does this mean that the theory was wrong? Discuss.

Web/eBook Questions

64. It has been claimed that the Dogon tribe in western Africa has known for thousands of years that Sirius is a binary star. Search the World Wide Web for information about these claims. What is the basis of these claims? Why are scientists skeptical, and how do they refute these claims?

65. Search the World Wide Web for recent information about SN 1993J. Has the shape of the supernova's light curve been adequately explained? Has the supernova produced any surprises?

66. Search the World Wide Web for information about SN 1994I, a supernova that occurred in the galaxy M51 (NGC 5194). Why was this supernova unusual? Was it bright enough to have been seen by amateur astronomers?

67. Convection Inside a Giant Star. Access and view the animation "Convection Inside a Giant Star" in Chapter 20 of the *Universe* Web site or eBook. Describe the motion of material in the interior of the star. In what ways is this motion similar to convection within the present-day Sun (see Section 16-2)? In what ways is it different? Is a dredge-up taking place in this animation? How can you tell?

68. Types of Supernovae. Access and view the animations "In the Heart of a Core-Collapse (Type II, Ib, or Ic) Supernova" and "A Thermonuclear (Type Ia) Supernova" in Chapter 20 of the *Universe* Web site or eBook. Describe how these two types of supernova are fundamentally different in their origin.

Activities

Observing Projects

Observing tips and tools

While planetary nebulae are rather bright objects, their brightness is spread over a relatively large angular size, which can make seeing them a challenge for the beginning observer. For example, the Helix Nebula (shown in the left-hand image on the page that opens this chapter) has the largest angular size of any planetary nebula but is also one of the most difficult to see. To improve your view, make your observations on a dark, moonless night from a location well shielded from city lights. Another useful trick, mentioned in Chapters 17 and 18, is to use "averted vision." Once you have the nebula centered in the telescope, you will get a brighter and clearer image if you look at the nebula out of the corner of your eye. The so-called Blinking Planetary in Cygnus affords an excellent demonstration of this effect; the nebula seems to disappear when you look straight at it, but it reappears as soon as you look toward the side of your field of view.

Another useful tip is to view the nebula through a green filter (a #58, or O III, filter available from telescope supply houses). Green light is emitted by excited, doubly ionized oxygen atoms, which are common in planetary nebulae but not in most other celestial objects. Using such a filter can make a planetary nebula stand out more distinctly against the sky. As a side benefit, it also helps to block out stray light from street lamps. The same tips also apply to observing supernova remnants.

69. Although they represent a fleeting stage at the end of a star's life, planetary nebulae are found all across the sky. Some of the brightest are listed in the accompanying table. Note that the distances to most of these nebulae are quite uncertain. Observe as many of these planetary nebulae as you can on a clear, moonless night using the largest telescope at your disposal. Note and compare the various shapes of the different nebulae. In how many cases can you see the central star? The central star in the Eskimo Nebula is supposed to be the "nose" of an Eskimo wearing a parka. Can you see this pattern?

70. Northern hemisphere observers with modest telescopes can see two supernova remnants, one in the winter sky and the other in the summer sky. Both are quite faint, however, so you should schedule your observations for a moonless night.

Planetary nebula	Distance (light-years)	Angular size	Constellation	Right ascension	Declination
Dumbbell (M27, NGC 6853)	490–3500	8.0 × 5.7	Vulpecula	19^{h} 59.6^{m}	+22° 43′
Ring (M57, NGC 6720)	1300–4100	1.4 × 1.0	Lyra	18 53.6	+33 02
Little Dumbbell (M76, NGC 650)	1700–15,000	2.7 × 1.8	Perseus	01 42.4	+51 34
Owl (M97, NGC 3587)	1300–12,000	3.4 × 3.3	Ursa Major	11 14.8	+55 01
Saturn (NGC 7009)	1600–3900	0.4 × 1.6	Aquarius	21 04.2	−11 22
Helix (NGC 7293)	450	41 × 41	Aquarius	22 29.6	−20 48
Eskimo (NGC 2392)	1400–10,000	0.5 × 0.5	Gemini	7 29.2	+20 55
Blinking Planetary (NGC 6826)	3300(?)	2.2 × 0.5	Cygnus	19 44.8	+50 31

The winter sky contains the Crab Nebula, which is discussed in detail in Chapter 23. The coordinates are R.A. = 5^{h} 34.5^{m} and Decl. = +22°00′, which places the object near the star marking the eastern horn of Taurus (the Bull). Whereas the entire Crab Nebula easily fits in the field of view of an eyepiece, the Veil or Cirrus Nebula in the summer sky is so vast that you can see only a small fraction of it at a time. The easiest way to find the Veil Nebula is to aim the telescope at the star 52 Cygni (R.A. = 20^{h} 45.7^{m} and Decl. = +30° 43′), which lies on one of the brightest portions of the nebula. If you then move the telescope slightly north or south until 52 Cygni is just out of the field of view, you should see faint wisps of glowing gas.

71. Use a telescope to observe the remarkable triple star 40 Eridani, whose coordinates are R.A. = 4^{h} 15.3^{m} and Decl. = −7° 39′. The primary, a 4.4-magnitude yellowish star like the Sun, has a 9.6-magnitude white dwarf companion, the most easily seen white dwarf in the sky. On a clear, dark night with a moderately large telescope, you should also see that the white dwarf has an 11th-magnitude companion, which completes this most interesting trio.

72. The red supergiant Betelgeuse in the constellation Orion will explode as a supernova at some time in the future. Use the *Starry Night Enthusiast™* program to investigate how the supernova might appear if this explosion were to happen tonight. Click the **Home** button in the toolbar to show the sky as seen from your location at the present time. (If the program does not place you at your true location, use the **Viewing Location . . .** command in the **Options** menu.) Use the **Find** pane to locate Betelgeuse. If Betelgeuse is below the horizon, allow the program to reset the time to when it is visible. (**a**) At what time does Betelgeuse rise on today's date? At what time does it set? (**b**) If Betelgeuse became a supernova today, would it be visible in the daytime? How would it appear at night? Do you think it would cast shadows? (**c**) Are Betelgeuse and the Moon both in the night sky tonight? (Use the **Find** pane to locate the Moon.) If they are, and Betelgeuse were to become a supernova, what kinds of shadows might they both cast?

73. Use the *Starry Night Enthusiast™* program to show the location of Supernova 1987A. In the menu, select **Favourites > Deep Space > Local Universe** to display the Milky Way and other nearby galaxies, conveniently labeled, against the background of distant galaxies, from a distance of 0.282 Mly from the Sun. (If the Milky Way does not appear immediately, click once on either of the **Zoom** buttons.) Remove the image of the astronaut's feet by clicking on **View > Feet**. You can rotate the Milky Way Galaxy and its neighbor galaxies by holding down both the **Shift** button and the mouse button while moving the mouse. (On a two-button mouse, hold down the left mouse button.) (**a**) Use the **Find** pane to locate and center the Sun in the field of view. Describe the position of the Large Magellanic Cloud (LMC), within which SN 1987A lies, relative to the Milky Way Galaxy and to our solar system. (**b**) Use the **Find** pane to center on the LMC. You should be able to locate the Tarantula Nebula, shown in Figure 20-17. Is SN 1987A near to the center or the edge of the LMC? (Note that, although *Starry Night Enthusiast™* depicts the LMC as being rather flat, it is thought to be an irregular blob of stars with some thickness.)

Collaborative Exercise

74. Imagine that a supernova originating from a close binary star system, both of whose stars have less than 4 solar masses, began (as seen from Earth) on the most recent birthday of the youngest person in your group. Using the light curves in Figure 20-22, what would its new luminosity be today and how bright would it appear in the sky (apparent magnitude) if it were located 10 parsecs (32.6 light-years) away? How would your answers change if you were to discover that the supernova actually originated from an isolated star with a mass 15 times greater than our Sun?

21

Neutron Stars

WEB LINK 21.1

R I V U X G

A composite image of the Crab Nebula using X-ray (blue) and visible (red) wavelengths. (X-ray: NASA/CXC/ASU/J. Hester et al.; optical: NASA/CXC/ASU/J. Hester et al.)

Almost all the light we see from objects in the night sky is due to thermonuclear fusion reactions. Such reactions are what make stars shine, and in turn make any surrounding nebulae glow. Even the light we see from the Moon and planets—which is simply reflected sunlight—can be traced back to thermonuclear reactions in the Sun's core.

By contrast, the exotic object shown here—the Crab Nebula in Taurus—is undergoing no thermonuclear fusion at all. Despite this, the object is so energetic that it is emitting copious amounts of X rays. This image shows tilted, glowing rings more than a light-year across, with oppositely directed jets flowing outward perpendicular to the plane of the rings. At the object's very center, too small to pick out in this image, is the source of all this dynamic activity: a neutron star, the leftover core of what was once a supergiant star.

Before 1967 neutron stars were considered pure speculation. But today we know of more than 1300 of these exotic stellar corpses, and hundreds of thousands more may be strewn across the Milky Way Galaxy. As we will discover in this chapter, neutron stars have powerful magnetic fields that are billions of times stronger than the Sun's field. As a neutron star rotates, its magnetic field sweeps beams of radiation across the sky. We detect these beams as pulsating radio signals.

Unlike normal stars, neutron stars actually have solid surfaces that can shift and fracture in a "starquake." When such a quake occurs on the surface of a neutron star with an unusually strong magnetic field, it releases a truly colossal burst of radiation that for a fraction of a second outshines an entire galaxy of stars.

We will find that neutron stars in close binary systems also display outrageous behavior, including bouts of explosive helium fusion that yield a vast outpouring of X rays. As we will see, a similar process involving a white dwarf in a close binary system produces a short-lived intense burst of visible light called a nova.

21-1 A neutron star is even more highly compressed than a white dwarf

On the morning of July 4, 1054, Yang Wei-T'e, the imperial astronomer to the Chinese court, made a startling discovery. Just a few minutes before sunrise, a new and dazzling object ascended above the eastern horizon. This "guest star," as Yang called it, was far brighter than Venus and more resplendent than any star he had ever seen.

Yang's records show that the "guest star" was so brilliant that it could easily be seen during broad daylight for the rest of July. Records from Constantinople (now Istanbul, Turkey) also

Learning Goals

By reading the sections of this chapter, you will learn

21-1 Why astronomers predicted the existence of neutron stars before they were discovered

21-2 What pulsars are and how they were discovered

21-3 The relationship between pulsars and neutron stars

21-4 Why pulsars slow down over time

21-5 What exotic states of matter are thought to exist inside neutron stars

21-6 How neutron stars with abnormally strong magnetic fields can produce amazingly intense bursts of radiation

21-7 How some pulsars are accelerated to truly dizzying rotation speeds

21-8 How pulsars in close binary systems can emit X-rays

21-9 How novae produced by white dwarfs are similar to X-ray bursters produced by neutron stars

21-10 What sets the upper limit on the mass of a neutron star

Figure 21-1 R I **V** U X G

A Supernova Pictograph? This drawing in an eleventh-century structure in New Mexico shows a ten-pointed star next to a crescent. It may depict the scene on the morning of July 5, 1054, when a "guest star" appeared next to the waning crescent moon. (Courtesy of National Parks Service)

describe this object, and works of art made by the Anasazi culture in the American Southwest suggest that they may have seen it as well (Figure 21-1). Over the next 21 months, however, the "guest star" faded to invisibility.

We now know that the "guest star" of 1054 was actually a remarkable stellar transformation: A massive star some 6500 light-years away perished in a supernova explosion, leaving behind both a supernova remnant (see Figure 1-8 and Section 20-10) and a bizarre object called a **neutron star**—an incredibly dense sphere composed primarily of neutrons. However, it took many centuries after Yang Wei-T'e's observations to understand what had happened; it was not until 1932 that the neutron itself was discovered by the English physicist James Chadwick. (We learned in Section 5-7 that the neutron has about the same mass as the positively charged proton, but has no electric charge.)

We saw in Section 20-6 that under very high pressure, a proton and an electron can combine to form a neutron (as well as a neutrino). Could such pressures be found within the core of a dying high-mass star? Within a year of Chadwick's discovery, astronomers Fritz Zwicky at the California Institute of Technology (Caltech) and his colleague Walter Baade at Mount Wilson Observatory made just this prediction—that a massive stellar core can transform into a sphere of neutrons. "With all reserve," Zwicky and Baade theorized, "we advance the view that supernovae represent the transition from ordinary stars into neutron stars, which in their final stages consist of extremely closely packed neutrons." In other words, there could be at least two types of stellar corpses—white dwarfs and neutron stars.

Zwicky and Baade realized that once a neutron star had formed in response to tremendous pressures, it would be able to resist any further compression. The neutrons themselves would provide a counterbalancing outward pressure. Motivated by the idea that white dwarfs are supported by degenerate electron pressure, Zwicky and Baade proposed that a highly compact ball of neutrons would similarly produce a **degenerate neutron pressure.** (Like electrons, neutrons obey the Pauli exclusion principle that we described in Section 19-3.) This powerful pressure could support a stellar corpse—perhaps one even more massive than a white dwarf, because degenerate neutron pressure can be much stronger than degenerate electron pressure. So, while a white dwarf collapses if its mass is above the Chandrasekhar limit (see Section 20-4), a neutron star might not.

Although neutron stars were first hypothesized in the 1930s, the idea seemed so strange that it was ignored for over thirty years

Although Zwicky and Baade's proposal would prove to be very close to the mark, most scientists politely ignored it for years. After all, a neutron star must be a rather weird object. If brought to the Earth's surface, a single thimbleful of neutron star matter would weigh 100 million tons!

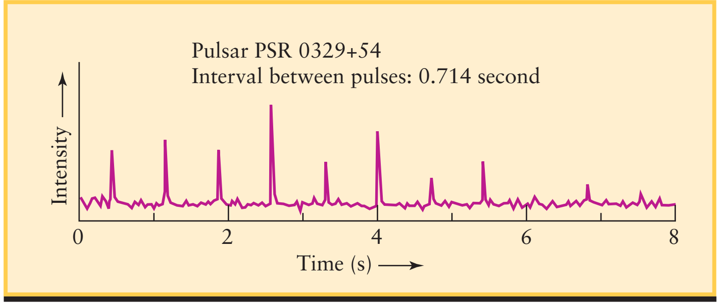
A star compacted to such densities must be very small. A $1.4\text{-}M_\odot$ neutron star would have a diameter of only 20 km (12 mi), about the size of a moderate-sized city on Earth. Its surface gravity would be so strong that an object would have to travel at one-half the speed of light to escape into space. These conditions seemed outrageous until the late 1960s, when astronomers discovered pulsating radio sources.

