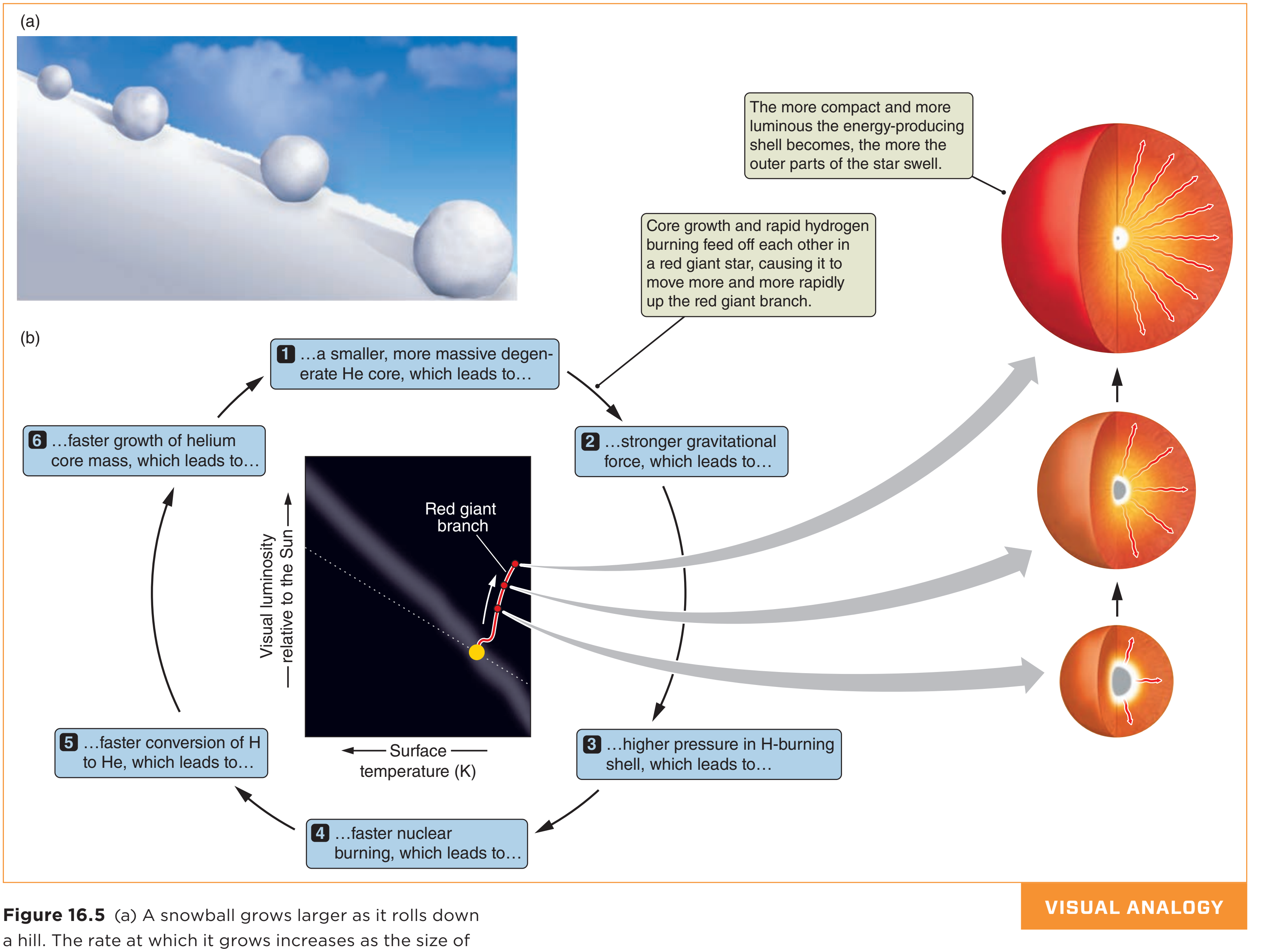


Figure 16.5 (a) A snowball grows larger as it rolls down a hill. The rate at which it grows increases as the size of the snowball increases. (b) Similarly, as a star moves up the red giant branch in the H-R diagram, the luminosity of the star grows faster and faster. The burning of hydrogen to helium in a shell surrounding a degenerate helium core feeds on itself, creating a cycle that speeds up as time goes on.

an ever-higher rate from 10 L_{Sun} to almost 1,000 L_{Sun}. The evolution of the star, illustrated in **Figure 16.5**, is reminiscent of the growth of a snowball rolling downhill. The larger the snowball becomes, the faster it grows; and the faster it grows, the larger it becomes. Growth and size feed off each other, and what began as a bit of snow at the top of the mountain soon becomes a huge ball.

CHECK YOUR UNDERSTANDING 16.2

When the main sequence Sun runs out of fuel in its core, the core will be: (a) empty; (b) filled with hydrogen; (c) filled with helium; (d) filled with carbon.

16.3 Helium Burns in the Degenerate Core

On the main sequence, a star burns hydrogen in the core. On the red giant branch, the star has an inert helium core but continues to burn hydrogen in a shell around this core. Once the hydrogen in the shell is used up, the core of the star contracts and heats. Eventually, the temperature and pressure rise enough to start the next stage of stellar evolution: helium burning.

Helium Burning and the Triple-Alpha Process

As the star evolves up the red giant branch, its helium core grows not only smaller and more massive but also hotter. This increase in temperature is due partly to the gravitational energy released as the core shrinks and partly to the energy released by hydrogen burning in the surrounding shell. The thermal motions of the atomic nuclei in the core become more and more energetic. Eventually, at a temperature of about 100 million (10^8) K, the collisions among helium nuclei in the core become energetic enough to overcome the electric repulsion. Helium nuclei are slammed together hard enough for the strong nuclear force to act, and helium burning begins.

Helium burns in a two-stage process called the **triple-alpha process**, which takes its name from the fact that it involves the fusion of three ^{4}He nuclei, which are called **alpha particles**. The process, illustrated in **Figure 16.6**, begins when two helium-4 (^{4}He) nuclei fuse to form a beryllium-8 (^{8}Be) nucleus consisting of four protons and four neutrons. The ^{8}Be nucleus is extremely unstable, with a very short lifetime before it decays. But if, in that short time, it collides with another ^{4}He nucleus, the two nuclei will fuse into a stable nucleus of carbon-12 (^{12}C) consisting of six protons and six neutrons. The reaction rate is very temperature-dependent: higher temperatures enable more reactions and increase the number of ^{8}Be nuclei that collide with a ^{4}He nucleus.

Recall that the core of the red giant star is electron-degenerate, which means that as many electrons are packed into that space as possible. These degenerate electrons prevent the core from collapsing. The atomic nuclei, however, behave like a normal gas, moving through the sea of degenerate electrons almost as if the electrons were not there. The atomic nuclei in the core move freely about as shown in **Figure 16.7**, just as they do throughout the rest of the star. We can understand the fusion of helium by treating the nuclei as matter in a normal state. Once the pressure and temperature are high enough, these nuclei begin

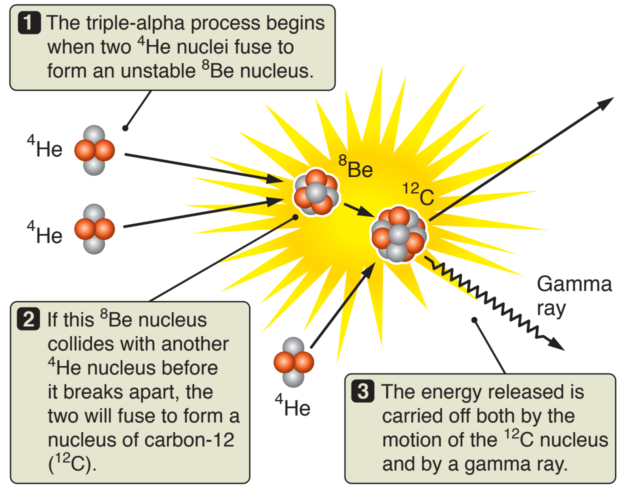
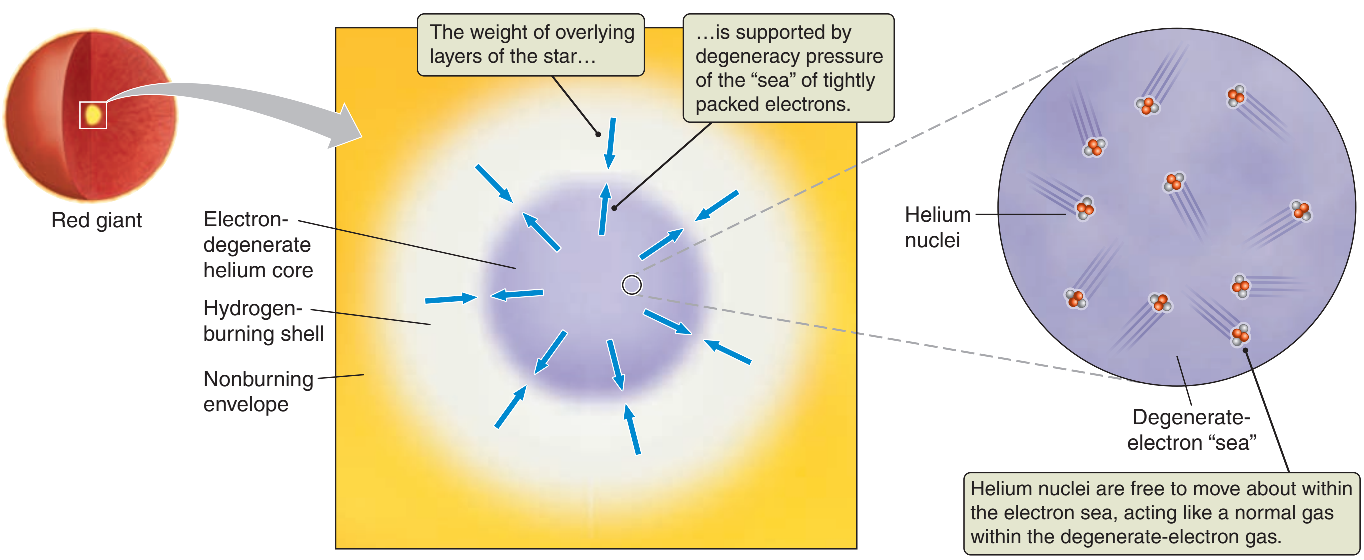


Figure 16.6 The triple-alpha process produces a stable nucleus of carbon-12. Two helium-4 (^{4}He) nuclei fuse to form an unstable beryllium-8 (^{8}Be) nucleus. If this nucleus collides with another ^{4}He nucleus before it breaks apart, the two will fuse to form a stable nucleus of carbon-12 (^{12}C). The energy produced is carried off both by the motion of the ^{12}C nucleus and by a high-energy gamma ray emitted in the second step of the process.

Figure 16.7 In a red giant star, the weight of the overlying layers is supported by electron degeneracy pressure in the core arising from the fact that electrons are packed together as tightly as quantum mechanics allows. Atomic nuclei in the core are able to move freely about within the sea of degenerate electrons, so they behave as a normal gas.



to fuse, as the hydrogen nuclei do in main-sequence stars. But the energy released will not affect the degenerate-electron core in the same way as in a main-sequence star.

The Helium Flash

Degenerate material is a very good conductor of thermal energy, so any differences in temperature within the core rapidly disperse. As a result, when helium burning begins at the center of the core, the energy released quickly heats the entire core. Within a few minutes, the entire core is burning helium into carbon by the triple-alpha process. Some of the carbon will fuse with an additional ^{4}He nucleus to form stable oxygen-16 (^{16}O).

In a normal gas such as the air around you or the core of a main-sequence star, the pressure of the gas comes from the random thermal motions of the atoms. Increasing the temperature of such a gas means that the motions of the atoms become more energetic, so the pressure of the gas increases. If the helium core of a red giant star were a normal gas, the increase in temperature that accompanies the onset of helium burning would lead to an increase in pressure. The core of the star would expand; the temperature, density, and pressure would decrease; nuclear reactions would slow; and the star would settle into a new balance between gravity and pressure. These are exactly the sorts of changes that are steadily occurring within the core of a main-sequence star like the Sun as the structure of the star shifts in response to the changing composition in the star's core.

However, the degenerate core of a red giant is not a normal gas. The pressure in a red giant's degenerate core comes from how tightly the electrons in the core are packed together. Heating the core does not change the number of electrons that can be packed into its volume, so the core's pressure does not respond to changes in temperature. Because the pressure does not increase, the core does not expand when heated, as a normal gas would.

Yet even though the higher temperature does not change the pressure, it does cause the helium nuclei to collide with more frequency and greater force, so the nuclear reactions become more vigorous. More vigorous reactions mean higher temperature, and higher temperature means even more vigorous reactions. The process begins to snowball again. Helium burning in the degenerate core runs wildly out of control as increasing temperature and increasing reaction rates feed each other. As long as the degeneracy pressure from the electrons is greater than the thermal pressure from the nuclei, this feedback loop continues.

Helium burning begins at a temperature of about 100 million K. By the time the temperature has climbed by just 10 percent, to 110 million K, the rate of helium burning has increased to 40 times what it was at 100 million K. By the time the core's temperature reaches 200 million K, the core is burning helium 460 million times faster than it was at 100 million K. As the temperature in the core grows higher and higher, the thermal motions of the electrons and nuclei become more energetic, and the pressure due to these thermal motions becomes greater and greater. Within seconds of helium ignition, the thermal pressure increases until it is no longer smaller than the degeneracy pressure. At this point, the helium core explodes in what astronomers call a helium flash, illustrated in **Figure 16.8**. Because the explosion is contained within the star, however, it cannot be seen outside the star. The energy released in this runaway thermonuclear explosion lifts the intermediate layers of the star, and as the core expands, the

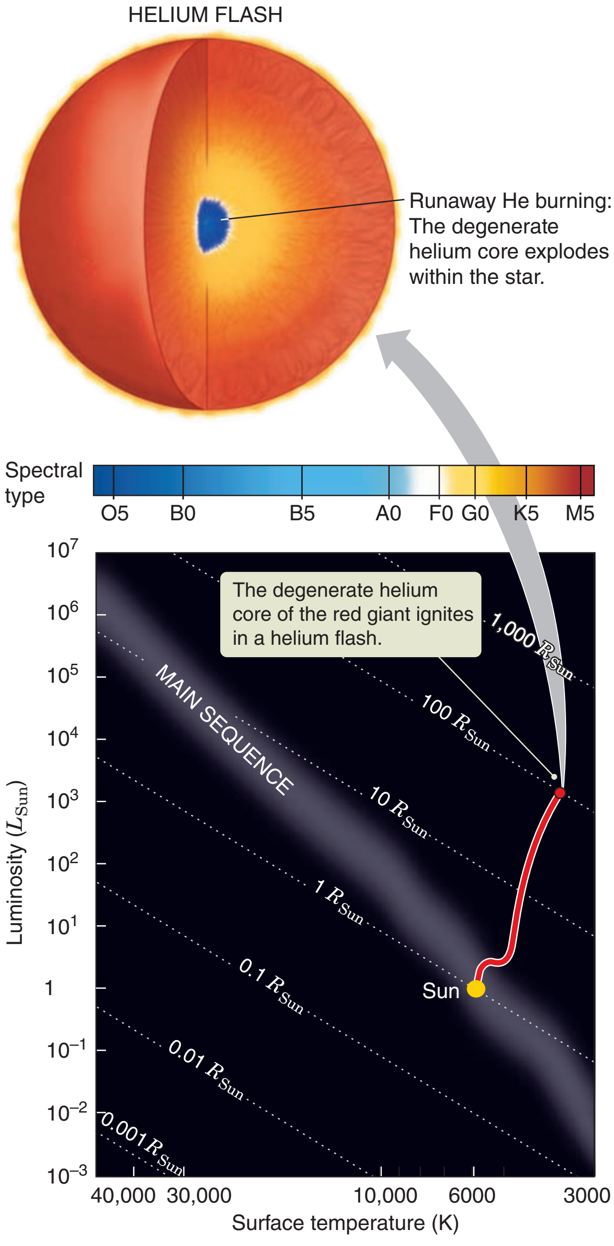


Figure 16.8 At the end of its life, a low-mass star travels a complex path on the H-R diagram. The first part of that path takes it up the red giant branch to a point where helium ignites in a helium flash. After a few hours, the core of the star begins to inflate, ending the helium flash.

electrons are able to spread out. The drama is over within a few hours because the expanded helium-burning core is no longer degenerate, and the star is on its way toward a new equilibrium.

Helium burning in the core does not cause the star to grow more luminous. The tremendous energy released during the helium flash goes into fighting gravity and puffing up the core. After the helium flash, the core (which is no longer degenerate) is much larger, so the force of gravity within it and the surrounding shell are much smaller. Weaker gravity means less weight pushing down on the core and the shell, which means lower pressure. Lower pressure, in turn, slows the nuclear reactions. The net result is that after the helium flash, core helium burning keeps the core of the star puffed up, and the star becomes less luminous than it was as a red giant.

The Horizontal Branch

The star takes about 100,000 years or so to settle into stable helium burning. It then spends about 100 million years burning helium into carbon in a normal, nondegenerate core while hydrogen burns to helium in a surrounding shell. The star is now about a hundred times less luminous than it was when the helium flash occurred. The lower luminosity means that the outer layers of the star are not as puffed up as they were when the star was a red giant. Gravity becomes stronger than the outward pressure of the escaping radiation and pulls the outer layers back in. The star shrinks, and its surface temperature climbs as gravitational energy is converted to thermal energy. The star moves horizontally across the H-R diagram, remaining at the same luminosity, but increasing in surface temperature. This portion of a star's path on the H-R diagram is called the horizontal branch, shown in **Figure 16.9**. At this point in their evolution, low-mass stars with chemical compositions similar to that of the Sun lie on the H-R diagram just to the left of the red giant branch. Stars that contain much less iron than the Sun tend to distribute themselves away from the red giant branch along the horizontal branch.

The structure and behavior of a star on the horizontal branch are similar to those of a main-sequence star. However, instead of burning hydrogen into helium in the core, the horizontal branch star burns helium into carbon. In addition, in the horizontal branch star, hydrogen continues to burn in a shell surrounding the core.

The star's time on the horizontal branch, however, is much shorter than its time on the main sequence. There is now less fuel to burn in its core. In addition, the star is more luminous than it was on the main sequence, so it is consuming fuel more rapidly. Helium is a much less efficient nuclear fuel than hydrogen, so the star has to burn fuel even faster to maintain equilibrium. Even so, for about a hundred million years the horizontal branch star remains stable, burning helium to carbon in its core and hydrogen to helium in a shell around the core.

The temperature at the center of a horizontal branch star is not high enough for carbon to burn, so carbon builds up in the heart of the star. When the horizontal branch star has burned all of the helium at its core, gravity once again begins to overwhelm the pressure of the escaping radiation. The nonburning carbon core is crushed by the weight of the layers of the star above it until once again the electrons in the core are packed together as tightly as possible, given its pressure. The carbon core is now electron-degenerate, with physical properties much like those of the degenerate helium core at the center of a red giant.

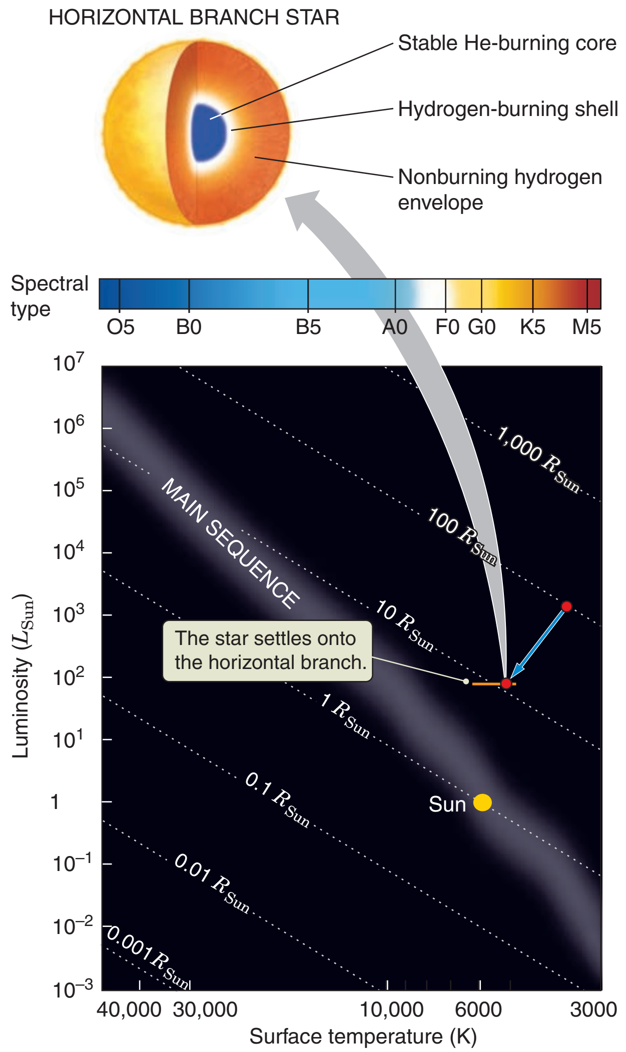


Figure 16.9 The star moves down from the red giant branch onto the horizontal branch. For about 100 million years or less, the star will remain on the horizontal branch and burn helium in its core and hydrogen in a surrounding shell.

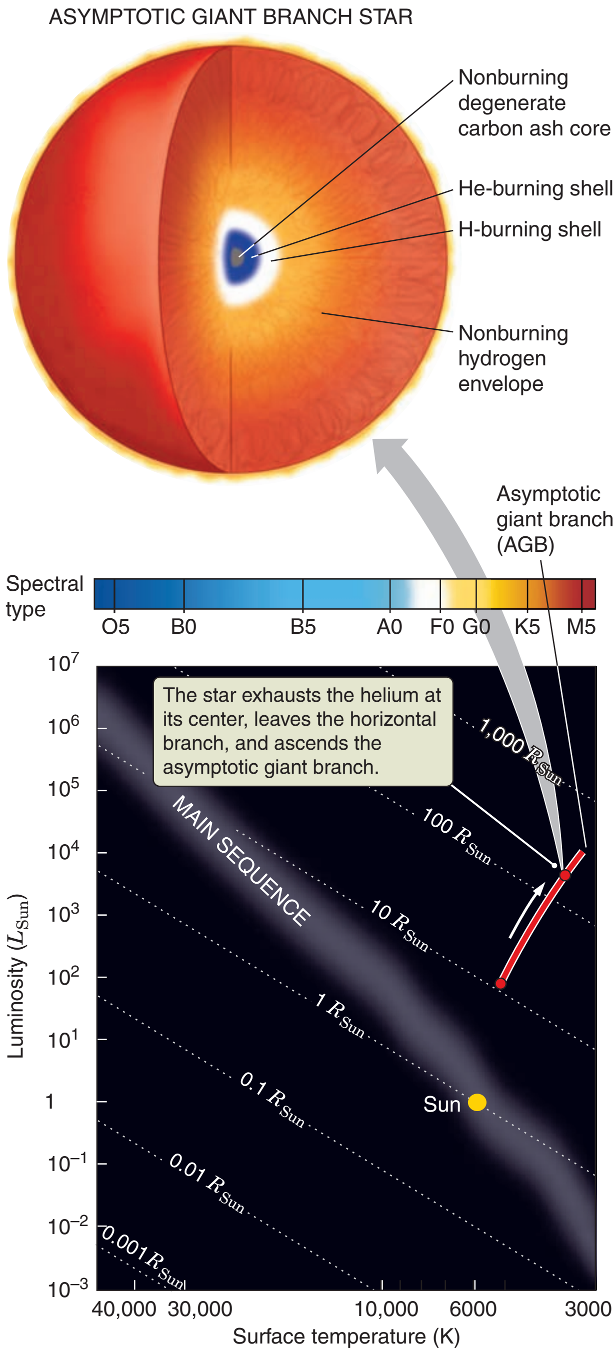


Figure 16.10 The star moves up from the horizontal branch onto the asymptotic giant branch (AGB). An AGB star consists of a degenerate carbon core surrounded by helium-burning and hydrogen-burning shells. As the carbon core grows, the star brightens, accelerating up the asymptotic giant branch just as it earlier accelerated up the red giant branch while its degenerate helium core grew.

The Lowest-Mass Stars

Brown dwarfs have masses less than 0.08 M_{Sun}. They do not become main-sequence stars, so they do not follow post-main-sequence steps. Small red dwarf stars with masses lower than about 0.4–0.5 M_{Sun} may be the majority of the stars in our galaxy, but they are small and faint and hard to detect. These stars have main-sequence lifetimes longer than the 13.8-billion-year age of the universe, so we have not actually observed any of these stars in post-main-sequence stages. There has not been enough time for these stars to evolve off of the main sequence.

CHECK YOUR UNDERSTANDING 16.3

Stars begin burning helium to carbon when the temperature rises in the core. This temperature increase is caused by (choose all that apply): (a) gravitational collapse; (b) fusion of hydrogen into helium in the core; (c) fusion of hydrogen into helium in a shell around the core; (d) electron degeneracy pressure.

16.4 Dying Stars Shed Their Outer Layers

After the horizontal branch, small changes in the properties of a star—mass, chemical composition, strength of the star's magnetic field, or even the rate at which the star is rotating—can lead to noticeable differences in how the star (and especially its outer envelope) evolves. In this section, we follow a 1-M_{Sun} star with solar composition as it concludes its evolutionary stages, losing its outer layers and leaving behind a cooling carbon core.

Stellar-Mass Loss and the Asymptotic Giant Branch

After a time on the horizontal branch, a small, dense, electron-degenerate carbon core remains. This core is very compact, causing the gravity in the inner parts of the star to be very high, which in turn drives up the pressure, which speeds up the nuclear reactions, which causes the degenerate core to grow more rapidly. The internal changes occurring within the star are similar to the changes that took place at the end of the star's main-sequence lifetime, and the path the star follows as it leaves the horizontal branch echoes that earlier phase of evolution. Just as the star accelerated up the red giant branch as its degenerate helium core grew, the star now leaves the horizontal branch and once again begins to grow larger, redder, and more luminous as its degenerate carbon core grows. As shown in **Figure 16.10**, the path that the star follows, called the **asymptotic giant branch (AGB)** of the H-R diagram, parallels the path it followed as a red giant, approaching the red giant branch as the star grows more luminous. An AGB star burns helium and hydrogen in nested concentric shells surrounding a degenerate carbon core as the star moves once again up the H-R diagram. Before the temperature in the carbon core becomes high enough for carbon to burn, the pressure of the escaping radiation exceeds the gravitational pull on the outer layers of the star. The star begins to lose these outer layers to space.

AGB stars are huge objects. When the Sun becomes an AGB star, its outer layers will swell to the point that they engulf the orbits of the inner planets, possibly including Earth and maybe even Mars. When a star expands to such a size, the gravitational force at its surface is only 1/10,000 as strong as the gravity at the surface of the present-day Sun. It takes little extra energy to push surface

16.2 Working It Out Escaping the Surface of an Evolved Star

Why are giant stars likely to lose mass? The escape velocity from the surface of a planet or star was given in Working It Out 4.2:

$$v_{esc} = \sqrt{2GM/R}$$

How does v_{esc} change when a star becomes a red giant? Let's look at the Sun as an example. When the Sun is on the main sequence, its escape velocity can be calculated using $M_{Sun} = 1.99 \times 10^{30}$ kg, $R_{Sun} = 6.96 \times 10^5$ km, and $G = 6.67 \times 10^{-20}\,\text{km}^3/(\text{kg s}^2)$:

$$v_{esc} = \sqrt{\frac{2 \times [6.67 \times 10^{-20}\,\text{km}^3/(\text{kg s}^2)] \times (1.99 \times 10^{30}\,\text{kg})}{6.96 \times 10^5\,\text{km}}}$$

$$v_{esc} = \sqrt{3.81 \times 10^5\,\text{km}^2/\text{s}^2} = 618\,\text{km/s}$$

What will the escape velocity be when the Sun becomes a red giant, with a radius 50 times greater than the radius it has today and a mass 0.9 times its current mass:

$$v_{esc} = \sqrt{\frac{2 \times [6.67 \times 10^{-20}\,\text{km}^3/(\text{kg s}^2)] \times 0.9 \times (1.99 \times 10^{30}\,\text{kg})}{50 \times (6.96 \times 10^5\,\text{km})}}$$

$$v_{esc} = \sqrt{6.86 \times 10^3\,\text{km}^2/\text{s}^2} = 83\,\text{km/s}$$

The escape velocity from the surface of a red giant star is only 13 percent that of a main-sequence star. This is part of the reason that red giant and AGB stars lose mass. The Sun may eventually lose half of its mass.

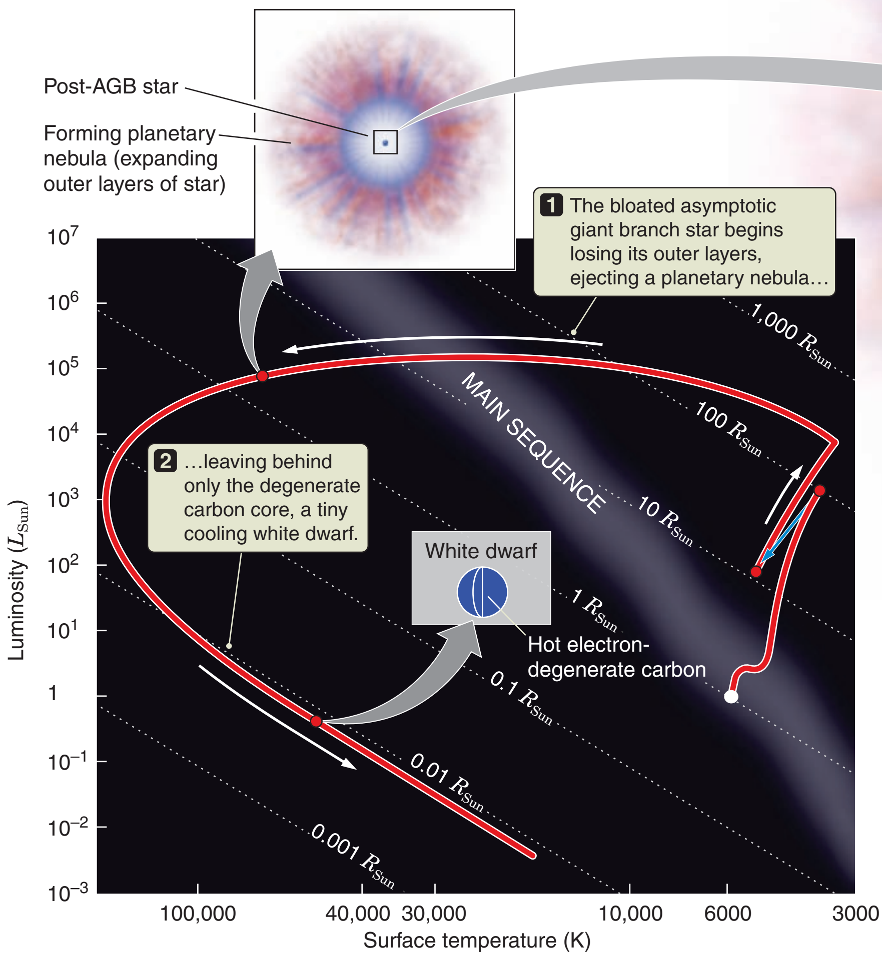
material away from the star. **Stellar-mass loss**—the loss of mass from the outer layers of the star as it evolves—actually begins when the star is still on the red giant branch: by the time a 1-M_{Sun} main-sequence star reaches the horizontal branch, it may have lost 10–20 percent of its total mass. As the star ascends the asymptotic giant branch, it loses another 20 percent or even more of its total mass. By the time it is well up on this branch, a star that began as a 1-M_{Sun} star may have lost more than half of its original mass. Stellar-mass loss is further explored in **Working It Out 16.2**.

Mass loss on the asymptotic giant branch can be spurred on by the star's unstable interior. The extreme sensitivity of the triple-alpha process to temperature in the core can lead to episodes of rapid energy release, which can provide the extra kick needed to expel material from the star's outer layers. Even stars that are initially quite similar can behave very differently when they reach this stage in their evolution.

Planetary Nebula

Toward the end of an AGB star's life, mass loss itself becomes a runaway process. When a star loses a bit of mass from its outermost layers, the weight pushing down on the underlying layers of the star is reduced. Without this weight holding them down, the remaining outer layers of the star puff up further. The post-AGB star, which is now both less massive and larger, is even less tightly bound by gravity, so less energy is needed to push its outer layers away. Mass loss leads to weaker gravity, which leads to faster mass loss, which leads to weaker gravity, and so on. When the end comes, much of the remaining mass of the star is ejected into space, typically at speeds of 20–30 kilometers per second (km/s).

After ejection of its outer layers, all that is left of the low-mass star is a tiny, very hot, electron-degenerate carbon core surrounded by a thin envelope in which hydrogen and helium are still burning. This star is now somewhat less luminous than when it was at the top of the asymptotic giant branch, but it is still much more luminous than a horizontal branch star. The remaining hydrogen and helium in the star rapidly burn to carbon, and as more and more of the mass of the star ends up in the carbon core, the star itself shrinks and becomes hotter and



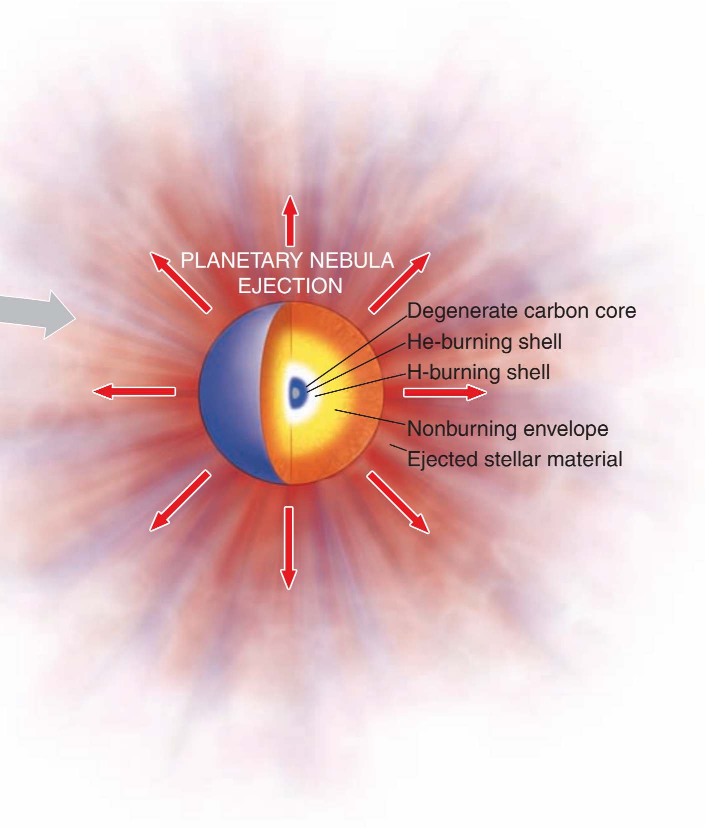


Figure 16.11 At the end of the AGB star's life, it ejects most of its mass in a planetary nebula, becoming a post-AGB star and ultimately leaving behind nothing but the star's degenerate core, which appears on the H-R diagram as a cooling white dwarf star.

hotter. Over the course of only about 30,000 years after the beginning of runaway mass loss, the star moves from right to left across the top of the H-R diagram as shown in **Figure 16.11**.