21-2 The discovery of pulsars in the 1960s stimulated interest in neutron stars

As a young graduate student at Cambridge University, Jocelyn Bell spent many months helping construct an array of radio antennas covering $4\frac{1}{2}$ acres of the English countryside. The instrument was completed by the summer of 1967, and Bell and her colleagues in Anthony Hewish's research group began using it to scrutinize radio emissions from the sky. They were looking for radio sources that "twinkle" like stars; that is, they looked for random small fluctuations in brightness caused by the motion of gas between the source and the observer. What they discovered was something far more exotic.

The Discovery of Pulsars

While searching for random flickering, Bell noticed that the antennas had detected regular pulses of radio noise from one particular location in the sky. These radio pulses were arriving at regular intervals of 1.3373011 seconds—much more rapid than those of any other astronomical object known at that time. Indeed, they were so rapid and regular that the Cambridge team at first suspected that they might not be of natural origin. Instead, it was proposed that these pulses might be signals from an advanced alien civilization.



WEB LINK 21.4

Figure 21-2

A Recording of a Pulsar This chart recording shows the intensity of radio emission from one of the first pulsars to be discovered. (The designation means "pulsar at a right ascension of 03^h 29^m and a declination of +54°") Note that the interval between pulses is very regular, even though some pulses are weak and others are strong. (Adapted from R. N. Manchester and J. H. Taylor)

That possibility had to be discarded within a few months after several more of these pulsating radio sources, which came to be called **pulsars,** were discovered across the sky. In all cases, the periods were extremely regular, ranging from about 0.25 second for the fastest to about 1.5 seconds for the slowest (Figure 21-2).

Ruling Out Possibilities

What could these objects possibly be? As is often done in science, astronomers first decided what pulsars could *not* be. It was clear that pulsars were not ordinary stars or nebulae, because while these objects emit some radio waves, their emissions do not pulsate (that is, their light output does not increase and decrease rhythmically). One type of star that *appears* to pulsate is an eclipsing binary star system, which appears dimmer when one star in the binary system passes in front of the other (see Section 17-11). However, in order to vary in brightness with a typical pulsar period of 1 second or less, an eclipsing binary would have to have an equally short orbital period. Newton's form of Kepler's third law (see Section 4-7) implies that the two stars would have to orbit only 1000 kilometers or so from each other. This is less than the diameter of any ordinary star, even a white dwarf star. Hence, the eclipsing binary hypothesis can be ruled out, because it leads to the highly implausible picture of two stars that overlap!

Astronomers also ruled out the idea that pulsars might be variable stars, which vary in luminosity as they increase and decrease in size (see Section 19-6). Pulsating variables have periods of days or weeks, while pulsars have periods of seconds or less. Just as the broad wings of an eagle could never flap as fast as a hummingbird's wings without breaking, an ordinary star or even a white dwarf is too large to pulsate in and out in less than a second.

Another possibility, however, was given serious consideration. Perhaps pulsars were rapidly rotating white dwarf stars with some sort of radio-emitting "hot spots" on their surfaces. But even this model had its limits, because a white dwarf that rotated on its axis in just 1 second would be on the verge of flying apart. What finally ruled out the spinning white dwarf model was the discovery of a pulsar in the middle of the Crab Nebula, the supernova remnant left behind by the supernova that Yang Wei-T'e saw in 1054 (see Figure 1-8). That discovery clarified the true nature of pulsars.

At the time of its discovery, the Crab pulsar was the fastest pulsar known to astronomers. Its period is 0.0333 second, which means that it flashes 1/0.0333 = 30 times each second. It was immediately apparent to astrophysicists that white dwarfs are too big and bulky to generate 30 signals per second; calculations demonstrated that a white dwarf could not rotate that fast without tearing itself apart. Hence, the existence of the Crab pulsar means that the stellar corpse at the center of the Crab Nebula has to be much smaller and more compact than a white dwarf.

The rapid flashing of radiation from a pulsar gave evidence that white dwarfs are not the only endpoint of stellar evolution

This was a truly unsettling realization, because most astronomers in the mid-1960s thought that *all* stellar corpses are white dwarfs. The number of white dwarfs in the sky seemed to account for all the stars that must have died since our Galaxy was formed. It was generally assumed that all dying stars—even the most massive ones, which produce supernovae—somehow manage to eject enough matter so that their corpses do not exceed the Chandrasekhar limit. But the discovery of the Crab pulsar showed these conservative opinions to be in error.

It was now clear that pulsars had to be something far more exotic than rotating white dwarfs. As Thomas Gold of Cornell University emphasized, astronomers would now have to seriously consider the existence of neutron stars.

21-3 Pulsars are rapidly rotating neutron stars with intense magnetic fields

Science is, by its very nature, based on skepticism. If a scientist proposes a new and exotic explanation for an unexplained phenomenon, that explanation will be accepted only if it satisfies two demanding criteria. First, there must be excellent reasons why more conventional explanations cannot work. Second, the new explanation must help scientists understand a range of other phenomena.

A good example is the Copernican idea that the Earth orbits the Sun, which gained general acceptance only after Galileo's observations (see Section 4-5). First, Galileo showed the older geocentric explanation of celestial motions to be untenable, and, second, it became clear that the Copernican model could also explain the motions of Jupiter's moons.

How the Neutron-Star Model Overcame Skepticism

We have seen how the discovery of the Crab pulsar ruled out the idea that pulsars were white dwarfs. But astronomers soon accepted the more exotic idea that pulsars were actually neutron stars. To understand why, we must assume the skeptical outlook of the scientist and ask some probing questions about this idea. Why should neutron stars emit radiation at all? In particular, why should they emit radio waves? Why should the emissions be pulsed? And how could the pulses occur as rapidly as 30 times per second, as in the Crab pulsar? The answers to all of these questions are intimately related to the small size of neutron stars.

LOOKING DEEPER 21.1

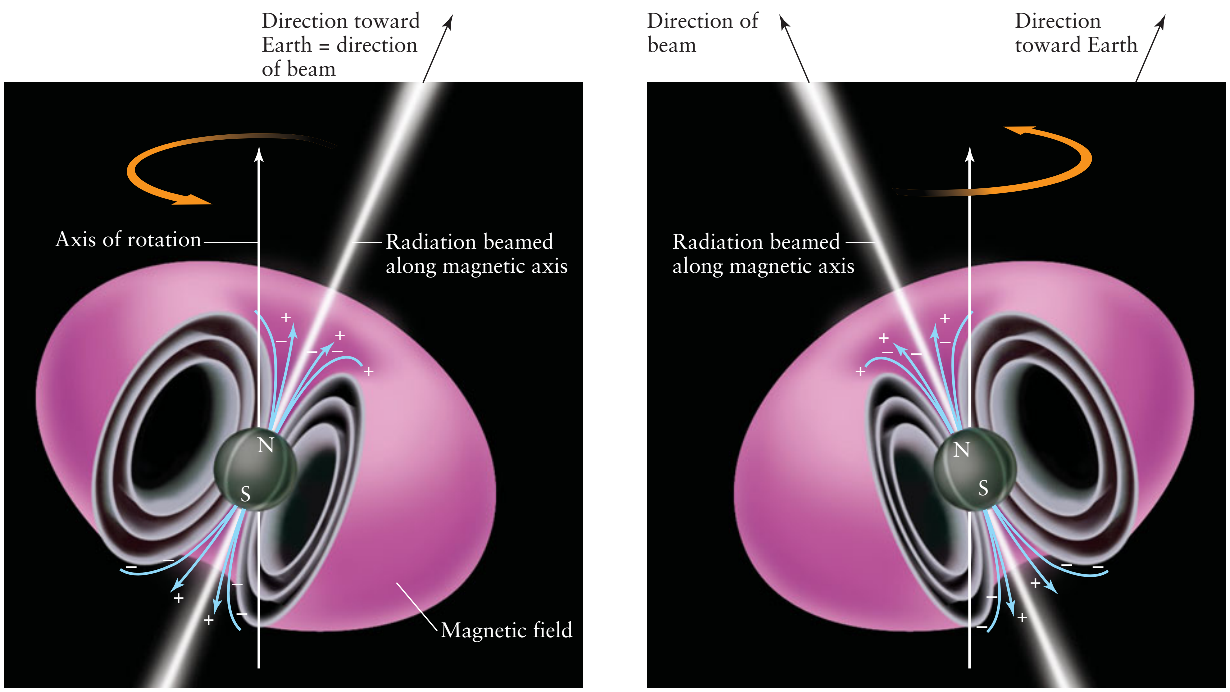
Because neutron stars are very small, they should also rotate rapidly. A typical star, such as our Sun, takes nearly a full month to rotate once about its axis. But just as an ice skater doing a pirouette speeds up when she pulls in her arms, a collapsing star also speeds up as its size shrinks. (We introduced this principle, called the *conservation of angular momentum,* in Section 8-4.) If our Sun were compressed to the size of a neutron star, it would spin about 1000 times per second! Because neutron stars are so small and dense, they can spin this rapidly without flying apart.

The neutron-star model of pulsars had to pass several stringent tests to be accepted by astronomers

The small size of neutron stars also implies that they have intense magnetic fields. It seems safe to say that every star possesses some magnetic field, but typically the strength of this field is quite low. The magnetic field of a main-sequence star is spread out over billions of square kilometers of the star's surface. However, if such a star collapses down to a neutron star, its surface area (which is proportional to the square of its radius) shrinks by a factor of about 10^{10}. The magnetic field, which is bonded to the star's ionized gases (see Section 16-9), becomes concentrated onto an area 10^{10} times smaller than before the collapse, and thus the field strength increases by a factor of 10^{10}.

A star's magnetic field splits one of its spectral lines into two or more lines whose spacing reveals the field's strength (see Figure 16-20). The strength of a magnetic field is usually measured in the unit called the *gauss* (G), named after the famous German mathematician and astronomer Karl Friedrich Gauss. (As an example, the Earth's magnetic field has a strength of about 0.5 G. The field at the surface of a refrigerator magnet is about 100 G, and the field used in an MRI scanner is about 15,000 G.) The magnetic splitting of lines in the spectra of main-sequence stars reveals field strengths in the range from about 1 G (the value for the Sun) to several thousand gauss. The splitting of lines in the spectra of certain white dwarfs corresponds to field strengths in excess of 10^6 G, comparable to the strongest fields ever produced in a laboratory on Earth. But the magnetic fields surrounding typical neutron stars are a million times stronger still, on the order of 1 trillion gauss (10^{12} G). As we will see in Section 21-6, some neutron stars may have fields as strong as 10^{15} G.

The magnetic field of a neutron star makes it possible for the star to radiate pulses of energy toward our telescopes. A number of different models have been proposed for how rotating neutron stars generate radiation; we'll discuss just one leading model. The idea behind this model is that the magnetic axis of a neutron star, the line connecting the north and south magnetic poles, is likely to be inclined at an angle to the rotation axis (Figure 21-3*a*). After



(a) One of the beams from the rotating neutron star is aimed toward Earth: we detect a pulse of radiation.

(b) Half a rotation later, neither beam is aimed toward Earth: we detect that the radiation is "off."

Figure 21-3

A Rotating, Magnetized Neutron Star Charged particles are accelerated near a magnetized neutron star's magnetic poles (labeled N and S), producing two oppositely directed beams of radiation. If the star's magnetic axis (a line that connects the north and south magnetic poles) is tilted at an angle from the axis of rotation, as shown here, the beams sweep around the sky as the star rotates. If the Earth happens to lie in the path of one of the beams, we detect radiation that appears to pulse **(a)** on and **(b)** off.

all, there is no fundamental reason for these two axes to coincide. (Indeed, these two axes do not coincide for any of the planets of our solar system.)

In 1969, Peter Goldreich at Caltech pointed out that the combination of such a powerful magnetic field and rapid rotation would act like a giant electric generator, creating very strong electric fields near the neutron star's surface. Indeed, these fields are so intense that part of their energy is used to create electrons and positrons out of nothingness in a process called **pair production.** The powerful electric fields push these charged particles out from the neutron star's surface and into its curved magnetic field, as sketched in Figure 21-3*a*. As the particles spiral along the curved field, they are accelerated and emit energy in the form of electromagnetic radiation. (A radio transmission antenna works in a similar way. Electrons are accelerated back and forth along the length of the antenna, producing radio waves.) The result is that two narrow beams of radiation pour out of the neutron star's north and south magnetic polar regions.

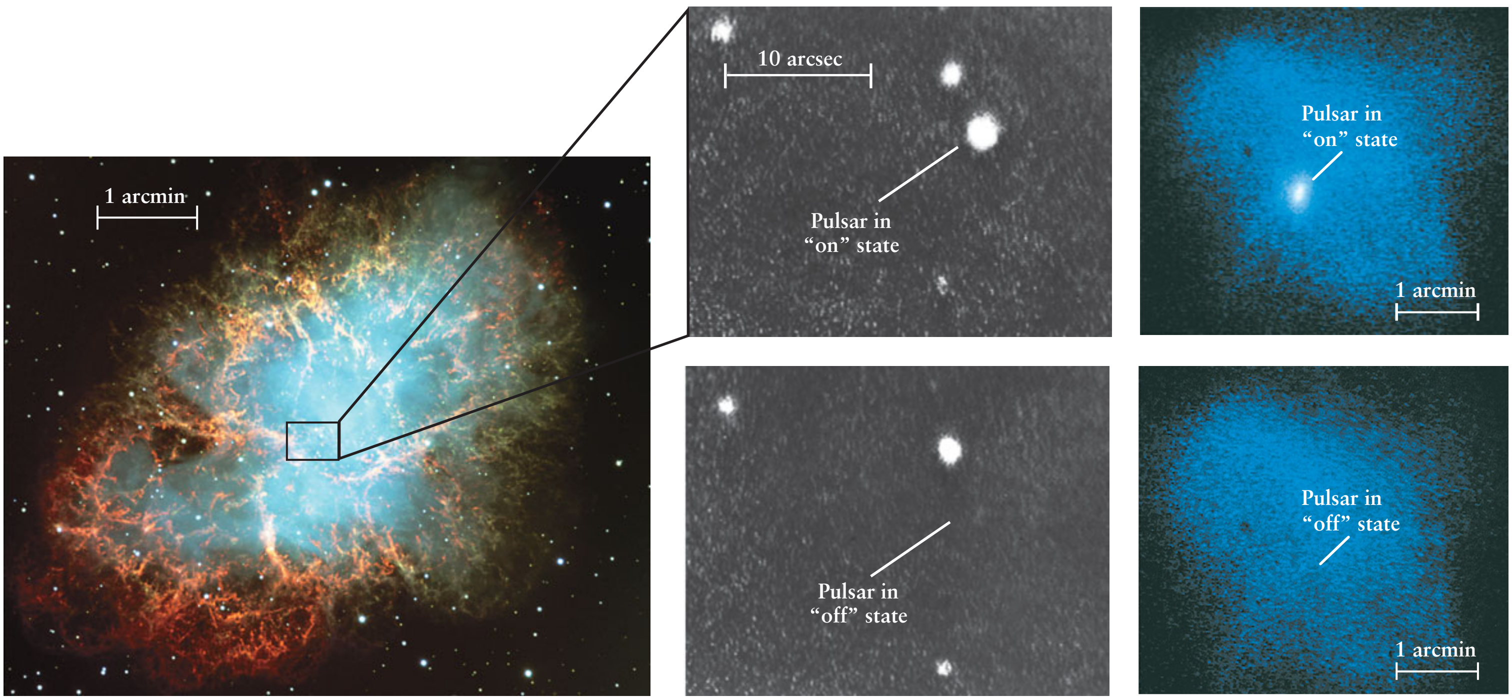
ANALOGY A rotating, magnetized neutron star is somewhat like a lighthouse beacon. As the star rotates, the beams of radiation sweep around the sky. If at some point during the rotation one of those beams happens to point toward the Earth, as shown in Figure 21-3*a,* we will detect a brief flash as the beam sweeps over us. At other points during the rotation the beam will be pointed away from Earth and the radiation from the neutron star will appear to have turned off (Figure 21-3*b*). Hence, a radio telescope will detect regular pulses of radiation, with one pulse being received for each rotation of the neutron star.

CAUTION! The name *pulsar* may lead you to think that the source of radio waves is actually pulsing. But in the model just described, this is not the case at all. Instead, beams of radiation are emitted continuously from the magnetic poles of the neutron star. The pulsing that astronomers detect here on the Earth is simply a result of the rapid rotation of the neutron star, which brings one of the beams periodically into our line of sight, as Figure 21-3 shows. In this sense, the analogy between a pulsar and a lighthouse beacon is a very close one.

The Crab Pulsar

WEB LINK 21.5

Soon after the rotating neutron star model of pulsars was proposed in the late 1960s, a team of astronomers at the University of Arizona began wondering if pulsars might also emit pulses at wavelengths other than the radio part of the spectrum. To investigate, they began a search for *visible-light* pulses from the Crab pulsar (Figure 21-4). They aimed a telescope at the center of the Crab Nebula (Figure 21-4*a*), then used a spinning disk with a slit in it to "chop," or interrupt,



(a) The Crab Nebula R I V U X G

(b) The Crab pulsar in visible light R I V U X G

(c) The Crab pulsar in X rays R I V U X G

VIDEO 21.1

Figure 21-4

The Crab Pulsar **(a)** A pulsar is located at the center of the Crab Nebula, which is about 2000 pc (6500 ly) from Earth and about 3 pc (10 ly) across. The boxed area is shown in closeup in part (b). **(b)** We see a flash of visible light when the rotating pulsar's beam is directed toward us (the "on" state). The pulsar fades (the "off" state) when the beam is aimed elsewhere. **(c)** The Crab pulsar also pulses "on" and "off" at X-ray wavelengths. The radio pulses, visual flashes, and X-ray flashes all have a period of 0.033 s. (a: The FORS Team, VLT, European Southern Observatory; b: Lick Observatory; c: Einstein Observatory, Harvard-Smithsonian Center for Astrophysics) incoming light at rapid intervals. To their surprise and delight, they found that one of the stars at the center of the nebula appears to be flashing on and off 30 times each second (Figure 21-4*b*)—exactly the same pulse rate as observed with radio waves.

This observation was strong evidence in favor of pulsars being rotating, magnetized neutron stars, because charged particles accelerating in a magnetic field can radiate strongly over a very wide range of wavelengths. By contrast, the blackbody radiation from an ordinary star, which has a surface temperature in thousands of degrees, is very weak at radio wavelengths. (The Sun's weak radio emissions come not from its glowing surface but rather from charged particles spiraling in the solar magnetic field.)

Since the pioneering days of the 1960s, periodic flashes of radiation have been detected from the Crab pulsar at still other wavelengths, including X rays (Figure 21-4*c*). The pulsar period is the same at all wavelengths, just as we would expect if the emissions are coming from a portion of the neutron star, which comes into view periodically as the star rotates.

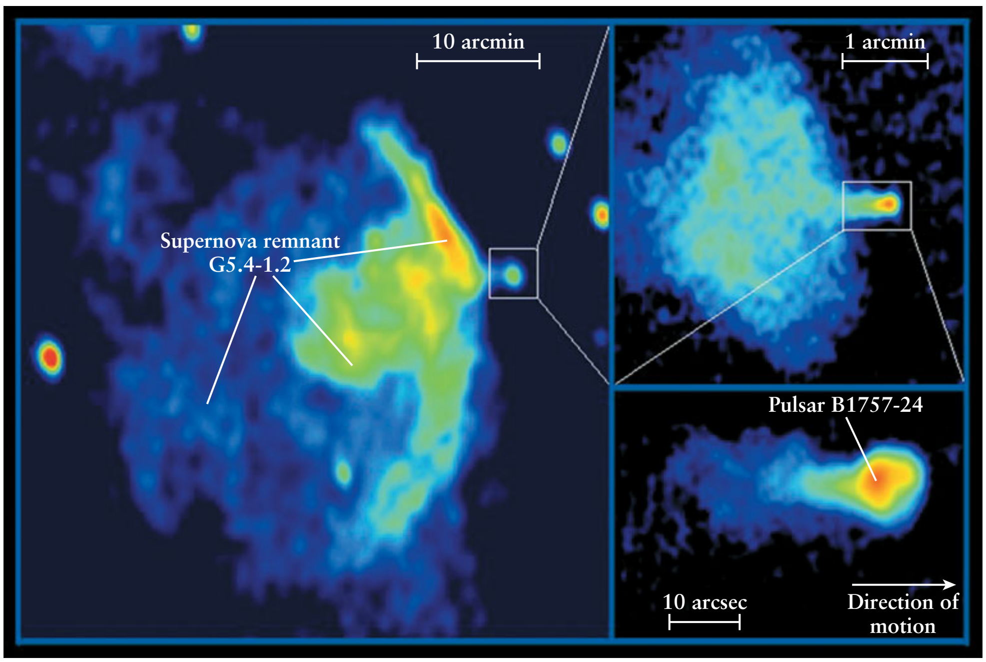
WEB LINK 21.6

The many successes of the neutron-star model of pulsars led to the rapid acceptance of the model. Since 1968 radio astronomers have discovered more than 1600 pulsars scattered across the sky, and it is estimated that perhaps 10^5 more are strewn around the disk of the Milky Way Galaxy. Each one is presumed to be the neutron star corpse of an extinct, massive star. Radio telescopes have detected pulsars with a wide variety of pulse periods, from a relatively sluggish 8.51 s to a blindingly fast 0.001396 s. In each case the pulse period is thought to be the same as the rotation period of the neutron star.

The Pulsar-Supernova Connection

The primary sources of pulsars are thought to be massive stars that become core-collapse supernovae (as described in Section 20-6). The supernova seen in 1054 that left behind the Crab Nebula and the Crab pulsar is thought to have been a Type II core-collapse supernova. However, not all supernovae leave pulsars amid their debris. For example, the supernovae observed by Tycho Brahe in 1572 and by Johannes Kepler in 1604 were probably thermonuclear supernovae in which a white dwarf in a binary system underwent a catastrophic explosion (see Section 20-9). It is difficult to imagine how the complete disruption of a white dwarf could produce a neutron star. Indeed, no pulsar has been found at the locations of either of those supernovae.

In many cases pulsars are found within the remnant of the supernova that spawned the pulsar. One example is the Crab pulsar, which is not far from the center of the Crab Nebula. This is what we might expect: When the progenitor star of the supernova of 1054 exploded, its core remained at the center of the explosion and became a rotating neutron star, while the star's outer layers were blasted into space to form the nebula. However, Figure 21-5 shows a pulsar that lies *outside* its supernova remnant and is moving away from the remnant at hundreds of kilometers per second. It appears that when the progenitor star collapsed and exploded into a supernova, the explosion was much more powerful on one side of the core than the other. As a result, the core was sent flying at immense speed. Many such fast-moving pulsars have been observed, which suggests that supernova explosions are often asymmetrical. In this way studies of pulsars have given us new insights into the nature of supernovae.



WEB LINK 21.7

Figure 21-5 R I V U X G

A Fast-Moving Pulsar When the progenitor star of this supernova remnant exploded, it ejected the star's core at such a high speed that the core overtook the blast wave and moved beyond the boundary of the supernova remnant. The ejected core became a pulsar, denoted B1757-24. Tens of thousands of years after the explosion, the pulsar is still moving at about 600 km/s (350 mi/s). In these false-color radio images, red denotes areas of strongest radio emission. (National Radio Astronomy Observatory)

The most intensely studied of all recent core-collapse supernovae is the great supernova of 1987 (SN 1987A, described in Section 20-7). Astronomers have been carefully observing its remains for signs of a pulsar. To date, however, no confirmed observations of pulses have been made, and the search for a pulsar in SN 1987A goes on.

21-4 Pulsars gradually slow down as they radiate energy into space

WEB LINK 21.8

As remarkable as the Crab pulsar is, the Crab Nebula that surrounds it is even more so. Although the nebula is 2000 pc (6500 ly) away, it can be spotted in dark skies with even a small telescope. It shines with a luminosity 75,000 times that of the Sun (and some 10^8 times the luminosity of the pulses from the Crab pulsar).

But why is the Crab Nebula so luminous? Certain other nebulae—the H II regions described in Section 18-2—can also be seen at great distances. But H II regions contain young stars that shine because of rapid thermonuclear reactions in their cores. These reactions heat the star's surfaces to such high temperatures that they give off prodigious amounts of ultraviolet radiation. This radiation ionizes the surrounding gas and provides the energy that the H II region emits into space. The Crab Nebula, however, is different from H II regions: It contains *no* young stars, and there are *no* thermonuclear reactions taking place within the neutron star that lies at the nebula's heart. What, then, is the source of this nebula's prodigious energy output? As we will see, the Crab pulsar provides the energy not only for radio pulses but for the entire Crab Nebula.

Rotational Energy as a Power Source

In 1966, two years before the discovery of pulsars, John A. Wheeler at Princeton University and Franco Pacini in Italy speculated that the ultimate source of the Crab Nebula's immense energy output might be the *spin* of a neutron star inside the nebula. In a kitchen blender, the energy of a rapidly spinning blade is used to chop food or mix beverages. Wheeler and Pacini proposed that the even greater energy of a rapidly spinning neutron star—which can rotate as rapidly as the blade in a blender, but has far more energy thanks to its much larger mass and radius—can provide the energy to light up the Crab Nebula.

The energy in a neutron star's rotation provides the power for its intense pulsar radiation

Wheeler and Pacini's speculation was quickly accepted as the correct explanation of the Crab Nebula's luminosity. What confirmed their idea were two discoveries: first, that there is a pulsar within the Crab Nebula, and second, that the Crab pulsar is slowing down.

Although pulsars were first noted for their very regular pulses, careful measurements by radio telescopes soon revealed that the rotation period of a typical pulsar—that is, the time for it to spin once on its axis—increases by a few billionths of a second each day. The Crab pulsar, which is one of the most rapidly rotating pulsars, is also slowing more quickly than most—its period increases by 3.8×10^{-8} second each day. Thirty billionths of a second may sound trivial, but there is so much energy in the rotation of a rapidly spinning neutron star that even the slightest slowdown corresponds to a tremendous loss of energy. In fact, the observed rate of slowing for the Crab pulsar corresponds to a rate of energy loss equal to the entire luminosity of the Crab Nebula.

Synchrotron Radiation

The interpretation is that the energy lost by the Crab pulsar as it slows down is transferred to the surrounding nebula. But *how* does this transfer of energy take place, and how is the energy radiated into space? The answer lies in the diffuse part of the Crab Nebula, which shines with an eerie bluish light (see Figure 21-4*a*). This type of light was first produced on the Earth in 1947 in a particle accelerator in Schenectady, New York. This machine, called a *synchrotron,* was built by the General Electric Company to accelerate electrons to nearly the speed of light for experiments in nuclear physics. (The electrons are said to be *relativistic,* because their velocities are near the speed of light, and Einstein's theory of relativity—which we discuss in Chapter 22—must be applied to understand their motions.)

WEB LINK 21.9

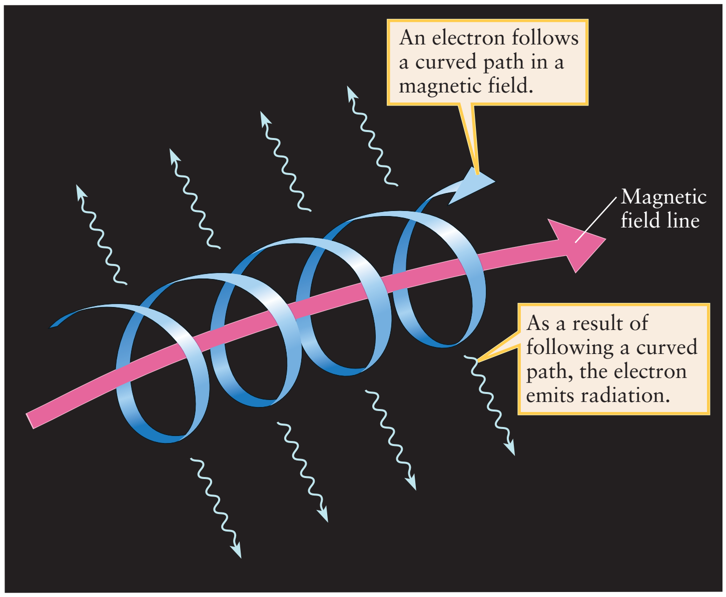
As a beam of electrons whirled along a circular path within the cyclotron, held in orbit by powerful magnets, scientists noticed that it emitted a strange light. It is now known that this light, called **synchrotron radiation,** is emitted whenever high-speed electrons move along curved paths through a magnetic field (Figure 21-6*a*). Synchrotron radiation has a continuous spectrum, but one that is distinctively different from the continuous spectrum emitted by a heated blackbody (Figure 21-6*b*). This difference helped astronomers confirm that the Crab Nebula emits primarily synchrotron radiation, not blackbody radiation.