The surface temperature of the star may eventually rise above 100,000 K. At such temperatures, the peak wavelength of the radiation is in the high-energy ultraviolet (UV) part of the spectrum, as determined by Wien's law. This intense UV light heats and ionizes the ejected, expanding shell of gas, causing it to glow in the same way that UV light from an O star causes an H II region to glow. When these glowing shells were first observed in small telescopes, they were named "planetary nebulae" because they appeared fuzzy like nebular clouds of dust and gas, but they were approximately round, like planets. Later imagery, such as the images in **Figure 16.12**, showed that these objects are not like planets at all. Rather, a **planetary nebula** is the remaining outer layers of a star, ejected into space at the end of the star's ascent of the asymptotic giant branch. A planetary nebula may be visible for 50,000 years or so before the gas ejected by the star disperses so far that the nebula is too faint to be seen. Not all stars form planetary

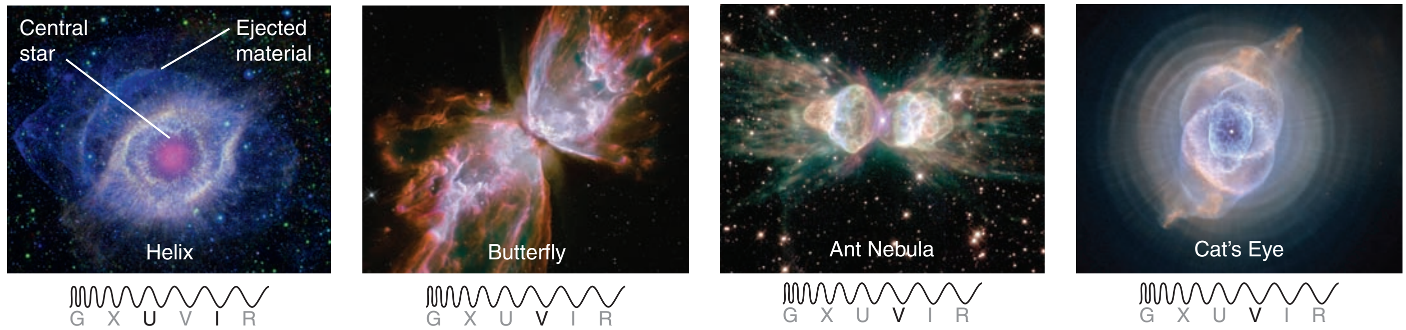


Figure 16.12 At the end of its life, a low-mass star ejects its outer layers and may form a planetary nebula consisting of an expanding shell of gas surrounding the white-hot remnant of the star. Planetary nebulae are not all simple spherical shells around their parent stars. These images of planetary nebulae from the Hubble Space Telescope and the Spitzer Space Telescope show the wealth of structures that result from the complex processes by which low-mass stars eject their outer layers.

nebulae. Stars more massive than about 8 M_{Sun} pass through the post-AGB stage too quickly. Stars with insufficient mass take too long in the post-AGB stage, so their envelope evaporates before they can illuminate it. Astronomers do not know if our own Sun will retain enough mass during its post-AGB phase to form a planetary nebula.

The detailed structure of a planetary nebula may contain concentric rings of varied density, indicating that the stellar-mass loss was slower and faster at various times. In the image of the Ring Nebula in the chapter-opening photograph, the colored rings are emission lines from different ions in different places. During the formation of this nebula, a lot of mass was lost nearly all at once. Then the mass loss ceased, resulting in a hollow shell around the central star. In other planetary nebula, the material may be concentrated parallel to the equator or poles of the star, indicating that the stellar-mass loss was blocked in some directions. You can see this in the images of the Butterfly, Ant, and Cat's Eye nebulae in Figure 16.12.

The gas in a planetary nebula carries the chemical elements enriching the star's outer layers off into interstellar space. Planetary nebulae often show an overabundance of elements such as carbon, nitrogen, and oxygen compared to the abundance of these elements in the outer layers of the Sun. These elements are by-products of nuclear burning either from the star that produced the planetary nebula or from the stars of earlier generations. Once this chemically enriched material leaves the star, it mixes with interstellar gas, increasing the chemical diversity of the universe.

White Dwarfs

Within about 50,000 years, a post-AGB star burns all of the fuel remaining on its surface, leaving behind a nonburning ball of carbon with a mass less than 70 percent of the mass of the original star. In the process, the post-AGB star becomes smaller and fainter, and its position on the H-R diagram falls down the left side. Within a few thousand years, the burned-out core shrinks to about the size of Earth, at which point it has become fully electron-degenerate and can shrink no further. This remnant of stellar evolution is called a **white dwarf**.

The white dwarf, composed of nonburning electron-degenerate carbon and maybe some oxygen, continues to radiate energy away into space. As it does so it cools, just like the heating coil on an electric stove once it is turned off. Because the white dwarf is electron-degenerate, its size does not change much as it cools, so it moves down and to the right on the H-R diagram, following a line of constant radius. The white dwarf will remain very hot for 10 million years or so, but its tiny size means it is a thousand times less luminous than a main-sequence star like the Sun. Many white dwarfs are known, but none can be seen without a telescope. Sirius, the brightest star in Earth's sky, has a faint white dwarf as a binary companion.

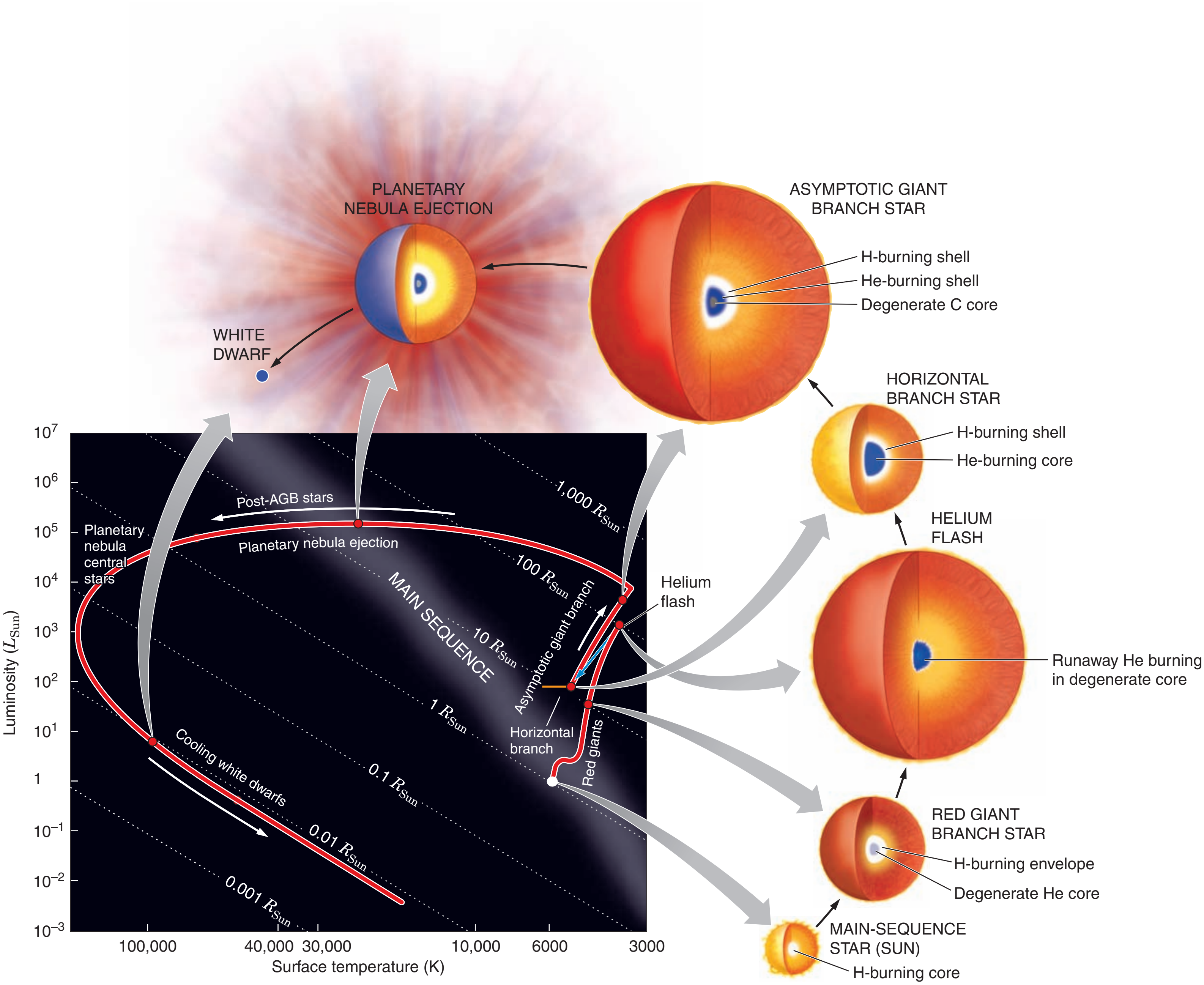


Figure 16.13 This H-R diagram summarizes the stages in the post-main-sequence evolution of a 1-M_{Sun} star.

Figure 16.13 summarizes the evolution of a solar-type, 1-M_{Sun} main-sequence star through to its final existence as a 0.6-M_{Sun} white dwarf. The star leaves the main sequence, climbs the red giant branch, falls to the horizontal branch, climbs back up the asymptotic giant branch, takes a left across the top of the diagram while ejecting a planetary nebula, and finally falls to its final resting place in the bottom left of the diagram. This process is representative of the fate of low-mass stars. Although every low-mass star forms a white dwarf at the end point of its evolution, the precise path a low-mass star follows from core hydrogen burning on the main sequence to white dwarf depends on many details particular to the star.

Some stars less massive than the Sun may become white dwarfs composed largely of helium rather than carbon. Conversely, temperatures in the cores of evolved 2- to 3-M_{Sun} stars are high enough to allow additional nuclear reactions to occur, leading to the formation of somewhat more massive white dwarfs composed of materials such as oxygen, neon, and magnesium. Differences in chemical composition of a star can also lead to differences in its post-main-sequence evolution.

In our discussion of stellar evolution, we have focused on what happens after a star leaves the main sequence. The spectacle of a red giant or AGB star is ephemeral. Once it leaves the main sequence, the Sun will travel the path from red giant to white dwarf in less than one-tenth of the time it spent on the main sequence steadily burning hydrogen to helium in its core. Stars spend most of their luminous lifetimes on the main sequence, which is why most of the stars in the sky are main-sequence stars. The fainter white dwarfs constitute the final resting place for the vast majority of stars that have been or ever will be formed.

The Fate of the Planets

What happens to the planets orbiting a low-mass star as it goes through these post-main-sequence stages? Because the first decade of extrasolar planet discoveries primarily yielded planets with orbits that are very close to their respective stars, astronomers assumed that any planets closer than 1–2 astronomical units (AU) would not survive the post-main-sequence expansion of the star. Therefore, they did not expect to find planets in orbit around evolved stars. However, many planets have been discovered in orbit around red giants, AGB stars, and horizontal branch stars.

Astronomers cannot be sure whether the extrasolar planets they observe have remained in the same orbital locations as when their stars were on the main sequence or if the planets migrated to different orbits. The surface gravity of a red giant is low, and some of the mass in its outermost layers will blow away. This decrease in mass reduces the gravitational force between the star and its planets, which could lead to planetary orbits evolving outward away from the star. There are some models of planetary migration in which the tidal forces between planets or between a star and a planet are significant factors too, suggesting that some planetary orbits could evolve inward.

As a star loses mass during the evolutionary process, stellar or planetary companions may change their orbits. Rocky asteroids or smaller planets could migrate inward past the Roche limit (see Chapter 4) of the white dwarf and break up. Some of the material could remain in orbit around the white dwarf as a debris disk, similar to those seen in main-sequence stars with a planetary system. These dusty debris disks have been observed around white dwarf stars with the Hubble Space Telescope and the Spitzer Space Telescope. Some of this dusty or rocky material may also fall onto the white dwarf, "polluting" its spectrum with heavy elements that were not produced in the stellar core.

What will happen to Earth? The age of the Sun from radioactive dating of meteorites indicates that the Solar System formed 4.6 billion years ago. In Chapter 14, we estimated how long the Sun might last by calculating its rate of hydrogen burning. The Sun's luminosity is about 30 percent higher now than it was early in the history of the Solar System, and it will continue to increase at a steady rate over the rest of the Sun's main-sequence lifetime of another 5 billion years or so. Models estimate that the Sun's luminosity may increase enough—even while the Sun is a main-sequence star—that Earth will heat up to the point where the oceans evaporate, perhaps as soon as 1 billion to 2 billion years from now. By the time the Sun leaves the main sequence, the habitable zone may have moved out to Mars and no longer include Earth.

It is not certain whether the radius of the red giant Sun will extend past 1 AU so that Earth becomes completely engulfed. The red giant Sun will have low surface gravity and will lose mass, which could cause Earth's orbit to expand, thereby enabling Earth to escape the encroaching solar surface. Alternatively, as the Sun expands in radius and its rotation rate slows (see Working It Out 7.1), tidal forces might pull Earth inward. The habitable zone of a red giant Sun might be in the vicinity of Jupiter or Saturn. Eventually, depending on how much mass the Sun loses in the red giant and possibly AGB stages, the outer layers may or may not form a planetary nebula. The solar core will become a white dwarf, perhaps with a dusty disk and a "polluted" atmosphere as the only remaining evidence of our rocky planet.

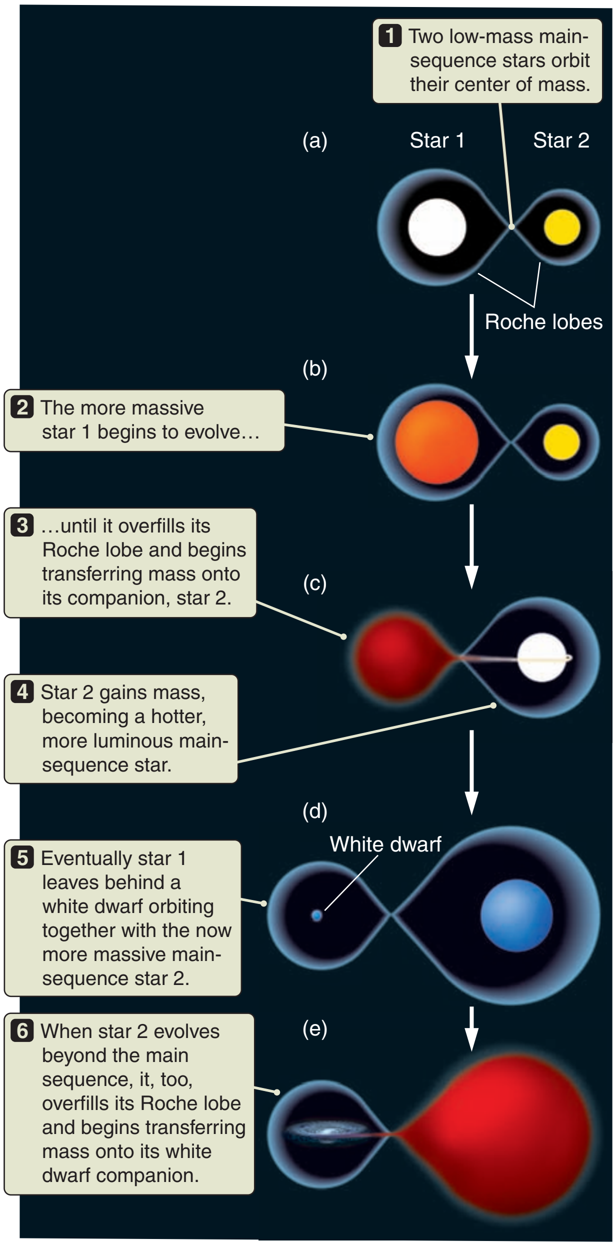


Figure 16.14 A compact binary system consisting of two low-mass stars passes through a sequence of stages as the stars evolve and mass is transferred back and forth.

CHECK YOUR UNDERSTANDING 16.4

A planetary nebula forms from: (a) the ejection of mass from a low-mass star; (b) the collision of planets around a dying star; (c) the collapse of the magnetosphere of a high-mass star; (d) the remainders of the original star-forming nebula.

16.5 Binary Star Evolution

So far in this chapter, we have discussed the evolution of single stars in isolation. But many stars are members of binary systems. How is evolution different for stars in these systems? In some cases, if the stars are close together and one star is more massive than the other, their evolution after the main sequence may be linked. In this section, we will trace the steps from binary star system through to the end point of their evolution: nova or supernova.

Mass Flows from an Evolving Star onto Its Companion

Think for a moment about what would happen if you were to travel in a spacecraft from Earth toward the Moon. When you are still near Earth, the force of Earth's gravity is far stronger than that of the Moon. As you move away from Earth and closer to the Moon, the gravitational attraction of Earth weakens, and the gravitational attraction of the Moon becomes stronger. You eventually reach an intermediate zone where neither body has the stronger pull. If you continue beyond this point, the lunar gravity begins to dominate until you find yourself firmly in the gravitational grip of the Moon. The regions surrounding the two objects—their gravitational domains—are called the **Roche lobes** of the system.

Exactly the same situation exists between two stars, as shown in **Figure 16.14**. Gas near each star clearly belongs to that star. When one star leaves the main sequence and swells up, its outer layers may cross that gravitational dividing line separating the star from its companion. Once a star expands past the boundary of its Roche lobe, its material begins to fall onto the other star. This exchange of material from one star to the other is called **mass transfer**.

Evolution of a Close Binary System

The best way to understand how mass transfer affects the evolution of stars in a binary system is to apply what is known from studying the evolution of single low-mass stars. Figure 16.14a shows a binary system consisting of two low-mass stars of somewhat different mass. Star 1 is the more massive of the two stars, and star 2 is the less massive of the two stars. This is an ordinary binary system, and each of these stars is an ordinary main-sequence star for most of the system's lifetime.

More massive main-sequence stars evolve more rapidly than less massive main-sequence stars. Therefore, star 1 will be the first to use up the hydrogen at its center and begin to evolve off the main sequence, as shown in Figure 16.14b. If the two stars are close enough to each other, star 1 will eventually grow to overfill its Roche lobe, and material will transfer onto star 2, as shown in Figure 16.14c. The transfer of mass between the two stars can result in a sort of "drag" that causes the orbits of the two stars to shrink, bringing the stars closer together and further enhancing mass loss. In addition, as star 1 loses mass, its Roche lobe shrinks, further enhancing the mass transfer. The two stars can even reach the point where they are effectively two cores sharing the same extended envelope of material.

Despite these complexities, star 2 probably remains a basically normal main-sequence star throughout this process, burning hydrogen in its core. However, over time, the mass of star 2 increases because of the accumulation from its companion. As it does so, the structure of star 2 must change to accommodate its new status as a higher-mass star. If we plotted star 2's position on the H-R diagram during this period, we would see it move up and to the left along the main sequence, becoming larger, hotter, and more luminous.

Star 1, because it is losing mass to star 2, never grows larger than its Roche lobe, so it does not become an isolated red giant or AGB star at the top of the H-R diagram. Yet star 1 continues to evolve, burning helium in its core on the horizontal branch, proceeding through a stage of helium shell burning, and finally losing its outer layers and leaving behind a white dwarf. Figure 16.14d shows the binary system after star 1 has completed its evolution. All that remains of star 1 is a white dwarf, orbiting about star 2, its bloated main-sequence companion.

Novae

As star 2 begins to evolve off the main sequence, it expands to fill its Roche lobe, as shown in Figure 16.14e. Like star 1 before it, star 2 grows to fill its Roche lobe: material from star 2 begins to pour through the "neck" connecting the Roche lobes of the two stars. However, this time the mass is not being added to a normal star but is drawn toward the tiny white dwarf left behind by star 1. Because the system is revolving and the white dwarf is so small, the infalling material generally misses the star, instead landing on an accretion disk around the white dwarf. This disk is similar to the accretion disk that forms around a protostar. As in the process of star formation, the accretion disk accumulates material that has too much angular momentum to hit the white dwarf directly.

A white dwarf has a large mass and a small radius; therefore, it has strong gravity. The material streaming toward the white dwarf in the binary system falls into an incredibly deep gravitational "well." The depth of this well affects the amount of energy with which matter impacts the white dwarf. A kilogram of material falling from space onto the surface of a white dwarf releases 100 times more energy than a kilogram of material falling from the outer Solar System onto the surface of the Sun. All of this energy is turned into thermal energy. The spot where the stream of material from star 2 hits the accretion disk can be heated to millions of kelvins, where it glows in the far-ultraviolet and X-ray parts of the electromagnetic spectrum.

The infalling material accumulates on the surface of the white dwarf (**Figure 16.15a**), where it is compressed by the enormous gravitational pull of the white dwarf to a density close to that of the white dwarf itself. As more and more material builds up on the surface of the white dwarf, the white dwarf shrinks (just as

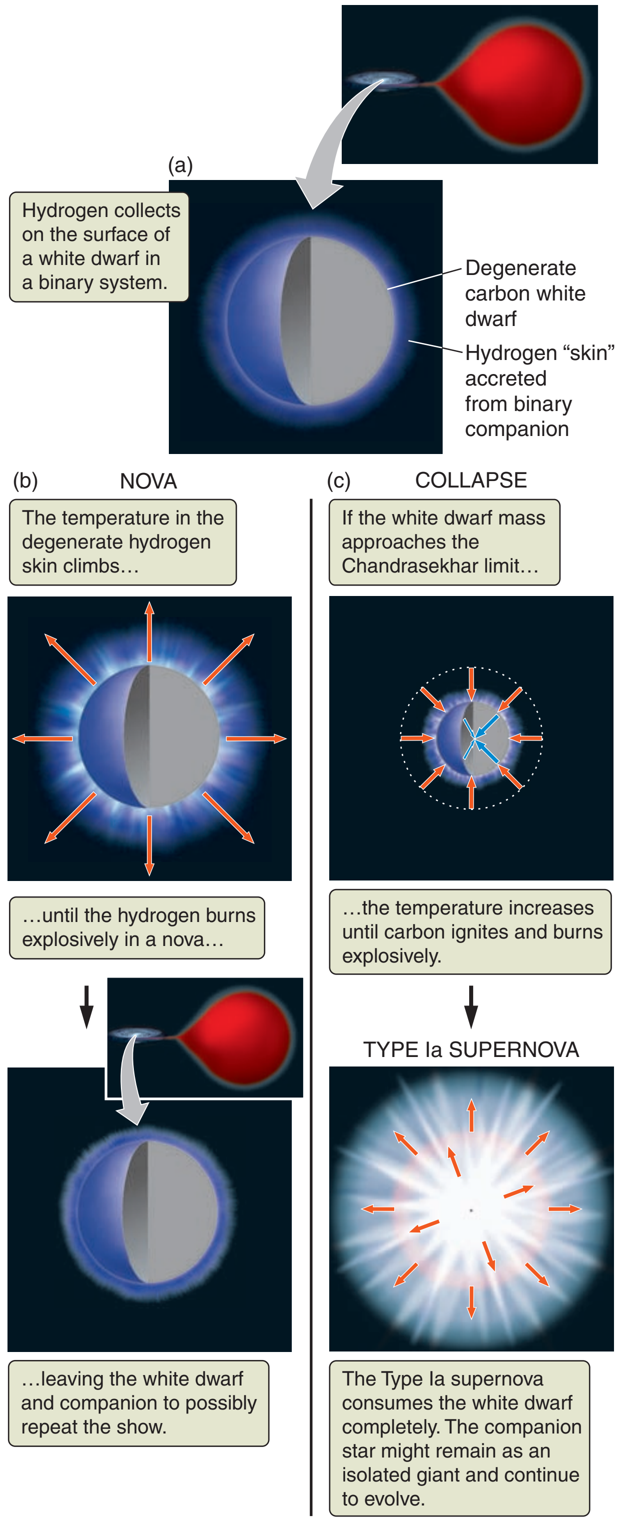


Figure 16.15 (a) In a binary system in which mass is transferred onto a white dwarf, a layer of hydrogen builds up on the surface of the degenerate white dwarf. (b) If hydrogen burning ignites on the surface of the white dwarf, the result is a nova. (c) If enough hydrogen accumulates to raise the core temperature high enough, carbon ignites and the result is a Type Ia supernova.

the core of a red giant shrinks as it grows more massive). The density increases more and more, and at the same time the release of gravitational energy drives the temperature of the white dwarf higher and higher. The infalling material comes from the outer, unburned layers of star 2, so it is composed mostly of hydrogen. Hydrogen, the best nuclear fuel around, is being compressed to higher and higher densities and heated to higher and higher temperatures on the surface of the white dwarf.

Once the temperature at the base of the white dwarf's surface layer of hydrogen reaches about 10 million K, this hydrogen begins to fuse to helium. But this is not the contained hydrogen burning that takes place in the center of the Sun; this is explosive hydrogen burning in a degenerate gas. Energy released by hydrogen burning drives up the temperature. Because the surface is degenerate, this rising temperature does not cause an expansion, as in normal matter, but instead drives up the rate of hydrogen burning. This runaway thermonuclear reaction is much like the runaway helium fusion that takes place during the helium flash, except now there are no overlying layers of a star to absorb the energy liberated by fusion. The result is a tremendous explosion that blows part of the layer covering the white dwarf out into space at speeds of thousands of kilometers per second, as shown in Figure 16.15b. An exploding white dwarf of this kind is called a **nova**.

The explosion of a nova does not destroy the underlying white dwarf star. In fact, much of the material that had built up on the white dwarf may remain behind after the explosion. Afterward, the binary system is in much the same configuration as before: material from star 2 is still pouring onto the white dwarf, as seen in Figure 16.15a. The nova can repeat many times, as material builds up and ignites again and again on the surface of the white dwarf. If the underlying white dwarf is old and cooler, or the mass accumulates slowly on its surface, outbursts are separated by thousands of years, so most novae have been seen only once in historical times. But if the white dwarf is hotter, or mass is transferring quickly, such explosions can happen every few years or even every few months. These recurring novae are called dwarf novae because they are generally less luminous than those that occur less frequently.

About 50 novae occur in our galaxy each year, but we observe only some of them because dust and gas obscure the view from Earth. Novae get bright quickly—typically reaching their peak brightness in only a few hours—and for a brief time they can be several hundred thousand times more luminous than the Sun. Although the brightness of a nova sharply declines in the weeks after the outburst, it can sometimes still be seen for years. During this time, the glow from the expanding cloud of ejected material is powered by the decay of radioactive isotopes created in the explosion. Large bursts of gamma rays have been observed to be associated with several novae.

Supernovae

Novae are spectacular events, but they pale in comparison with another observed phenomenon: the supernova. A **supernova** is an extraordinary nova—a newly visible star that is much more luminous than a nova. Tycho Brahe observed one in 1572, and Kepler observed one in 1604. Supernovae occur in several types, which are distinguished by both their spectra and the way in which they brighten and fade over time. The peak luminosity of a **Type Ia supernova** can be inferred from the rate at which the brightness changes after the explosion. This peak luminosity can then be combined with the peak apparent brightness to find the distance to the supernova. Type Ia supernovae occur in a galaxy the size of the Milky Way about once a century on average. For a brief time they can shine with a luminosity billions of times that of the Sun, in some cases outshining the galaxy itself. Type Ia supernovae are so luminous that they can be seen at great distances, so they can be useful for estimating distances to very distant galaxies. The origins of Type Ia supernovae lie in the evolution of binary systems, and so we discuss them here. Type II supernovae will be discussed in the next chapter.

There is a limit to how massive a white dwarf can be, called the **Chandrasekhar limit**. Above this mass, gravity is stronger than the pressure supplied by degenerate electrons, and the white dwarf collapses. As a white dwarf in a binary system gains mass and approaches the Chandrasekhar limit, the increase in pressure and density causes the temperature of the core to rise. When the core temperature reaches about 6×10^8 K, carbon fusion begins. The star may then have a "simmering" phase with a growing central convective region that prevents thermonuclear runaway for a while. After about 1,000 years, when the temperature reaches about 8×10^8 K, runaway carbon burning begins throughout the entire white dwarf, and the star explodes (Figure 16.15c). The white dwarf is consumed within about a second. This likely happens at about 1.38 M_{Sun}, just before the white dwarf actually reaches the ultimate mass limit. The luminosity of the explosion is about 5 billion times the luminosity of the Sun. Runaway fusion reactions convert a large fraction of the mass of the star into elements such as iron and nickel, and the explosion blasts the remains of the white dwarf into space at speeds in excess of 20,000 km/s, enriching the interstellar medium with these heavier elements. Explosive carbon burning in a white dwarf is the leading theory used to explain Type Ia supernovae.

In this scenario, the explosion completely destroys star 1 of the binary system, but star 2 can be left behind to continue its evolution. Surveys at different wavelengths have found companions to very few Type Ia supernovae, so some scientists think that another process may be responsible for as many as 80 percent of these supernovae. For example, eventually star 2 will go on to form a white dwarf too, leaving behind a binary system consisting of two degenerate white dwarfs. These two white dwarfs might orbit around each other quickly, getting closer and closer. Eventually, tidal forces will disrupt the smaller one, and material will fall onto the larger one until they merge, creating one object with a larger mass, as shown in **Figure 16.16**. In this case, the massive white dwarf came from the merger of two smaller white dwarfs, and so a supernova explosion destroys both stars. This could be why astronomers see so few companions remaining. However, it is also possible that the companions are "missing" because the explosion of one star will eject the other from the system, making it difficult to find. This topic is still an active area of exploration in astronomy (**Process of Science Figure**).

These explosions leave behind expanding shells of dust and gas called **supernova remnants**, as shown in **Figure 16.17**. The material is heated by the expanding blast wave from the supernova and glows in X-rays.

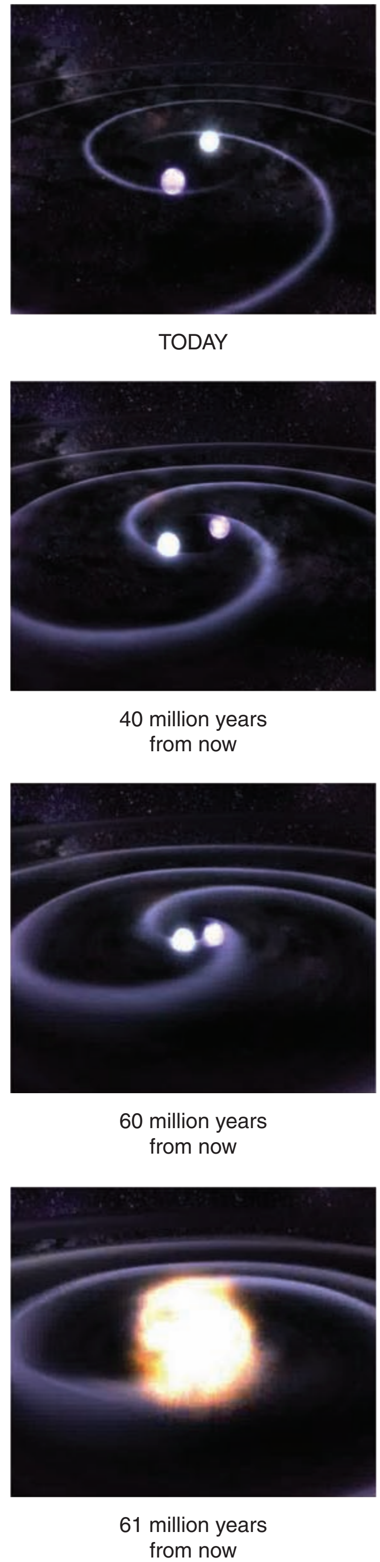


Figure 16.16 Two white dwarfs in close orbit about each other get closer and move faster over time, eventually merging into a single, more massive white dwarf. If the mass of the combined white dwarfs is above the Chandrasekhar limit, it will explode as a Type Ia supernova.

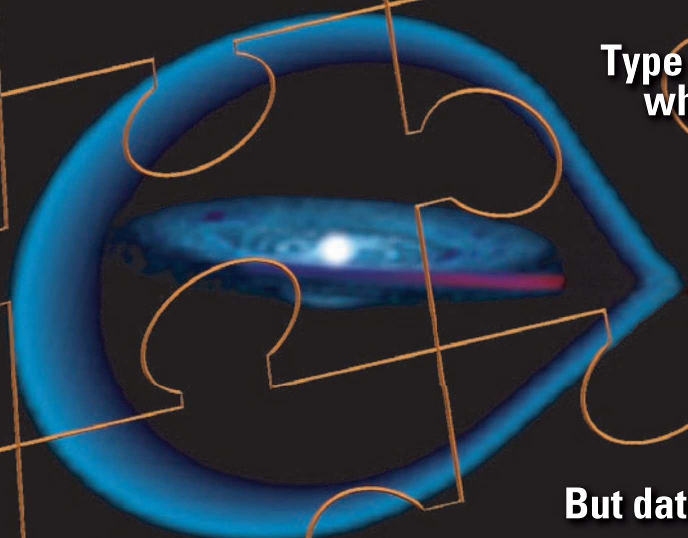
CHECK YOUR UNDERSTANDING 16.5

A white dwarf will become a supernova if: (a) the original star was more than 1.38 M_{Sun}; (b) it accretes an additional 1.38 M_{Sun} from a companion; (c) some mass falls on it from a companion; (d) enough mass accretes from a companion to give the white dwarf a total mass of 1.38 M_{Sun}.

SCIENCE IS NOT FINISHED

Understanding Type Ia supernova is critical for measuring distances to the farthest galaxies and for our conclusions about the universe as a whole. Yet the observational evidence remains incomplete.

Supernovae have been observed since antiquity. As technology developed, it became clear there are several types of supernovae.



Type Ia supernova are thought to come from white dwarfs which accumulate mass.

?

But data suggest…

If the white dwarf accretes mass from a red giant or other large star, and explodes at exactly 1.38 M_{Sun}, then the peak luminosity of the explosion should be the same for all Type Ia supernovae.

But searches for potential companions that survived the explosion have found only a few candidates

If the explosion arises from 2 white dwarfs merging together, then the mass of the explosion is variable and the peak luminosity may also vary.

But searches for binary white dwarfs suggest they are not very common

When observational evidence is inconclusive, each scientist adds a piece to the puzzle. Some will conduct larger observational studies, and others will create new theoretical models.