The total energy output of the Crab Nebula in synchrotron radiation is 3×10^{31} watts, compared to the relatively paltry 4×10^{26} watts emitted by the Sun. Thus, the Crab Nebula must contain quite a large number of relativistic electrons spiraling in an extensive magnetic field. Figure 21-7 shows where these electrons come from: They are ejected from the neutron star's vicinity, forming immense jets that emanate from the star's poles and rings that spread outward from its equator. (Figure 18-16 shows a related process that produces jets from young stars.) As the Crab pulsar slows, the energy of its rotation is transferred to these electrons through the magnetic field and is then emitted by them in the form of synchrotron radiation. So much energy is transferred in this way that the Crab Nebula shines very brightly.

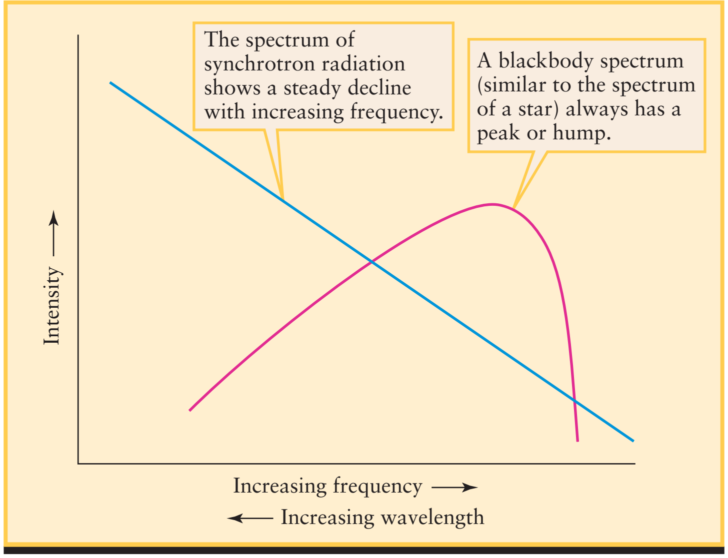
This model for the luminosity of the Crab Nebula also provides a check on our ideas about the magnetic fields of neutron stars (Section 21-3). The stronger the magnetic field of a neutron star, the greater the rate at which it emits energy in the form of radiation and hence the more rapidly it will slow down. From the slowdown rate of the Crab pulsar, astronomers calculate that its magnetic field must be about 5×10^{12} G. This is very close to the estimate of 10^{12} G that we obtained in Section 21-3.

Aging Pulsars

Because a spinning neutron star slows down as it radiates away its rotational energy, it follows that an old neutron star should be spinning more slowly than a young one. It is thought that slower pulsars were formed in supernova explosions that occurred hundreds of thousands of years ago. Over the ages, the supernova



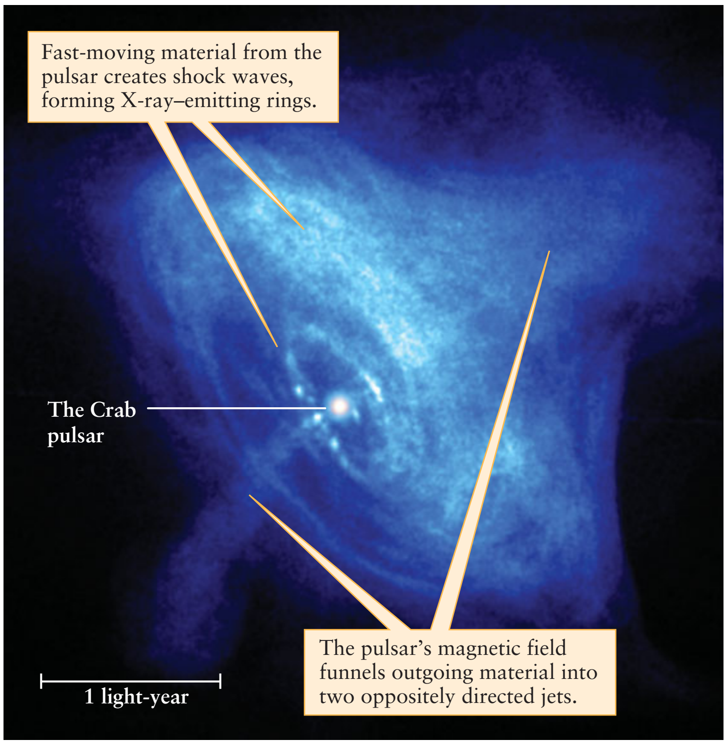
(a)



(b)

Figure 21-6

Synchrotron Radiation **(a)** Synchrotron radiation is emitted by electrons that move at high speed in a magnetic field. **(b)** Astronomers can distinguish synchrotron radiation from blackbody radiation because the two kinds of emission have very different spectra.



VIDEO 21.2 ANIMATION 21.1

Figure 21-7 R I V U X G

The Dynamic Heart of the Crab Nebula This image from the Chandra X-ray Observatory shows material being expelled from the Crab pulsar into the surrounding nebula. (The image that opens this chapter combines this X-ray view with a visible-light view.) Some of the ejected material travels at more than half the speed of light. (NASA/CXC/ASU/J. Hester et al.)

remnants have dispersed into the interstellar medium, and the pulsars left behind have slowed to periods of a second or more. (More than 400 such old pulsars are known.) Only relatively young, rapidly rotating pulsars like the Crab pulsar have enough energy to emit flashes of visible light along with radio pulses (see Figure 21-4*b*). We can summarize these ideas as a general rule:

An isolated pulsar slows down as it ages, so its pulse period increases.

From a pulsar's period (which increases as it ages) and the rate at which its period is increasing (which decreases as it ages), astronomers can estimate how old a pulsar is.

Like all general rules, the one relating a pulsar's age and its period has exceptions. In Section 21-7 we will see how some old pulsars that are *not* isolated—that is, are members of binary star systems—can be reaccelerated to truly dizzying rotation speeds.

21-5 Superfluidity and superconductivity are among the strange properties of neutron stars

The radiation from pulsars gives us information about the surroundings of neutron stars, as well as how they rotate. To deduce what conditions are like *inside* a neutron star, however, astrophysicists must construct theoretical models. We saw examples of such models for the Sun in Section 16-2 and for post–main-sequence stars in Sections 20-1 and 20-5.

Superfluidity and Pulsar Glitches

The models for neutron stars are very different. For one thing, neutron stars are thought to have a solid crust on their surface.

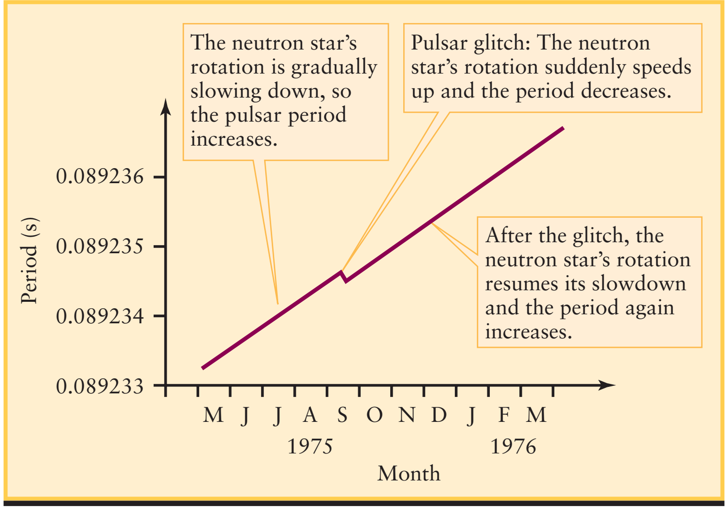


Figure 21-8

A Pulsar Glitch This graph shows how the period of the Vela pulsar varied during 1975 and 1976. Most of the time the neutron star's rotation slowed as its rotational energy was converted into radiation. In September 1975, however, a sudden speedup (glitch) occurred. A number of pulsars have been observed to undergo multiple glitches at intervals of a few years.

Furthermore, the interior of a neutron star is a sea of densely packed, degenerate neutrons, with properties quite unlike those of ordinary gases or even degenerate electrons. Detailed calculations strongly suggest that degenerate neutron matter can flow without any friction whatsoever, a phenomenon called **superfluidity.**

Material deep inside a neutron star may flow completely without friction

Superfluidity can be observed in the laboratory by cooling liquid helium to temperatures near absolute zero. Because it is frictionless, superfluid liquid helium exhibits strange properties such as being able to creep up the walls of a container in apparent defiance of gravity. Within a neutron star, elongated, friction-free whirlpools of superfluid neutrons may form. Their interaction with the star's crust may be the cause of sudden changes in a pulsar's rotation.

In addition to the general slowing down of pulsars, astronomers have found that they sometimes exhibit a sudden speedup, sometimes called a **glitch.** For example, Figure 21-8 shows measurements of the period of the Vela pulsar (named for the constellation in which it lies) during 1975 and 1976. On this graph, the pulsar's gradual slowdown is shown as a steady increase in its period. In September 1975, however, there was an abrupt speedup, after which the pulsar continued to slow down at its usual rate. Similar glitches are seen for other pulsars.

Current opinion among pulsar theorists is that glitches are caused by the superfluid neutrons within the neutron star. As a rotating neutron star radiates energy into space, the rotation of its crust slows down, but the neutron whirlpools in the star's interior continue to rotate with the same speed. Some of these whirlpools cling to the crust, as though they were bungee cords with one end attached to the crust and the other to the star's interior. As the crust slows down relative to the interior, these superfluid "bungee cords" stretch out. When the tension in them gets too great, they deliver a sharp jolt that makes the crust speed up suddenly. (An older model, in which pulsar glitches were caused by a "starquake" in the neutron star's settling crust, does not appear to agree with the accumulated observational data on glitches.)

The Ultimate Electric Conductor

Superfluidity is not the only exotic property of neutron star interiors. Models of the internal structure of a neutron star strongly suggest that the protons in the core can move around without experiencing any electrical resistance whatsoever. This phenomenon, called **superconductivity,** also occurs on the Earth with certain substances at low temperatures. In a normal conductor of electricity such as the copper wires in a flashlight, an electric current only flows if there is a battery in the flashlight circuit to overcome the resistance of the wires. Once the battery goes "dead," the current stops and the flashlight turns off. Once you start an electric current moving in a superconducting material, however, it keeps moving forever.

CAUTION! You may have been surprised to read in the preceding paragraph about *protons* in a *neutron* star. While a neutron star is made up predominantly of neutrons, some protons and electrons must be scattered throughout the star's interior. Indeed, a pulsar's magnetic field must be anchored to the neutron star by charged particles. Neutrons are electrically neutral, so without the protons and electrons in its interior a neutron star would rapidly lose its magnetic field.

As Figure 21-9 shows, the structure of a neutron star probably consists of a core with superfluid neutrons and superconducting

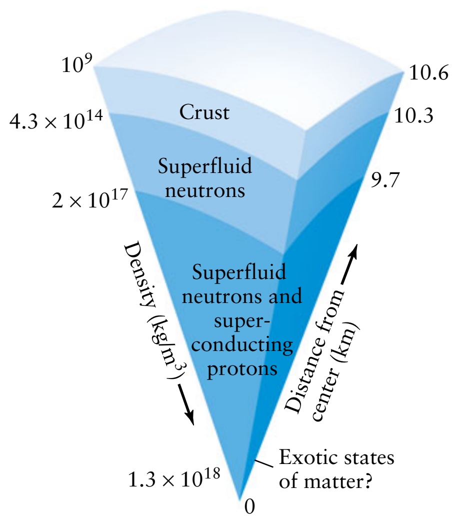
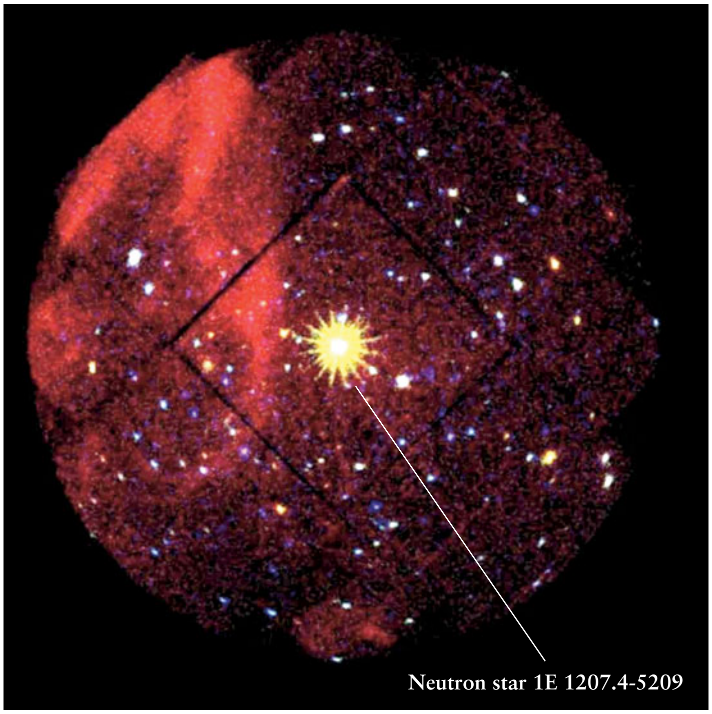
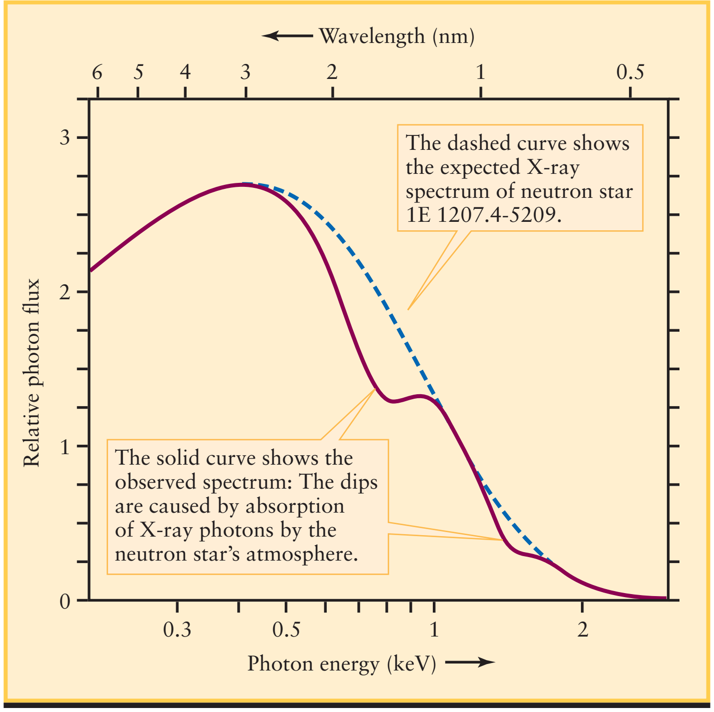


Figure 21-9

A Model of a Neutron Star This theoretical model for a 1.4-$M_{\odot}$ neutron star has a superconducting, superfluid core 9.7 km in radius. The core is surrounded by a 0.6-km thick mantle containing nuclei, electrons, and superfluid neutrons. The star's crust is only 0.3 km thick (the length of three football fields) and is composed of heavy nuclei (such as iron) and free electrons. (The thicknesses of the layers are not shown to scale.)