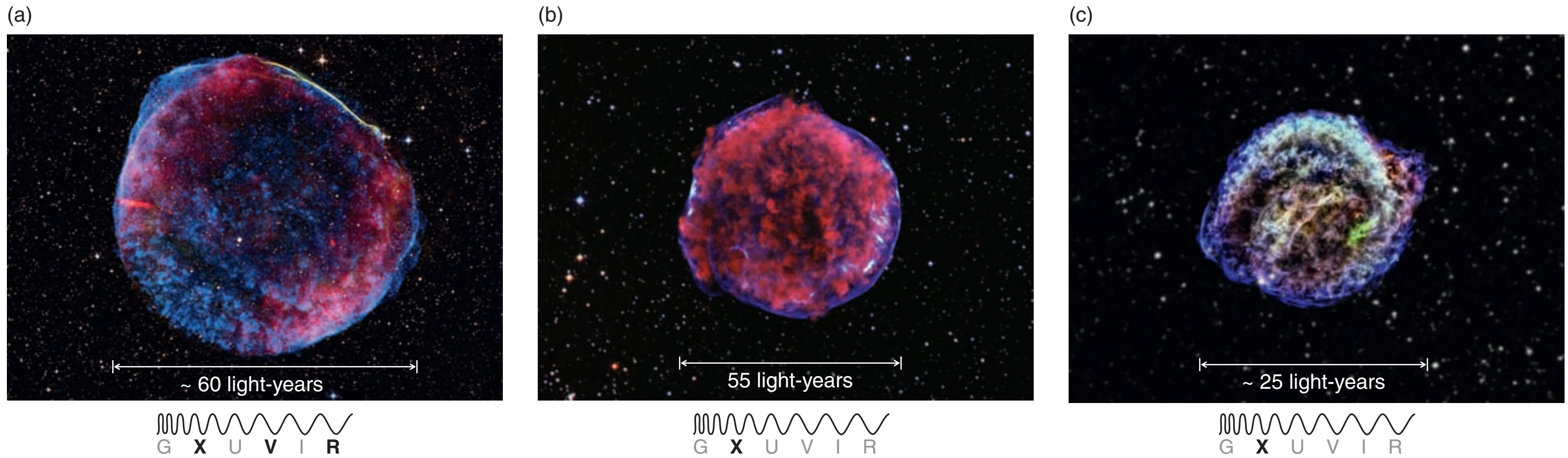


Figure 16.17 These images show Type Ia supernovae. The material heated by the expanding blast wave from a supernova glows in the X-ray portion of the spectrum. (a) SN 1006 is the brightest recorded supernova and was observed in China, Japan, Europe, and the Middle East in 1006. X-ray data are shown in blue; radio data are shown in red. (b) Tycho's supernova was observed in 1572. Low-energy X-rays are red; high-energy X-rays are blue. (c) Kepler's supernova, observed in 1604, is shown in five different X-ray wavelengths.

Origins

Stellar Lifetimes and Biological Evolution

From fossil records and DNA analysis, scientists estimate that life appears to have taken hold on Earth within 1 billion years after the Solar System and Earth formed 4.6 billion years ago. It took another 1.5 billion years for more complex cells to develop and another billion years to develop multicellular life. The first animals didn't appear on Earth until 600 million years ago, 4 billion years after the formation of the Sun and the Solar System.

The only known example of biology is the life on Earth. It is always risky to extrapolate from one data point, and it is not known whether biology is widespread in the universe or whether Earth's biological timeline is "typical." Still, reasoning from this one example is the only way to begin thinking about life in the universe. How does the preceding timeline of the evolution of life on Earth compare with the lifetimes of main-sequence stars? Table 16.1 indicates that the lifetime of an O5 star is less than half a million years; of a B5 star, about 120 million years; and of an A0 star, about 700 million years. These stars would have run out of hydrogen in the core and started post-main-sequence evolution in less time than it took for Earth to settle down after its periods of heavy bombardment by debris early in the Solar System.

The 1-billion-year main-sequence lifetime of an A5 star corresponds to the amount of time it took for the simplest life-forms to develop on Earth. The 3-billion-year lifetime for F0 stars corresponds to the amount of time it took for photosynthetic bacteria to develop on Earth. Only stars cooler and less massive than F5 stars have main-sequence lifetimes longer than the 4 billion years it took for life to evolve into animals on Earth. Thus, searches for extrasolar planets survey stars that are F5 or cooler—because hotter and more massive stars probably don't live long enough on the main sequence for complex life to develop.

After a star leaves the main sequence, the helium-burning red giant stage is estimated to last for about 1/10 of the main-sequence lifetime, so this doesn't help stars with short lifetimes to last long enough for complex biology. Could life survive the transition of its star to a red giant? As noted earlier in the chapter, even if a planet is not destroyed, its orbit, temperature, and atmospheric conditions will drastically change, and any life might have to relocate if it is to survive.

ARTICLES QUESTIONS

Scientists report that the spectra of some dead stars show evidence of rocky material left over from planetary systems.

Scientists Solve Riddle of Celestial Archaeology

By **University of Leicester Press Office**

A decades old space mystery has been solved by an international team of astronomers led by Professor Martin Barstow of the University of Leicester and President-elect of the Royal Astronomical Society.

Scientists from the University of Leicester and University of Arizona investigated hot, young, white dwarfs—the super-dense remains of Sun-like stars that ran out of fuel and collapsed to about the size of the Earth. Their research is featured in MNRAS—the *Monthly Notices of the Royal Astronomical Society*, published by Oxford University Press.

It has been known that many hot white dwarf atmospheres, essentially of pure hydrogen or pure helium, are contaminated by other elements—like carbon, silicon, and iron. What was not known, however, was the origins of these elements, known in astronomical terms as metals.

"The precise origin of the metals has remained a mystery and extreme differences in their abundance between stars could not be explained," said Professor Barstow, a Pro-Vice-Chancellor at the University of Leicester whose research was assisted by his daughter Jo, a coauthor of the paper, during a summer work placement in Leicester. She has now gone on to be an astronomer working in Oxford—on extrasolar planets.

"It was believed that this material was 'levitated' by the intense radiation from deeper layers in the star," said Professor Barstow.

Now the researchers have discovered that many of the stars show signs of contamination by rocky material, the leftovers from a planetary system.

The researchers surveyed 89 white dwarfs, using the Far Ultraviolet Spectroscopic Explorer to obtain their spectra (dispersing the light by color) in which the "fingerprints" of carbon, silicon, phosphorus, and sulfur can be seen when these elements are present in the atmosphere.

"We found that in stars with polluted atmospheres the ratio of silicon to carbon matched that seen in rocky material, much higher than found in stars or interstellar gas.

"The new work indicates that around one-third of all hot white dwarfs are contaminated in this way, with the debris most likely in the form of rocky minor planet analogs. This implies that a similar proportion of stars like our Sun, as well as stars that are a little more massive like Vega and Fomalhaut, build systems containing terrestrial planets. This work is a form of celestial archaeology where we are studying the 'ruins' of rocky planets and/or their building blocks, following the demise of the main star.

"The mystery of the composition of these stars is a problem we have been trying to solve for more than 20 years. It is exciting to realize that they are swallowing up the leftovers from planetary systems, perhaps like our own, with the prospect that more detailed follow-up work will be able to tell us about the composition of rocky planets orbiting other stars," said Professor Barstow.

The study also points to the ultimate fate of the Earth billions of years from now—ending up as a contamination within the white dwarf Sun.

ARTICLES QUESTIONS

1. This communication refers to the white dwarfs in the study as "hot, young, white dwarfs." What does "young" mean in this context?
2. The spectra described are compared to fingerprints. In what ways are white dwarf spectra like fingerprints?
3. Why were scientists surprised to find elements other than hydrogen and helium in the atmospheres of white dwarf stars? Where do they think these other elements originate?
4. Why does a telescope need to be in space to observe far-ultraviolet wavelengths?
5. How common are contaminated white dwarfs in the sample in this study? Compare this to the percentage of Sun-like stars with planets (22 percent). Does the finding in this communication seem like a sensible number?

Summary

Stars with masses similar to the Sun form planetary nebulae like the one shown in the chapter-opening photograph. This is a likely fate for our own Sun at the end of its evolution. After leaving the main sequence, low-mass stars follow a convoluted path along the H-R diagram that includes the red giant branch, the horizontal branch, the asymptotic giant branch, and a path across the top and then down to the lower left of the diagram. These stages of evolution are dominated by the balance between gravity and energy production from various fusion processes that "turn on and off" in the core of the star. This entire process takes much less time than the main-sequence lifetime of the star. Stars that follow this path have main-sequence lifetimes comparable to the evolutionary timescales of life on Earth. If life on Earth is typical, then more massive stars with shorter lifetimes will not be stable long enough for complex life to evolve.

LG 1 **Estimate the main-sequence lifetime of a star from its mass.** All stars eventually exhaust their nuclear fuel as hydrogen fuses to helium in the cores of main-sequence stars. Less massive stars exhaust their fuel more slowly and have longer lifetimes than more massive stars.

LG 2 **Explain why low-mass stars grow larger and more luminous as they run out of fuel.** When a low-mass star uses up the hydrogen in its core, it begins to burn hydrogen in a shell around the core, heating the gaseous interior. The star expands in size and becomes a red giant.

LG 3 **Sketch post-main-sequence evolutionary tracks on a Hertzsprung-Russell (H-R) diagram, and list the stages of evolution for low-mass stars.** After exhausting its hydrogen, a low-mass star leaves the main sequence and swells to become a red giant, with a helium core made of electron-degenerate matter. The red giant fuses helium to carbon via the triple-alpha process, and quickly the core ignites in a helium flash. The star then moves onto the horizontal branch. A horizontal branch star accumulates carbon and sometimes oxygen in its core and then moves up the asymptotic giant branch. The star may become an AGB star and lose some of its mass.

LG 4 **Describe how planetary nebulae and white dwarfs form.** In their dying stages, some stars eject their outer layers to form planetary nebulae. All low-mass stars eventually become white dwarfs, which are very hot but very small.

LG 5 **Explain how some close binary systems evolve differently than single stars.** Transfer of mass within some binary systems can lead to a nuclear explosion. A nova occurs when hydrogen collects and ignites on the surface of a white dwarf in a binary system. If the mass of the white dwarf approaches 1.38 M_{Sun}, the entire star explodes in a Type Ia supernova.

? UNANSWERED QUESTIONS

- Why do planetary nebulae have different shapes? Some are not simply chaotic but are well organized, with varying types of symmetry. Some are spherically symmetric (like a ball), some have bipolar symmetry (like a long hollow tube, pinched in the middle), and some are even point-symmetric (like the letter *S*). How can an essentially spherically symmetric object such as a star produce such beautifully organized outflows? Because these are three-dimensional objects, and astronomers can view each object from only one direction, it is difficult to determine how much of this variation is due to orientation and how much is due to actual differences in the shape of the object. For example, a bipolar nebula, viewed from one end, would appear spherically symmetric. This orientation effect, among other problems, complicates efforts to understand how these shapes are formed. No single explanation has yet satisfactorily covered all the object types.
- Could Earth be moved farther from the Sun to accommodate the Sun's inevitable changes in luminosity, temperature, and radius? One proposal suggests that Earth could capture energy from a passing asteroid and migrate outward, thus staying in the habitable zone while moving farther from the Sun as the Sun ages. Or a huge, thin "solar sail" could be constructed so that radiation pressure from the Sun would slowly push Earth into a larger orbit. These feats of "astronomical engineering" are not feasible anytime in the near future, but perhaps in millions of years (when they will be needed) this could be accomplished.

Questions and Problems

Test Your Understanding

1. Place the main-sequence lifetimes of the following stars in order from shortest to longest.
   a. the Sun: mass 1 M_{Sun}, luminosity 1 L_{Sun}
   b. Capella Aa: mass 3 M_{Sun}, luminosity 76 L_{Sun}
   c. Rigel: mass 24 M_{Sun}, luminosity 85,000 L_{Sun}
   d. Sirius A: mass 2 M_{Sun}, luminosity 25 L_{Sun}
   e. Canopus: mass 8.5 M_{Sun}, luminosity 13,600 L_{Sun}
   f. Achernar: mass 7 M_{Sun}, luminosity 3,150 L_{Sun}
2. Place the following steps in the evolution of a low-mass star in order.
   a. main-sequence star
   b. planetary nebula ejection
   c. horizontal branch
   d. helium flash
   e. red giant branch
   f. asymptotic giant branch
   g. white dwarf
3. If a star follows a horizontal path across the H-R diagram, the star
   a. maintains the same temperature.
   b. stays the same color.
   c. maintains the same luminosity.
   d. keeps the same spectral type.
4. Degenerate matter is different from normal matter because as the mass goes up,
   a. the radius goes down.
   b. the temperature goes down.
   c. the density goes down.
   d. the luminosity goes down.
5. The most massive stars have the shortest lifetimes because
   a. the temperature is higher in the core, so they burn their fuel faster.
   b. they have less fuel in the core when the star forms.
   c. their fuel is located farther from the core.
   d. the temperatures are lower in the core, so they burn their fuel slower.
6. If a main-sequence star suddenly started burning hydrogen at a faster rate in its core, it would become
   a. larger, hotter, and more luminous.
   b. larger, cooler, and more luminous.
   c. smaller, hotter, and more luminous.
   d. smaller, cooler, and more luminous.
7. It is rare to see a helium flash because
   a. few stars go through this stage.
   b. stars that go through this stage are all far away.
   c. the flash only glows at infrared wavelengths.
   d. the flash does not take very long.
8. Post-main-sequence stars lose up to 50 percent of their mass because
   a. jets from the poles release material at an increasing rate.
   b. the mass of the star drops because of mass loss from fusion.
   c. the magnetic field causes increasing numbers of coronal mass ejections.
   d. the star swells until the surface gravity is too weak to hold material.
9. A planetary nebula glows because
   a. it is hot enough to emit UV radiation.
   b. fusion is happening in the nebula.
   c. it is heating up the interstellar medium around it.
   d. light from the central star causes emission lines.
10. As an AGB star evolves into a white dwarf, it runs out of nuclear fuel, and one might guess that the star should cool off and move to the right on the H-R diagram. Why does the star move instead to the left?
    a. It becomes larger.
    b. More of the star is involved in fusion.
    c. As outer layers are lost, deeper layers are exposed.
    d. The temperature of the core rises.
11. When compressed, ordinary gas heats up but degenerate gas does not. Why, then, does a degenerate core heat up as the star continues shell burning around it?
    a. It is heated by the radiation from fusion.
    b. It is heated by the gravitational collapse of the shell.
    c. It is heated by the weight of helium falling on it.
    d. It is insulated by the shell.
12. All Type Ia supernovae
    a. are at the same distance from Earth.
    b. always involve two stars of identical mass.
    c. have identical peak luminosities.
    d. always release the same amount of energy in fusion.
13. In Latin, *nova* means "new." This word is used for novae and supernovae because they are
    a. newly formed stars.
    b. newly dead stars.
    c. newly visible stars.
    d. new main-sequence stars.
14. When the Sun runs out of hydrogen in its core, it will become larger and more luminous because
    a. it will start fusing hydrogen in a shell around a helium core.
    b. it will start fusing helium in a shell and hydrogen in the core.
    c. infalling material will rebound off the core and puffs up the star.
    d. energy balance will no longer hold, and the star will drift apart.

15. A white dwarf is located in the lower left of the H-R diagram. From this information alone, you can determine that
    a. it is very massive.
    b. it is very dense.
    c. it is very hot.
    d. it is very bright.

Thinking about the Concepts

16. Is it possible for a star to skip the main sequence and immediately begin burning helium in its core? Explain your answer.

17. Suppose a main-sequence star suddenly started burning hydrogen at a faster rate in its core. How would the star react? Discuss changes in size, temperature, and luminosity.

18. Describe some possible ways in which the temperature in the core of a star might increase while the density decreases.

19. Astronomers typically say that the mass of a newly formed star determines its destiny from birth to death. However, there is a common environmental circumstance for which this statement is not true. Identify this circumstance and explain why the birth mass of a star might not fully account for the star's destiny.

20. Study the Process of Science Figure. Suppose that a new mechanism is found to explain Type Ia supernovae. In this mechanism, all Type Ia supernovae are more luminous than previously thought. Would the derived distances to galaxies be larger or smaller than we currently understand them to be?

21. Do stars change structure while on the main sequence? Why or why not?

22. Suppose Jupiter is not a planet but a G5 main-sequence star with a mass of 0.8 M_{Sun}.
    a. How will life on Earth be affected, if at all?
    b. How will the Sun be affected as it comes to the end of its life?

23. Explain the similarity in the paths that a star follows along the H-R diagram as it forms from a protostar and as it leaves the main sequence to climb the red giant branch?

24. Why is a horizontal branch star (which burns helium at a high temperature) less luminous than a red giant branch star (which burns hydrogen at a lower temperature)?

25. Suppose the core temperature of a star is high enough for the star to begin fusing oxygen. Predict how the star will continue to evolve, including its path on the H-R diagram.

26. Why does a white dwarf move down and to the right along the H-R diagram?

27. Why does fusion in degenerate material always lead to a runaway reaction?

28. Suppose the more massive red giant star in a binary system engulfs its less massive main-sequence companion, and their nuclear cores combine. What structure will the new star have? Where will the star lie on the H-R diagram?

29. T Coronae Borealis is a well-known recurrent nova.
    a. Is it a single star or a binary system? Explain.
    b. What mechanism causes a nova to flare up?
    c. How can a nova flare-up happen more than once?

30. Why do astronomers prefer to search for planets around low-mass stars?

Applying the Concepts

31. Figure 16.1 contains the text "Straight-line (exponential) approximation." What does this text tell you about the axes on the graph: are they linear or logarithmic? Explain why these data are plotted this way.

32. Use Figure 16.2 to estimate the percentage of the Sun's mass that is turned from hydrogen into helium over its lifetime.

33. Study Figure 16.13. How many times brighter is a star at the top of the giant branch than the same star (a) when it was on the main sequence? and (b) when it was on the horizontal branch?

34. Study Figure 16.13. Make a graph of surface temperature versus time for the evolutionary track shown—from the time the star leaves the main sequence until it arrives at the dot showing that it is a white dwarf. Your time axis may be approximate, but it should show that the star spends different amounts of time in the different phases.

35. For most stars on the main sequence, luminosity scales with mass as $M^{3.5}$ (see Working It Out 16.1). What luminosity does this relationship predict for (a) 0.5-M_{Sun} stars, (b) 6-M_{Sun} stars, and (c) 60-M_{Sun} stars? Compare these numbers to values given in Table 16.1.

36. Calculate the main-sequence lifetimes for (a) 0.5-M_{Sun} stars, (b) 6-M_{Sun} stars, and (c) 60-M_{Sun} stars. (See Working It Out 16.1.) Compare them to the values given in Table 16.1.

37. What will the escape velocity be when the Sun becomes an AGB star with a radius 200 times greater and a mass only 0.7 times that of today? How will these changes in escape velocity affect mass loss from the surface of the Sun as an AGB star?

38. Each form of energy generation in stars depends on temperature.
    a. The rate of hydrogen fusion (proton-proton chain) near 10^7 K increases with temperature as T^4. If the temperature of the hydrogen-burning core is raised by 10 percent, how much does the hydrogen fusion energy increase?
    b. Helium fusion (the triple-alpha process) at 10^8 K increases with an increase in temperature at a rate of T^{40}. If the temperature of the helium-burning core is raised by 10 percent, how much does the helium fusion energy increase?

39. A planetary nebula has an expansion rate of 20 km/s and a lifetime of 50,000 years. Roughly how large will this planetary nebula grow before it disperses?

40. Suppose a companion star transferred mass onto a white dwarf at a rate of about $10^{-9}\ M_{\text{Sun}}$ per year. Roughly how long after mass transfer begins will the white dwarf explode as a Type Ia supernova? How does this length of time compare to the typical lifetime of a low-mass star? Assume that the white dwarf started with a mass of 0.6 M_{Sun}.

41. Use Kepler's third law to estimate how fast material in an accretion disk orbits around a white dwarf.

42. A white dwarf has a density of approximately 10^9 kilograms per cubic meter (kg/m^3). Earth has an average density of 5,500 kg/m^3 and a diameter of 12,700 km. If Earth were compressed to the same density as a white dwarf, what would its radius be?

43. What is the density of degenerate material? Calculate how large the Sun would be if all of its mass were degenerate.

44. Recall from Chapter 5 that the luminosity of a spherical object at temperature T is given by $L = 4\pi R^2 \sigma T^4$, where R is the object's radius. If the Sun became a white dwarf with a radius of 10^7 meters, what would its luminosity be at the following temperatures: (a) 10^8 K; (b) 10^6 K; (c) 10^4 K; (d) 10^2 K?

45. According to current astronomical evidence, our universe is approximately 13.8 billion years old. If this is correct, what is the least mass that a star could possess in order to have already evolved into a red giant? What spectral type of star is this?

USING THE WEB

46. Go to the website for the Katzman Automatic Imaging Telescope at Lick Observatory (http://astro.berkeley.edu/bait/public_html/kait.html). What is this project? Why can a search for supernovae be automated? Pick a recent year. How many supernovae were discovered? Look at some of the images. How bright do the supernovae look compared to their galaxies?

47. Go to the American Association of Variable Star Observers (AAVSO) website (http://www.aavso.org). What does this 100-year-old organization do? Read about the types of intrinsic variable stars. Click on "Getting Started." If you have access to dark skies, you can contribute to the study of variable stars. Go to the page for observers (http://www.aavso.org/observers) and click on each item in the "For New Observers" list, including the list of stars that are easy to observe. Assemble a group and observe a variable star from this list.

48. Go to the University of Washington's "Properties of Planetary Nebulae" Web page (https://sites.google.com/a/uw.edu/introductory-astronomy-clearinghouse/assignments/labs-exercises/properties-of-planetary-nebulae) and download and complete the lab exercise.

49. Go to the Hubble Space Telescope's planetary nebula gallery (http://hubblesite.org/gallery/album/nebula/planetary). For each of the three types of symmetry, find an example of a nebula that shows clearly the type of symmetry: spherical (being symmetric in every direction, like a circle), bipolar (having an axis about which they are symmetric, like a person's face), and point-symmetric (being symmetric about a point, like the letter *S*). Print each of the three images you chose, and label the type of symmetry each one represents. For all three nebulae, identify the location of the central star. For bipolar symmetry, draw a line that shows the axis about which the nebula is symmetric. For point symmetry, identify several features that are symmetric across the location of the central star.

50. In the Hubble telescope news archive, look up press releases on planetary nebulae (http://hubblesite.org/newscenter/archive/releases/nebula/planetary) and white dwarf stars (http://hubblesite.org/newscenter/archive/releases/star/white-dwarf). Pick a story for each. What observations were reported, and why were they important?

smartwork5

If your instructor assigns homework in Smartwork5, access your assignments at digital.wwnorton.com/astro5.

digital.wwnorton.com/astro5

The evolution of a low-mass star, as discussed in this chapter, corresponds to many twists and turns on the H-R diagram. In this Exploration, we return to the "H-R Diagram Explorer" interactive simulation to investigate how these twists and turns affect the appearance of the star.

Open the "H-R Explorer" simulation, linked from the Chapter 16 area on the Student Site of the Digital Landing Page. The box labeled "Size Comparison" shows an image of both the Sun and the test star. Initially, these two stars have identical properties: the same temperature, the same luminosity, and the same size.

Examine the box labeled "Cursor Properties." This box shows the temperature, luminosity, and radius of a test star located at the "X" in the H-R diagram. Before you change anything, answer the first three questions.

1 What is the temperature of the test star?

2 What is the luminosity of the test star?

3 What is the radius of the test star?

As a star leaves the main sequence, it moves up and to the right on the H-R diagram. Move the cursor (the X on the diagram) up and to the right.

4 What changes about the image of the test star next to the Sun?

5 What is the test star's temperature? What property of the image of the test star indicates that its temperature has changed?

6 What is the test star's luminosity?

7 What is the test star's radius?

8 Ordinarily, the hotter an object is, the more luminous it is. In this case, the temperature has gone down, but the luminosity has gone up. How can this be?

The star then moves around quite a lot in that part of the H-R diagram. Look at Figure 16.13, and then use the cursor to approximate the motion of the star as it moves up the red giant branch, back down and onto the horizontal branch, and then back to the right and up the asymptotic giant branch.

9 Are the changes you observe in the image of the star as dramatic as the ones you observed for question 4?

10 What is the most noticeable change in the star as it moves through this portion of its evolution?

Next, the star begins moving across the H-R diagram to the left, maintaining almost the same luminosity. Drag the cursor across the top of the H-R diagram to the left, and study what happens to the image of the star in the "Size Comparison" box.

11 What changed about the star as you dragged it across the H-R diagram?

12 How does the star's size now compare to that of the Sun?

Finally, the star drops to the bottom of the H-R diagram and then begins moving down and to the right. Move the cursor toward the bottom of the H-R diagram, where the star becomes a white dwarf.

13 What changed about the star as you dragged it down the H-R diagram?

14 How does its size now compare to that of the Sun?

To thoroughly cement your understanding of stellar evolution, press the "Reset" button and then move the star from main sequence to white dwarf several times. This exercise will help you remember how this part of a star's life appears on the H-R diagram.

17 Evolution of High-Mass Stars

The vast majority of stars are smaller and less massive than the Sun and live a relatively long time. The most massive stars, the O and B stars, are rarer and live a much shorter time. When these massive stars die, the result is far more spectacular than when lower-mass stars die. In this chapter, we will look at the lives of these high-mass stars.

LEARNING GOALS

By the conclusion of this chapter, you should be able to:

LG 1 Describe how the death of high-mass stars differs from that of low-mass stars.

LG 2 List the sequence of stages for evolving high-mass stars.

LG 3 Explain the origin of chemical elements up to and heavier than iron.

LG 4 Identify how Hertzsprung-Russell (H-R) diagrams of clusters enable astronomers to measure the ages of stars and test theories of stellar evolution.

The Crab Nebula is the remains of a massive star whose explosion was observed in 1054. The center is an X-ray image of the pulsar. ▶▶▶

What caused this star to explode?

17.1 High-Mass Stars Follow Their Own Path

We have seen that low-mass stars live for billions of years. High-mass stars have masses greater than about 8 M_{Sun}. These stars have more mass, but they also have luminosities thousands or even millions of times greater than the Sun's luminosity. Even though they have more fuel to begin with, high-mass stars use it up faster than low-mass stars and therefore have shorter lives. High-mass stars live only hundreds of thousands to millions of years. In this section, we will discuss the stages in the evolution of high-mass (greater than 8 M_{Sun}) and medium-mass (3–8 M_{Sun}) stars.

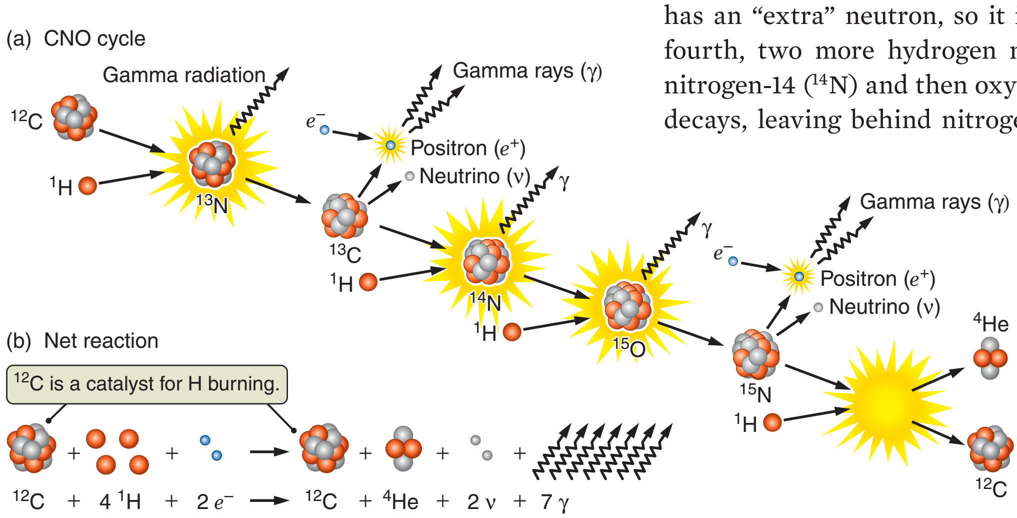
The CNO Cycle

A high-mass star has greater gravitational force pressing down on the interior than that of a low-mass star. This greater force leads to higher temperature and pressure, increasing the rate of nuclear reactions, and therefore generating greater luminosities. In addition, at the much higher temperatures at the center of a high-mass star, additional nuclear reactions become possible. Recall from Chapter 14 that the hydrogen nucleus has only one proton—a single positive charge—so hydrogen fuses at lower temperatures (a few million kelvins) than any other atomic nucleus. However, the probability that any two hydrogen atoms will fuse is low. The low probability of this first step in the proton-proton chain limits how rapidly the entire process can move forward.

Nebraska Simulation: CNO Cycle Animation

If the star contains elements from a previous generation of stars that lived and died, it will have carbon and other elements mixed in with its hydrogen. At the high temperatures in the core of a massive star, the hydrogen and carbon nuclei can interact in a series of reactions called the carbon-nitrogen-oxygen (CNO) cycle. The **CNO cycle** is a nuclear fusion process that converts hydrogen to helium in the presence of carbon. This process is illustrated in **Figure 17.1a**, which shows each step. Through most of the process, only one of two reactions happens at each step: either a hydrogen nucleus fuses with another nucleus to create a new element with higher atomic number or a proton spontaneously decays to a neutron to create a new element with lower atomic number. First, a hydrogen nucleus fuses with a carbon-12 (^{12}C) nucleus to form nitrogen-13 (^{13}N). Second, a proton in this ^{13}N nucleus decays, so that the atom is once again carbon. This carbon nucleus has an "extra" neutron, so it is now carbon-13 (^{13}C), not carbon-12. Third and fourth, two more hydrogen nuclei then fuse with this ^{13}C nucleus, creating nitrogen-14 (^{14}N) and then oxygen-15 (^{15}O). Fifth, a proton in the oxygen nucleus decays, leaving behind nitrogen-15 (^{15}N). One more proton enters the nucleus, causing the ejection of a helium-4 (^{4}He) nucleus, and leaving behind a ^{12}C nucleus, which can participate in the cycle again. Along the way, several by-products are produced: a positron and a neutrino are ejected each time a proton decays to form a neutron. Each positron subsequently annihilates with an electron to produce a gamma ray. Additional gamma rays are released each time fusion occurs. Figure 17.1b shows the net reaction: a carbon-12 nucleus and four hydrogen nuclei

Figure 17.1 (a) In high-mass stars, carbon serves as a catalyst for the fusion of hydrogen to helium. This process is the carbon-nitrogen-oxygen (CNO) cycle. (b) The CNO cycle takes carbon-12, hydrogen, and electrons as inputs and produces carbon-12, helium, neutrinos, and gamma rays.



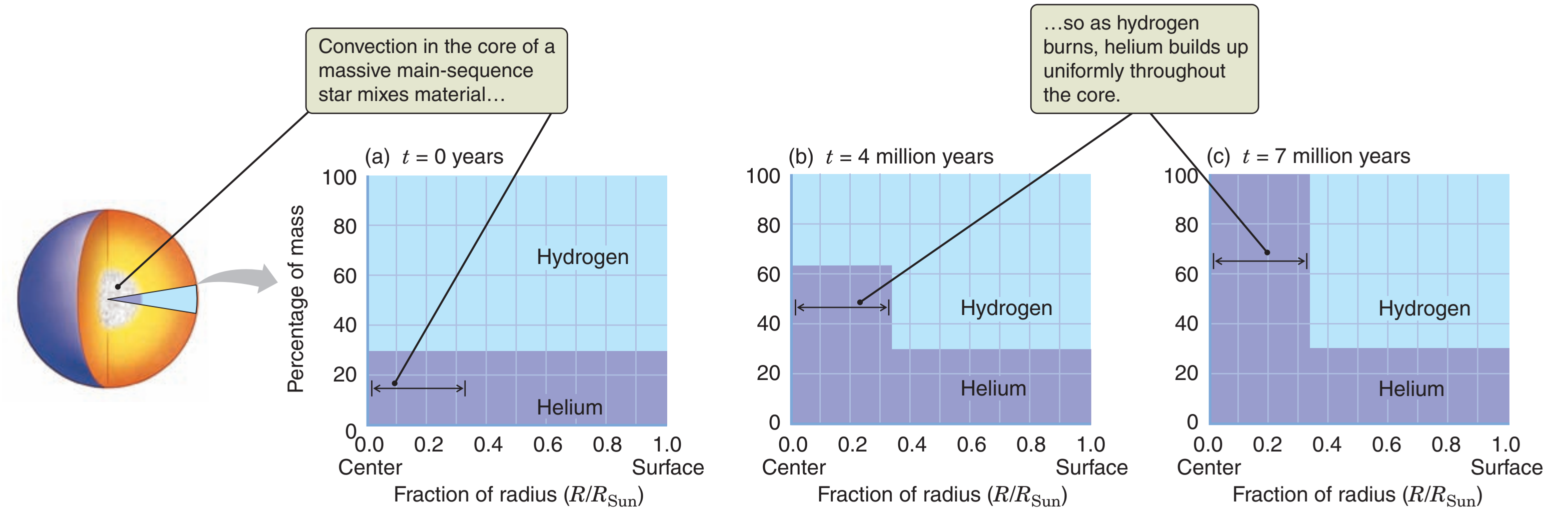


Figure 17.2 Convection keeps the core of a high-mass main-sequence star well mixed, so the composition remains uniform throughout the core as it evolves from zero age in the top graph to age 7 million years in the bottom graph. (Evolution times are for a 25-M_{Sun} star.)

combine with two electrons to produce a carbon-12 nucleus, a helium-4 nucleus, two neutrinos, and seven gamma rays. It takes a lot of energy to get a hydrogen nucleus past the electric barrier set up by a carbon nucleus, with its six protons, but when this barrier can be overcome, fusion is much more probable than it is in the interaction between two hydrogen nuclei. The CNO cycle is far more efficient than the proton-proton chain in stars more massive than about 1.3–1.5 M_{Sun}.

The different ways that hydrogen burning takes place in high-mass stars and low-mass stars are reflected in the different core structures of the two types of stars. The temperature gradient in the core of a high-mass star is so steep that convection sets in within the core itself, "stirring" the core like the water in a boiling pot. As shown in **Figure 17.2**, helium ash spreads uniformly throughout the core of a high-mass star as the star consumes its hydrogen. This is unlike the case for low-mass stars, in which case the helium ash builds up from the center outward (see Figure 16.2).

The High-Mass Star Leaves the Main Sequence

As the high-mass star runs out of hydrogen in its core, the weight of the overlying star compresses the core, just as it does in a low-mass star. Yet long before the core of the high-mass star becomes electron-degenerate, the pressure and temperature in the core become high enough (10^8 K) for helium burning to begin. This rapid increase in pressure and temperature prevents the growth of a degenerate core in the high-mass star. The star makes a fairly smooth transition from hydrogen burning to helium burning as it leaves the main sequence.

Recall that when a low-mass star leaves the main sequence, the path it follows on the Hertzsprung-Russell (H-R) diagram is nearly vertical (see Figure 16.4), going to higher and higher luminosities at roughly constant temperature. But as a high-mass star leaves the main sequence, it grows in size while its surface temperature falls, so it moves nearly horizontally on the H-R diagram, as shown in **Figure 17.3**. The massive star has the same structure as a low-mass horizontal branch star, burning helium in its core and hydrogen in a surrounding shell. Stars more massive than 10 M_{Sun} become red supergiants during their helium-burning phase. They have very cool surface temperatures (about 4000 K) and radii as much as 1,000 times that of the Sun.

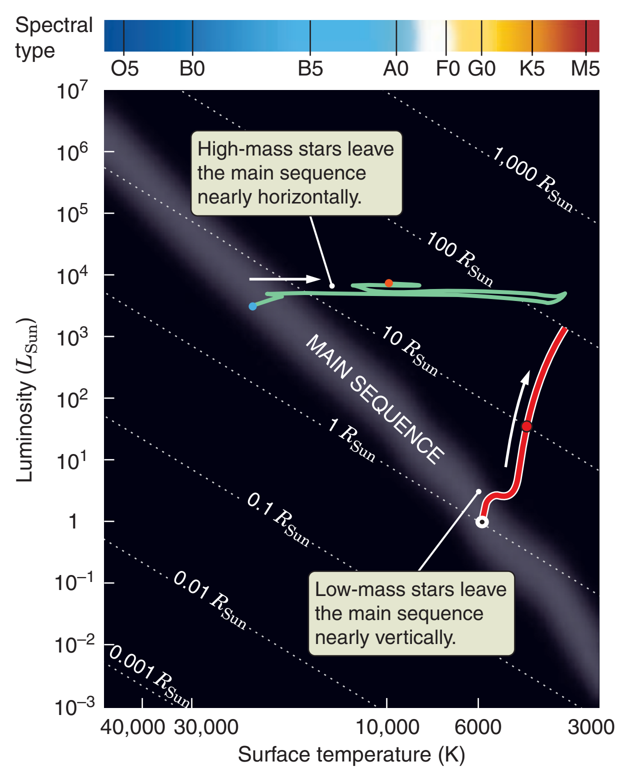


Figure 17.3 When high-mass stars leave the main sequence, they move horizontally across the H-R diagram, unlike low-mass stars, which move nearly vertically.

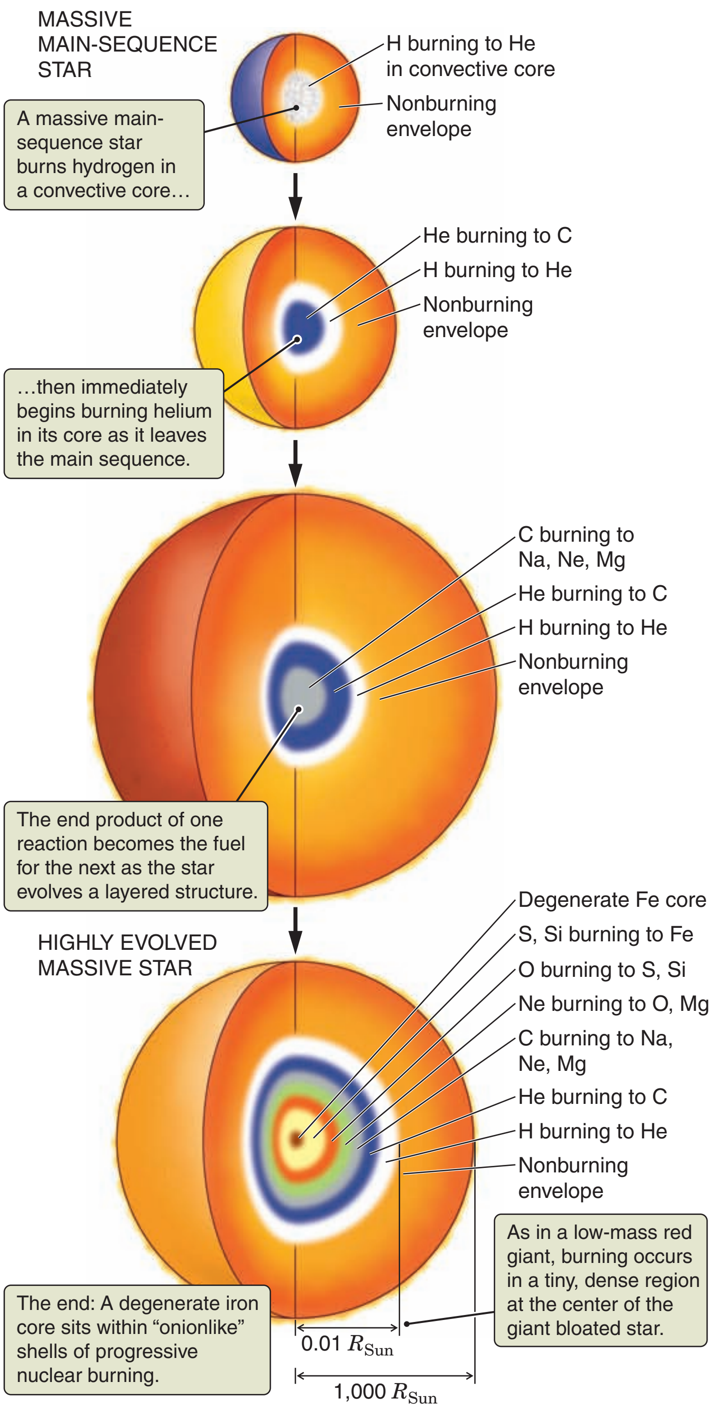


Figure 17.4 As a high-mass star evolves, it builds up a layered structure like that of an onion, with progressively more advanced stages of nuclear burning found deeper and deeper within the star. Note that the bottom image has been reduced in size in order to fit on the page.

The next stage in the evolution of a high-mass star has no analog in low-mass stars. When the high-mass star exhausts the helium in its core, the core again begins to collapse. Carbon fusion begins when the core reaches temperatures of 8×10^8 K or higher. This produces more massive elements, including oxygen, neon, sodium, and magnesium. The star at this time consists of a carbon-burning core surrounded by a helium-burning shell surrounded by the outward-moving, hydrogen-burning shell. When carbon is exhausted as a nuclear fuel at the center of the star, neon breaks down to oxygen and helium or fuses to magnesium; and when neon is exhausted, oxygen begins to fuse. The structure of the evolving high-mass star, shown in **Figure 17.4**, is like an onion, with many concentric layers.

Medium-mass stars with masses between 3 and 8 M_{Sun} burn hydrogen via the CNO cycle like massive stars. They also leave the main sequence as massive stars do, burning helium in their cores immediately after their hydrogen is exhausted and skipping the helium flash phase of low-mass star evolution. When helium burning in the core is complete, however, the temperature at the center of a medium-mass star is too low for carbon to burn. From this point on, the star evolves more like a low-mass star, ascending the asymptotic giant branch, burning helium and hydrogen in shells around a degenerate core, and then ejecting its outer layers and leaving behind a white dwarf.

Stars on the Instability Strip

As a star undergoes post-main-sequence evolution, it may make one or more passes through a region of the H-R diagram known as the **instability strip**, as shown by the dashed-line region in **Figure 17.5**. **Variable stars** are stars that vary in brightness over time. In this region of the H-R diagram, pulsating variable stars can be found. **Pulsating variable stars** do not achieve a steady balance between pressure and gravity; rather, they alternately grow larger and smaller. The pulsating variable star repeatedly overshoots the equilibrium point, shrinking too far before being pushed back out by pressure or expanding too far before being pulled back in by gravity.

The most luminous pulsating variable stars are **Cepheid variables**, named after the prototype star Delta Cephei. Type I, or Classical, Cepheids are massive and luminous yellow supergiants. A Cepheid variable completes one cycle of its pulsation in anywhere from about 1 to 100 days, depending on its luminosity, as shown in **Figure 17.6**. This type of relationship is called a **period-luminosity relationship**; in this case, the more luminous the star, the longer its period of variation. This period-luminosity relationship for Cepheid variables, first discovered experimentally by Henrietta Leavitt (1868–1921) in 1912, allows astronomers to use Cepheid variables to find the distances to galaxies beyond our own. By observing the period, astronomers can determine the luminosity. Combining this luminosity with the brightness in the sky gives the distance to the star.

Thermal energy powers the pulsations of stars like Cepheid variables. Ionization of helium atoms in the star alternately traps and releases thermal energy, causing the star to expand and contract. An ionized gas is nearly opaque, because ions can interact with light of all wavelengths, scattering many of them back toward the

Nebraska Simulation: H-R Explorer

interior. A neutral gas can absorb only light with energies that match the electron transitions. As the helium atoms alternate between ionized and neutral states, the atmosphere of the star alternates between opaque and transparent. This alternately traps and releases light from the star. It is much like a lid on a pot of boiling water. The pot builds up pressure enough to pop open the lid and let steam escape. Then gravity pulls the lid back down, and pressure builds again to repeat the process. These pulsations do not affect the nuclear burning in the star's interior. However, they do affect the light escaping from the star. From Chapter 13, recall the luminosity-temperature-radius relationship for stars: both the luminosity and the temperature (and hence color) of the star change periodically as the star expands and shrinks. The star is at its brightest and bluest while it expands through its equilibrium size and at its faintest and reddest while it falls back inward.

Type I Cepheid variables are not the only type of variable star. The instability strip on the H-R diagram also intersects the low-mass horizontal branch. Low-mass (~0.8-M_{Sun}) stars can be Type II Cepheid variables or **RR Lyrae variables**, which have periods of less than a day. RR Lyrae stars pulsate by the same mechanism as Cepheid variables but are typically hundreds of times less luminous. They, too, follow a period-luminosity relationship, illustrated in Figure 17.6b.

High-mass stars also change their composition by expelling a significant percentage of their mass back into space throughout their lifetimes. Even while on

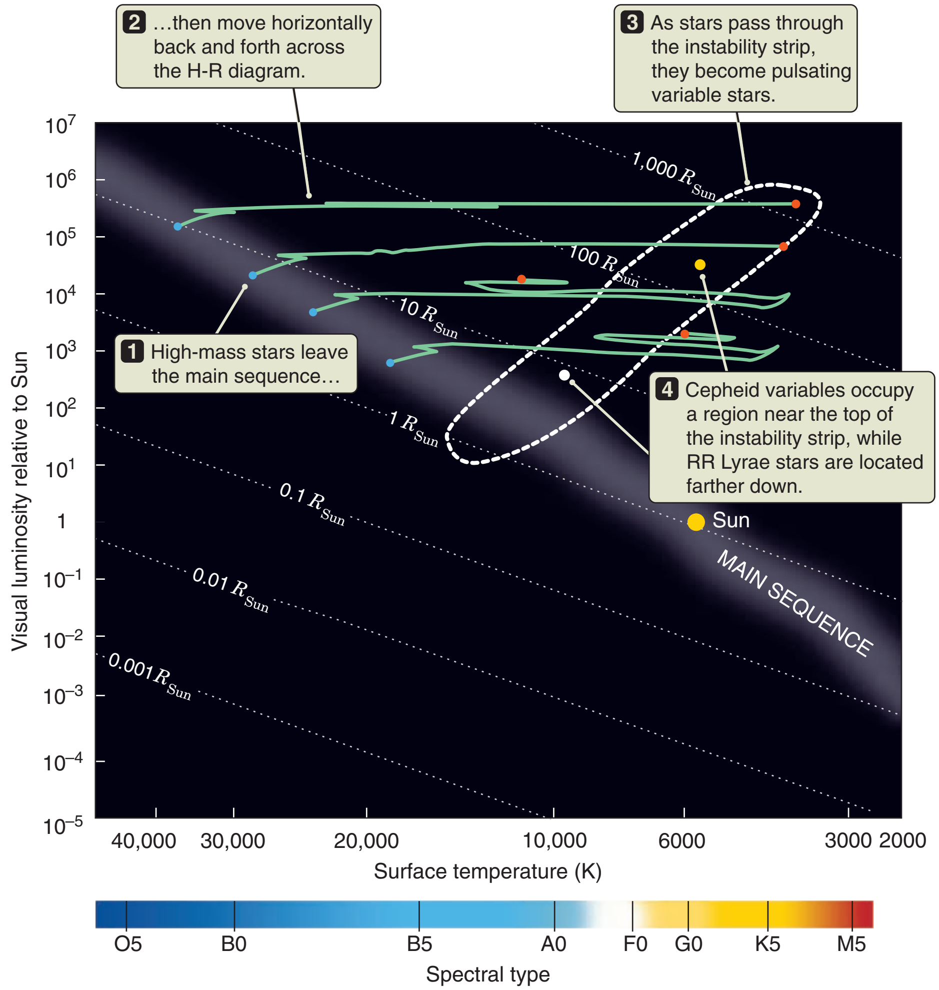


Figure 17.5 The paths of high-mass stars along the H-R diagram takes them through a region known as the instability strip, shown by the region surrounded by the white dashed line. Pulsating variable stars such as Cepheid variables and RR Lyrae stars are found in this region of the H-R diagram.

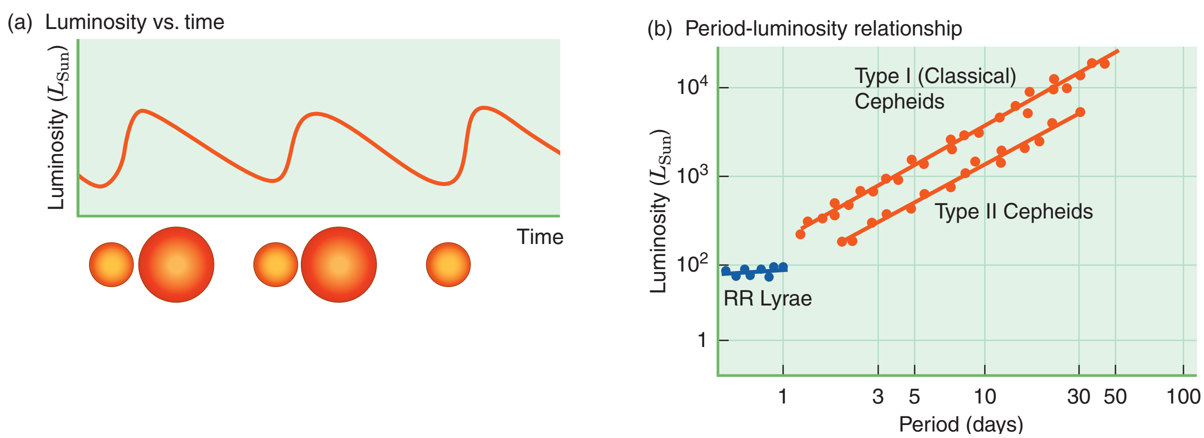


Figure 17.6 (a) Cepheid variable stars pulsate in size and therefore their luminosity changes with time in a periodic way. The shape of the resulting curve is distinctive. (b) The length of the period of pulsation is related to the star's luminosity at maximum.

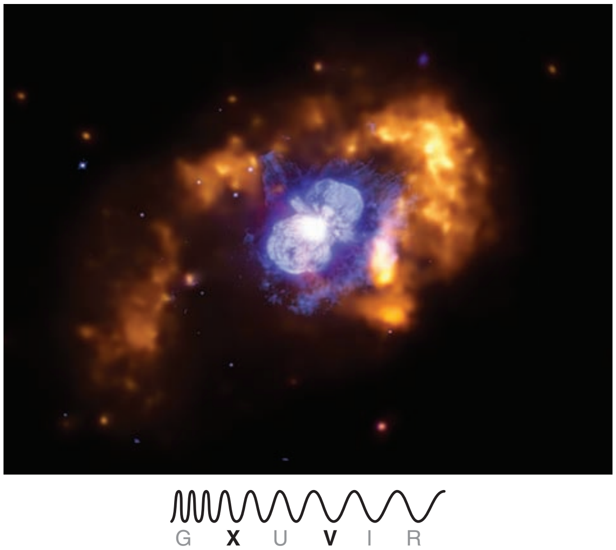


Figure 17.7 In this image of the luminous blue variable (LBV) star Eta Carinae, an expanding cloud of ejected dusty material is seen in optical (blue) and X-ray (yellow) light. The star itself, which is largely hidden by the surrounding dust, has a luminosity of 5 million L_{Sun} and a mass probably in excess of 120 M_{Sun}. Dust is created when volatile material ejected from the star condenses.

the main sequence, massive O and B stars have low-density winds with velocities as high as 3,000 kilometers per second (km/s). These winds are pushed outward by the pressure of the radiation from the star. The pressure of the intense radiation at the surface of a massive star overcomes the star's gravity and drives away material in its outermost layers. Main-sequence O and B stars lose mass at variable rates ranging from 10^{-7} up to 10^{-5} M_{Sun} of material per year. The fastest mass loss occurs in the most massive stars. These numbers may sound tiny, but over millions of years mass loss plays a prominent role in the evolution of high-mass stars. O stars with masses of 20 M_{Sun} or more may lose about 20 percent of their mass while on the main sequence and possibly more than 50 percent of their mass over their entire lifetimes. Even an 8-M_{Sun} star may lose 5–10 percent of its mass.

In general, stars with masses below 15 M_{Sun} go through a red supergiant stage, perhaps with a Cepheid variable phase. Stars under 30 M_{Sun} move back and forth on the H-R diagram—becoming red supergiants, then blue supergiants, then red supergiants again (sometimes with a brief period as a yellow supergiant), depending on what is burning in their cores. Luminous blue variable (LBV) stars are hot, luminous, extremely rare stars that may be as massive as 150 M_{Sun}. An example is Eta Carinae (**Figure 17.7**), a binary system with a 120-M_{Sun} star and a luminosity (summed over all wavelengths) of 5 million L_{Sun}. Currently, Eta Carinae is losing mass at a rate of 1 M_{Sun} every 1,000 years. However, during a 19th century eruption, when Eta Carinae became the second brightest star in the sky, it shed 10 M_{Sun} of material in only 20 years. Eta Carinae is expected to explode in the astronomically near future.

CHECK YOUR UNDERSTANDING 17.1

How does energy production in a high-mass, main-sequence star differ from energy production in the Sun? (Choose all that apply.) (a) High-mass stars get a lot of energy through non-nuclear processes. (b) High-mass stars produce energy at a faster rate. (c) High-mass stars burn carbon on the main sequence. (d) High-mass stars use carbon in a process that fuses hydrogen to helium.

17.2 High-Mass Stars Go Out with a Bang

You have already seen that a low-mass star approaches the end of its life relatively slowly and gently, ejecting its outer parts into nearby space and leaving behind a degenerate core. In contrast, the end for a high-mass star comes suddenly and explosively. To understand why the life of a high-mass star ends this way, we need to understand the concept of binding energy in atomic nuclei and how it affects nuclear fusion in the center of the star.

Binding Energy

An evolving high-mass star builds up its onionlike structure as nuclear burning in its interior proceeds to more and more advanced stages (see Figure 17.4). Hydrogen burns to helium, helium burns to carbon and oxygen, carbon burns to magnesium, oxygen burns to sulfur and silicon, and then silicon and sulfur burn to iron. Many different types of nuclear reactions occur up to this point, forming almost all of the different stable isotopes of elements less massive than iron. The chain of nuclear fusion stops with iron.

Recall that when four hydrogen atoms combine to form a helium atom, the resulting helium atom has less mass than the sum of the four individual hydrogen atoms. This difference in mass is converted to energy, which maintains the temperature of the gas at the high levels needed to sustain the reaction. The same happens in the triple-alpha process when three helium nuclei combine to form a ^{12}C nucleus: The net energy produced by the reaction is the difference between the mass of the three alpha particles and the mass of the carbon nucleus. This energy helps sustain the nuclear reactions in the core.

The **binding energy** of an atomic nucleus is the energy required to break the nucleus into its constituent parts. A nuclear reaction that increases a nucleus's binding energy releases energy. Conversely, decreasing the binding energy absorbs energy. **Figure 17.8** shows the binding energy per nucleon (that is, per each proton or neutron in the nucleus) for different atomic nuclei. Moving up the plot from helium to carbon increases the binding energy, so fusing helium to carbon releases energy. Iron is at the peak of the binding-energy curve, so moving up the plot from lighter elements to iron (Fe) also releases energy; conversely, moving down the plot from iron to heavier elements absorbs energy. Iron fusion absorbs energy in the reaction. This reaction, therefore, does not sustain itself (**Working It Out 17.1**).

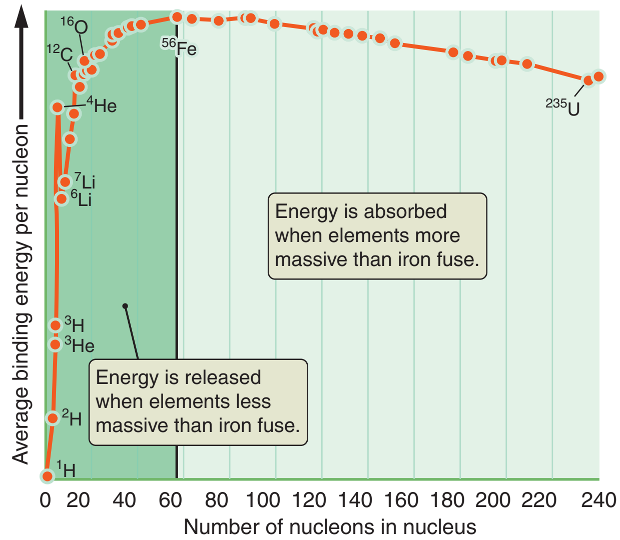


Figure 17.8 The binding energy per nucleon is plotted against the number of nucleons for each element. This is the energy it would take to break the atomic nucleus apart into protons and neutrons. Energy is released by nuclear fusion only if the resulting element is higher on the curve.

The Final Days in the Life of a Massive Star

Fusion of helium to carbon produces about one-quarter as much energy per reaction as the fusion of hydrogen to helium. This lower energy production rate results in lower pressure inside the star. As gravity compresses the star, this less efficient nuclear fuel is consumed more rapidly. Although fusion of hydrogen into helium can provide the energy needed to support the high-mass star against the force of gravity for millions of years, helium fusion can support the star for only a

17.1 Working It Out Binding Energy of Atomic Nuclei

The net energy released by a nuclear reaction is the difference between the binding energy of the products and the binding energy of the reactants:

$$\text{Net energy} = \begin{pmatrix}\text{Binding energy}\\ \text{of products}\end{pmatrix} - \begin{pmatrix}\text{Binding energy}\\ \text{of reactants}\end{pmatrix}$$

In the example of the triple-alpha process, the binding energy of a helium nucleus is 6.824×10^{14} joules (J) per kilogram (kg) of helium, and the binding energy of the produced ^{12}C nuclei is 7.402×10^{14} J. The amount of energy available from fusing 1 kg of helium nuclei into carbon is given by

$$\begin{pmatrix}\text{Net energy from}\\ \text{burning 1 kg He}\end{pmatrix} = \begin{pmatrix}\text{Binding energy}\\ \text{of C formed}\end{pmatrix} - \begin{pmatrix}\text{Binding energy}\\ \text{of He burned}\end{pmatrix}$$

$$= (7.402 \times 10^{14}\text{ J}) - (6.824 \times 10^{14}\text{J})$$

$$= 5.780 \times 10^{13}\text{ J}$$

This release of net energy indicates that helium is a good nuclear fuel, as Figure 17.8 shows.

What about fusing iron into more massive elements? Because iron is at the peak of the binding-energy curve, the products of iron burning will have less binding energy than the initial reactants. Going from iron to more massive elements means moving down on the binding-energy curve in Figure 17.8, so the net energy in the reaction will be negative. Rather than producing energy, fusion of iron uses energy.

TABLE 17.1 Burning Stages in High-Mass Stars

Core Burning Stage	15-M_{Sun} Star	25-M_{Sun} Star	Typical Core Temperatures
Hydrogen (H) burning	11 million years	7 million years	$(3–10) \times 10^7$ K
Helium (He) burning	2 million years	800,000 years	$(1–7.5) \times 10^8$ K
Carbon (C) burning	2,000 years	500 years	$(0.8–1.4) \times 10^9$ K
Neon (Ne) burning	8 months	11 months	$(1.4–1.7) \times 10^9$ K
Oxygen (O) burning	2.6 years	5 months	$(1.8–2.8) \times 10^9$ K
Silicon (Si) burning	18 days	0.7 day	$(2.8–4) \times 10^9$ K

few hundred thousand years. **Table 17.1** shows that the star proceeds from helium burning to the end of its life at a much faster pace.

Maintaining the balance in a star is like trying to keep a leaky balloon inflated. The larger the leak, the more rapidly air must be pumped into the balloon to keep it inflated. A star that is fusing hydrogen or helium is like a balloon with a slow leak. At the temperatures generated by hydrogen or helium fusion, energy leaks out of the interior of a star primarily by radiation and convection. Neither of these processes is very efficient, because the outer layers of the star act like a thick, warm blanket. Much of the energy is kept in the star, so nuclear fuels do not need to burn very fast to support the weight of the outer layers of the star while keeping up with the energy escaping from the surface.

Once carbon burning begins, this balance shifts. Energy is carried away primarily by neutrinos—referred to as neutrino cooling—rather than by radiation and convection. Recall from your study of the Sun in Chapter 14 that neutrinos escape easily, carrying energy away from the core. Like air pouring out through a huge hole in the side of a balloon, neutrinos produced in the interior of the star stream through the overlying layers of the star as if they were not even there, carrying the energy from the stellar interior out into space. As thermal energy pours out of the interior of the star, the outer layers of the star fall inward, driving up the density and temperature and increasing the rate of nuclear reactions.

Once this process of neutrino cooling becomes significant, the star begins evolving much more rapidly. Carbon fusion supports the star for less than about a thousand years. Oxygen fusion holds the star up for only about a year. Silicon fusion lasts only a few days. A silicon-burning star is not much more luminous than it was while burning helium. But because of neutrino cooling, the silicon-burning star actually releases about 200 million times more energy per second than it did while it was fusing helium.

The Collapse of the Core and Subsequent Explosion

As shown in Figure 17.4, an evolving high-mass star builds up its onionlike structure as hydrogen fuses to helium, helium fuses to carbon, carbon fuses to sodium, neon, and magnesium, oxygen fuses to sulfur and silicon, and silicon and sulfur fuse to iron. Many different types of nuclear reactions occur up to this point, forming almost all of the different stable isotopes of elements less massive than iron. However, because iron does not release energy when it fuses but rather absorbs energy, the chain of nuclear fusion stops with iron. After silicon fuses to form an iron core in the star, the end comes suddenly and dramatically. For iron, once the reaction starts, energy is absorbed. No source of nuclear energy remains to replenish the energy that is being taken away by escaping neutrinos. The high-mass star's life balancing gravity and controlled nuclear energy production is over. No longer supported by thermonuclear fusion, the iron core of the massive star begins to collapse.

Figure 17.9 shows the stages a high-mass star passes through at the end of its life. The early stages of collapse of the iron core of an evolved massive star are much the same as in the collapse of a nonburning core in a low-mass star. As the core collapses, the force of gravity increases and the density and temperature skyrocket. The gas in the core becomes electron-degenerate when it is about the size of Earth. Unlike the electron-degenerate core of a low-mass red giant, however, the weight bearing down on the interior of the iron core is too great to be held up by electron degeneracy pressure (step 1 in Figure 17.9). As the core collapses, the core temperature climbs to 10 billion K (10^{10} K) and higher, while the density exceeds 10^{10} kilograms per cubic meter (kg/m^3)—10 times the density of an electron-degenerate white dwarf.

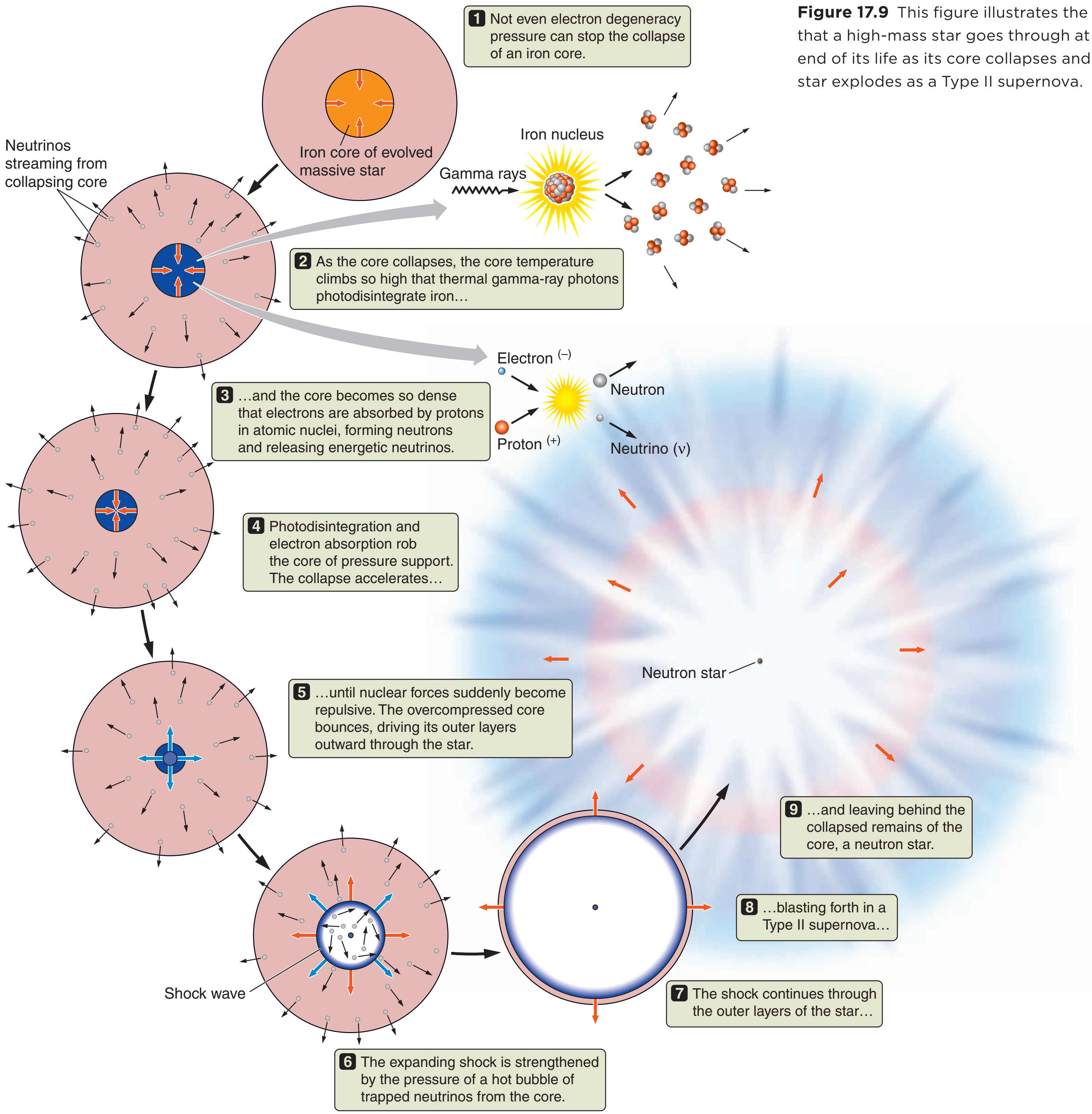


Figure 17.9 This figure illustrates the stages that a high-mass star goes through at the end of its life as its core collapses and the star explodes as a Type II supernova.

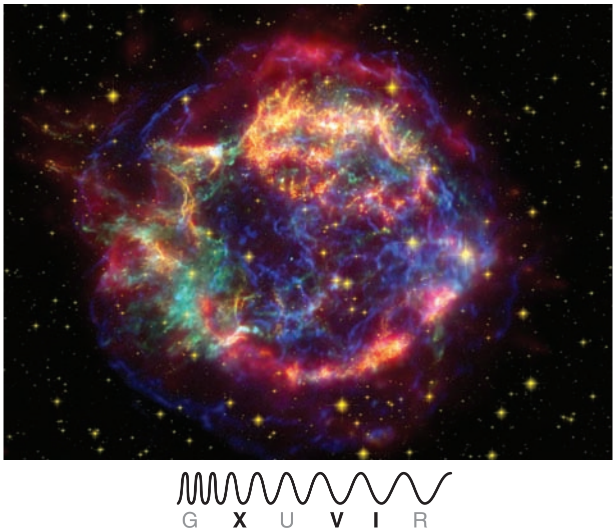


Figure 17.10 When an evolved star explodes, it forms a supernova remnant like Cassiopeia A. X-ray observations such as these suggest that the star may have turned itself inside out as it spit out elements from its core.

Astronomy in Action: Type II Supernova

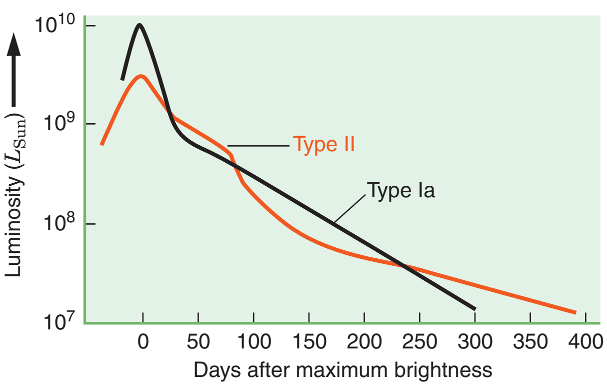


Figure 17.11 These light curves show the changes in brightness of average Type Ia and Type II supernovae.

These phenomenal temperatures and pressures trigger fundamental changes in the core. The laws describing thermal radiation say that at these temperatures, the nucleus of the star is awash in extremely energetic thermal radiation. This radiation is so energetic that thermal gamma-ray photons are produced with enough energy to break iron nuclei apart into helium nuclei (step 2 in Figure 17.9). This process, called photodisintegration, absorbs thermal energy and begins reversing the results of nuclear fusion. At the same time, the density of the core is so great that electrons are forced into atomic nuclei, where they combine with protons to produce neutrons and neutrinos (step 3 in Figure 17.9). Both this process and photodisintegration absorb much of the energy that was holding up the dying star. Neutrinos continue to take more energy with them as they leave the star. The collapse of the core accelerates, reaching a speed of 70,000 km/s, or almost one-fourth the speed of light on its inward fall (step 4 in Figure 17.9). All of these events together take place remarkably quickly—in less than a second.

As material in the collapsing core exceeds the density of an atomic nucleus, the strong nuclear force actually becomes repulsive (step 5 in Figure 17.9). About half of the collapsing core suddenly slows its inward fall. The remaining half slams into the innermost part of the star at a significant fraction of the speed of light and "bounces," sending a tremendous shock wave back out through the star (step 6 in Figure 17.9).