(a) X-ray image R I V U X G



(b) X-ray spectrum

Figure 21-10

Evidence for a Neutron-Star Atmosphere **(a)** This image from XMM-Newton (see Figure 6-30*b*) shows a neutron star roughly 2000 pc (7000 ly) from Earth in the constellation Centaurus. Like the Crab pulsar (Figure 21-4*a*), this neutron star is an X-ray-emitting pulsar. **(b)** The dashed curve shows the X-ray spectrum that would be expected if the neutron star had no atmosphere. The observed dips are evidence that an atmosphere is present. (a: ESA/CESR/G. Bignami et al.)

protons, a mantle of superfluid neutrons in which some of the neutrons combine with protons to form nuclei, and a brittle crust less than a kilometer thick. At the center of a neutron star, the density is on the order of 10^{18} kg/m^3. This is some 10 trillion (10^{13}) times greater than the central density of our present-day Sun and approximately twice the density of an atomic nucleus. It has been speculated that under these conditions the neutrons and protons dissolve into more fundamental particles called *quarks*. To date, however, astronomers have not found conclusive observational evidence of such extreme states of matter within neutron stars.

Neutron Star Atmospheres

While the deep interiors of neutron stars remain mysterious, astronomers have recently found evidence that the crusts of neutron stars are surrounded by a highly unusual atmosphere (Figure 21-10). The evidence comes from the neutron star 1E 1207.4-5209, which like the Crab pulsar emits X rays and lies at the center of a supernova remnant (Figure 21-10*a*). The dashed curve in Figure 21-10*b* shows the expected spectrum of the emitted X rays. The actual spectrum, shown by the solid curve in Figure 21-10*b*, has two pronounced dips. The interpretation of these dips is that they are caused by absorption of certain X-ray wavelengths by an atmosphere around the neutron star. (Compare Figure 21-10*b* with Figure 7-3.) The atoms that make up such an atmosphere must be dramatically different from those around you: In the intense magnetic field of a young neutron star, ordinarily spherical atoms are deformed into elongated cigar shapes and have radically altered spectral lines.

The bizarre properties of neutron star interiors and atmospheres are of great interest to astronomers and physicists precisely because they are so bizarre. Just as physicians learn about health by studying disease, scientists gain insight into the fundamental nature of matter by studying its behavior under extreme conditions.

21-6 The most highly magnetized pulsars can emit colossal bursts of radiation

We learned in Section 6-7 that a number of spacecraft observe the universe at X-ray and gamma-ray wavelengths. Most of the objects that these spacecraft examine are rather faint, so the detectors that these spacecraft carry are quite sensitive. But on December 27, 2004, the detectors on seven such spacecraft were overloaded by an intense burst of radiation that lasted a mere 0.2 second.

By comparing the times at which spacecraft at different locations in space detected the burst, astronomers were able to triangulate the point in space from which the radiation emanated. The source proved to be a pulsar designated SGR 1806-20 that lies about 15,000 parsecs (50,000 light-years) from Earth in the constellation Sagittarius. For the spacecraft to detect such intense radiation from a source so far away, the pulsar's luminosity during the time it was emitting the burst must have been a staggering 10^{14} times greater than the luminosity of the Sun. It would take the Sun 300,000 years to radiate as much energy as SGR 1806-20 emitted in just 0.2 second!

How could a pulsar release such a stupendous amount of energy in such a brief amount of time? If the energy source were the rotation of a neutron star, as it is for the Crab Nebula, the neutron star's spin would have slowed appreciably after emitting the radiation burst. But observations of the pulsar made before and after the burst showed the *same* pulse period of 7.47 seconds. Since the pulse period and rotation rate are the same (see Figure 21-3), this means that the rotation did not slow at all. What is more, the total rotational energy of a neutron star spinning once every 7.47 seconds is only about 1% of the energy that was released as radiation in the 0.2-second burst. Hence, the neutron star's rotation could not have supplied the immense release of energy seen on December 27, 2004.

Short bursts of radiation that were nearly as intense as that seen from SGR 1806-20 have been detected from two other pulsars, one (SGR 0526-66) in 1979 and one (SGR 1900+14) in 1998. Like SGR 1806-20, neither of these pulsars showed any appreciable slowing after emitting their bursts, so their rotational energy could not have been the source of the bursts. To explain all three of these exotic events, astronomers have turned to another distinguishing characteristic of a pulsar: the neutron star's intense magnetic field. As we will see, the three pulsars in question appear to have the strongest magnetic fields of any object in the universe, far stronger than those of "ordinary" pulsars.

Magnetars and Starquakes

The magnetic field of an ordinary star such as the Sun is produced by the motion of its electrically conducting gas, or plasma (see Section 16-9). The same is true for a supergiant star. We saw in Section 21-3 that when thermonuclear reactions cease in a supergiant star and it begins to collapse, the star's magnetic field moves with the infalling material and becomes greatly amplified. This amplified magnetic field remains with the resulting neutron star, which typically ends up with a field strength of about 10^{12} G.

However, a neutron star can generate an additional magnetic field as it forms. For about the first 10 seconds after a neutron star forms from the core of a collapsing supergiant, it has an extremely high internal temperature—about 10^{11} K, as compared to a relatively frigid 1.55×10^7 K in the present-day Sun's core—and there is furious convection of the neutron star material. This material includes electrically charged protons and electrons, and the convective motions produce a magnetic field. If the newly formed neutron star spins relatively slowly, so that the time it takes a convection cell to rise and fall (about 0.005 s) is less than the time required for the neutron star to make a complete rotation, the magnetic fields caused by different convection cells remain uncorrelated and do not contribute much to the neutron star's magnetism. This seems to be what happens for the vast majority of neutron stars. But if a brand-new neutron star has a rotation period of less than 0.005 s, its rapid rotation causes the convective motions in different parts of the neutron star to organize together. When the neutron star has completely formed, it is left with a magnetic field that can be as strong as 10^{15} G, or about 1000 times stronger than in an ordinary neutron star. The magnetism of such a star could pull the keys out of your pocket and demagnetize your credit cards at a distance of 200,000 km, or half the distance from the Earth to the Moon.

Astronomers Robert Duncan and Christopher Thompson developed the theory of highly magnetized neutron stars, or **magnetars,** in 1992. In their theory the magnetic field of a magnetar exerts tremendous stresses on the electrically conducting material at the star's surface (see Figure 21-9). From time to time these stresses can cause a localized fracture of the surface called a **starquake.** The magnetic field in the vicinity of the starquake responds by rearranging itself in a process that resembles magnetic reconnection in the Sun's magnetic field (see Figure 16-25) but is much more energetic. The energy released by such a localized starquake produces a moderate burst of gamma rays. But on rare occasions the crust can fracture simultaneously over much of the magnetar's surface. In such a catastrophic event the star's magnetic field undergoes a dramatic rearrangement that releases an immense amount of magnetic energy. The result is a short-lived but titanic blast of gamma rays and X rays.

A starquake on the surface of a magnetar releases magnetic energy in the form of an intense blast of X rays and gamma rays

Observing Magnetars

The theoretical model of a magnetar turns out to be an excellent description of the three pulsars that produced the energetic blasts observed in 1979, 1998, and 2004. Each of these pulsars emits moderate gamma-ray bursts at seemingly random intervals, which is why they are also called *soft gamma repeaters*. This is the origin of the acronym "SGR" in the pulsars' designations. (The adjective "soft" means that the gamma rays emitted are of relatively low energy.) According to Duncan and Thompson's model, each such burst is caused by a localized starquake on a magnetar; the truly stupendous bursts are the result of rare quakes that span much of the magnetar's surface.

How can we test these radical ideas? As we saw in Section 21-4, the rate at which a pulsar slows down is a measure of its magnetic field. In 1998 the Greek astronomer Chryssa Kouveliotou determined that over the preceding 5 years the pulse period of SGR 1806-20 had increased by 0.008 seconds. (This is a small change, but is about 40,000 times greater than the increase in the Crab pulsar's period over the same time interval.) Calculations show that the neutron star magnetic field needed to produce this rate of slowdown is 8×10^{14} G, in close agreement with the magnetar model.

Although magnetars begin their existence with dizzyingly fast rotation, their rapid slowdown soon makes them among the

COSMIC CONNECTIONS

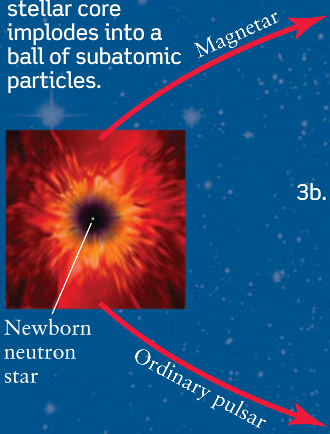
Magnetars and Ordinary Pulsars

A magnetar is a neutron star with an extraordinarily strong magnetic field of 10^{15} gauss, a thousand times greater than the 10^{12}-gauss field of an ordinary neutron star. The magnetic energy stored in this intense field can be released in the form of powerful bursts of X rays and gamma rays. (After C. Kouveliotou, R.C. Duncan, and C. Thompson, "Magnetars", *Scientific American*, February 2003)

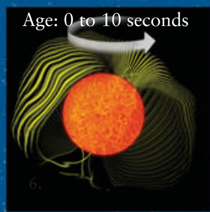
How magnetars form

1. Most neutron stars are thought to begin as massive but otherwise ordinary stars, between 8 and 20 times as massive as the Sun.

2. Massive stars die in a core-collapse supernova explosion, as the stellar core implodes into a ball of subatomic particles.

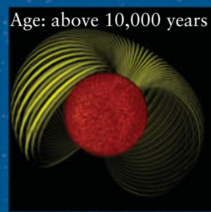


3a. If the newborn neutron star spins fast enough, it generates an intense magnetic field. Field lines inside the star get twisted.

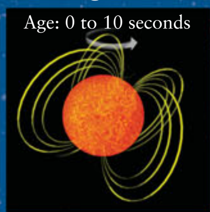


4a. The magnetar settles into neat layers, with twisted field lines inside and smooth lines outside. It might emit a narrow radio beam.

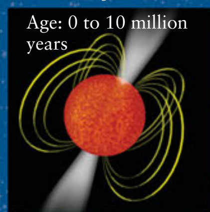
5a. The old magnetar has cooled off, and much of its magnetism has decayed away. It emits very little energy.



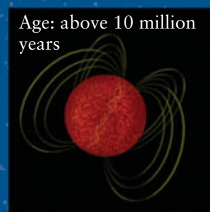
3b. If the newborn neutron star spins slowly, its magnetic field, though strong by everyday standards, does not reach magnetar levels.



4b. The mature pulsar is cooler than a magnetar of equal age. It emits a broad radio beam, which radio telescopes can readily detect.



5b. The old pulsar has cooled off and no longer emits a radio beam.

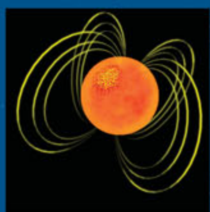


How magnetar bursts happen

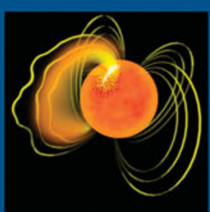
The magentic field of a magnetar is so strong that the rigid crust sometimes breaks and crumbles, releasing a huge surge of energy.

1. Most of the time the magnetar is quiet. But magnetic stresses are slowly building up.

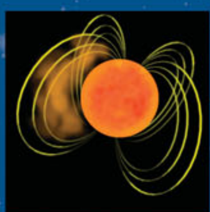
2. At some point the solid crust is stressed beyond its limit. It fractures, probably into many small pieces.



3. This “starquake” creates a surging electric current, which decays and leaves behind a hot fireball.



4. The fireball cools by releasing X rays from its surface. It evaporates in minutes or less.



(NASA)

- A burst from a magnetar 15,000 parsecs away was observed on December 27, 2004.
- This burst was so intense that it caused a major disturbance of the Earth’s upper atmosphere.
- During the 0.2-second duration of the burst, the magnetar’s luminosity was more than 100 times the total luminosity of the Milky Way Galaxy.
- It is estimated that this blast ripped away a layer of the magnetar’s crust about 50 meters thick.

slowest of all pulsars. For example, in the few thousand years since SGR 1806-20 first formed from the core collapse of a supergiant star, its rotation period has increased from less than 0.005 s to 7.47 s. With the passage of time, a magnetar's rotation will continue to slow down and starquakes will eventually relieve the crust of all of its stresses. After perhaps 10,000 years, a magnetar will cease its activity altogether. It is estimated that there may be millions of these old, inert magnetars in our Galaxy alone.

As yet, only a handful of active magnetars are known. But because magnetars represent such an extreme state of matter, with magnetic effects stronger than those found anywhere else in the universe, astronomers are studying them with great interest. The *Cosmic Connections* figure depicts the key differences between magnetars and ordinary neutron stars.