Under the extreme conditions in the center of the star, neutrinos are produced at an enormous rate. Over the next second or so, almost a fifth of the core mass is converted into neutrinos. Most of these neutrinos pour outward through the star; but at the extreme densities found in the collapsing core of the massive star, not even neutrinos pass with complete freedom. The dense material behind the expanding shock wave traps a few tenths of a percent of the energy of the neutrinos streaming out of the core of the dying star. The energy of these trapped neutrinos drives the pressure and temperature in this region higher, inflating a bubble of extremely hot gas and intense radiation around the core of the star. The pressure of this bubble adds to the strength of the shock wave moving outward through the star. Within about a minute, the shock wave has pushed its way out through the helium shell within the star. Within a few hours it reaches the surface of the star itself, heating the stellar surface to 500,000 K and blasting material outward at velocities of up to about 30,000 km/s. The evolved massive star has exploded, becoming more than a billion times as luminous as the Sun and leaving behind a cloud of dust and gas (**Figure 17.10**). This type of supernova, which is triggered by the collapse of the core, is called a **Type II supernova**.

The difference between a Type Ia supernova (discussed in the past chapter) and a Type II supernova is subtle. Both become suddenly very luminous. Both leave behind expanding clouds of dust and gas. However, in the first year or so after the explosion, it is possible to distinguish between the types by their light curves, as shown in **Figure 17.11**. Type II supernovae have more complicated light curves, with a peak luminosity less than a Type Ia supernova and a light curve that falls off less rapidly. Additional distinguishing characteristics appear in the spectra of these objects.

In 1987, astronomers observed the explosion of a massive star in the Large Magellanic Cloud (LMC; a companion galaxy to the Milky Way 160,000 light-years away, shown in **Figure 17.12**). Astronomers working in all parts of the electromagnetic spectrum pointed their telescopes at Supernova 1987A—the first naked-eye supernova since the invention of the telescope. Astronomers were ultimately surprised to discover from looking at old photographs that the star that blew up was not a red supergiant, but a 20-M_{Sun} B3 I blue supergiant now classified as a luminous blue variable star. Neutrino telescopes recorded a burst of neutrinos passing through Earth from this tremendous stellar explosion that had occurred in the LMC. The detection of neutrinos from SN 1987A provided astronomers with a rare and crucial glimpse of the very heart of a massive star at the moment of its death, confirming a fundamental prediction of theories about the collapse of the core and its effects.

CHECK YOUR UNDERSTANDING 17.2

What causes a high-mass star to explode as a Type II supernova? (a) The high-mass star merges with another star. (b) Iron absorbs energy when it fuses. (c) The high-mass star runs out of mass in the core. (d) The CNO cycle uses up all the carbon.

17.3 The Spectacle and Legacy of Supernovae

Supernova explosions leave a rich and varied legacy in the universe. Explosions that occurred thousands of years ago have created huge expanding bubbles of million-kelvin gas that glow in X-rays and ultraviolet radiation and drive visible shock waves into the surrounding interstellar medium. The energy and matter flowing out from supernova explosions also compresses nearby clouds, triggering the initial collapse that begins star formation. In this section, we discuss what happens after a massive star explodes as a supernova.

The Energetic and Chemical Legacy of Supernovae

The energy carried away by light from a supernova represents only about 1 percent of the kinetic energy being carried away by the outer parts of the star. This ejected material contains approximately 10^{47} J of kinetic energy—enough energy to accelerate the entire Sun to a speed of 10,000 km/s. The kinetic energy of the material ejected from both Type Ia and Type II supernovae heats the hottest phases of the interstellar medium and pushes around the clouds in the interstellar medium. Yet even this amount of energy is small by comparison with the energy carried away from the supernova explosion by neutrinos—an amount of energy at least 100 times larger.

Perhaps even more important is the chemical legacy left behind by supernova explosions. Only the least massive chemical elements were present at the beginning of the universe: hydrogen, helium, and trace amounts of lithium and beryllium. All of the rest of the chemical elements formed in stars, either during nuclear burning in the core or in the rapid nuclear reactions that occur during a supernova explosion. These elements were released to the interstellar medium when stars died. The process of forming more massive atomic nuclei from less massive nuclei is called **nucleosynthesis**. Nucleosynthesis is responsible for the progressive chemical enrichment of the universe.

Elements up to carbon, oxygen, and small amounts of neon and magnesium form from nuclear fusion in the cores of low-mass stars and travel to the interstellar medium in asymptotic giant branch (AGB) winds and planetary nebulae. A look at the periodic table of the elements (see Appendix 3) shows that many naturally occurring elements are more massive than iron. Recall that fusion up to iron

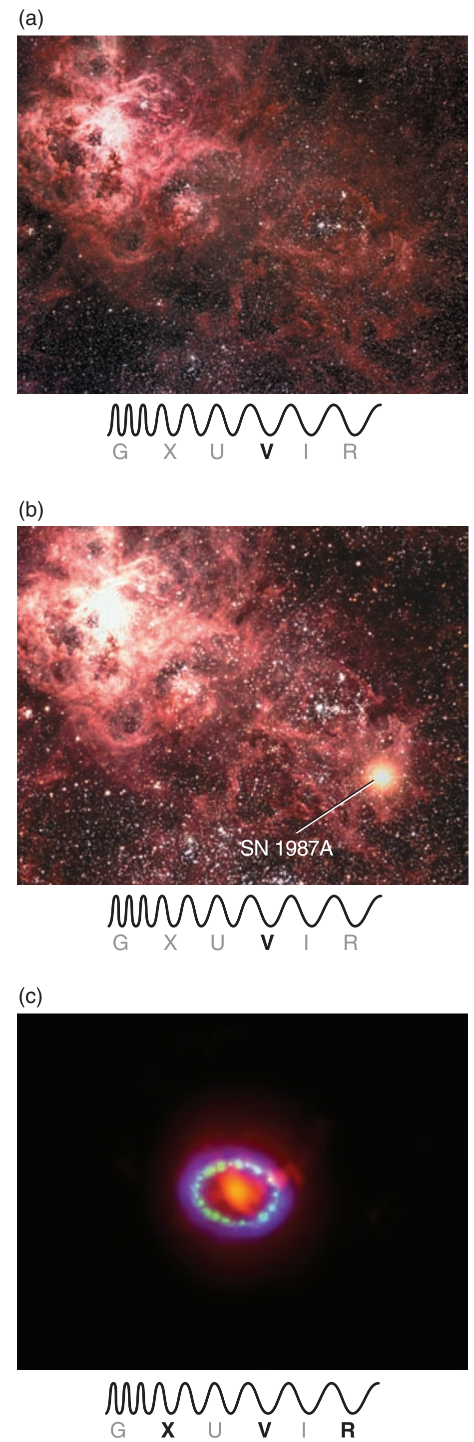


Figure 17.12 Supernova 1987A was a supernova that exploded in a small companion galaxy of the Milky Way called the Large Magellanic Cloud (LMC). These images show the LMC before the explosion (a) and while the supernova was near its peak (b). Notice the “new” bright star at lower right. A 2013 image (c) from the Atacama Large Millimeter/submillimeter Array (ALMA) telescope shows freshly formed dust inside the glowing rings of gas.

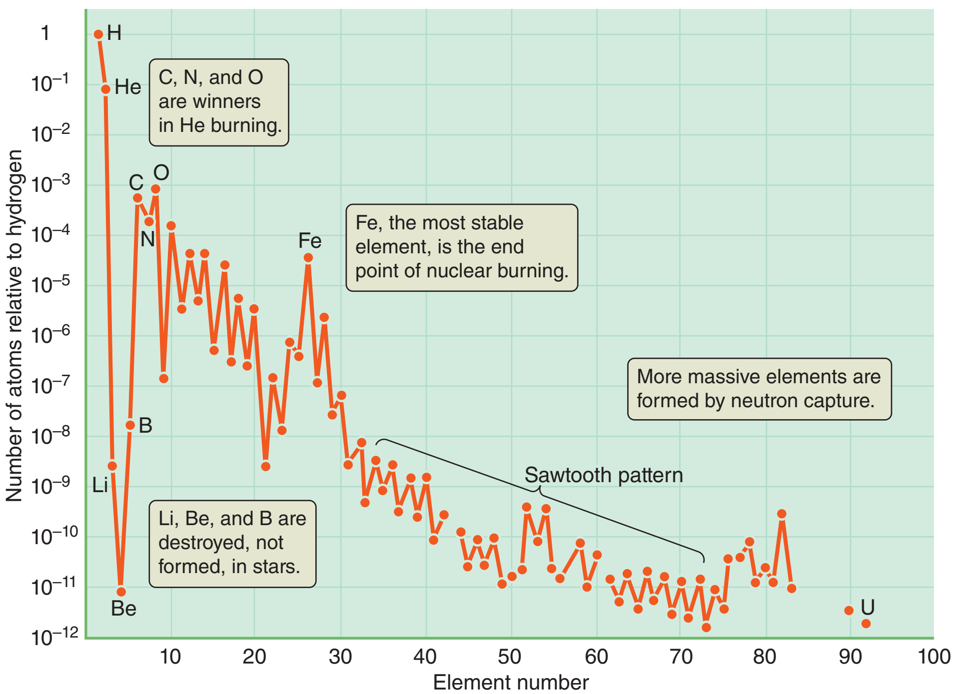


Figure 17.13 This graph plots the observed relative abundances of different elements in the Solar System against the element number of each element's nucleus. This pattern can be understood as a result of nucleosynthesis in stars. The periodic table of the elements in Appendix 3 identifies individual elements by their element number (the number of protons).

creates energy, but fusion beyond iron absorbs energy (see Figure 17.8). So elements heavier than iron fuse only under conditions in which there is abundant energy to be absorbed—such as in the enormous energies of supernova explosions. The naturally occurring elements heavier than iron were all produced in the deaths of high-mass stars.

Normally, electric repulsion keeps positively charged atomic nuclei far apart. Extreme temperatures are needed to slam nuclei together hard enough to overcome this electric repulsion. Free neutrons, however, have no net electric charge, so there is no electric repulsion to prevent them from simply colliding with an atomic nucleus, regardless of how many protons that nucleus contains. Under normal conditions in nature, free neutrons are rare. In the interiors of evolved stars, however, a number of nuclear reactions produce free neutrons, and under some circumstances—including those shortly before and during a supernova—free neutrons are produced in very large numbers. Free neutrons are easily captured by atomic nuclei and later decay to become protons. This process of neutron capture and decay forms the elements with atomic numbers and masses higher than those of iron.

Nuclear physics predicts the abundances of the elements. These predictions agree with abundances that have been measured on Earth, in the Solar System (**Figure 17.13**), and in the atmospheres of stars and their remains. Less massive elements are far more abundant than more massive elements because more massive elements are progressively built up from less massive elements. An exception to this pattern is the dip in the abundances of the light elements lithium (Li), beryllium (Be), and boron (B). Nuclear burning easily destroys these elements, and they are not produced by the common reactions involved in burning hydrogen (H) and helium (He). Conversely, carbon (C), nitrogen (N), and oxygen (O) are produced in quantity in the triple-alpha process, so they are very abundant. The spike in the abundances of elements near iron is evidence of processes that favor these tightly bound nuclei. Even the sawtooth pattern observed in the plot of the abundances of even-numbered and odd-numbered elements is a consequence of the way atomic nuclei form in stars. By comparing the predictions of nuclear physics with observations of elemental abundances, astronomers repeatedly test the theory of stellar evolution.

Neutron Stars and Pulsars

In the explosion of a Type II supernova, the outer parts of the star are blasted back into interstellar space—but what remains of the core that was left behind? The matter at the center of the massive star has collapsed to the point where it has about the same density as the nucleus of an atom. For cores less than about 3 M_{Sun}, the collapse is halted when neutrons are packed as tightly together as the rules of quantum mechanics permit. The *neutron-degenerate* core left behind by the explosion of a Type II supernova is a **neutron star**. It has a radius of 10–20 km, making it roughly the size of a small city, but into that volume is packed a mass between 1.4 and 3 M_{Sun}. (Cores greater than 3 M_{Sun} cannot be supported by

17.2 Working It Out Gravity on a Neutron Star

Neutron stars are incredibly dense objects. As a result, the surface gravity and escape velocity of neutron stars are very high. For example, let's look at a typical neutron star with a radius of 15 km and a mass of 2 M_{Sun}.

Recall from Working It Out 4.1 that the acceleration due to gravity on the surface—in this case the surface of a neutron star (NS)—is given by

$$g = \frac{GM_{\text{NS}}}{R_{\text{NS}}^2}$$

$$g = 6.67 \times 10^{-20} \frac{\text{km}^3}{\text{kg s}^2} \times \frac{2.0 \times (1.99 \times 10^{30}\,\text{kg})}{(15\,\text{km})^2}$$

$$g = 1.2 \times 10^9 \frac{\text{km}}{\text{s}^2}$$

Dividing this number by the gravitational acceleration on Earth, 9.8 m/s^2 = 0.0098 km/s^2, shows that the gravitational acceleration on a neutron star is more than 100 billion times as large as that on Earth.

What about the escape velocity from a neutron star? From Working It Out 16.2, we know that the escape velocity is given by

$$v_{\text{esc}} = \sqrt{2GM/R}$$

Putting in the above numbers for a typical neutron star yields

$$v_{\text{esc}} = \sqrt{\frac{2 \times [6.67 \times 10^{-20}\,\text{km}^3/(\text{kg s}^2)] \times 2.0 \times (1.99 \times 10^{30}\,\text{kg})}{(15\,\text{km})}}$$

$$v_{\text{esc}} = 190{,}000\,\text{km/s}$$

Dividing this result by the speed of light gives

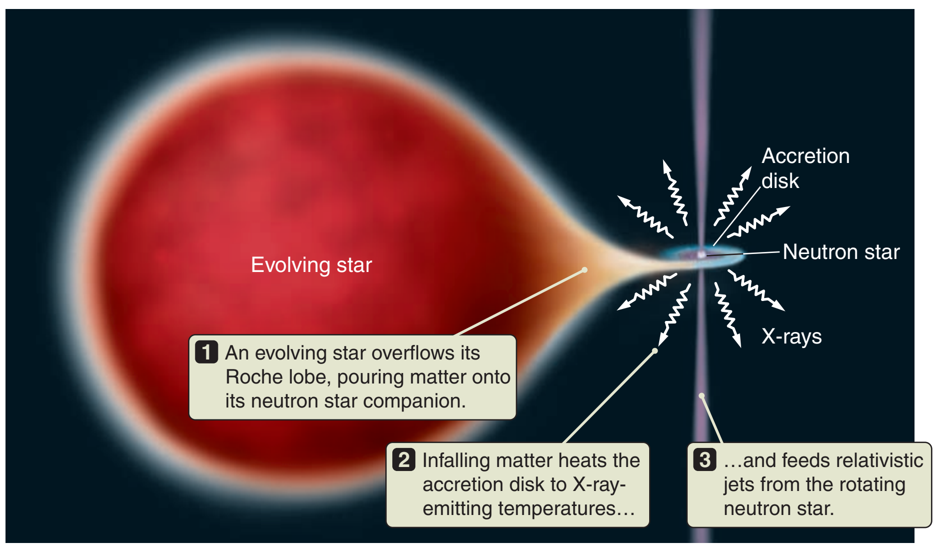
$$\frac{v_{\text{esc}}}{c} = \frac{190{,}000\,\text{km/s}}{300{,}000\,\text{km/s}} = 0.63$$

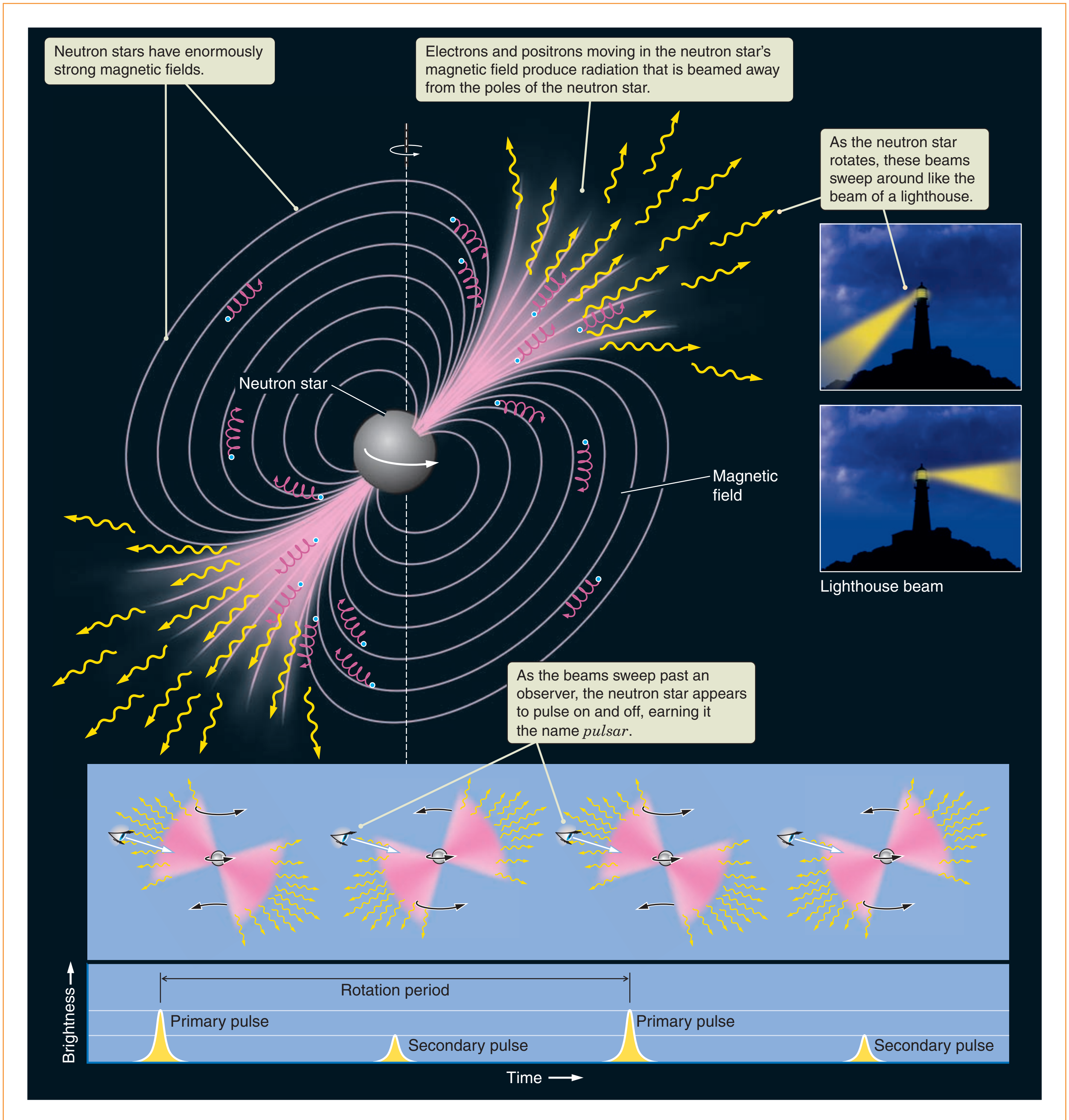
The escape velocity from this neutron star is more than 60 percent of the speed of light and almost 17,000 times greater than the escape velocity from Earth (11.2 km/s). The physicist Albert Einstein showed that strange things happen at velocities near the speed of light (including modifications to Newton's equations), which we will discuss in detail in Chapter 18.

neutron degeneracy and become black holes. These will be discussed in the next chapter.) At a density of about 10^{18} kg/m^3, the neutron star is a billion times denser than a white dwarf and a thousand trillion (10^{15}) times denser than water. Imagine the entire Earth were crushed down to the size of a football stadium. Earth would then have about the same density as a neutron star. Neutron stars are so compact that the acceleration due to gravity on the surface is more than 100 billion times the acceleration on Earth, as shown in **Working It Out 17.2**. This extremely high surface gravity implies a very large escape velocity. A spacecraft would need to be traveling at 0.63*c* to escape the gravity of a typical neutron star.

The massive star from which the neutron star formed may have been part of a binary system. Unless the massive star loses so much mass that gravity no longer holds the two stars together, then the neutron star is left with a binary companion. Processes like those in the white dwarf binary systems responsible for novae and Type Ia supernovae are possible. **Figure 17.14** illustrates an **X-ray binary**, a binary system in which mass from an evolving star spills over onto a collapsed companion such as a white dwarf, neutron star, or black hole. As the lower-mass star in such a binary system evolves and overfills its Roche lobe, matter falls toward the accretion disk around the neutron star, heating it to millions of kelvins and causing it to glow brightly in X-rays. X-ray binaries sometimes develop powerful jets of material that are perpendicular to the accretion disk and carry material away at speeds close to the speed of light.

Figure 17.14 X-ray binaries are systems consisting of a normal evolving star with a white dwarf, a neutron star, or a black hole. As the evolving star overflows its Roche lobe, mass falls toward the collapsed object. The gravitational well of the collapsed object is so deep that when the material hits the accretion disk, it is heated to such high temperatures that it radiates away most of its energy as X-rays.





VISUAL ANALOGY

Figure 17.15 When a highly magnetized neutron star rotates rapidly, light is given off, much like the beams from a rotating lighthouse lamp. As these beams sweep past Earth, the star will appear to pulse on and off, earning it the name *pulsar*.

Recall from Chapter 15 that the conservation of angular momentum requires a collapsing molecular cloud to spin faster as it shrinks into a protostar. Similarly, as the core of a massive star collapses, it spins faster because angular momentum is conserved. A massive main-sequence O star rotates perhaps once every few days. As a neutron star, it might rotate tens or even hundreds of times each second. The collapsing star also concentrates the magnetic field to strengths trillions of times greater than the magnetic field at Earth's surface. A neutron star has a magnetosphere just like Earth and several other planets do, except that the neutron star's magnetosphere is much stronger and is whipped around many times a second by the spinning star. As in planets, the magnetic axis in stars is often not aligned with the rotation axis.

Electrons and positrons circle around the magnetic-field lines of the neutron star and are "funneled" along the field toward the magnetic poles of the system. Any accelerating charged particle produces radiation, so these particles produce beams of radiation along the magnetic poles of the neutron star as shown in **Figure 17.15**. As the neutron star rotates, these beams sweep through space like the rotating beams of a lighthouse. When Earth is located in the paths of these beams, the neutron star appears to flash on and off with a regular period equal to the period of rotation of the star (or half the rotation period, if both beams are seen).

Astronomy in Action: Pulsar Rotation

Many of the unusual objects discussed in this and the previous chapter—such as pulsating stars, supernovae, and planetary nebulae—puzzled astronomers when they were first observed, but were later understood to be associated with the end points of stellar evolution. In contrast, neutron stars were predicted in 1934, not long after neutrons themselves were discovered. Astronomers Walter Baade and Fritz Zwicky proposed that supernova explosions could lead to the formation of a neutron star. But neutron stars were not actually observed for another 30 years. In 1967, rapidly pulsing objects were first discovered by people observing in radio wavelengths (**Process of Science Figure**). These objects, which blinked like very fast, regularly ticking clocks, puzzled astronomers. Today, these rotating neutron stars are called **pulsars**. More than 2,000 pulsars are known, and more are being discovered all the time.

The Crab Nebula

In 1054, Chinese astronomers noticed a "guest star" in the direction of the constellation Taurus. The new star was so bright that it could be seen during the daytime for 3 weeks, and it did not fade away altogether for many months. From the Chinese description of the changing brightness and color of the object, modern-day astronomers concluded that the guest star of 1054 was a fairly typical Type II supernova. Today, an expanding cloud of debris from the explosion occupies this place in the sky—forming an object called the Crab Nebula (see the chapter-opening photograph).

The Crab Nebula has filaments of glowing gas expanding away from the central star at 1,500 km/s—50 times faster than the expansion rate of a planetary nebula. These filaments contain anomalously high abundances of helium and other more massive chemical elements—the products of the nucleosynthesis that took place in the supernova and its progenitor star.

The Crab pulsar at the center of the nebula flashes 60 times a second: first with a main pulse associated with one of the "lighthouse" beams, then with a

Process of Science

OCCAM'S RAZOR

Occam's razor is a guiding principle in science: when scientists consider two hypotheses that explain a phenomenon equally well, they should give preference to the simpler theory and prioritize testing it. Simpler does not mean that the math is easier, or even that the concept is easy to understand. It means that the fewest number of other, new assumptions need to be made.

Jocelyn Bell, a student at Cambridge, has a "mystery signal" in her data.

Her adviser, Anthony Hewish, half-jokingly suggests "little green men" as the cause of the signal.

Bell and Hewish find four more such signals. It is unlikely that the same "little green men" would be sending the same signal from four separate locations in the sky.

They suggest the signals are from pulsating white dwarfs or neutron stars.

Franco Pacini and Thomas Gold each develop a detailed explanation involving rotating neutron stars.

This explanation relies entirely on previously understood physical phenomena: rotation, magnetic fields, and neutron stars.

It does not require assumptions about the existence of extra-terrestrials.

The neutron star explanation is "simpler."

Because simpler theories are easier to rule out, if they survive testing then scientists are more likely to consider them more seriously.

fainter secondary pulse associated with the other beam. As the Crab pulsar spins 30 times a second, it whips its powerful magnetosphere around with it. A few thousand kilometers from the pulsar, material in its magnetosphere must move at almost the speed of light to keep up with this rotation. Like a tremendous slingshot, the rotating pulsar magnetosphere flings particles away from the neutron star in a powerful wind moving at nearly the speed of light. This wind fills the space between the pulsar and the expanding shell. The Crab Nebula is almost like a big balloon, but instead of being filled with hot air, it is filled with a mix of very fast particles and strong magnetic fields. The energy that accelerates these particles is exactly equal to the energy lost as the pulsar's rotation slows down. Images of the Crab Nebula show this bubble as a glow from synchrotron radiation—a type of beamed radiation that is emitted as very fast moving particles spiral around the magnetic field.

CHECK YOUR UNDERSTANDING 17.3

One reason astronomers think neutron stars were formed in supernova explosions is that: (a) all supernova remnants contain pulsars; (b) pulsars are made of heavy elements, such as those produced in supernova explosions; (c) pulsars spin very rapidly, as did the massive star just before it exploded; (d) pulsars sometimes have material around them that looks like the ejecta from supernovae.

17.4 Star Clusters Are Snapshots of Stellar Evolution

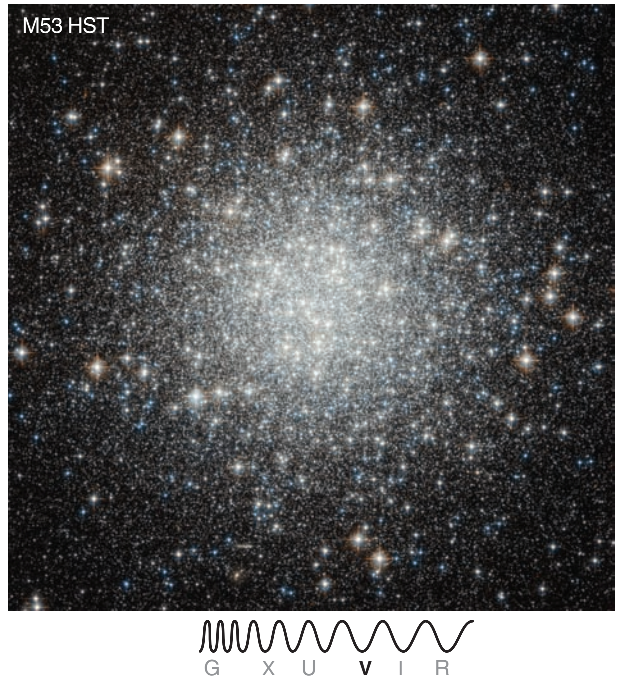
Recall from Chapter 15 that when an interstellar cloud collapses, it breaks into pieces, forming not one star but many stars of different masses. These large groups of gravitationally bound stars are called star clusters. **Globular clusters** are densely packed collections of hundreds of thousands to millions of stars (**Figure 17.16a**). **Open clusters** are much less tightly bound collections of a few dozen to a few thousand stars (Figure 17.16b). Because stars in a cluster are formed together at nearly the same time, observations of star clusters at different ages provide evidence for the evolution of stars of different masses.

Cluster Ages

In the 1920s, astronomers plotted the observed brightness versus the spectral type for as many stars as possible in each cluster. The resulting cluster H-R diagrams showed stars of all the categories in the "textbook" H-R diagram (see Figure 13.15). Because all of the stars in a cluster are at approximately the same distance from Earth, the effect of distance on the brightness of each star is the same. By matching the main sequence on the observed cluster H-R diagram to the main sequence on the "textbook" H-R diagram, astronomers could estimate the distance to a cluster.

Astronomers also realized that the cluster H-R diagrams offered clues to the newly developing theories of stellar evolution. All of the stars in a cluster formed together at nearly the same time, so a look at a cluster that is 10 million years old shows what the stars of different masses evolve into during the first 10 million years after they form. A look at a cluster 10 billion years after it formed shows

(a)



(b)

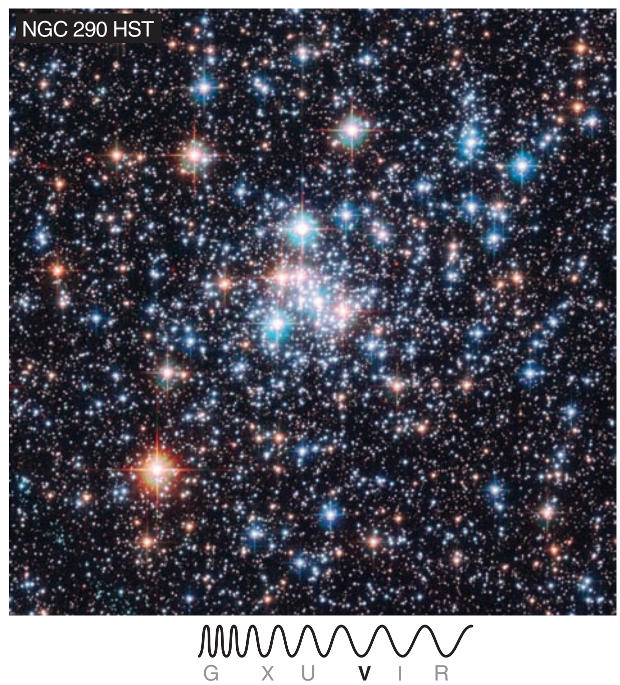


Figure 17.16 (a) Globular clusters can have hundreds of thousands of stars. (b) Open clusters have up to hundreds or thousands of stars.

what becomes of stars of different masses after 10 billion years. **Figure 17.17a** shows an H-R diagram of a very young cluster, NGC 6530. O, B, and A stars are on the main sequence; F through M stars with lower masses are still evolving *to* the main sequence. There are no red giants or white dwarfs. In contrast, Figure 17.17b shows the H-R diagram of a very old cluster, M55. There are no high-mass stars on the main sequence because they have evolved off of it, but there are stars on the horizontal, red giant, and asymptotic giant branches and in the lower part of the main sequence. This globular cluster is about 12 billion years old.

Astronomers cannot watch an individual cluster age over millions of years, but they can observe different clusters of different ages. Astronomers explore cluster evolution by examining H-R diagrams of a *simulated* cluster of 40,000 stars as it would appear at several different ages and then comparing it to observed H-R diagrams of actual clusters. In **Figure 17.18a**, most stars have not yet reached the main sequence. Star formation in a molecular cloud is spread out over several million years, and it takes considerable time for lower-mass stars to contract to reach the main sequence. The H-R diagram of a very young cluster typically shows many lower-mass stars located well above the main sequence; eventually, they move onto the main sequence.

The more massive a star is, the shorter its life on the main sequence will be. Figure 17.18b shows that after only 4 million years, all stars with masses greater

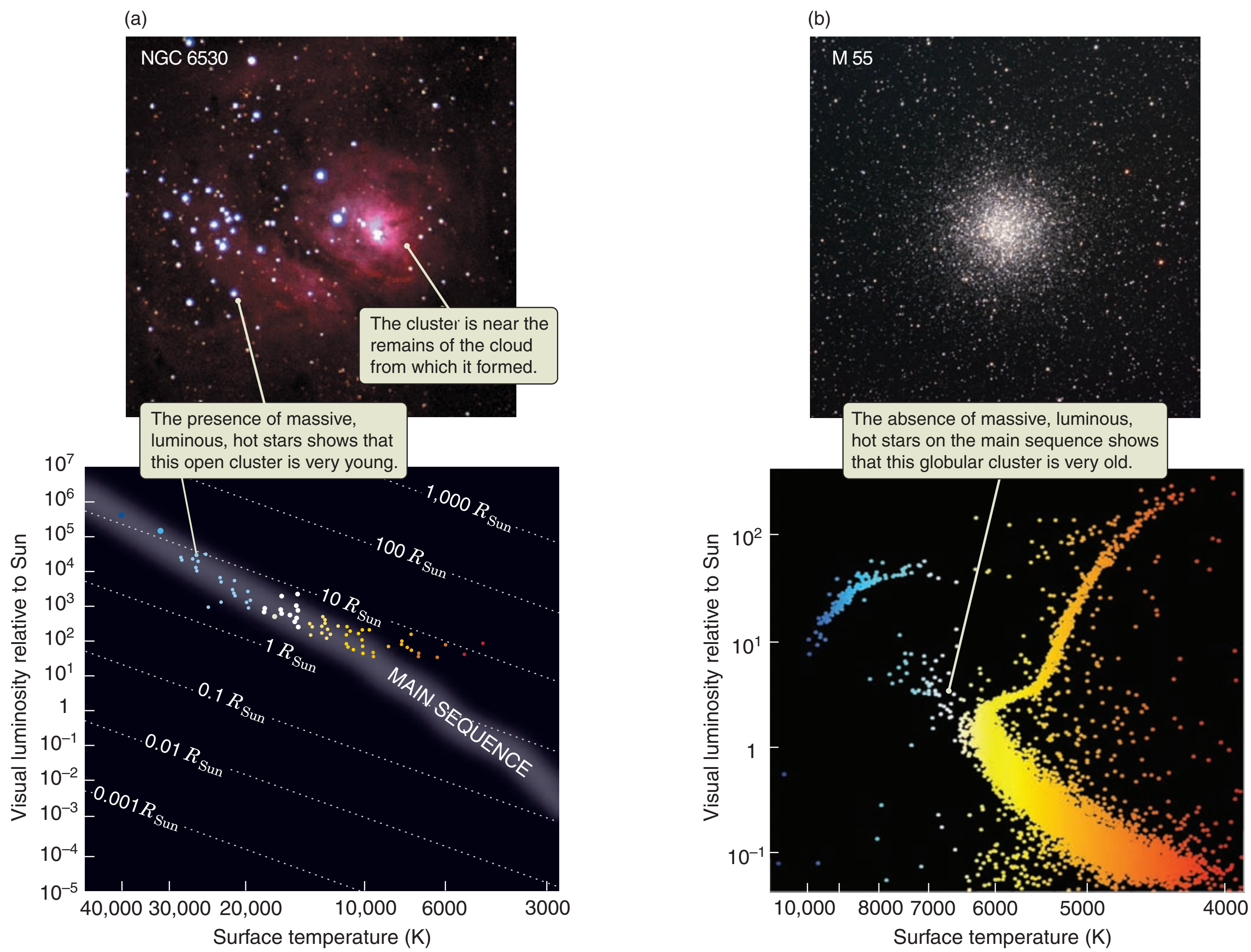


Figure 17.17 This figure shows H-R diagrams of (a) a very young cluster (2 million years old), NGC 6530, and (b) a very old cluster (12 billion years old), M55. In (a), some of the stars haven't yet arrived on the main sequence. In (b), more than the top half of the main sequence has already evolved. Note that the vertical scales are logarithmic, (b) is zoomed in compared with (a).

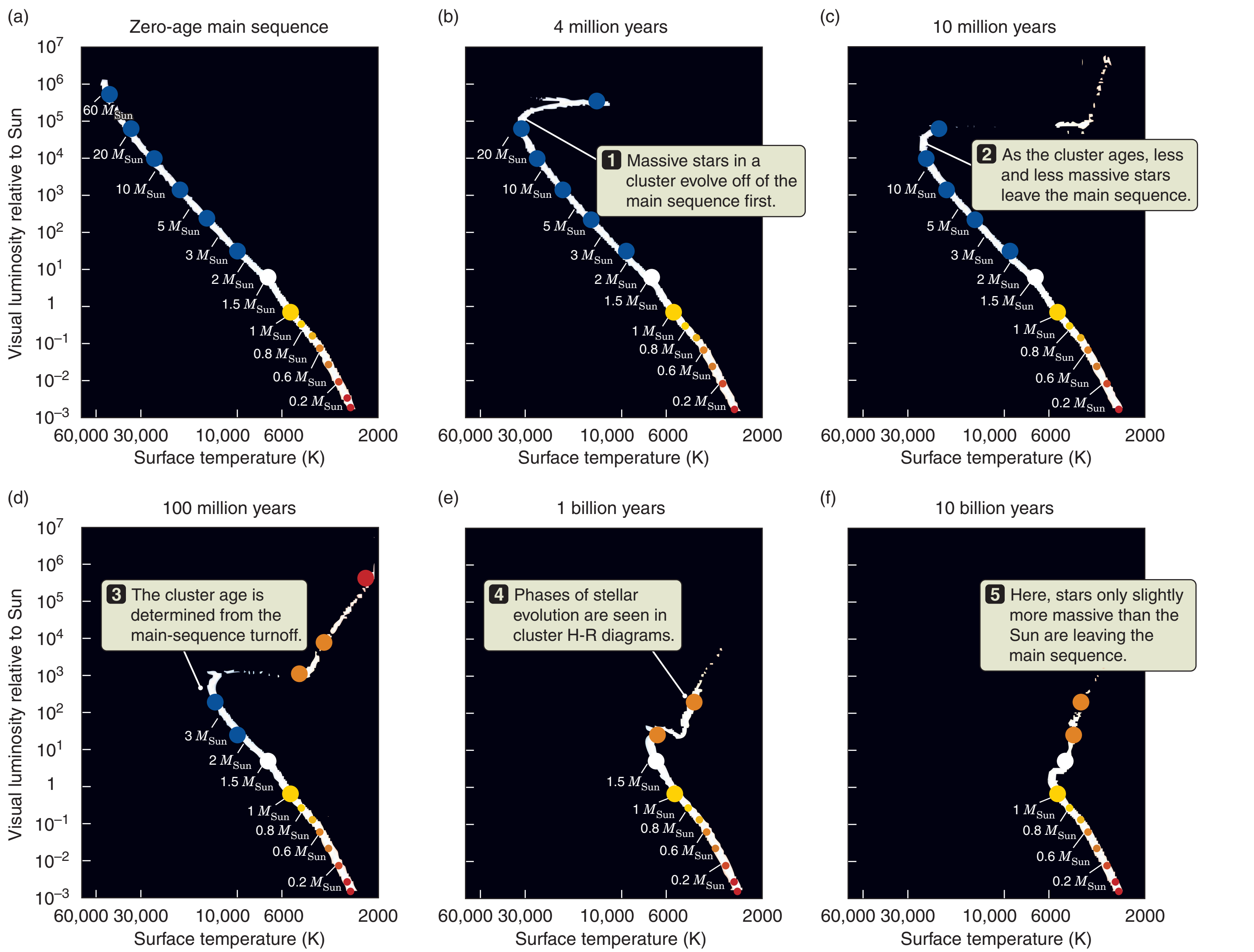


Figure 17.18 H-R diagrams of star clusters are snapshots of stellar evolution. These H-R diagrams of a simulated cluster of 40,000 stars of solar composition are shown at different times after the birth of the cluster. Note the progression of the main-sequence turnoff to lower and lower masses. For the purposes of the simulation, the stars are all placed on the main sequence at zero age. In reality, however, the lowest-mass stars have not yet reached the main sequence by the time the most massive stars have left it.

than about 20 M_{Sun} have evolved off the main sequence and are now spread out across the top of the H-R diagram. The most massive stars have already disappeared from the H-R diagram entirely, having vanished in supernovae. As time goes on, stars of lower and lower mass evolve off the main sequence, and the turnoff point moves toward the bottom right in the H-R diagram. By the time the cluster is 10 million years old, illustrated in Figure 17.18c, only stars with masses less than about 15 M_{Sun} remain on the main sequence. The location of the most massive star that remains on the main sequence is called the **main-sequence turnoff**. As the cluster ages, the main-sequence turnoff moves farther and farther down the main sequence to stars of lower and lower mass.

As a cluster ages further, shown in Figures 17.18d and e, we see the details of all stages of stellar evolution. By the time the star cluster is 10 *billion* years old (Figure 17.18f), stars with masses of only 1 M_{Sun} are beginning to die. Stars slightly more massive than this are seen as giant stars of various types. Note how few supergiant and giant stars are present in any of the cluster H-R diagrams. The supergiant, giant, horizontal, and asymptotic giant branch phases in the evolution of stars pass so quickly in comparison with a star's main-sequence lifetime

Nebraska Simulation: H-R Diagram Star Cluster Fitting Explorer

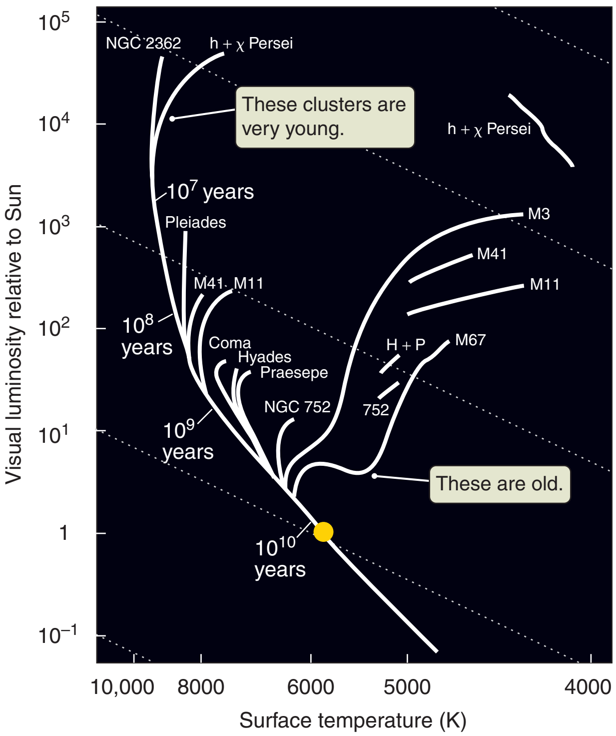


Figure 17.19 This figure shows H-R diagrams for clusters having a range of different ages. The ages associated with the different main-sequence turnoffs are indicated.

that even though this simulated cluster started with 40,000 stars, only a handful of stars are seen in these phases of evolution. Similarly, even though the majority of evolved stars in an old cluster are white dwarfs, all but a few of these stars will have cooled and faded into obscurity at any given time. More of these evolved stars are seen in the larger globular cluster M55 in Figure 17.17b.

To astronomers observing a star cluster, the location of the main-sequence turnoff immediately indicates the age of the cluster. **Figure 17.19** traces the observed H-R diagrams for several real star clusters. Once you know what to look for, the difference between young and old clusters is obvious. NGC 2362 is clearly a young cluster. Its complement of massive, young stars on the main sequence shows it to be only a few million years old. In contrast, cluster M3 has a main-sequence turnoff that indicates its cluster age is about 11 billion years. When the H-R diagrams of open clusters are studied, a wide range of ages is observed. Some open clusters contain the short-lived O and B stars and are therefore very young. Other open clusters contain stars that are somewhat older than the Sun. But even the youngest globular clusters are several billion years older than the oldest open clusters. Open clusters tend to be young because their stars are loosely bound and gradually leave the cluster. Globular clusters, in contrast, are tightly bound by gravity, allowing them to survive for billions of years.

Age, Color, and Different Chemical Composition

This understanding of stellar evolution applies even when groups of stars are so far away that individual stars cannot be seen. The light from a star cluster is dominated by its most luminous stars: massive, short-lived blue O and B main-sequence stars and evolved post-main-sequence supergiant and giant stars. If the cluster is young, then most of the light comes from luminous hot, blue stars and some red supergiants. If the cluster is old, then the light from the cluster has the color of red giants and red dwarf stars. This can be complicated by chemical composition—stars with lower amounts of massive elements in their atmospheres often look significantly bluer than stars that formed from chemically enriched material.

Astronomers usually can figure out something about the properties of a group of stars from its overall color. A group of stars with similar ages and other shared characteristics is called a **stellar population**. The link between color and specific characteristics will be useful when we begin to discuss the much larger collections of stars called galaxies. An especially bluish color to a galaxy or a part of a galaxy often signifies that the galaxy contains a young stellar population that still includes hot, luminous, blue stars that formed recently. In contrast, a galaxy or part of a galaxy that has a reddish color is usually composed primarily of an old stellar population.

All elements more massive than boron are formed in stars and are expelled into the interstellar medium when a star dies. New stars form from this material, thus the abundance of massive elements in the atmosphere of a star provides a snapshot of the chemical composition of the interstellar medium *at the time the star formed*. In main-sequence stars, material from the core does not mix with material in the atmosphere, so the abundances of chemical elements inferred from the spectrum of a star are the same as the abundances in the interstellar gas from which the star formed. Thus, the abundance of massive elements in the interstellar medium provides a record of the cumulative amount of star formation that has taken place up to the present time. Gas that shows large abundances of massive elements has gone through a great deal of stellar processing—so it contains more "recycled" material. Gas with low abundances of massive elements is more pristine.

The chemical composition of a star's atmosphere reflects the cumulative amount of star formation that has occurred up to that moment. Stars in globular clusters contain only very small amounts of massive elements; some globular-cluster stars contain only 0.5 percent as much of these massive elements as the Sun does, indicating that they were among the earliest stars to form. Open clusters are younger and contain stars that formed from a more enriched interstellar medium, and therefore they have higher amounts of the more massive elements.

Even the very oldest globular-cluster stars contain *some* amount of massive chemical elements. There must have been at least one generation of massive stars that lived and died, ejecting newly synthesized massive elements into space, before even the oldest globular clusters formed. Further, every red dwarf star less massive than about 0.5 M_{Sun} that ever formed is still around today. Note, though, that even a chemically rich star like the Sun, which is made of gas processed through approximately 9 billion years of previous generations of stars, is still composed of less than 2 percent massive elements. Luminous matter in the universe is still dominated by hydrogen and helium formed before the first stars. In upcoming chapters, you will learn that these variations in the chemical content of stars indicate a lot about the chemical evolution of galaxies.

CHECK YOUR UNDERSTANDING 17.4

If the main-sequence turnoff of a globular cluster occurs near the very top of the main sequence, then the cluster is: (a) very old; (b) very young; (c) very hot; (d) very dense.

Origins

Seeding the Universe with New Chemical Elements

On Earth, massive elements are everywhere. The surfaces of rocky planets contain silicon, oxygen, magnesium, and sodium. The iron-and-nickel solid inner core and liquid outer core of Earth are responsible for Earth's magnetic field. The most common chemical elements in biological molecules are carbon (C), hydrogen (H), nitrogen (N), oxygen (O), phosphorus (P), and sulfur (S)—all but hydrogen created in a dying star. The fact that these elements are here means that the Sun is not a first-generation star: it formed from material in the interstellar medium enriched by material from dying massive stars.

In supergiant, giant, and AGB stars, more massive chemical elements that formed from nuclear fusion deep within their interiors are carried upward and mixed with material in the outer parts of the star. As a star ages, its core grows hotter and hotter, and the temperature gradient within the star grows steeper. Under certain circumstances, convection can spread so deep into a star that chemical elements formed by nuclear burning within the star are dredged up and carried to the star's surface. For example, some AGB stars show an overabundance of carbon and other by-products of nuclear burning in their spectra. This extra carbon originated in each star's helium-burning shell and was carried to the surface by convection. For stars with lower masses, stellar winds and planetary nebulae carry the enriched outer layers off into interstellar space. The nuclear burning that occurs in supergiant stars goes well beyond the formation of elements such as carbon. Supernova explosions seed the universe with much more massive atoms, from iron and nickel up to uranium.

The oxygen atoms in the air you breathe and the water you drink were created by nucleosynthesis in dying stars. The iron atoms that are a key element of hemoglobin, which makes up the red blood cells that carry oxygen from your lungs to the rest of your body, formed in the explosions of massive stars. The nickel, copper, and zinc atoms in the coins in your pocket and the rare-earth atoms in your electronics were created in exploding massive stars. The Sun, the planets (including Earth), and all life on Earth are made of recycled stars. Supernovae are in you.

ARTICLES QUESTIONS

Astronomers report observations of a nearby series of supernovae.

We Are Swimming in a Superhot Supernova Soup

By **IAN O'NEILL, News.discovery.com**

Approximately 10 million years ago, a nearby cluster of stars erupted as a violent series of supernovae and, according to new observations, the million-degree plasma from these powerful detonations surround the Solar System today.

Although astronomers have known about this tenuous "Local Bubble" of gas for some time, a suborbital NASA sounding rocket launched by scientists in 2012 has removed any doubt about the origins of this 300 light-year wide feature we are cocooned within.

At a temperature in excess of a million degrees kelvin, the supernova "soup" is actually very tenuous—around 0.001 atoms per cubic centimeter—and very different from other particles of matter that occupy the interstellar medium.

The idea is that several neighboring massive stars went supernova after reaching the ends of their lives. Although supernovae are very powerful, apparently they weren't close enough to Earth to sterilize our planet of life. The supernovae occurred during the early stages of human evolution.

Evidence for the Local Bubble has been brewing for several decades and its presence was inferred from X-ray observations of the local galaxy—a background glow of X-ray radiation was detected in all directions. Although the evidence seemed strong for an ancient supernova soup, there was other possible interpretation.

"Within the last decade, some scientists have been challenging the (supernova) interpretation, suggesting that much or all of the soft X-ray diffuse background is instead a result of charge exchange," said F. Scott Porter of the Goddard Space Flight Center, Greenbelt, Maryland.

Charge exchange can occur between solar wind ions (charged particles that lack electrons) and neutral gases. When the two gases come into contact within the Solar System—between the solar wind and a comet's coma, for example—electrons can be stripped from the neutral particles, generating X-ray emissions.

Therefore, many astronomers argued that this diffuse X-ray glow observed in all directions may be a phenomenon *inside* the Solar System and not superhot particles from 10 million year–old supernovae *outside* the Solar System.

To put this debate to bed, an international team of scientists constructed an instrument called the Diffuse X-ray emission from the Local Galaxy (or DXL), which could distinguish between the two scenarios. After launching the instrument 160 miles in altitude atop a sounding rocket, well above the Earth's atmosphere, the instrument specifically detected the amount of charge exchange that is occurring within interplanetary space.

After 5 minutes of observations on December 12, 2012, the DXL was able to determine that only 40 percent of the diffuse soft-X-ray emissions were generated by charge exchange. The rest must be coming from the Local Cloud from outside the Solar System—so therefore the Local Bubble is real and we did indeed get rumbled by a series of supernovae 10 million years ago.

"This is a significant discovery," said Massimiliano Galeazzi, of the University of Miami in Coral Gables, who led the team. "(It) affects our understanding of the area of the galaxy close to the Sun, and can, therefore, be used as a foundation for future models of the galaxy structure."

ARTICLES QUESTIONS

1. Why would a group of stars in a cluster go supernovae at nearly the same time?
2. Why was a rocket needed to observe these X-rays instead of observing them from the ground on Earth?
3. What is the evidence that the X-rays came from outside of the Solar System?
4. How hot is the X-ray gas? Could you safely stick your hand in it—explain.
5. Go to NASA News (http://science.nasa.gov/science-news/science-at-nasa/2014/26aug_localbubble/) and watch the 4-minute ScienceCast video on this discovery. Which part of this story is explained better with video? Search at Nasa.gov—was the next DXL rocket launched in December 2015 as planned?

Summary

As high-mass stars evolve, their interiors form concentric shells of progressive nuclear fusion. Once they leave the main sequence, they may pass through the instability strip and become pulsating variable stars. High-mass stars eventually explode as Type II supernovae, which eject newly formed massive elements into interstellar space. The supernova explosion that ends the life of a massive star leaves behind a neutron star that contains between 1.4 and 3 M_{Sun} of neutron-degenerate matter packed into a sphere 10–20 km in diameter. Accretion of mass onto neutron stars produces X-rays in some binary systems. Pulsars are rapidly spinning, magnetized neutron stars. The Sun, the Solar System, Earth, and all life on Earth contain heavy elements created inside of earlier generations of short-lived massive stars.

LG 1 **Describe how the death of high-mass stars differs from that of low-mass stars.** The larger masses of high-mass stars allow them to fuse heavier elements than those produced in low-mass stars. This leads to a more violent death that leaves massive cores behind.

LG 2 **List the sequence of stages for evolving high-mass stars.** Evolving high-mass stars leave the main sequence as they burn heavier elements. Once an iron core is produced, the star becomes unstable and the core collapses, heating the material to cause photodisintegration of the iron nuclei and the merging of protons and electrons into neutrons. The outer layers bounce off of the dense core and produce a shock wave that travels outward. This shock wave causes neutrons to fuse with atomic nuclei and form more massive elements.

LG 3 **Explain the origin of chemical elements up to and heavier than iron.** The chain of nuclear fusion reactions consists of increasingly shorter stages of burning, resulting in more massive elements up to iron. However, these atoms are destroyed in the set of processes that define the core collapse. Heavy elements in the universe today were formed during the rebound explosion of massive stars as high-energy neutrons penetrated atomic nuclei. These neutrons then decay to protons, creating new elements with higher atomic numbers.

LG 4 **Identify how Hertzsprung-Russell (H-R) diagrams of clusters enable astronomers to measure the ages of stars and test theories of stellar evolution.** Clusters are groups of stars that were born together and are all at about the same distance from Earth. H-R diagrams of clusters show stars leaving the main sequence in a progression from the highest-mass stars to the lowest-mass stars, confirming theories of stellar evolution. The location of the main-sequence turnoff indicates the age of the cluster.

? UNANSWERED QUESTIONS

- Blue straggler stars, found in clusters, are bluer and brighter than the stars at the main-sequence turnoff point. How do they fit into the picture of stellar evolution? They may have resulted from mass transfer in a binary pair or from the merger of two single or two binary stars, either of which could have resulted in a more massive star than what might be expected from the age of the cluster. Astronomers study the environments of these stars by estimating the likelihood of collisions and the number of binary systems, which may be different in clusters where the density of stars is high.
- What creates magnetars? There is a class of pulsars called magnetars, which are characterized by extremely large magnetic fields. These objects are observed to produce bursts of lower energy gamma rays. The origin of their huge magnetic fields is not well understood. These fields may originate from a dynamo in the interior of a superconducting region of the neutron star, but we do not know whether ordinary pulsars go through a magnetar phase.

Questions and Problems

Test Your Understanding

1. Why does the interior of an evolved high-mass star have layers like an onion?
   a. Heavier atoms sink to the bottom because stars are not solid.
   b. Before the star formed, heavier atoms accumulated in the centers of clouds because of gravity.
   c. Heavier atoms fuse closer to the center because the temperature and pressure are higher there.
   d. Different energy transport mechanisms occur at different densities.

2. Arrange the following elements in the order they burn inside the nucleus of a high-mass star during the star's evolution.
   a. helium
   b. neon
   c. oxygen
   d. silicon
   e. hydrogen
   f. carbon

3. Elements heavier than iron originated
   a. in the Big Bang.
   b. in the cores of low-mass stars.
   c. in the cores of high-mass stars.
   d. in the supernova explosions of high-mass stars.

4. A pulsar pulses because
   a. its spin axis crosses Earth's line of sight.
   b. it spins.
   c. it has a strong magnetic field.
   d. its magnetic axis crosses Earth's line of sight.

5. Study Figure 17.5. If it were possible to watch a high-mass star move to the right, along one of these post-main-sequence lines, what would you observe happening to the star's color?
   a. It would become redder.
   b. It would become bluer.
   c. It would remain the same.

6. Study Figure 17.5. If it were possible to watch a high-mass star move to the right, along the topmost of these post-main-sequence lines, what would you observe happening to the star's size?
   a. It would become much larger.
   b. It would become much smaller.
   c. It would remain the same size.

7. Study Figure 17.18. If the Sun were a member of a globular cluster, that cluster's H-R diagram would fall between
   a. (a) and (b)
   b. (b) and (c)
   c. (c) and (d)
   d. (d) and (e)
   e. (e) and (f)

8. In a high-mass star, hydrogen fusion occurs via
   a. the proton-proton chain.
   b. the CNO cycle.
   c. gravitational collapse.
   d. spin-spin interaction.

9. The layers in a high-mass star occur roughly in order of
   a. atomic number.
   b. decay rate.
   c. magnetic field strength.
   d. spin state.

10. Eta Carinae is an extreme example of
    a. a massive star.
    b. a planetary nebula.
    c. a supernova remnant.
    d. an ancient star.

11. Iron fusion cannot support a star because
    a. iron oxidizes too quickly.
    b. iron absorbs energy when it fuses.
    c. iron emits energy when it fuses.
    d. iron is not dense enough to hold up the layers.

12. The start of photodisintegration of iron in a star sets off a process that *always* results in a
    a. supernova.
    b. neutron star.
    c. supergiant.
    d. pulsar.

13. The Crab Nebula is a test of our ideas about supernova explosions because
    a. the system contains an x-ray binary.
    b. the nebula is slowly expanding.
    c. the supernova was observed in 1054 and now astronomers see a pulsar in the nebula.
    d. the original star was like the Sun before exploding.

14. What mechanism provides the internal pressure inside a neutron star?
    a. ordinary pressure from hydrogen and helium gas
    b. degeneracy pressure from neutrons
    c. degeneracy pressure from electrons
    d. rapid rotation

15. Very young star clusters have main-sequence turnoffs
    a. that drop below the main sequence.
    b. at the top left of the main sequence.
    c. at the bottom right of the main sequence.
    d. in the middle of the main sequence.

Thinking about the Concepts

16. Explain the differences between the ways that hydrogen is converted to helium in a low-mass star (proton-proton chain) and in a high-mass star (CNO cycle). What is the catalyst in the CNO cycle, and how does it take part in the reaction?

17. How does a high-mass star begin fusing helium in its core? How is this process different from what happens in low-mass stars?

18. Why does the core of a high-mass star not become degenerate, as the core of a low-mass star does?

19. List the two reasons why each post-helium-fusion cycle for high-mass stars (carbon, neon, oxygen, silicon, and sulfur) becomes shorter than the preceding cycle.

20. Cepheids are highly luminous, variable stars in which the period of variability is directly related to luminosity. Why are Cepheids good indicators for determining stellar distances?

21. Identify and explain two important ways in which supernovae influence the formation and evolution of new stars.

22. Study the Process of Science Figure. Why is this pulsar explanation "simpler"?

23. Describe what an observer on Earth will witness when Eta Carinae explodes.

24. Recordings show that neutrinos from SN 1987A were detected on February 23, 1987. About 3 hours later it was detected in optical light. What was the reason for the time delay?

25. Why can the accretion disk around a neutron star release so much more energy than the accretion disk around a white dwarf, even though the two stars have approximately the same mass?

26. In Section 17.2, you learned that Type II supernovae blast material outward at 30,000 km/s. The material in the Crab Nebula described in Section 17.3 is expanding at only 1,500 km/s. What explains the difference?

27. An experienced astronomer can take one look at the H-R diagram of a star cluster and immediately estimate its age. How is this possible?

28. Explain how astronomers know that there was an even earlier generation of stars before the oldest observed stars.

29. What is the binding energy of an atomic nucleus? How does this quantity help astronomers calculate the energy given off in nuclear fusion reactions?

30. Explain how Earth is made up of material from supernova.

Applying the Concepts

31. Study Figure 17.2. What fraction of the star is helium at time $t = 0$ and at time $t = 7$ million years?

32. Study Figure 17.5. How much hotter, larger, and more luminous than the Sun is the uppermost main-sequence star on this H-R diagram?

33. Study Figure 17.4. Are the radius of the core and the radius of the star represented to scale in this figure? What fraction of the star's radius is the core's radius?

34. If the Crab Nebula has been expanding at an average velocity of 1500 km/s since the year 1054, what was its average radius in the year 2014? (Note: There are approximately 3×10^7 seconds in a year.)

35. Suppose you observe a Classical Cepheid variable with a period of 10 days. What is the luminosity of this star? What other piece of information would you need to find out how far away this star is?

36. Figure 17.13 shows the relative abundance of the elements. Is this a log or a linear plot? Explain what it means that oxygen lies on the y-axis at 10^{-3}.

37. Use Working It Out 17.2 to find the surface gravity on a neutron star with radius 10 km and mass 2.8 M_{Sun}.

38. Use Working It Out 17.2 to find the escape velocity from a neutron star with radius 10 km and mass 2.8 M_{Sun}.

39. The Milky Way has about 50,000 stars of average mass (0.5 M_{Sun}) for every main-sequence star of 20 M_{Sun}. But 20-M_{Sun} stars are about 10,000 times as luminous as the Sun, and 0.5-M_{Sun} stars are only 0.08 times as luminous as the Sun.
    a. How much more luminous is a single massive star than the total luminosity of the 50,000 less massive stars?
    b. How much mass is in the lower-mass stars compared to the single high-mass star?
    c. Which stars—lower-mass or higher-mass stars—contain more mass in the galaxy, and which produce more light?

40. In a large outburst in 1841, the 120-M_{Sun} star Eta Carinae was losing mass at the rate of 0.1 M_{Sun} per year.
    a. The mass of the Sun is 2×10^{30} kg. How much mass (in kilograms) was Eta Carinae losing each minute?
    b. The mass of the Moon is 7.35×10^{22} kg. How does Eta Carinae's mass loss per minute compare with the mass of the Moon?

41. An O star can lose 20 percent of its mass during its main-sequence lifetime. Estimate the average mass loss rate (in solar masses per year) of a 25-M_{Sun} O star with a main sequence lifetime of 7 million years.

42. The approximate relationship between the luminosity and the period of Cepheid variables is L_{star} (in L_{Sun}) = 335 P (in days). Delta Cephei has a cycle period of 5.4 days and a parallax of 0.0033 arcsecond (arcsec). A more distant Cepheid variable appears 1/1,000 as bright as Delta Cephei and has a period of 54 days.
    a. How far away (in parsecs) is the more distant Cepheid variable?
    b. Could the distance of the more distant Cepheid variable be measured by parallax? Explain.

43. For a pulsar that rotates 30 times per second, at what radius in the pulsar's equatorial plane would a co-rotating satellite (rotating about the pulsar 30 times per second) have to be positioned to be moving at the speed of light? Compare this to the pulsar radius of 1 km.

44. Verify the claim in Section 17.3 that Earth would be roughly the size of a football stadium if it were as dense as a neutron star.

45. Estimate the size of a neutron star with the mass of the Sun.

USING THE WEB

46. Go to the Chandra X-ray Observatory's "Variable Stars" Web page (http://chandra.harvard.edu/edu/formal/variable_stars/index.html). Do the two exercises on Cepheid variable stars, which ask you to estimate their changes in brightness. You might want to look at Appendix 7 to review apparent magnitudes before you do the projects.

47. The International Astronomical Union's "List of Recent Supernovae" (http://cbat.eps.harvard.edu/lists/RecentSupernovae.html) includes all recently discovered supernovae. Pick a few of the most recent ones. What type of supernova is each one? How bright is it? Why are these so much fainter than the novae you looked at in Chapter 16? Are Type Ia or Type II supernovae more common?

48. What method is used by the Intermediate Palomar Transient Factory (iPTF) survey (http://www.ptf.caltech.edu/iptf) to find supernova? What kinds of supernovae has this study found? How will the replacement camera, the Zwicky Transient Factory (ZTF), improve the search for supernova? Has ZTF found any yet?

49. Go to the website for the Gaia mission (http://esa.int/science/gaia). How is this mission contributing to the study of variable stars? How is it contributing to the study of novae and supernovae?

50. Go to the "Einstein@Home" website (http://einsteinathome.org). In this distributed computing project, volunteers use their spare computer processing power to help search for new pulsars. Look over the "News" section on the right. Have any pulsars been found lately? Join the project, create an account, download BOINC, and follow directions to look for pulsars.

smartwork5

If your instructor assigns homework in Smartwork5, access your assignments at digital.wwnorton.com/astro5.

EXPLORATION

The CNO Cycle

digital.wwnorton.com/astro5

Nuclear reactions usually involve many steps. In the Exploration for Chapter 14, you investigated the proton-proton chain. In this Exploration, you will study the CNO cycle, which is even more complex. Visit the Student Site at the Digital Landing Page, and open the "CNO Cycle" interactive simulation in Chapter 17.

First, press "Play Animation" and watch the animation all the way through. Press "Reset Animation" to clear the screen, and then press "Play Animation" again, allowing the animation to proceed past the first collision before pressing "Pause."

1 Which atomic nuclei are involved in this first collision?

2 What color is used to represent the proton (hydrogen nucleus)?

3 What does the blue squiggle represent?

4 What atomic nucleus is created in the collision?

5 The resulting nucleus is not the same type of element as either of the two that entered the collision. Why not?

Press "Play Animation" again, and then pause as soon as the yellow ball and the dashed line appear.

6 Is this a collision or a spontaneous decay?

7 What does the yellow ball represent?

8 What does the dashed line represent?

9 The resulting nucleus has the same number of nucleons (13), but it is a different element. What happened to the proton that was in the nitrogen nucleus but is not in the carbon nucleus?

Proceed past the next two collisions, to "^{15}O."

10 Study the pattern that is forming. When a blue ball comes in, what happens to the number of nucleons and the type of the nucleus (that is, what happens to the "12" and the "C," or the "14" and the "N")?

11 What is emitted in these collisions?

Proceed until "^{15}N" appears.

12 Is this a collision or a spontaneous decay?

13 Which previous reaction is this most like?

Now proceed to the end of the animation.

14 After the final collision, a line is drawn back to the beginning, telling you what type of nucleus the upper red ball represents. What is this nucleus?

15 How many nucleons are not accounted for by the upper red ball? (Hint: Don't forget the ^{1}H that came into the collision.) These nucleons must be in the nucleus represented by the bottom red ball.

16 Carbon has six protons. Nitrogen has seven. How many protons are in the nucleus represented by the bottom red ball?

17 How many neutrons are in the nucleus represented by the bottom red ball?

18 What element does the bottom red ball represent?

19 What is the net reaction of the CNO cycle? That is, which nuclei are combined and turned into the resulting nucleus?

20 Why is ^{12}C not considered to be part of the net reaction?

18 Relativity and Black Holes

Some stars leave behind a black hole at the end of their lives. Black holes have such extreme conditions that the laws of Newtonian physics are inadequate to describe them. To discuss black holes, we must understand how Albert Einstein changed the way physicists thought about the nature of space and time in the early 20th century. Einstein's special theory of relativity shows that matter behaves differently when it is traveling near the speed of light. Einstein's general theory of relativity shows that space itself is warped near very massive objects. This warping of space is so extreme at a black hole, it is as if there is a hole in space. In this chapter, we will move beyond Newtonian ideas of space and time in order to understand black holes.

LEARNING GOALS

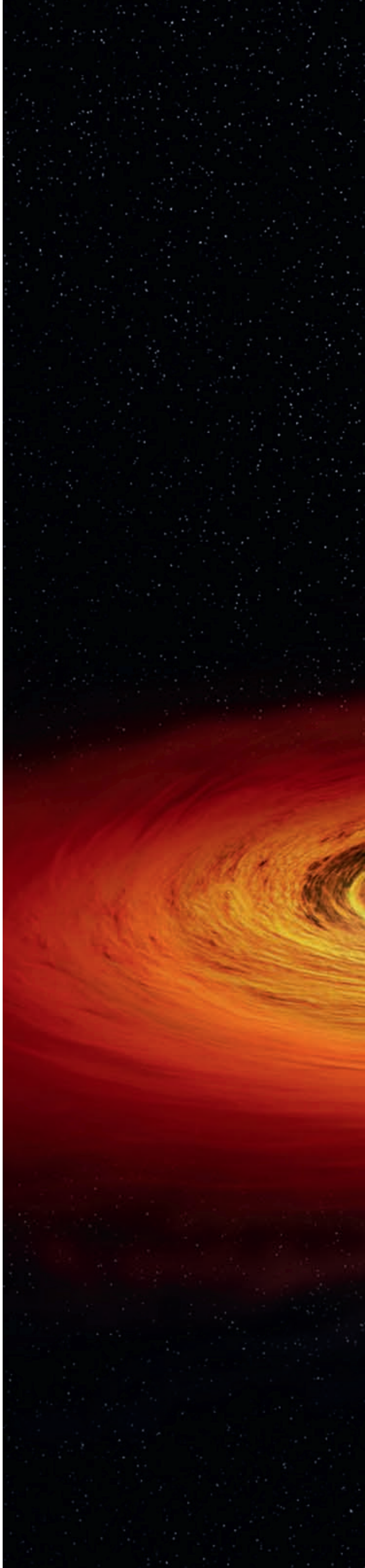
By the conclusion of this chapter, you should be able to:

LG 1 Describe how the motion of the observer affects the observed velocity of objects.