Even neutron stars that lack the intense magnetic field of a magnetar can display exotic behavior. As we will see in the remainder of this chapter, this happens when a neutron star is one member of a binary star system in which the two stars orbit close to each other.

21-7 The fastest pulsars were probably created by mass transfer in close binary systems

In 1982, radio astronomers discovered a pulsar whose period is only 1.558 ms. (One millisecond, or 1 ms, equals a thousandth of a second). This remarkable pulsar, called PSR 1937+21 after its coordinates in the sky, is a neutron star spinning 642 times per second—about 3 times faster than the blades of a kitchen blender! Such incredibly rapid rotation is a clue to this pulsar's unusual history.

The fastest pulsars spin nearly a thousand times per second, more than twice as fast as a kitchen blender

As we have seen, pulsars should slow down as they age. The rapid spin rate of PSR 1937+21 suggests that it has hardly aged at all and so must be very young. A young pulsar should also be slowing down very rapidly. The faster it rotates, the more rapidly it can transfer rotational energy to its surroundings and the faster the slowdown. But PSR 1937+21 is actually slowing down at a very *gradual* rate, millions of times more gradually than the Crab pulsar. Such a gradual slowdown is characteristic of a pulsar hundreds of millions of years old. But if PSR 1937+21 is so old, how can it be spinning so rapidly?

PSR 1937+21 is not the only pulsar with unusually rapid rotation. Since 1982, astronomers have discovered more than 180 very fast pulsars, which are now called **millisecond pulsars.** All have periods between 1 and 10 ms, which means that these neutron stars are spinning at rates of 100 to 1000 rotations per second. (As of this writing the fastest known pulsar is PSR J1748-2446ad, which rotates 716 times a second.)

The Origin of Millisecond Pulsars

An important clue to understanding millisecond pulsars is that the majority are in close binary systems, with only a small separation between the spinning neutron star and its companion. (We know the separation must be small because the orbital periods of these binary systems are short, between 10 and 100 days. Kepler's third law for binary star systems, which we described in Section 17-9, tells us that the shorter the period, the smaller the distance between the two stars.) This fact suggests a scenario for how millisecond pulsars acquired their very rapid rotation.

Imagine a binary system consisting of a high-mass star and a low-mass star. The high-mass star evolves more rapidly than the low-mass star, and within a few million years becomes a Type II supernova that creates a neutron star. Like most newborn neutron stars, it spins several times per second, and, thus, initially we see a pulsar rather like the Crab or Vela pulsar.

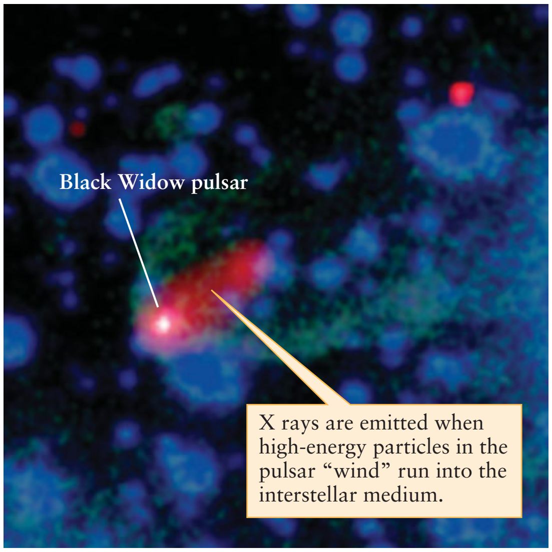
ANIMATION 21.3

Over the next few billion years, the pulsar slows down as it radiates energy into space. Meanwhile, the slowly evolving low-mass star begins to expand as it evolves away from the main sequence to become a red giant. When the red giant gets big enough to fill its Roche lobe, it starts to spill gas over the inner Lagrangian point onto the neutron star. (See Section 19-7 for a review of close binary systems.) The infalling gas strikes the neutron star's surface at high speed and at an angle that causes the star to spin faster. In this way, a slow, aging pulsar is "spun up" by mass transfer from its bloated companion.

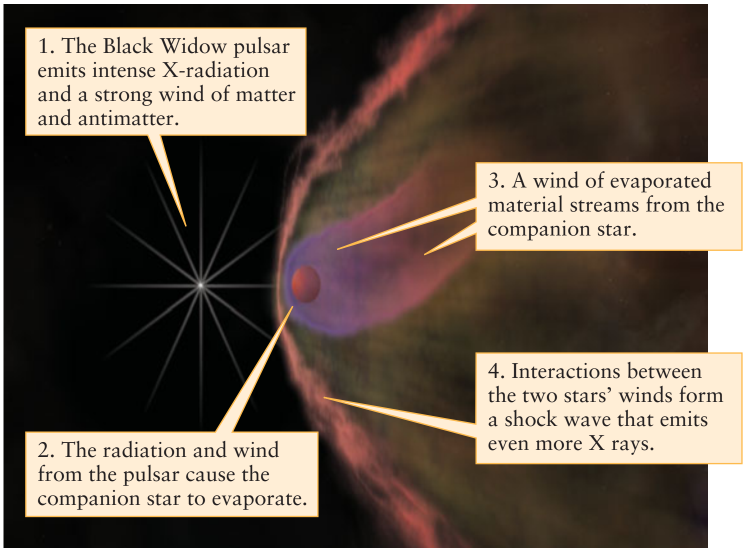
What about the few millisecond pulsars, like PSR 1937+21, that are *not* members of close binaries? It may be that solitary millisecond pulsars were once part of close binary systems, but the companion stars have been eroded away by the high-energy particles emitted by the pulsar after it was spun up. The Black Widow pulsar—named for the species of spider whose females kill and eat their mates—may be caught in the act of destroying its companion in just such a process (Figure 21-11).

The Black Widow pulsar (a neutron star spinning 622 times per second) and its companion orbit each other with a period of 9.16 hours. The orbital plane of the system is nearly edge-on to our line of sight, so this system is an eclipsing binary (described in Section 19-11): The companion blocks out the pulsar's radio signals for 45 minutes during each eclipse. Just before and after an eclipse, however, the pulsar's signals are delayed substantially, as if the radio waves were being slowed as they pass through a cloud of ionized gas surrounding the companion star. This circumstellar material is probably the star's outer layers, dislodged by radiation and by fast-moving particles emitted from the pulsar. In a few hundred million years or so, the companion of the pulsar will have completely disintegrated, leaving behind only a spun-up, solitary millisecond pulsar.

It is also possible that both stars in a close binary system might evolve into neutron stars, forming a pair of pulsars. Several such *binary pulsars* are known. Because the two pulsars are so massive and move around each other in tight orbits, they lose energy through a process called *gravitational radiation.* (We will discuss this process, which is a prediction of Einstein's general theory of relativity, in more detail in Chapter 22.) As a result, the two neutron stars spiral toward each other and eventually collide. Computer simulations of such a collision show that some of the neutrons are ejected into space, where many of them decay into protons and combine to form various heavy elements. In addition to the heavy elements produced directly by supernova collisions,



(a) The Black Widow pulsar R I V U X G



(b) An illustration of the pulsar and its companion

Figure 21-11

The Black Widow Pulsar (a) This false-color composite of optical and X-ray images shows pulsar B1957+20, called the Black Widow. As the pulsar moves through space at over 200 km/s, it emits a wind that streams behind it. **(b)** This artist's conception of the pulsar and its doomed companion shows an area far smaller than the white dot in (a) labeled "Black Widow pulsar." (a: NASA/CXC/ASTRON/B. Stappers et al. (X-ray), AAO/J. Bland-Hawthorn & H. Jones (visible); b: CXC/M. Weiss)

the elements produced by such neutron star collisions may play an important role in seeding the interstellar medium with the building blocks of planets (see Section 20-6).

21-8 Neutron stars in close binary systems can also be pulsating X-ray sources

Millisecond pulsars are not the only result of having a neutron star in a close binary system with an ordinary star. Even more exotic are **pulsating X-ray sources,** in which material from the companion star is drawn onto the magnetic poles of the neutron star, producing intense hot spots that emit breathtaking amounts of X rays. As the neutron star spins and the X-ray beams are directed alternately toward us and away from us, we see the X rays flash on and off. Such pulsating X-ray sources, while similar in many ways to ordinary pulsars, are far more luminous.

Pulsating X-ray Sources and Binary Systems

In an X-ray binary pulsar, gas from a companion star is funneled to the magnetic poles of the neutron star, forming hot spots that emit X rays rather than radio waves

Pulsating X-ray sources were discovered in 1971 by the *Uhuru* spacecraft, the first to give astronomers a comprehensive look at the X-ray sky. (As we saw in Section 6-7, the Earth's atmosphere is opaque to X rays.) The first X-ray source found to emit pulses was Centaurus X-3, the third X-ray source found in the southern constellation Centaurus (Figure 21-12). The X-ray pulses have a regular period of 4.84 s. A few months later, similar pulses were discovered coming from a source designated Hercules X-1, which had a period of 1.24 s. Because the periods of these two X-ray sources are so short, astronomers began to suspect that they were actually observing rapidly rotating neutron stars.

It soon became clear that Centaurus X-3 and Hercules X-1 are members of binary star systems. One piece of evidence is that every 2.087 days, Centaurus X-3 appears to turn off for almost 12 hours. Apparently, Centaurus X-3 is a neutron star in an eclipsing binary system, and it takes nearly 12 hours for the X-ray source to pass behind its companion star. Hercules X-1, too, appears to turn off periodically, corresponding to a 6-hour eclipse every 1.7 days by its visible companion, the star HZ Herculis. Moreover, careful timing of its X-ray pulses shows a periodic Doppler shifting every 1.7 days, which is direct evidence of orbital motion about a companion star. When the X-ray source is approaching us, the pulses from Hercules X-1 are separated by slightly less than 1.24 seconds. When the source is receding from us, slightly more than 1.24 seconds elapse between the pulses.

Putting all the pieces together, astronomers now realize that pulsating X-ray sources like Centaurus X-3 and Hercules X-1 are binary systems in which one of the stars is a neutron star. All these binary systems have very short orbital periods, which means that the distance between the two stars is quite small. The neutron star can therefore capture gases escaping from its ordinary companion (Figure 21-13). More than 20 such exotic binary systems, also called *X-ray binary pulsars,* are known. A typical rate

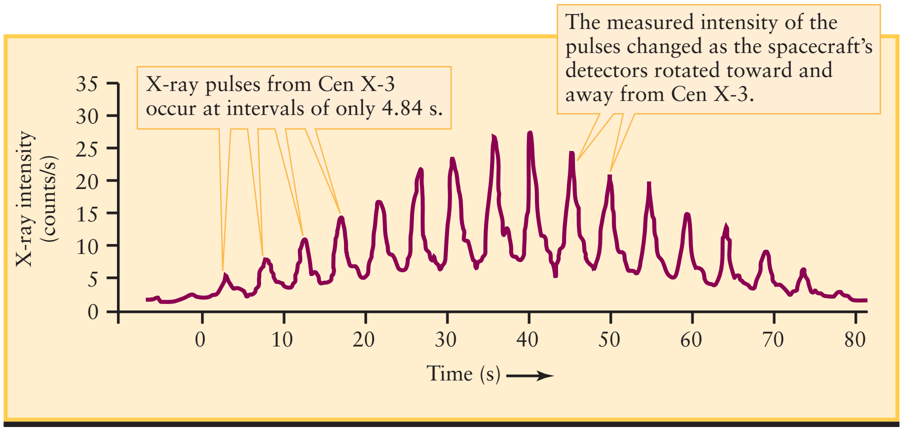


Figure 21-12

X-Ray Pulses from Centaurus X-3 This graph shows the intensity of X rays detected by the *Uhuru* spacecraft as Centaurus X-3 (often abbreviated as Cen X-3) moved across the satellite's field of view. Because the pulse period is so short, Centaurus X-3 is in all probability a rotating neutron star. (Adapted from R. Giacconi and colleagues)

of mass transfer from the ordinary star to the neutron star is roughly 10^{-9} solar masses per year.

Because of its strong gravity, a neutron star in a pulsating X-ray source easily captures much of the gas escaping from its companion. But like an ordinary pulsar, the neutron star is rotating rapidly and has a powerful magnetic field inclined to the axis of rotation (recall Figure 21-3). As the gas falls toward the neutron star, its magnetic field funnels the incoming matter down onto the star's north and south magnetic polar regions. The neutron star's gravity is so strong that the gas is traveling at nearly half the speed of light by the time it crashes onto the star's surface. This violent impact creates spots at both poles with searing temperatures of about 10^8 K. Material raised to these temperatures emits abundant X rays, and because so much material is dumped onto these hot spots the X-ray luminosity reaches roughly 10^{31} watts. This is much greater than the X-ray luminosity of isolated pulsars like the Crab (see Figure 21-4*b*) or 1E 1207.4-5209 (see Figure 21-10*a*), whose surface temperatures are a relatively chilly 10^6 K or so. It is also nearly 10^5 times greater than the combined luminosity of the Sun at all wavelengths.