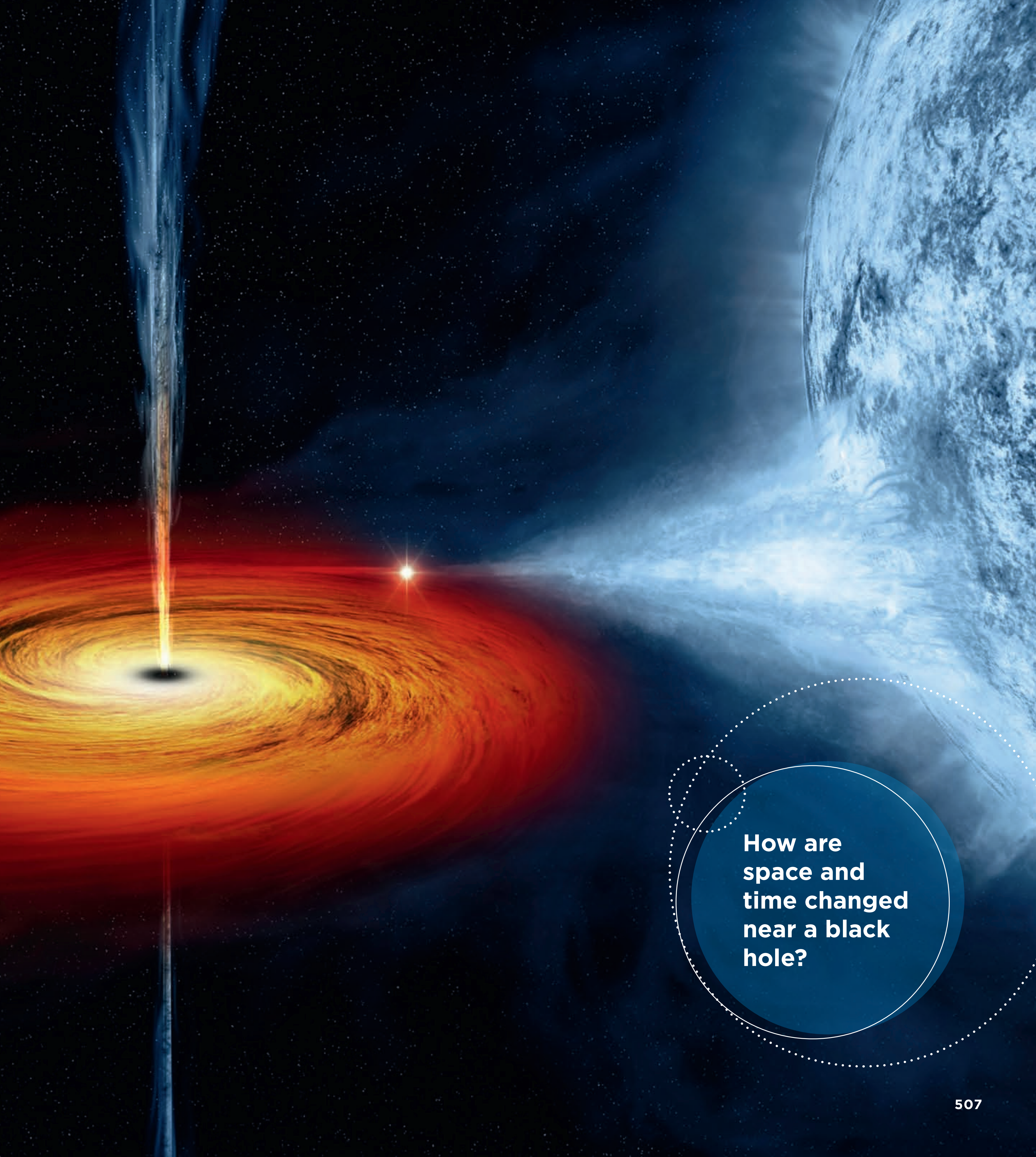
LG 2 Discuss the observable consequences of the relationship between space and time.

LG 3 Explain how gravity is a consequence of the way mass distorts the very shape of spacetime.

LG 4 Explain why the most massive stars end as black holes, and describe the key properties of these stellar black holes.



This artist's rendition of the black hole Cygnus X-1 shows the accretion disk and the massive star that feeds it. ▶▶▶



How are space and time changed near a black hole?

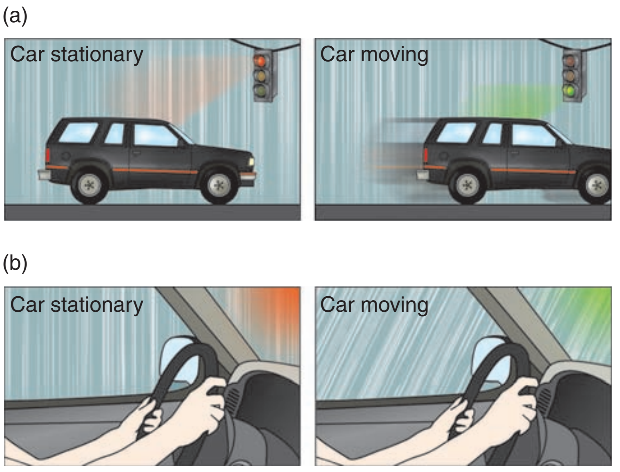


Figure 18.1 On a windless day, the direction in which rain falls depends on the frame of reference in which it is viewed. (a) From outside the car, the rain is seen to fall vertically downward whether the car is stationary or moving. (b) From inside the car, the rain is seen to fall vertically downward if the car is stationary; but if the car is moving, the rain is seen to fall at an angle determined by the speed and direction of the car's motion.

18.1 Relative Motion Affects Measured Velocities

All observers, whatever their motion, measure the same speed of light. This observed fact has profound implications for relative motion, space, and time. In this section, we will lay the groundwork for our discussion of relativity by considering how relative motion affects measurements and how these effects change at very high speeds.

Aberration of Starlight

The first direct measurement of the effect of Earth's motion around the Sun was made in the 18th century. Imagine that you are sitting in a car in a windless rainstorm, as shown in **Figure 18.1**. If the car is stationary and the rain is falling vertically, when you look out your side window you see raindrops falling straight down. When the car is moving forward, however, the situation is different. Between the time a raindrop appears at the top of your window and the time it disappears beneath the bottom of your window, the car has moved forward. The raindrop disappears beneath the window behind the point at which it appeared at the top of the window, which means the raindrop looks as if it is falling at an angle, even though in reality it is falling straight down. As you go faster, the apparent front-to-back slant of the raindrops increases, and their apparent paths become more slanted. An observer by the side of the road would say the raindrops are coming from directly overhead, but to you in the moving car they are coming from a direction in front of the car. You are observing this apparent motion of the raindrops from within your own unique frame of reference.

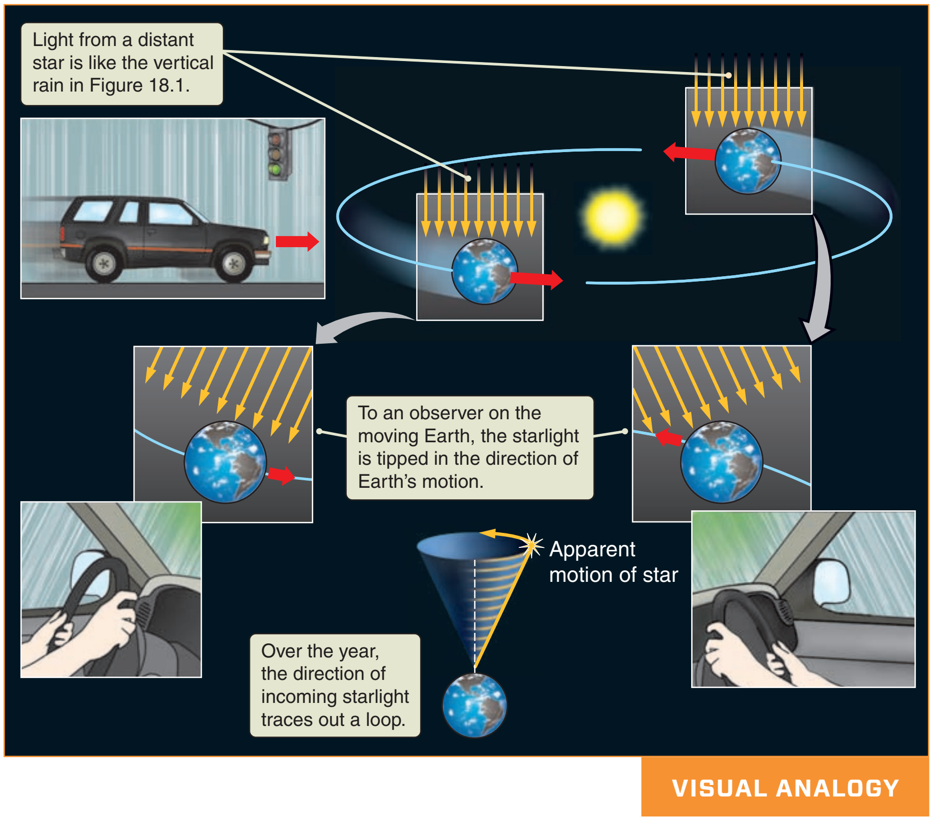


Figure 18.2 When we observe stars, their apparent positions are deflected slightly toward the direction in which Earth is moving. As Earth orbits the Sun, stars appear to trace out small loops in the sky. This effect is called *aberration of starlight*.

The light from a distant star arrives at Earth from the direction of the star, as shown in **Figure 18.2**. However, just as the raindrops appeared to be coming from in front of the moving car in Figure 18.1b, an observer on the moving Earth sees the starlight coming from a slightly different direction. Because the direction of Earth's motion around the Sun continually changes during the year, the apparent position of a star in the sky moves in a small loop, a phenomenon known as the **aberration of starlight**. This shift in apparent position was first detected in the 1720s by two astronomers—Samuel Molyneux and James Bradley. Measurement of the aberration of starlight shows that Earth moves on its path about the Sun with an average speed of just under 30 kilometers per second (km/s). Because distance equals speed multiplied by time, astronomers were able to use this measurement to determine the circumference, and therefore the radius, of Earth's orbit. The speed of Earth (29.8 km/s) multiplied by the number of seconds in 1 year (3.16×10^7 seconds) gives a circumference of 9.42×10^8 km. Astronomers used this circumference to estimate the radius of Earth's orbit (1.5×10^8 km). The aberration of starlight is an astronomical example of the effects of relative motion on a measurement.

Relative Speeds Close to the Speed of Light

Every observer inhabits a reference frame, a set of coordinates in which the observer measures distances and speeds. A nonaccelerating reference frame, in which no net force is acting, is known as an **inertial reference frame**. Newton's laws of motion predict how the motion of an object will be measured by observers in different inertial reference frames. For example, the observed motion of a ball thrown from a moving car depends on the reference frame of the observer (recall the analogy for the Coriolis effect from Figure 2.11). If light behaved like other objects, the speed of light should differ from one observer to the next as a result of the observer's motion, just like the speed of a ball thrown from a moving car. However, the results of laboratory experiments with light in the late 19th and early 20th centuries conflicted with these predictions. Scientists found instead that all observers measure exactly the same value for the speed of light, regardless of the frame of reference of each observer.

As shown in **Figure 18.3a**, imagine that you are in a red car traveling at 100 kilometers per hour (km/h) on a highway, and you throw a ball at a speed of 50 km/h out the window at an oncoming green car, also traveling at 100 km/h. An observer standing by the side of the road watches the entire event. In *your* reference frame (Figure 18.3a, top panel), the red car is stationary; that is, you and the car are moving together, so the car does not move relative to you. So in your reference frame, you measure the ball traveling at 50 km/h. However, to the observer standing by the road (Figure 18.3a, middle panel), the ball is moving at 150 km/h—50 km/h from throwing it plus 100 km/h from the motion of the car. To passengers in the oncoming green car moving at 100 km/h (Figure 18.3a, bottom panel), the speed of the ball is 250 km/h, because the ball approaches them at 150 km/h, and they approach the ball at 100 km/h. This example shows that the speed of the ball depends on how the observer, the cars, and the ball are moving relative to one another. The velocities are added together to find the velocity of one object relative to another. This is Galilean relativity, which you use in everyday life.

Figure 18.3b demonstrates how light differs from the ball in the previous example. Imagine that you are riding in the yellow spaceship at half the speed of light ($0.5c$) and you shine a beam of laser light forward (Figure 18.3b, top panel). You measure the speed of the beam of light to

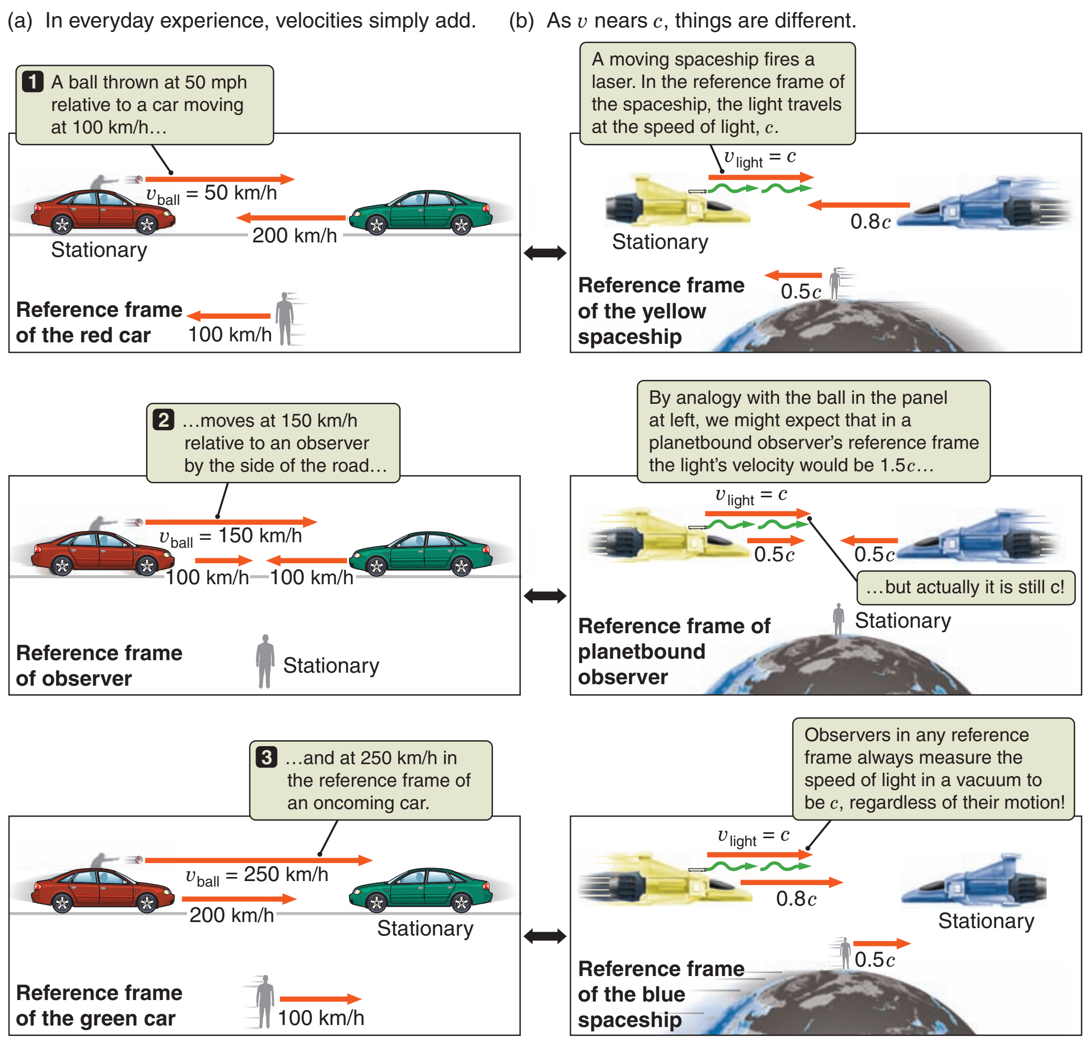


Figure 18.3 (a) The Newtonian rules of motion apply in daily life; however, these rules break down when speeds approach the speed of light. (b) The fact that light itself always travels at the same speed for any observer is the basis of special relativity, and it implies that velocities don't simply add together, as they do in the Newtonian world.

be c, or 3×10^8 meters per second (m/s), as expected because you are holding the source of the light. But the observer on the planet also measures the speed of the passing beam of light to be 3×10^8 m/s, not the sum of the speed of light plus the speed of the spaceship, 1.5c (Figure 18.3b, middle panel). Even a passenger in an oncoming blue spacecraft traveling at 0.5c, as shown in the bottom panel of Figure 18.3b, finds that the beam from your light is traveling at exactly c in her own reference frame, and not the 2.0c sum of the speeds of the spacecraft and the laser light. At speeds close to the speed of light, speeds do not simply add. This is true not just for light, but for all ordinary objects moving at nearly the speed of light. Using Einstein's relativistic formulas, the relative speed between the two spacecraft in the top and bottom panels of Figure 18.3b ($0.5c + 0.5c$) adds to 0.8c, not 1.0c! At **relativistic speeds** (speeds close to the speed of light), everyday experience no longer holds true. Every observer always finds that light in a vacuum travels at exactly the same speed, c, regardless of his or her own motion or the motion of the light source. This directly conflicts with Newtonian theory and Galilean relativity.

CHECK YOUR UNDERSTANDING 18.1

Which beam of light is moving faster: (a) that from the headlight of a parked car; (b) that from the headlight of a moving car; (c) that from the headlight of a moving spaceship; (d) all the beams are moving at c regardless of reference frame.

18.2 Special Relativity Explains How Time and Space Are Related

Albert Einstein's first scientific paper, written when he was a 16-year-old student, was about traveling along with a light wave, moving in a straight line at constant speed. Einstein reasoned that according to Newton's laws of motion, it should be possible to "keep up" with light so that you are moving right along with it. In this inertial reference frame, the light is stationary: an oscillating electric and magnetic wave that does not move. This was impossible according to Maxwell's equations for electromagnetic waves (discussed in Chapter 5). Einstein later took a radical new approach to resolve the conflict between experiments and Newton's laws of motion. Rather than starting with preconceived ideas about space and time, Einstein started with the observed fact that light always travels at the same speed, and then he reasoned backward to find out what that must imply about space and time. This led to the 1905 publication of his **special theory of relativity**, sometimes called *special relativity*, which describes the effects of traveling at constant speeds close to the speed of light. In this section, we explore special relativity and some of its consequences.

Time and Relativity

In developing special relativity, Einstein focused his thinking on pairs of *events*. An **event** is something that happens at a particular location in space at a particular time. Snapping your fingers is an event, because that action has both a time and a place. Everyday experience indicates that the distance between any two events depends on the reference frame of the observer. Imagine you are sitting in a car that is traveling on the highway in a straight line at a constant 60 km/h. You snap your fingers (event 1), and a minute later you snap your fingers again (event 2). In your reference frame you are stationary, and the two events happened at exactly the same place—in the car. The events were, however, separated by a minute in time. This is very different from what happens in the reference frame of an observer sitting by the road. This observer agrees that the second snap of your fingers (event 2) occurred a minute after the first snap of your fingers (event 1), but to this observer the two events were separated from each other in space by a kilometer, the distance your car traveled in the minute. In this "Newtonian" view, the distance between two events depends on the motion of the observer, but the *time* between the two events does not. Special relativity instead finds that both the distance and the time between events varies depending on the motion of the observer.

The notion that different observers will measure time differently is a *very* counterintuitive idea, but it is central to special relativity and therefore to our scientific understanding of the universe as well. To see how Einstein arrived at the concept of relative time, consider his thought experiment known as the boxcar experiment. In this experiment, observer 1 is in a boxcar of a train moving to the right. Observer 1 has a lamp, a mirror (mounted to the roof of the boxcar), and a clock. Observer 2 is standing on the ground outside. The clock is based on a value that everyone can agree on—such as the speed of light.

Figure 18.4a shows the experimental setup as seen by observer 1, who is stationary with respect to the clock. At time t_1, event 1 happens: the lamp gives off a pulse of light. The light bounces off a mirror at a distance l meters away and then heads back toward its source. At time t_2, event 2 happens: the light arrives at the clock and is recorded by a photon detector. (Note that the light beam is shown leaving and arriving at two different locations. This was done by the artist so that you can see both events in the figure. Both events occur at the same location.) The time between events 1 and 2 is just the distance the light travels ($2l$ meters) divided by the speed of light: $t_2 - t_1 = 2l/c$.

Figure 18.4b shows the experiment as seen by observer 2, stationary on the ground outside the train, which is moving at speed v. In observer 2's reference frame, the clock moves to the right between the two events, so the light has farther to go because of the horizontal distance. The time between the two events is still the distance traveled divided by the speed of light, but now that distance is *longer* than $2l$ meters. Because the speed of light is the same for all observers, the time between the two events must be longer as well.

The two events are the same two events, regardless of the reference frame from which they are observed. Because the speed of light is the same for all observers, there must be more time between the two events when viewed from a reference frame in which the clock is moving (observer 2). The seconds of a moving clock are stretched. That is, it takes a moving clock more time than a stationary clock to complete one "tick." Therefore, the passage of time must depend on an observer's frame of reference. Because both frames of reference are equally good places to do physics, both time measurements are valid in their own frames, even though they differ from one another.

In this experiment, light travels farther between events in a moving boxcar than between events in a stationary boxcar and consequently takes a longer time to travel between the events in the moving boxcar. Einstein realized that the *only* way the speed of light can be the same for all observers is *if the passage of time is different from one observer to the next*. For moving observers, the time is stretched out, so that each second is longer, a phenomenon known as **time dilation**.

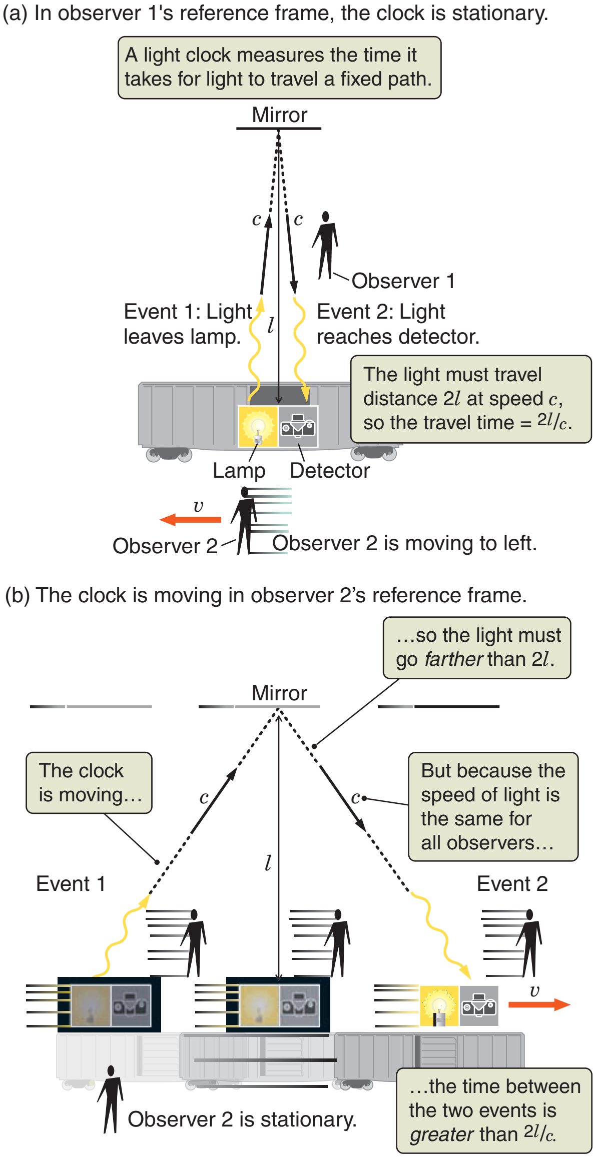


Figure 18.4 The "tick" of a light clock is different when seen in two different reference frames: (a) stationary as in the reference frame of observer 1 in the boxcar, and (b) moving as in the reference frame of observer 2 on the tracks. As Einstein's thought experiment demonstrates, if the speed of light is the same for every observer, then moving clocks *must* run slowly compared to stationary clocks.

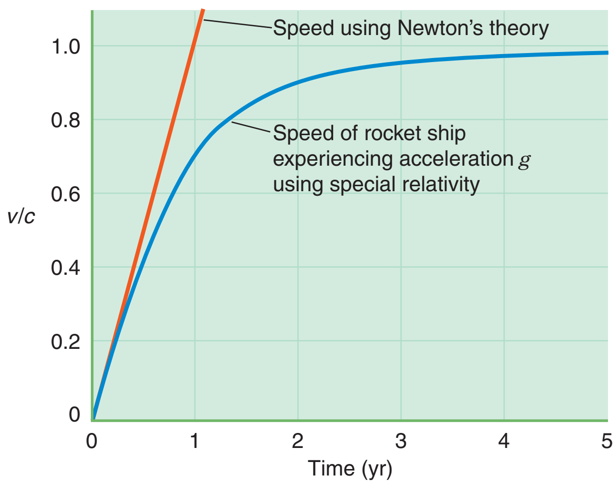


Figure 18.5 The speed of a rocket ship experiencing an acceleration equal to Earth's acceleration of gravity (g) increases with time. The red curve shows the speed using Newtonian theory. The blue curve shows the effects of special relativity, where the rocket ship approaches the speed of light but never gets there. The blue line correctly accounts for relativistic effects; the red line does not.

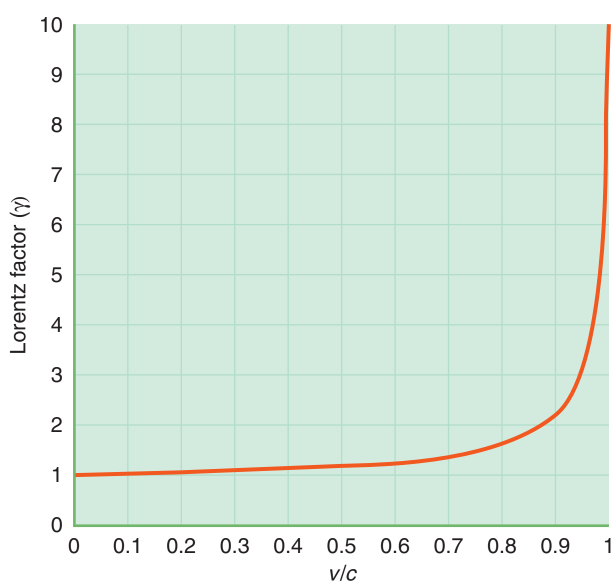


Figure 18.6 In this graph, the Lorentz factor γ is plotted against v/c. This factor doesn't become significant until velocities are about 50 percent the speed of light.

The Newtonian view of the world describes a three-dimensional space through which time marches steadily onward. But Einstein discovered that time flows differently for different observers. He reshaped the three dimensions of space and the one dimension of time into a four-dimensional combination called **spacetime**. Events occur at specific locations within this four-dimensional spacetime, but how much of a spacetime distance is measured in space and how much is measured in time depends on the observer's reference frame.

Einstein did not "disprove" Newtonian physics. At speeds much less than the speed of light, Einstein's equations become identical to the equations of Newtonian physics, so that Newtonian physics is contained within special relativity. Only when objects approach the speed of light do our observations begin to depart measurably from the predictions of Newtonian physics. These departures are called **relativistic** effects. In our everyday lives, we never encounter relativistic effects because we never travel at speeds that approach the speed of light. Even the fastest object ever made by humans, the *Helios II* spacecraft, traveled at only about $0.00023c$.

Einstein's ideas remained controversial well into the 20th century, and his 1921 Nobel Prize in Physics was awarded for his work on the photoelectric effect (see Chapter 5) and not for his work on relativity because not all physicists were yet convinced that his theory was correct. But as one experiment after another confirmed the strange and counterintuitive predictions of relativity, scientists came to accept its validity.

The Implications of Relativity

Today, special relativity shapes our thinking about the motions of both the tiniest subatomic particles and the most distant galaxies. In this subsection, we discuss only a few of the essential insights that come from Einstein's work.

Mass and Energy What we think of as "mass" and what we think of as "energy" are actually closely related. The energy of an object depends on its speed: the faster it moves, the more energy it has. But Einstein's famous equation, $E = mc^2$, says that even a stationary object has an intrinsic "rest" energy that equals the mass (m) of the object multiplied by the speed of light (c) squared. The speed of light is a very large number, so a small mass has a very large rest energy. A single tablespoon of water has a rest energy equal to the energy released in the explosion of more than 300,000 tons of TNT. We used this relationship between mass and energy in Chapter 14 when discussing the nuclear fusion that makes stars shine.

In Chapter 3, we connected the *mass* of an object to its inertia—its resistance to changes in motion. At relativistic speeds, it becomes clear that adding to the energy of motion of an object increases its inertia. For example, a proton in a high-energy particle accelerator may travel so close to the speed of light that its total energy is 1,000 times greater than its rest energy. Such an energetic proton is harder to "push around" (in other words, it has more inertia) than a proton at rest. It also more strongly attracts other masses through gravity.

The Ultimate Speed Limit We already discussed Einstein's insight that if it were possible to travel at the speed of light, then in that reference frame light would cease to be a traveling wave, and the laws of electromagnetism wouldn't be valid. You can also think about this limit in terms of the equivalence of mass and energy just discussed. As the speed of an object gets closer and closer to the speed of light, the energy of that object, and therefore its mass, become greater and greater, so it becomes increasingly resistant to further changes in its motion. Only a photon or other massless particle can travel at the speed of light. We rely on the fact that light travels at a constant speed in a vacuum whenever we use the travel time of light to describe astronomical distances.

Adding energy to an object will cause its velocity to get closer and closer to the speed of light, but it will never actually reach the speed of light. It would take an *infinite* amount of energy to accelerate an object with a nonzero rest mass to the speed of light. There is not enough energy in the entire universe to accelerate even one electron to the speed of light. The electron can get arbitrarily close to that number—0.9999999999999999999999 . . . × *c* is possible (at least in principle)—but there is not enough energy available to accelerate the electron beyond that to the speed of light. **Figure 18.5** shows how a rocket ship, which experiences a constant acceleration equal to that of gravity on Earth (so that its occupants will feel "normal" gravity), moves faster and faster but never reaches the speed of light. In the Newtonian view shown by the red curve, this limit is not present, and the speed of the spaceship would continually increase at the same rate.

Time Time passes more slowly in a moving reference frame: for moving objects, the seconds are stretched out in time dilation. No inertial reference frame is special. If you compared clocks with an observer moving at nine-tenths the speed of light (0.9*c*) relative to you, you would find that the other observer's clock was running 2.29 times slower than your clock. The other observer would find instead that *your* clock was running 2.29 times slower. To you, the other observer may be moving at 0.9*c*, but to the other observer, *you* are the one who is moving. Either frame of reference is equally valid, so you would each find the other's clock to be slow compared to your own. This time dilation effect increases with speed, and this symmetry holds as long as neither frame accelerates. **Figure 18.6** and **Table 18.1** show that how much the time is stretched depends on the object's speed. The factor of 2.29 by which time is stretched in the above example is called the Lorentz factor and is usually denoted by the symbol γ.

A scientific observation demonstrates time dilation in nature. As illustrated in **Figure 18.7**, fast particles called cosmic-ray muons are produced 15 km up in Earth's atmosphere when high-energy **cosmic rays**—elementary particles moving at nearly the speed of light—strike atmospheric atoms or molecules. Muons at rest decay very rapidly into other particles. This decay happens so quickly that even if they could move at the speed of light, virtually all muons would have decayed long before traveling the 15 km to reach Earth's surface. However, time dilation slows the muons' clocks, so the particles live longer and can travel farther and reach the ground. The faster muons move, the slower their clocks run and the more of them that can reach the ground. The same general principle is observed in particle accelerators, where particles that are traveling at speeds near the speed of light live longer before decaying. **Working It Out 18.1** shows some examples of time dilation.

Length An object appears shorter in motion than it is at rest. Moving objects are compressed in the direction of their motion by a factor of $1/\gamma$, where γ is the same Lorentz factor introduced in the discussion of time dilation. This phenomenon is called length contraction. A meter stick moving at 0.9*c*

TABLE 18.1 Lorentz Factor

v/c	γ
0.10	1.005
0.20	1.02
0.30	1.05
0.40	1.09
0.50	1.15
0.60	1.25
0.70	1.40
0.80	1.67
0.90	2.29
0.95	3.20
0.99	7.09
0.995	10.01
0.999	22.37
0.9999	70.71

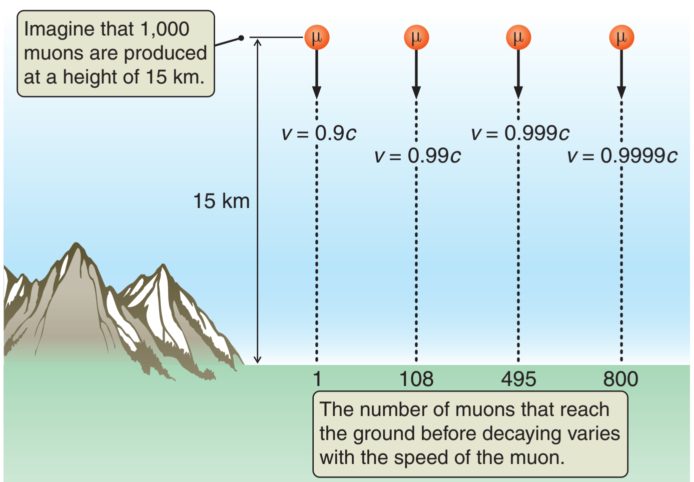


Figure 18.7 Muons created by cosmic rays high in Earth's atmosphere decay long before reaching the ground if they are not traveling at nearly the speed of light. Here, we show what happens to 1,000 muons produced at an altitude of 15 km for a variety of speeds. Faster muons have slower clocks, so more of them survive long enough to reach the ground—many more than would be expected simply due to the faster speed.

18.1 Working It Out Time Dilation

Physicist Hendrik Lorentz (1853–1928) derived the equation for how much time is dilated and how much space is contracted when something is traveling at velocities near the speed of light. This *Lorentz factor* (abbreviated γ) is given by

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Figure 18.6 shows the Lorentz factor plotted against velocity, and Table 18.1 gives the calculated value of γ for different values of velocities v/c. You can see that for something moving at half the speed of light, the Lorentz factor is 1.15. But for something moving at 90 percent of c, the factor is 2.29, and it goes up quickly from there, becoming arbitrarily large as the velocity approaches but never quite reaches the speed of light.

Let's look at a few examples. The first one is a thought experiment. Suppose you take a trip to a star 20 light-years away in order to study the "super-Earth" planets in orbit around it. You travel at $0.99c$, while your twin stays behind on Earth. Our common Newtonian experience on Earth would tell you how long the trip takes using the familiar equation that time equals distance divided by speed:

$$\text{Time passed} = \frac{20\,\text{light-years}}{0.99c} = 20.2\,\text{years}$$

The round trip would take 40.4 years. When you return, you will find that 40.4 years have passed on Earth.

To you on the spaceship moving at $0.99c$, special relativity says time will pass slower, by the Lorentz factor of 7.09. Your trip to the star would take $1/7 \times 20.2$ years $= 2.8$ years and another 2.8 years for the return trip. So as you traveled to this star and back, 5.6 years would have passed for you in the spaceship, but *40.4* years would have passed on Earth. Your twin on Earth would be almost 35 years older than you are!