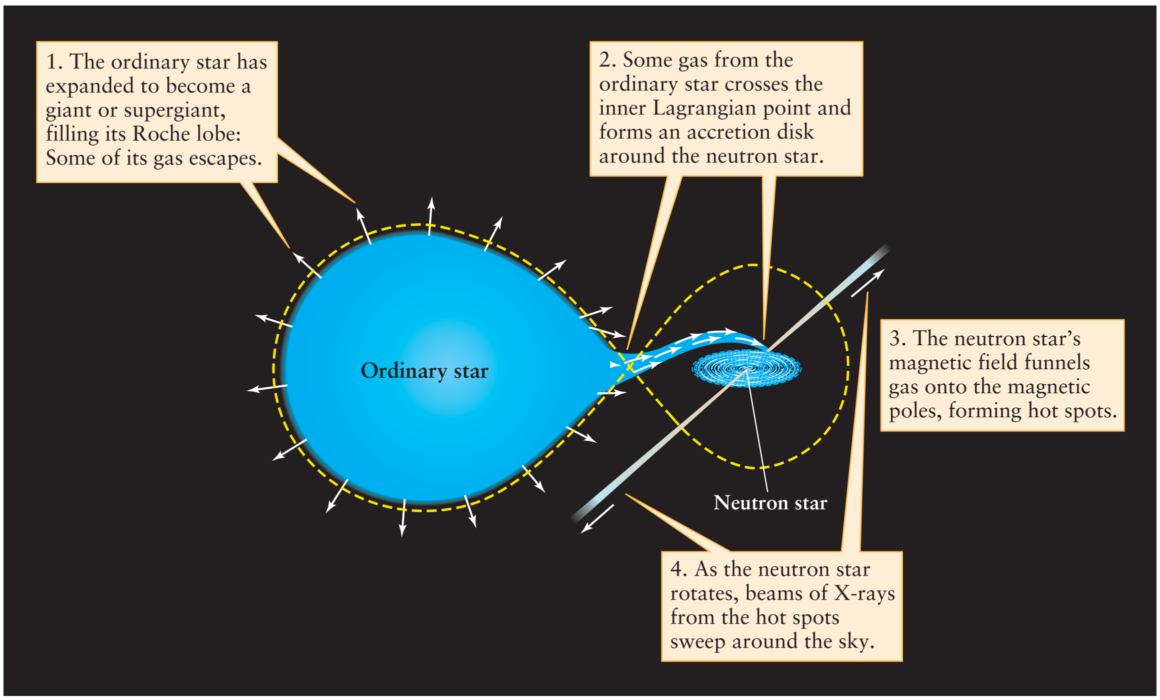


Figure 21-13

A Model of a Pulsating X-Ray Source Pulsating X-ray sources are close, semidetached binary systems (compare with Figure 19-20*b* and Figure 19-21*b*) in which one member is a rotating neutron star. Gas from the ordinary star flows onto the neutron star and creates hot spots that emit beams of X rays. If one of the beams is oriented so that it sweeps over the Earth as the neutron star rotates, we see a pulse of X rays on each sweep.

As the neutron star rotates, the beams of X rays from its magnetic poles sweep around the sky. If the Earth happens to be in the path of one of the two beams, we can observe a pulsating X-ray source. We detect a pulse once per rotation period, which means that the neutron star in Hercules X-1 spins at the rate of once every 1.24 seconds. If the orbital plane of the binary system is nearly edge-on to our line of sight, as for Centaurus X-3 and Hercules X-1, the pulsating X rays appear to turn off when the neutron star is eclipsed by the companion star. Other pulsating X-ray sources, such as Scorpius X-1, do not undergo this sort of turning off. The orbital planes of these systems are oriented more nearly face-on to our line of sight, so no eclipses occur.

The neutron stars in most pulsating X-ray sources have rotation periods (and hence pulse periods) of a few seconds. But as time passes, the neutron star accretes more and more mass from its companion. In the process, it will spin up and may eventually become a millisecond pulsar (Section 21-7). In 1998 the *Rossi X-ray Timing Explorer* spacecraft first observed an evolutionary "missing link" between pulsating X-ray sources and millisecond pulsars. This object, designated SAX J1808.4-3658, is an X-ray–emitting millisecond pulsar in a close binary system with a pulse period of just 2.5 ms. Apparently the pulsar has accreted a substantial amount of mass from its companion star. Like the Black Widow pulsar (see Figure 21-11), it has also blown much of the companion's mass into space. Within a billion years, the companion star may disappear altogether.

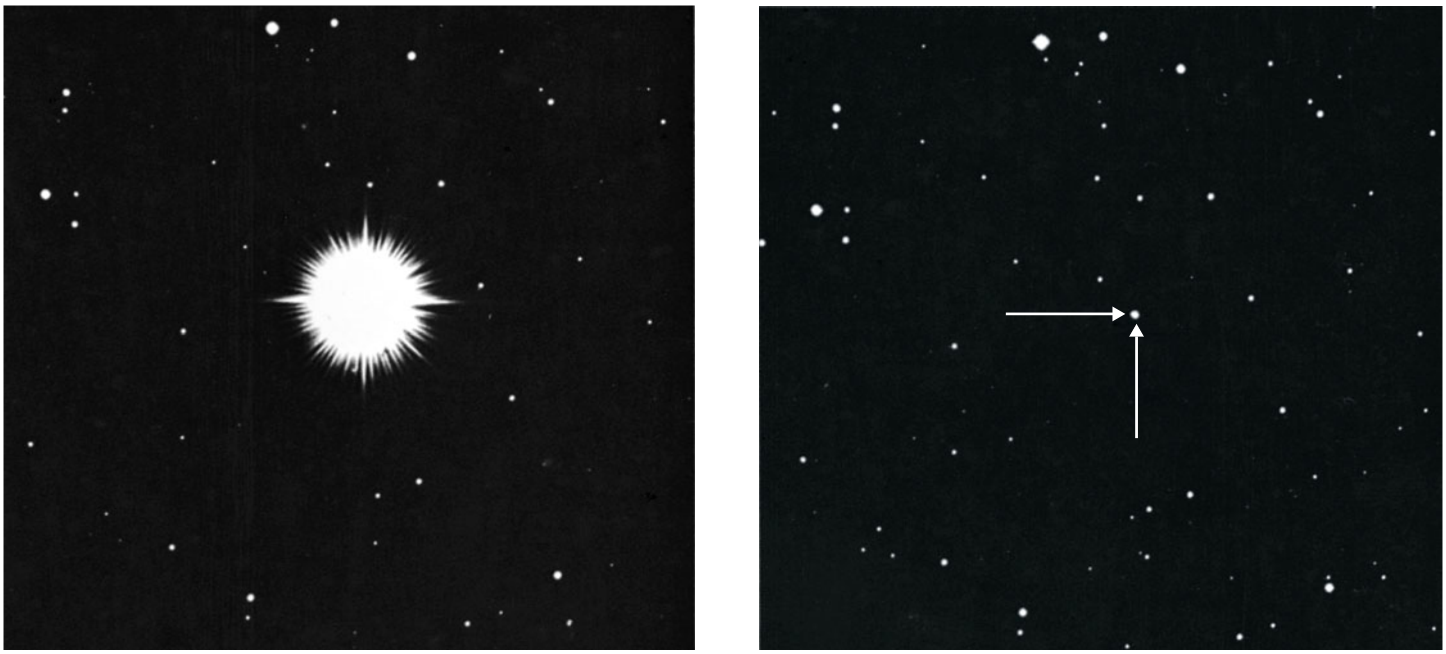
21-9 Explosive thermonuclear processes on white dwarfs and neutron stars produce novae and bursters

Still other exotic phenomena occur when a stellar corpse is part of a close binary system. One example is a **nova** (plural **novae**), in which a faint star suddenly brightens by a factor of 10^4 to 10^8 over a few days or hours, reaching a peak luminosity of about $10^5\ L_\odot$. By contrast, a *superova* has a peak luminosity of about $10^9\ L_\odot$.

A nova's abrupt rise in brightness is followed by a gradual decline that may stretch out over several months or more (Figure 21-14). Every year two or three novae are seen in our Galaxy, and several dozen more are thought to take place in remote regions of the Galaxy that are obscured from our view by interstellar dust.

Novae and White Dwarfs

In the 1950s, painstaking observations of numerous novae by Robert Kraft, Merle Walker, and their colleagues at the University of California's Lick Observatory led to the conclusion that all novae are members of close binary systems containing a white dwarf. Gradual mass transfer from the ordinary companion star, which presumably fills its Roche lobe, deposits fresh hydrogen



(a) Nova Herculis 1934 shortly after peak brightness

(b) Two months later

Figure 21-14 R I **V** U X G

Nova Herculis 1934 These two pictures show a nova **(a)** shortly after peak brightness as a star of apparent magnitude +3, bright enough to be seen easily with the naked eye, and **(b)** two months later, when it had faded by a factor of 4000 in brightness to apparent magnitude +12. (See Figure 17-6 for a description of the apparent magnitude scale.) Novae are named after the constellation and year in which they appear. (Lick Observatory)

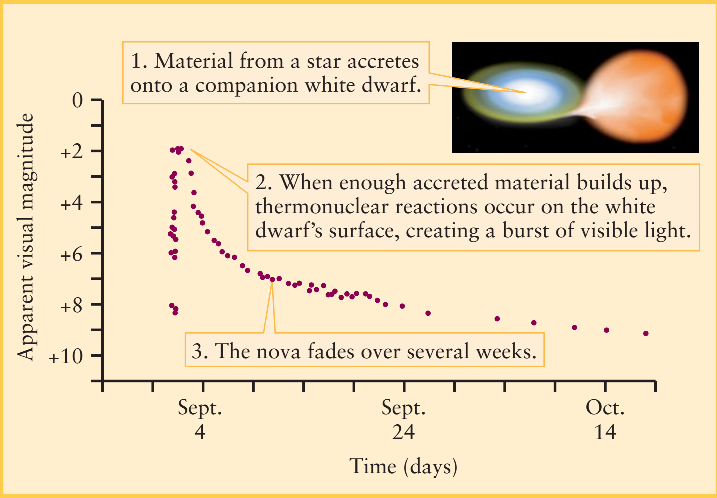


Figure 21-15

The Light Curve of a Nova This illustration and graph show the history of Nova Cygni 1975, a typical nova. Its rapid rise and gradual decline in apparent brightness are characteristic of all novae. This nova, also designated V1500 Cyg, was easily visible to the naked eye (that is, was brighter than an apparent magnitude of +6) for nearly a week. (Illustration courtesy CXC/M. Weiss)

onto the white dwarf (see Figure 19-21*b* for a schematic diagram of this sort of mass transfer).

Because of the white dwarf's strong gravity, this hydrogen is compressed into a dense layer covering the hot surface of the white dwarf. As more gas is deposited and compressed, the temperature in the hydrogen layer increases. When the temperature reaches about 10^7 K, hydrogen fusion ignites throughout the gas layer, embroiling the white dwarf's surface in a thermonuclear holocaust that we see as a nova (Figure 21-15).

CAUTION! It is important to understand the similarities and differences between novae and the thermonuclear (Type Ia) supernovae that we described in Section 20-9. Both kinds of celestial explosions are thought to occur in close binary systems where one of the stars is a white dwarf. But, as befits their name, supernovae are much more energetic. A Type Ia supernova explosion radiates 10^{44} joules of energy into space, while the corresponding figure for a typical nova is 10^{37} joules. (To be fair to novae, this relatively paltry figure is as much energy as our Sun emits in 1000 years.) The difference is thought to be that in a Type Ia supernova, the white dwarf accretes much more mass from its companion. This added mass causes so much compression that nuclear reactions can take place *inside* the white dwarf. Eventually, these reactions blow the white dwarf completely apart. In a nova, by contrast, nuclear reactions occur only within the accreted material. The reaction is more sedate because it takes place only on the white dwarf's surface (perhaps because the accretion rate is less or because the white dwarf had less mass in the first place).

Novae and thermonuclear supernovae both occur in close binary systems with a white dwarf, but a nova can recur while a supernova is a one-shot event

Because the white dwarf itself survives a nova explosion, it is possible for the same star to undergo more than one nova. As an example, the star RS Ophiuchi erupted as a nova in 1898, then put in repeat performances in 1933, 1958, 1967, 1985, and 2006. By contrast, a given star can only be a supernova once.

X-ray Bursters and Neutron Stars

A surface explosion similar to a nova also occurs with neutron stars. In 1975 it was discovered that some objects in the sky emit sudden, powerful bursts of X rays. Figure 21-16 shows the record of a typical burst. The source emits X rays at a constant low level until suddenly, without warning, there is an abrupt increase in X rays, followed by a gradual decline. An entire burst typically lasts for only 20 s. Unlike pulsating X-ray sources, there is a fairly long interval of hours or days between bursts. Sources that behave in this fashion are known as **X-ray bursters.** Several dozen X-ray bursters have been discovered in our Galaxy.

X-ray bursters, like novae, are thought to involve close binaries whose stars are engaged in mass transfer. With a burster, however, the stellar corpse is a neutron star rather than a white dwarf. Gases escaping from the ordinary companion star fall onto the neutron star. The X-ray burster's magnetic field is probably not strong enough to funnel the falling material toward the magnetic poles, so the gases are distributed more evenly over the surface of the neutron star. The energy released as these gases crash down onto the neutron star's surface produces the low-level X rays that are continuously emitted by the burster.

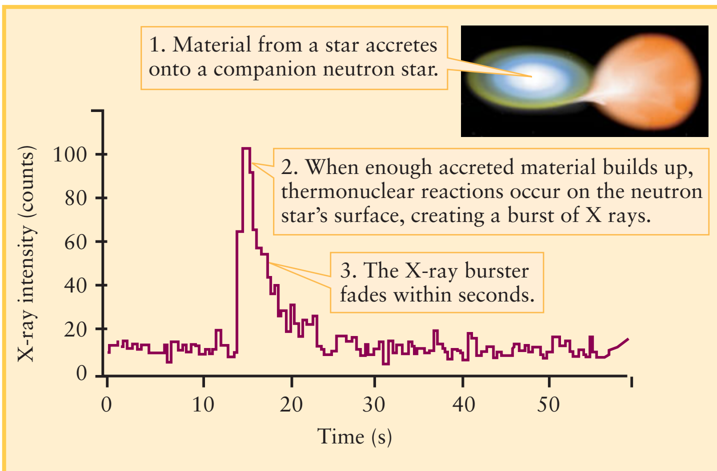


Figure 21-16

The Light Curve of an X-Ray Burster This illustration and graph show the history of a typical X-ray burster. A burster emits a constant low intensity of X rays interspersed with occasional powerful X-ray bursts. This burst was recorded on September 28, 1975, by an Earth-orbiting X-ray telescope. Contrast this figure with Figure 21-15, which shows a typical nova. (Data adapted from W. H. G. Lewin; illustration courtesy CXC/M. Weiss)

Most of the gas falling onto the neutron star is hydrogen, which the star's powerful gravity compresses against its hot surface. In fact, temperatures and pressures in this accreting layer become so high that the arriving hydrogen is converted into helium by hydrogen fusion. As a result, the accreted gases develop a layered structure that covers the entire neutron star, with a few tens of centimeters of hydrogen lying atop a similar thickness of helium. The structure is reminiscent of the layers within an evolved giant star (see Figure 20-2), although the layers atop a neutron star are much more compressed, thanks to the star's tremendous surface gravity.

When the helium layer is about 1 m thick, helium fusion ignites explosively and heats the neutron star's surface to about 3×10^7 K. At this temperature the surface predominantly emits X rays, but the emission ceases within a few seconds as the surface cools. Hence, we observe a sudden burst of X rays only a few seconds in duration. New hydrogen then flows onto the neutron star, and the whole process starts over. Indeed, X-ray bursters typically emit a burst every few hours or days.

Whereas explosive *hydrogen* fusion on a white dwarf produces a nova, explosive *helium* fusion on a neutron star produces an X-ray burster. In both cases, the process is explosive, because the fuel is compressed so tightly against the star's surface that it becomes degenerate, like the star itself. As with the helium flash inside red giants (described in Section 19-3), the ignition of a degenerate thermonuclear fuel involves a sudden thermal runaway. This is because an increase in temperature does not produce a corresponding increase in pressure that would otherwise relieve compression of the gases and slow the nuclear reactions.

CAUTION! Be careful not to confuse X-ray bursters with magnetars (Section 21-6), which are also neutron stars that emit powerful bursts of X rays. An X-ray burster is a member of a binary system that accretes matter from its companion star, then releases *nuclear* energy when the accreted matter undergoes thermonuclear reactions. By contrast, a burst from a magnetar is a release of *magnetic* energy stored in its extraordinarily powerful field; no thermonuclear reactions are involved, and no accretion from a companion star is required. There is also a huge difference in the strength and duration of the two kinds of burst. An X-ray burster releases about 10^{32} joules in a typical 20-second burst, while a magnetar burst can release more than 10^{39} joules of X-ray and gamma-ray radiation in a mere 0.2 second.

One of the great puzzles in modern astronomy has been the nature of an even more remarkable class of events called *gamma-ray bursters*. As the name suggests, these objects emit sudden, intense bursts of high-energy gamma rays. During the short duration of its burst, a gamma-ray burster can be 10^6 to 10^9 times brighter than a supernova at the same distance! We will discuss these exotic objects further in Chapter 22.

21-10 Like a white dwarf, a neutron star has an upper limit on its mass

A white dwarf will collapse if its mass is greater than the Chandrasekhar limit of 1.4 $M_\odot$. At that point, degenerate electron pressure cannot support the overpowering weight of the star's matter, which presses inward from all sides (see Section 20-4). The mass of a neutron star also has an upper limit. However, the pressure within a neutron star is harder to analyze, because it comes from *two* sources. One is the degenerate nature of the neutrons, and the other is the strong nuclear force that acts between the neutrons themselves.

The strong nuclear force is what holds protons and neutrons together in atomic nuclei. Neutrons exert strong nuclear forces on one another only when they are almost touching. This force behaves somewhat like the force that billiard balls exert on one another when they touch: It strongly resists further compression. (Try squeezing two billiard balls together and see how much success you have.) Hence, the strong nuclear force is a major contributor to the star's internal pressure. Unfortunately, there is a good deal of uncertainty about the details of this force. This translates into uncertainties about how much weight the neutron star's internal pressure can support—that is, the neutron star's maximum mass. Theoretical estimates of this maximum mass range from 2 to 3 $M_\odot$.

For a neutron star to collapse, gravity must overwhelm both degeneracy pressure and the short-range repulsion between neutrons

Before pulsars were discovered, most astronomers believed all dead stars to be white dwarfs. Dying stars were thought to somehow eject enough material so that their corpses could be below the Chandrasekhar limit. The discovery of neutron stars proved this idea incorrect. Inspired by this lesson, astronomers soon began wondering what might happen if a dying massive star failed to eject enough matter to get below the upper limit for a neutron star. For example, what might a 5-$M_\odot$ stellar corpse be like?

The gravity associated with a neutron star is so strong that the escape speed from it is roughly one-half the speed of light. But if a stellar corpse has a mass greater than 3 $M_\odot$, so much matter is crushed into such a small volume that the escape speed actually *exceeds* the speed of light. Because nothing can travel faster than light, nothing—not even light—can leave this dead star. Its gravity is so powerful that it leaves a hole in the fabric of space and time. Thus, the discovery of neutron stars inspired astrophysicists to examine seriously one of the most bizarre and fantastic objects ever predicted by modern science, the black hole. We take up its story in the next chapter.

Key Words

Key Ideas

Neutron Stars: A neutron star is a dense stellar corpse consisting primarily of closely packed degenerate neutrons.

• A neutron star typically has a diameter of about 20 km, a mass less than 3 $M_\odot$, a magnetic field 10^{12} times stronger than that of the Sun, and a rotation period of roughly 1 second.

• A neutron star consists of a superfluid, superconducting core surrounded by a superfluid mantle and a thin, brittle crust.

• Intense beams of radiation emanate from regions near the north and south magnetic poles of a neutron star. These beams are produced by streams of charged particles moving in the star's intense magnetic field.

Pulsars: A pulsar is a source of periodic pulses of radio radiation. These pulses are produced as beams of radio waves from a neutron star's magnetic poles sweep past the Earth.

• The pulse rate of many pulsars is slowing steadily. This reflects the gradual slowing of the neutron star's rotation as it radiates energy into space. Sudden speedups of the pulse rate, called glitches, may be caused by interactions between the neutron star's crust and its superfluid interior.

Magnetars: A magnetar is a pulsar with an extraordinarily strong magnetic field. This field is produced by convection inside the pulsar when it first forms.

• The solid crust of a magnetar is under tremendous magnetic stress. When the surface rearranges in a starquake, the released magnetic energy produces a powerful burst of X-rays and gamma rays.

Neutron Stars in Close Binary Systems: If a neutron star is in a close binary system with an ordinary star, tidal forces will draw gas from the ordinary star onto the neutron star.

• The transfer of material onto the neutron star can make it rotate extremely rapidly, giving rise to a millisecond pulsar.

• Magnetic forces can funnel the gas onto the neutron star's magnetic poles, producing hot spots. These hot spots then radiate intense beams of X rays. As the neutron star rotates, the X-ray beams appear to flash on and off. Such a system is called a pulsating X-ray variable.

Novae and Bursters: Material from an ordinary star in a close binary can fall onto the surface of the companion white dwarf or neutron star to produce a surface layer in which thermonuclear reactions can explosively ignite.

• Explosive hydrogen fusion may occur in the surface layer of a companion white dwarf, producing the sudden increase in luminosity that we call a nova. The peak luminosity of a nova is only 10^{-4} of that observed in a supernova.

• Explosive helium fusion may occur in the surface layer of a companion neutron star. This produces a sudden increase in X-ray radiation, which we call a burster.

Questions

Review Questions

1. What are neutron stars? What led scientists to propose their existence?
2. What is degenerate neutron pressure? How does it make it possible for a neutron star to be more massive than the Chandrasekhar limit for white dwarfs?
3. How were pulsars discovered? How do they differ from variable stars?
4. Why did astronomers rule out the idea that pulsars are rapidly rotating white dwarfs?
5. Why do astronomers think that pulsars are rapidly rotating neutron stars?
6. During the weeks immediately following the discovery of the first pulsar, one suggested explanation was that the pulses might be signals from an extraterrestrial civilization. Why did astronomers soon discard this idea?
7. Why do neutron stars rotate so much more rapidly than ordinary stars? Why do they have such strong magnetic fields?
8. How are rotating neutron stars able to produce pulses of radiation as seen by an observer on Earth?
9. Do all supernova remnants contain pulsars? Are all pulsars found within supernova remnants? For each question, explain why or why not.
10. Is our Sun likely to end up as a neutron star? Why or why not?
11. Why are some neutron stars seen moving through space at hundreds of kilometers per second?
12. What is synchrotron radiation? How is it involved in making the Crab Nebula glow?
13. Why does an isolated pulsar rotate more slowly as time goes by?
14. Astronomers have deduced that the Vela pulsar is about 11,000 years old. How do you suppose they did this?
15. What is the difference between superconductivity and superfluidity?
16. Does a neutron star contain only neutrons? If not, what else does it contain?
17. What is a pulsar glitch? How does a glitch affect a pulsar's period? What is thought to be the cause of glitches?
18. Compare the internal structure of a white dwarf to that of a neutron star. What are the similarities? What are the differences?
19. What is the evidence that neutron stars have atmospheres?
20. How does a magnetar differ from an ordinary pulsar? What determines whether a pulsar becomes a magnetar?
21. How are magnetars able to emit bursts of X-rays and gamma rays?
22. Why do astronomers think that millisecond pulsars are very old?
23. If millisecond pulsars are formed in close binary systems, why are some found without companion stars?
24. Describe a pulsating X-ray source like Hercules X-1 or Centaurus X-3. What produces the pulsation?
25. What is the connection between pulsating X-ray sources and millisecond pulsars?
26. What are the similarities between a nova and a Type Ia supernova? What are the differences?
27. What are the similarities between novae and X-ray bursters? What are the differences?
28. What are the similarities between pulsating X-ray sources and X-ray bursters? What are the differences?
29. (**a**) Rank the following explosive phenomena in order of the amount of energy released, from smallest to largest: (i) a nova; (ii) a Type Ia supernova; (iii) an X-ray burster; (iv) a major burst from a magnetar. (**b**) For each of the phenomena

listed in part (a), explain what the source of the released energy is.

30. Why is the maximum mass of neutron stars not known as accurately as the Chandrasekhar limit for white dwarfs?

Advanced Questions

Problem-solving tips and tools

The volume of a sphere of radius r is $4\pi r^3/3$. The small-angle formula is given in Box 1-1. Section 5-4 describes how to relate the temperature of a blackbody to its wavelength of maximum emission. Section 17-6 gives the formula relating a star's luminosity, surface temperature, and radius (see Box 17-4 for worked examples). Appendix 6 gives the conversion between seconds and years as well as the radius of the Sun.

31. Using a diagram like Figure 21-3, explain why the number of pulsars that we observe in nearby space is probably quite a bit less than the number of rotating, magnetized neutron stars in nearby space.

32. There are many more main-sequence stars of low mass (less than 8 $M_\odot$) than of high mass (8 $M_\odot$ or more). Use this fact to explain why white dwarf stars are far more common than neutron stars.

33. The distance to the Crab Nebula is about 2000 parsecs. In what year did the star actually explode? Explain your answer.

34. How do we know that the Crab pulsar is really embedded in the Crab Nebula and not simply located at a different distance along the same line of sight?

35. The Crab Nebula has an apparent size of about 5 arcmin, and this size is increasing at a rate of 0.23 arcsec per year. (**a**) Assume that the expansion rate has been constant over the entire history of the Crab Nebula. Based on this assumption, in what year would Earth observers have seen the supernova explosion that formed the nebula? (**b**) Does your answer to part (a) agree with the known year of the supernova, 1054 A.D.? If not, can you point to assumptions you made in your computations that led to the discrepancies? Or do you think your calculations suggest additional physical effects are at work in the Crab Nebula, over and above a constant rate of expansion?

36. Emission lines in the spectrum of the Crab Nebula exhibit a Doppler shift, which indicates that gas in the part of the nebula closest to us is moving toward us at 1450 km/s. (**a**) Assume that the expanding gas has been moving at the same speed since the original supernova explosion, observed in 1054 A.D., and calculate what radius and what diameter (in light-years) we should observe the nebula to have today. (**b**) Compare your result in part (a) to the actual size of the nebula, given in the caption to Figure 21-4.

37. The supernova remnant G5.4-1.2 shown in Figure 21-5 lies about 5000 pc (16,000 ly) from Earth in the constellation Sagittarius. (**a**) The green arc in the large left-hand image in Figure 21-5 represents part of the outer edge of this spherical supernova remnant. Estimate the diameter of this remnant in parsecs. (*Hint:* To calculate this, you will need to make measurements on the image.) (**b**) How far (in parsecs) did the neutron star travel from where it was formed (at the position of the supernova's progenitor star, presumably at the center of the present-day remnant) to the position shown in Figure 21-5? Explain your answer.

38. To determine accurately the period of a pulsar, astronomers must take into account the Earth's orbital motion about the Sun. (**a**) Explain why. (**b**) Knowing that the Earth's orbital velocity is 30 km/s, calculate the maximum correction to a pulsar's period because of the Earth's motion. Explain why the size of the correction is greatest for pulsars located near the ecliptic.

39. If a pulsar has period P (in seconds) and its period is increasing at a rate R (in seconds per second), an approximate formula for the age T of the pulsar (in seconds) is $T = P/2R$. For the Crab pulsar, $P = 0.0333$ s and $R = 4.21 \times 10^{-13}$ s/s. (**a**) Calculate the approximate age of the Crab pulsar in years. (**b**) Based on the information given in Section 21-1, is your result in (a) an underestimate or an overestimate? Explain your answer.

40. (See Advanced Question 39.) (**a**) Magnetar SGR 1806-20 was the source of the intense burst observed on December 27, 2004. Based on the information given in Section 21-6, calculate the rate R (in seconds per second) at which the period of SGR 1806-20 is increasing. (**b**) The ordinary pulses from SGR 1806-20 have a period of 7.47 s. Calculate the approximate age of SGR 1806-20 in years. (**c**) Theory predicts that a magnetar becomes inactive after about 10,000 years. Is your result in part (b) consistent with this prediction? Explain why or why not.

41. A neutron has a mass of about 1.7×10^{-27} kg and a radius of about 10^{-15} m. (**a**) Compare the density of matter in a neutron with the average density of a neutron star. (**b**) If the neutron star's density is more than that of a neutron, the neutrons within the star are overlapping; if it is less, the neutrons are not overlapping. Which of these seems to be the case for average neutrons within the star? Which do you think is the case at the center of the neutron star, where densities are higher than average?

42. The total luminosity at all wavelengths of the magnetar burst observed on December 27, 2006, was approximately 10^{14} $L_\odot$. At what distance from the magnetar would the brightness of the burst have been equal to the brightness of the Sun as seen on Earth? Give your answer in AU and in parsecs.

43. X-ray pulsars are speeding up but ordinary (radio) pulsars are slowing down. Propose an explanation for this difference.

44. If the model for Hercules X-1 discussed in the text is correct, at what orientation of the binary system do we see its maximum optical brightness? Explain your answer.

45. Explain why heavy elements that are produced by neutron star collisions can still be regarded as having been processed through a supernova.

46. In an X-ray burster, the surface of a neutron star 10 km in radius is heated to a temperature of 3×10^7 K. (**a**) Determine the wavelength of maximum emission of the heated surface (which you may treat as a blackbody). In what part of the electromagnetic spectrum does this lie? (See Figure 5-7.) (**b**) Find the luminosity of the heated neutron star. Give your answer in watts and in terms of the luminosity of the Sun, given in Table 16-1. How does this compare with the peak luminosity of a nova? Of a Type Ia supernova?