For a more practical experiment that can be performed on Earth, we've noted that time dilation can be seen with subatomic particles. Suppose one type of particle, called a pion, can "live" for 20 nanoseconds (ns) before it decays into different particles. If pions are produced in a particle accelerator at a speed of $0.999c$, how long will physicists see the pions before they decay? Here $v = 0.999c$, so from Table 18.1, $\gamma = 22.37$. In the reference frame of the moving pions, they are still living for 20 ns: they are like the twin that traveled into space at high speed. But in the reference frame of the physicists, special relativity predicts that the particle will last longer, 20 ns $\times$ 22.37, or 447 ns. Indeed, physicists observe that pions moving at nearly the speed of light "live" longer and travel farther before decay (like the muons in Figure 18.7).

The same factor γ applies for mass and length. If physicists measured a ruler moving at $0.999c$, the length of the ruler would be 22.37 times shorter compared to its length when at rest. In the earlier example, the pions traveling at $0.999c$ will behave as if their mass were 22.37 times larger: if the high-speed pions collided with other particles, the energy of the collision would be as if they had the higher mass. Particles at high speeds have provided the experimental evidence for special relativity.

appears to be 0.44 meters long. This also explains our muon experiment from the perspective of the muons themselves. In the reference frame of the fast-moving muon produced at a height of 15 km, Earth's atmosphere is moving fast and appears to be much shorter than 15 km; indeed, it is so compressed from the muon's perspective that the muon may be able to reach the ground before decaying. This length contraction effect also increases with speed.

Twin Paradox One interesting consequence of relativity is the *twin paradox*. Suppose you take a trip into space and leave your identical twin back on Earth. You accelerate to nearly the speed of light as you leave Earth. After you arrive at your destination, you return to Earth, again traveling at a speed close to c. To your twin on Earth, you were the one in a moving reference frame, so your twin measures your time as running much slower (recall Working It Out 18.1), and you should return younger than your twin. However, from your perspective, the spaceship didn't move. Instead, Earth receded from you, stopped, and returned. Your twin is the one who moved at just under the speed of light, and your twin's time ran slowly compared to yours, so your twin has aged much less. Both you and your twin cannot be correct, and that is the nature of the paradox.

To resolve the paradox, you must realize that you experienced acceleration during your trip, while your twin did not. Accelerated motion is not uniform motion.

As a result, you changed reference frames during your trip. You changed reference frames when you left Earth, changed again when you stopped at your destination, changed a third time when you left your destination to return home, and changed reference frames one final time when you arrived back at Earth. Your twin, however, remained in Earth's reference frame. Upon your return, you would find that in fact more time has passed for your twin on Earth than for you, and your twin has aged more than you have.

Space Travel What is the actual reality of human travel in space? Some American astronauts have been to the Moon and back, and a number of robotic spacecraft have been sent to explore objects throughout the Solar System. (Only one robotic spacecraft has actually left the Solar System.) But it would be very difficult for humans to visit and explore other planetary systems in the Milky Way Galaxy because of the constraints of energy and the ultimate speed limit of light. With current technology, engineers can construct rockets able to travel at speeds of up to 20,000 m/s. At such a speed, a one-way trip to Earth's nearest neighbor star, Proxima Centauri, at a distance of 4.2 light-years, would take well over 50,000 years. Travel to more distant stars would take even longer.

In principle, travel just under the speed limit c is possible, so one could take advantage of relativistic time dilation to make such adventures well within the lifetime of a space traveler. In the example in Working It Out 18.1, an astronaut experiences a round-trip travel time to planets 20 light-years away in just 5.8 years at a speed of $0.99c$. Or an astronaut could travel to the center of the Milky Way Galaxy and back in just 2 years by traveling at $0.9999999992c$.

Although theoretically possible, travel at these speeds in practice would require an impractical amount of energy. If M is the mass of the astronauts, rocket ship, and fuel, it would take γMc^2 of energy just to accelerate the rocket up to such a high speed, or $10Mc^2$ in the first example and $25{,}000Mc^2$ in the second example to the center of the Milky Way. For this second example, the energy to accelerate just the astronaut (not even including the spaceship) to such energies is more than that contained in 10 billion nuclear weapons. So, while not theoretically impossible, visits to other stars in our galaxy will not take place anytime soon.

CHECK YOUR UNDERSTANDING 18.2

Suppose that your friend flies past you in a spaceship, and both of you measure the time it takes the spaceship to pass your location. Which of the following is true? (a) The time you measure is longer than the time your friend measures. (b) The time you measure is shorter than the time your friend measures. (c) You both measure the same amount of time.

18.3 Gravity Is a Distortion of Spacetime

Our exploration of special relativity began with the observation that the speed of light is always the same regardless of the motion of an observer or the motion of the source of the light. We have seen that three-dimensional space and time are actually just the result of a particular, limited perspective on a four-dimensional spacetime that is different for each observer. This four-dimensional spacetime is itself warped and distorted by the masses it contains. As we discuss the properties of black holes—indeed, of all massive objects in the universe—the concepts

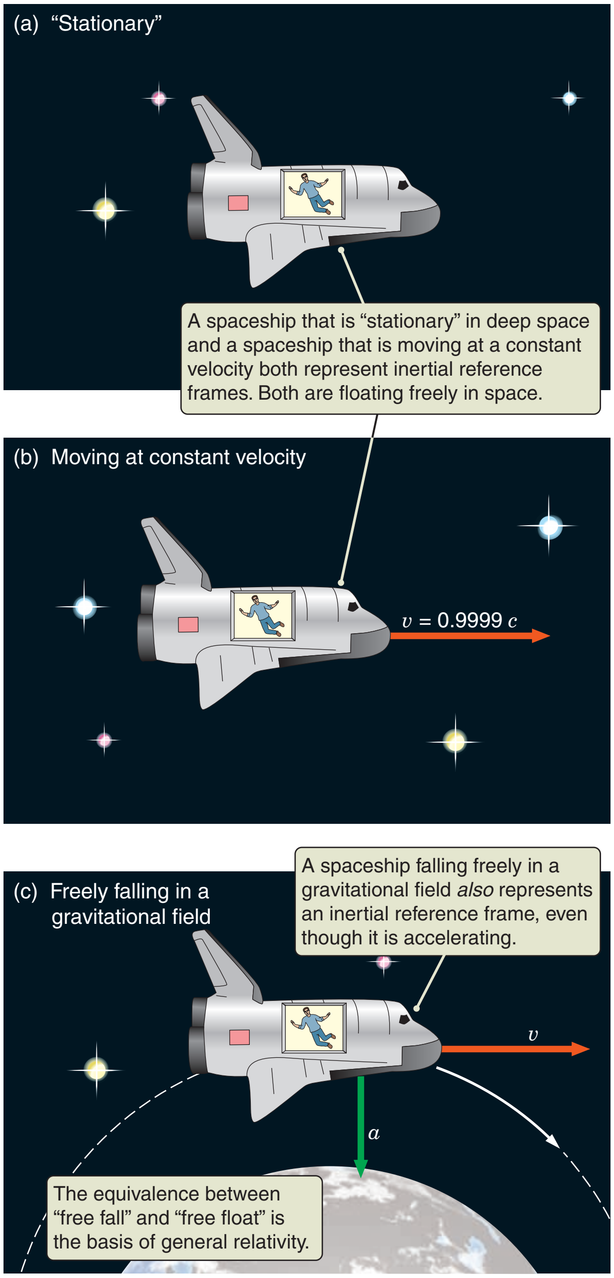


Figure 18.8 Special relativity says that there is no difference between (a) a reference frame that is floating stationary in space and (b) one that is moving through the galaxy at constant velocity. General relativity adds that there is no difference between these inertial reference frames and (c) an inertial reference frame that is falling freely in a gravitational field. Free fall is the same as free float, as far as the laws of physics are concerned.

of space and time will diverge even further from the absolutes of Newtonian physics. In this section, we will explore the general theory of relativity, which describes how mass affects space and time.

The Equivalence Principle

We have already discussed one fundamental connection between gravity and spacetime in Chapter 4 and showed that the *inertial mass* of an object—the mass appearing in Newton's equation $F = ma$—is *exactly* the same as the object's gravitational mass. In addition, any two objects at the same location and moving with the same velocity will follow the same path through spacetime, regardless of their masses. The astronaut in an orbiting spaceship falls around Earth, moving in lockstep with the spaceship itself. A feather dropped by an *Apollo* astronaut standing on the Moon falls toward the surface of the Moon at exactly the same rate as a dropped hammer does. Rather than thinking of gravity as a force that acts on objects, it is more accurate to think of gravity as a consequence of the warping of spacetime in the presence of a mass. *Gravitation is the result of the shape of spacetime that objects move through.* This is one of the key insights of the **general theory of relativity**, Einstein's theory of gravity.

Special relativity tells us that any inertial reference frame is as good as any other. No experiment can distinguish between sitting in an enclosed spaceship floating stationary in deep space, as seen in **Figure 18.8a**, and sitting in an enclosed spaceship traveling at 0.9999 times the speed of light (Figure 18.8b). These two situations feel the same because neither observer feels an acceleration. Each is an equally valid inertial reference frame. As long as nothing accelerates either spaceship, neither observer can distinguish between these two spacecraft.

But what if there is an acceleration? Recall that an acceleration can change either the speed or the direction of an object. Consider an astronaut inside a spaceship orbiting Earth, as shown in Figure 18.8c. This spaceship is accelerating, as the direction of its velocity is constantly changing as it orbits Earth. The astronaut is also accelerating. Because he feels weightless, the astronaut has no way to tell the difference between being inside the spaceship as it falls around Earth and being inside a spaceship floating through interstellar space. Even though its velocity is constantly changing as it falls, the inside of a spaceship orbiting Earth is an inertial frame of reference just as an object drifting along a straight line through interstellar space is an inertial frame of reference.

The idea that a freely falling reference frame is equivalent to a freely floating reference frame is called the **equivalence principle**. If you close your eyes and jump off a diving board, for the brief time that you are falling freely through Earth's gravitational field, the sensation you feel is exactly the same as the sensation you would feel floating in interstellar space.

The natural path that an object will follow through spacetime in the absence of other forces is called the object's **geodesic**. In the absence of a gravitational field, the geodesic of an object is a straight line, in accordance with Newton's first law: an object will move at a constant speed in a constant direction unless acted on by a net external force. However, the shape of spacetime becomes distorted in the presence of mass, so an object's geodesic becomes curved.

Figure 18.9a shows two examples of inertial frames: an astronaut in a spaceship coasting through space, and a person in a box falling toward the ground

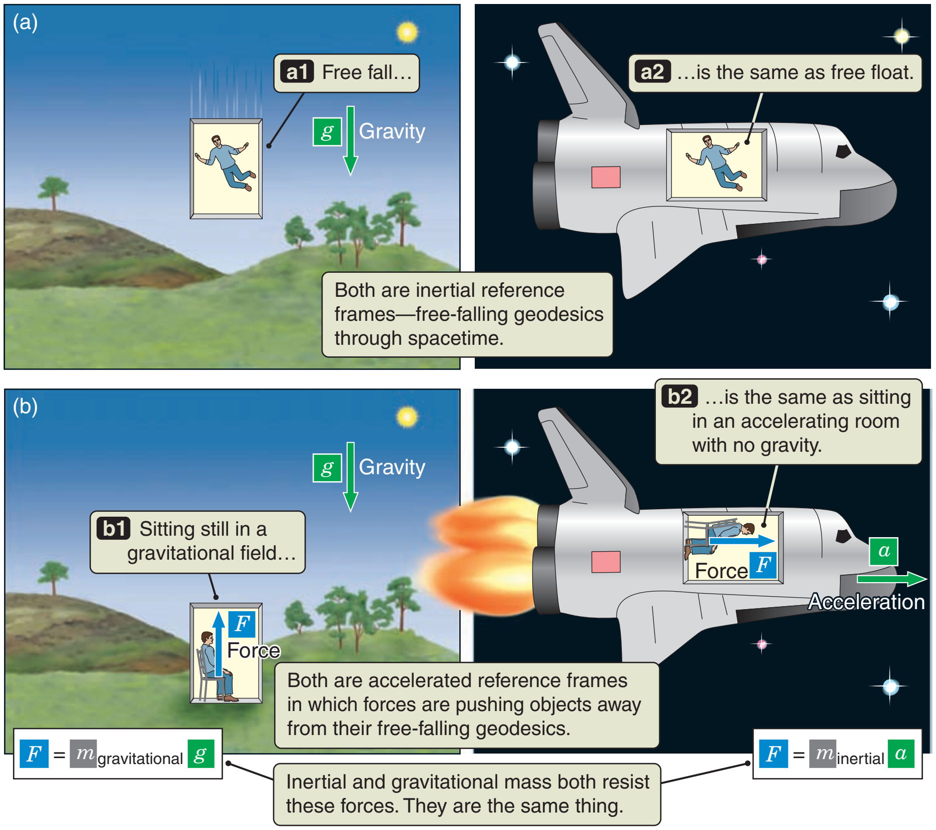


Figure 18.9 According to the equivalence principle, (a) an object falling freely in a gravitational field is in an inertial reference frame, and (b) an object at rest in a gravitational field is in an accelerated reference frame. According to the equivalence principle, sitting in a spaceship accelerating at 9.8 m/s^2 feels the same as sitting still on Earth.

with an acceleration g. Both of these people are following their geodesics, so these two inertial reference frames are equivalent, and neither observer can distinguish between them. The equivalence principle also applies in cases of accelerations that result in a change in speed. Imagine an observer sitting in a closed box on the surface of Earth, as in Figure 18.9b. The floor of the box pushes on him to keep him from following his geodesic, and he feels that force. Now imagine the box is inside a spaceship that is accelerating through deep space at a rate of 9.8 meters per second per second (m/s^2) in the direction of the arrow shown in Figure 18.9b. The floor of the box pushes on the observer to overcome his inertia and cause him to accelerate at 9.8 m/s^2, so he feels as though he is being pushed into the floor of the box. In both of these cases, a force acts on the observer so that he does not follow the same free-falling geodesic. The observer feels as though he is being pushed into the floor of the box, so he feels the acceleration, and his frame of reference is not inertial. According to the equivalence principle, sitting in an armchair in a spaceship traveling with an acceleration of 9.8 m/s^2 is equivalent to sitting in an armchair on the surface of Earth reading this book. In each case, it is the same mass—the mass that gives an object inertia—that resists the change. Gravitational mass and inertial mass are the same thing.

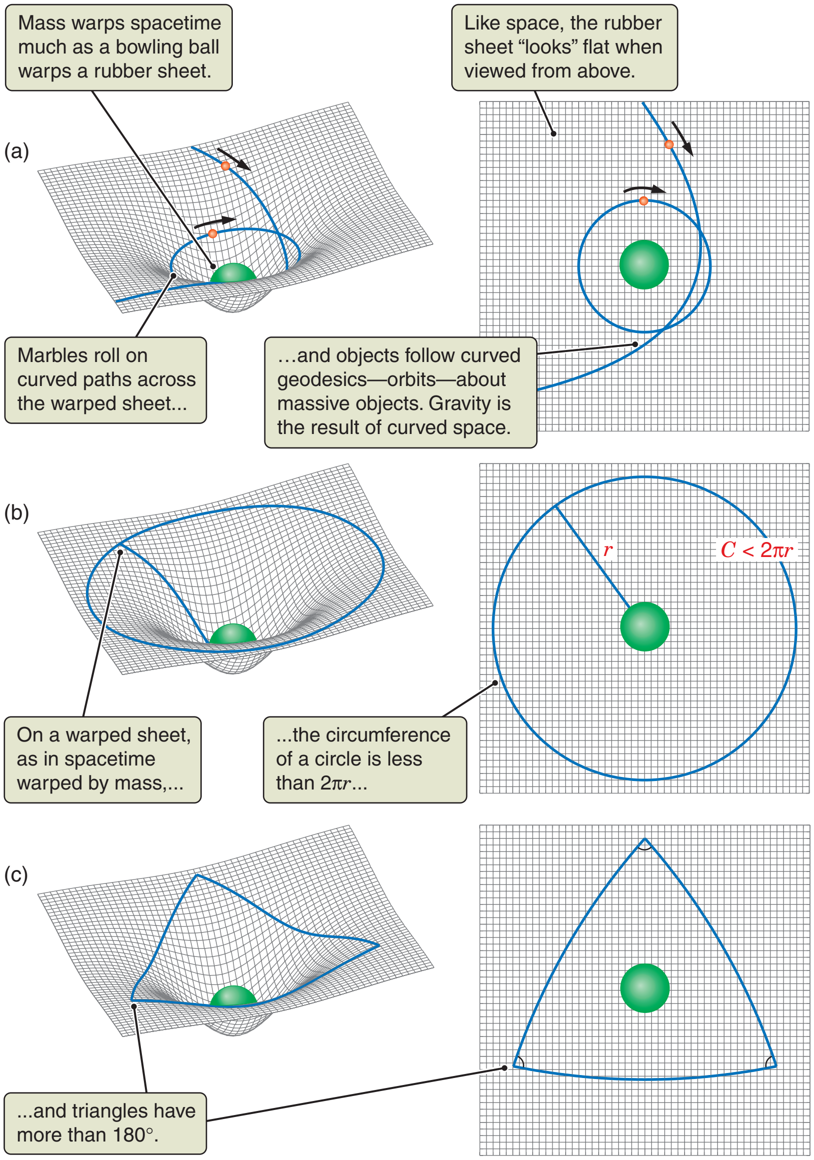


Figure 18.10 Mass warps the geometry of spacetime in much the same way that a bowling ball warps the surface of a stretched rubber sheet. This distortion of spacetime has many consequences; for example, (a) objects follow curved paths or geodesics through curved spacetime, (b) the circumference of a circle around a massive object is less than 2π times the radius of the circle, and (c) angles in triangles need not add to exactly 180 degrees.

There is an important caveat to the equivalence principle. In an accelerated reference frame such as an accelerating spaceship, the same acceleration (both magnitude and direction) is experienced everywhere. In contrast, the curvature of space by a massive object is weaker farther from the object. The effects of gravity and acceleration are equivalent only locally; that is, the equivalence principle is valid as long as attention is restricted to small-enough volumes of space so that changes in gravity can be ignored.

Mass Distorts Spacetime

The general theory of relativity describes how mass distorts the geometry of spacetime. Imagine the surface of a tightly stretched, flat rubber sheet. A marble will roll in a straight line across the sheet. Euclidean geometry, the geometry of everyday life, applies on the surface of the sheet: if you draw a circle, its circumference is equal to 2π times its radius, r; if you draw a triangle, the angles add up to 180 degrees; lines that are parallel anywhere are parallel everywhere.

Now place a bowling ball in the middle of the rubber sheet, creating a deep depression, or "well," as in **Figure 18.10**. The surface of the sheet is no longer flat, and Euclidian geometry no longer applies. If you roll a marble across the sheet, its path dips and curves, as shown in Figure 18.10a. You can roll the marble so that it moves around and around the bowling ball, like a planet orbiting about the Sun. Figure 18.10b shows that if you draw a circle around the bowling ball, the circumference of the circle is less than $2\pi r$. If you draw a triangle, the angles add up to more than 180 degrees (Figure 18.10c).

Mass has an effect on the fabric of spacetime that is like the effect of the bowling ball on the fabric of the rubber sheet. The bowling ball stretches the sheet, changing the distances between any two points on the surface of the sheet. Similarly, mass distorts spacetime, changing the distance between any two locations or events. Larger masses produce larger spacetime distortions. We can visualize how a rubber sheet with a bowling ball on it is stretched through a third spatial dimension, but it is impossible for most people to visualize what a curved four-dimensional spacetime would "look like." Yet experiments verify that the geometry of four-dimensional spacetime is distorted much like the rubber sheet, whether or not it can be easily pictured.

When One Physical Law Supplants Another

Earlier in this book, we described gravity as a force that obeys Newton's universal law of gravitation: $F = Gm_1m_2/r^2$. In this chapter, we have introduced the ideas of general relativity and asked you to change the way you think about gravity. If general relativity is correct, does that mean Newton's formulation of gravity is wrong? If so, then why does Newton's law continue to be used?

These questions go to the heart of how science progresses. As long as a gravitational field is not too strong, Newton's law of gravitation is a very close *approximation* to the results of a calculation using general relativity. In this context, even the gravitational field near the core of a massive main-sequence star would be considered "weak." An astronomer will obtain the same results if she uses a general relativistic formulation of gravity rather than Newton's laws to calculate the structure of a main-sequence star. Similarly, even though spacetime is curved by the presence of mass, this curvature near Earth is so slight that over small regions it can be ignored entirely, and flat Euclidean geometry can be used. This is exactly the kind of approximation people use when they navigate with a flat road map, despite the curvature of Earth.

Similarly, Newton's laws of motion are *approximations* of special relativity and quantum mechanics. In fact, Newton's laws can be mathematically derived from special relativity and quantum mechanics using the assumptions that speeds of objects are much less than the speed of light and that objects are much larger than the particles from which atoms are made. Newton's laws of motion and gravitation are used most of the time because they are far easier to apply than general relativity and because any inaccuracies introduced by using Newtonian approximations are usually far too tiny to measure. This was illustrated in Figure 18.5, where the behavior of an accelerated object moving at a velocity much less than *c* is the same whether or not relativity is used. Calculations based on general relativity are required only when conditions are very different from those of everyday life (for example, the behavior of an electron in an atom or the behavior of the gravitational field of a black hole) or in special cases when very high accuracy is needed (for example, the precise timing used by the GPS satellite network).

If a new theory is to replace an earlier, highly successful scientific theory, the new theory must hold the old theory within it—it must be able to reproduce the successes of the earlier theory. Special relativity contains Newton's laws of motion, and general relativity holds within it the successful Newtonian description of gravity that we have relied on throughout this book (**Process of Science Figure**).

The Observable Consequences of General Relativity

The curved spacetime of general relativity does have observable consequences. Indeed, there are observations that distinguish general relativity from Newtonian physics in our own Solar System. In Newton's theory, orbits are elliptical and fixed in space. In contrast, general relativity predicts that the long axis of an elliptical orbit slowly rotates, or precesses. **Figure 18.11** illustrates the difference between an orbit in Newton's theory, which remains stationary (left panel), versus one in general relativity, which precesses (right panel). In our Solar System, even after accounting for the effects of other planets on Mercury, there remains a very small shift in its axis equal to 43 arcseconds (arcsec) per century, which cannot be explained by Newton's laws alone. General relativity predicts exactly that precession for Mercury.

There are other unique implications of general relativity. A beam of light moving through empty space travels in a straight line, but a beam of light moving through the distorted spacetime around a massive object is bent by gravity, just as the lines in Figure 18.10 are bent by the curvature of the sheet. This bending of the light path by curved spacetime is called **gravitational lensing** because optical lenses also bend light paths. The first measurement of gravitational lensing came during the total solar eclipse of 1919. Several months before the eclipse, astrophysicist Sir Arthur Stanley Eddington (1882–1944) measured the positions of a number of stars in the direction of the sky where the eclipse would occur. He then

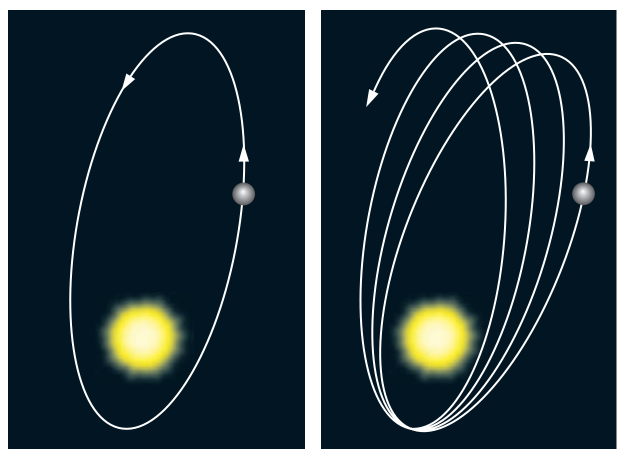
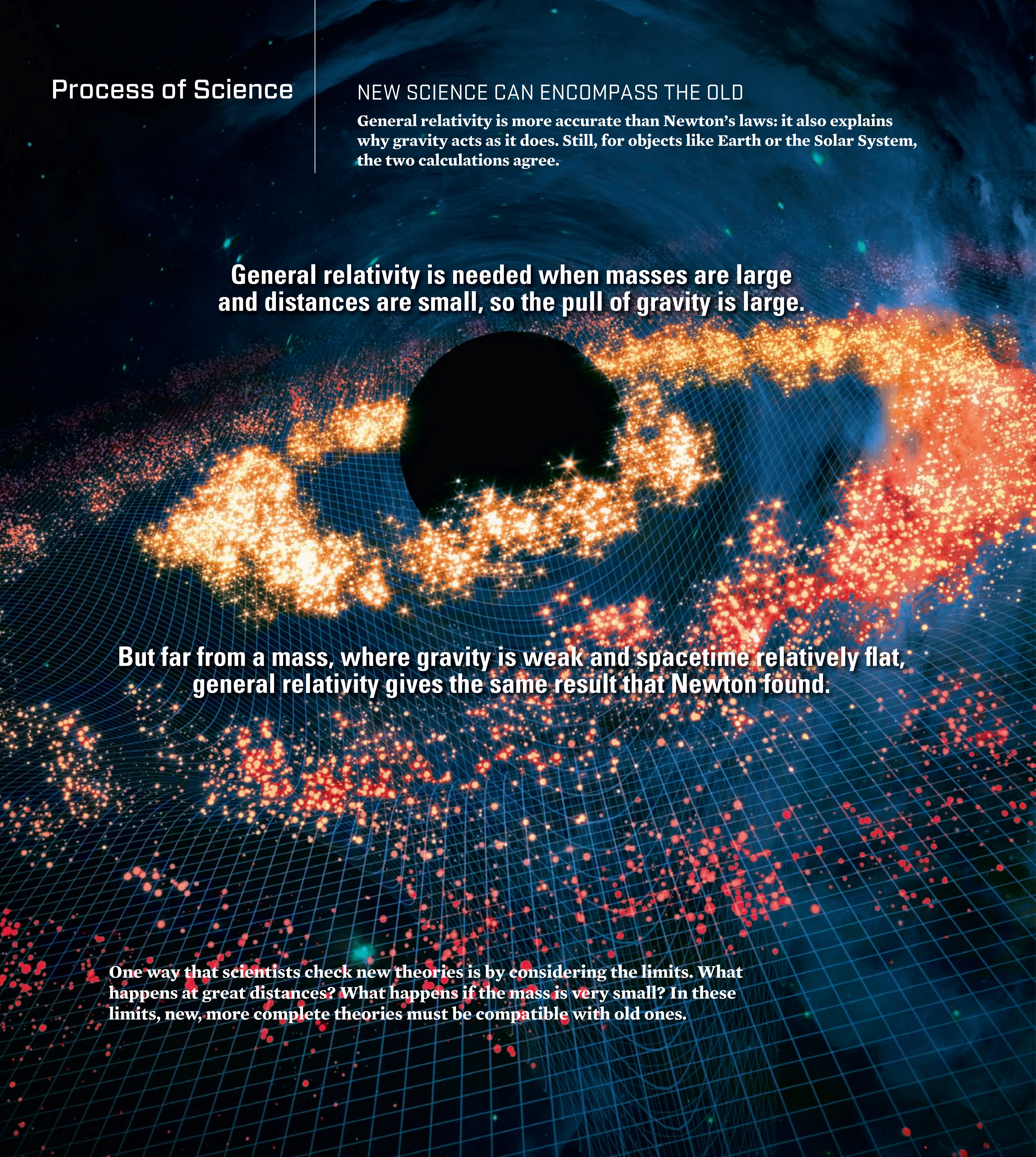


Figure 18.11 The left panel denotes an elliptical orbit about the Sun. In the Newtonian view, this elliptical orbit remains stationary. In the right panel, Mercury's orbit precesses due to the warped spacetime near the Sun.

Process of Science

NEW SCIENCE CAN ENCOMPASS THE OLD

General relativity is more accurate than Newton's laws: it also explains why gravity acts as it does. Still, for objects like Earth or the Solar System, the two calculations agree.



General relativity is needed when masses are large and distances are small, so the pull of gravity is large.

But far from a mass, where gravity is weak and spacetime relatively flat, general relativity gives the same result that Newton found.

One way that scientists check new theories is by considering the limits. What happens at great distances? What happens if the mass is very small? In these limits, new, more complete theories must be compatible with old ones.

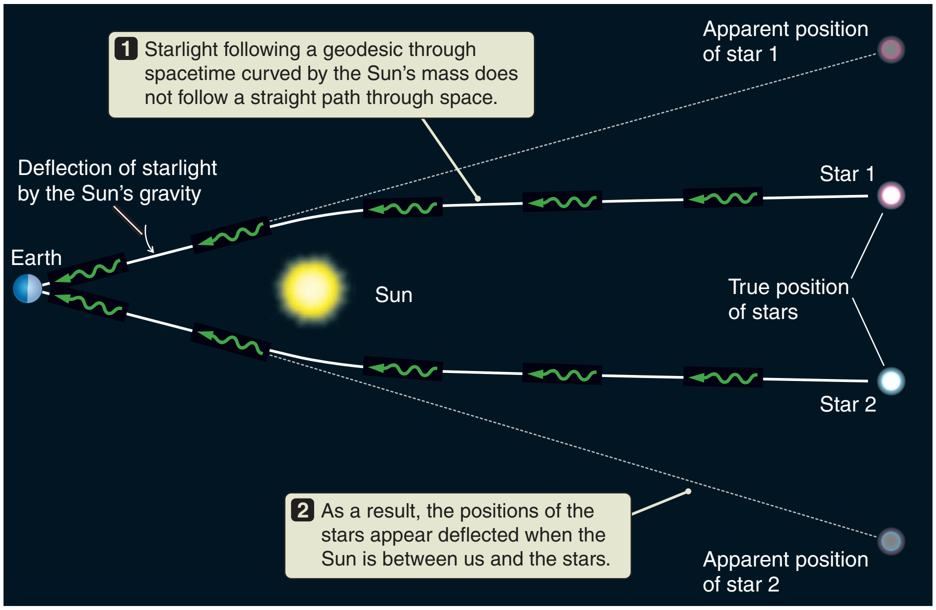


Figure 18.12 Measurements during the total solar eclipse of 1919 found that the gravity of the Sun bends the light from distant stars by the amount predicted by Einstein's general theory of relativity. This is an example of gravitational lensing. Note that the "triangle" formed by Earth and the two stars contains more than 180 degrees, just like the triangle in Figure 18.10c.

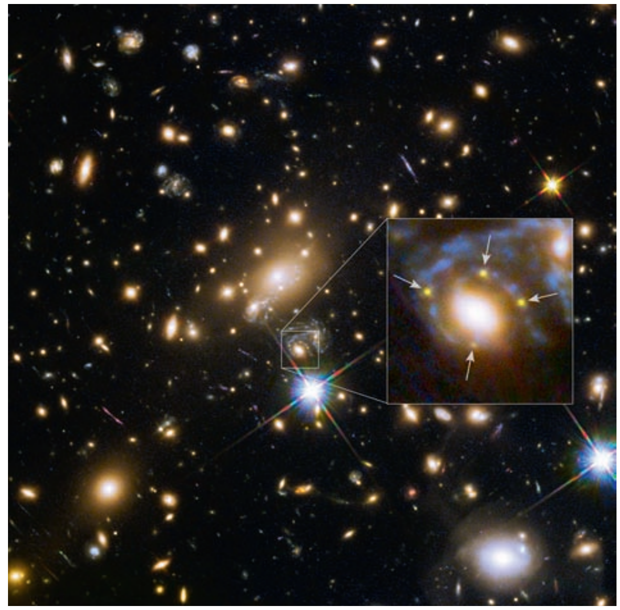


Figure 18.13 This Hubble Space Telescope image shows four images of the same distant supernova (arrows), where the images arise from the gravitational lensing due to a massive foreground galaxy.

repeated the measurements during the eclipse. **Figure 18.12** shows how the light from distant stars curved as it passed the Sun, causing the measured positions of the stars to shift outward. The stars appeared farther apart in Eddington's second measurement than in his first by the amount predicted by general relativity. This is another example of the effects of general relativity—in curved space, parallel lines actually can intersect.

Images of galaxies are sometimes distorted by gravitational lensing by other galaxies or clusters of galaxies. In the Hubble Space Telescope image shown in **Figure 18.13**, a supernova in a distant galaxy is lensed into four images by a nearby massive galaxy. What makes this case particularly interesting is that these four images appeared at different times because the light for each image took a different path, and some were longer than others. In an extreme case, a lensed galaxy image can be distorted into a ring called an **Einstein ring**, as illustrated in **Figure 18.14**.

Mass distorts not only the geometry of space but also the geometry of time. The deeper one descends into the gravitational field of a massive object, the more slowly clocks appear to run from the perspective of a distant observer, an effect called **general relativistic time dilation**. Suppose a light is attached to a clock sitting on the surface of a neutron star. The light is timed so that it flashes once a second. Because time near the surface of the star is dilated due to high gravity, an observer far from the neutron star perceives the light to be pulsing with a lower frequency—less than once a second. Now suppose there is an emission line source on the surface of the neutron star. Because time is running slowly on the surface of the neutron star, the light that reaches the distant observer will have a lower frequency than that when it was emitted. Recall that a lower frequency means a longer wavelength, so the light from the source will be seen at a longer, redder wavelength than the wavelength at which it was emitted. This shift in the wavelengths of light from objects deep within a gravitational well is called the

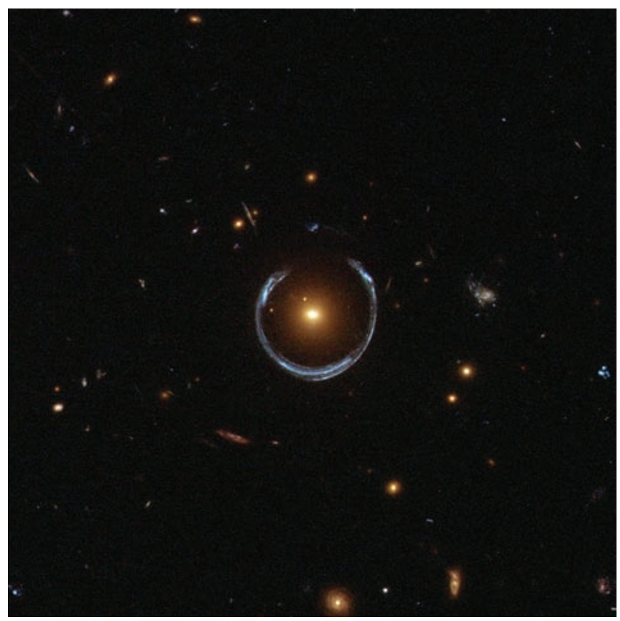


Figure 18.14 This photograph shows an Einstein ring created by the gravity of a luminous red galaxy gravitationally lensing the light from a much more distant blue galaxy.

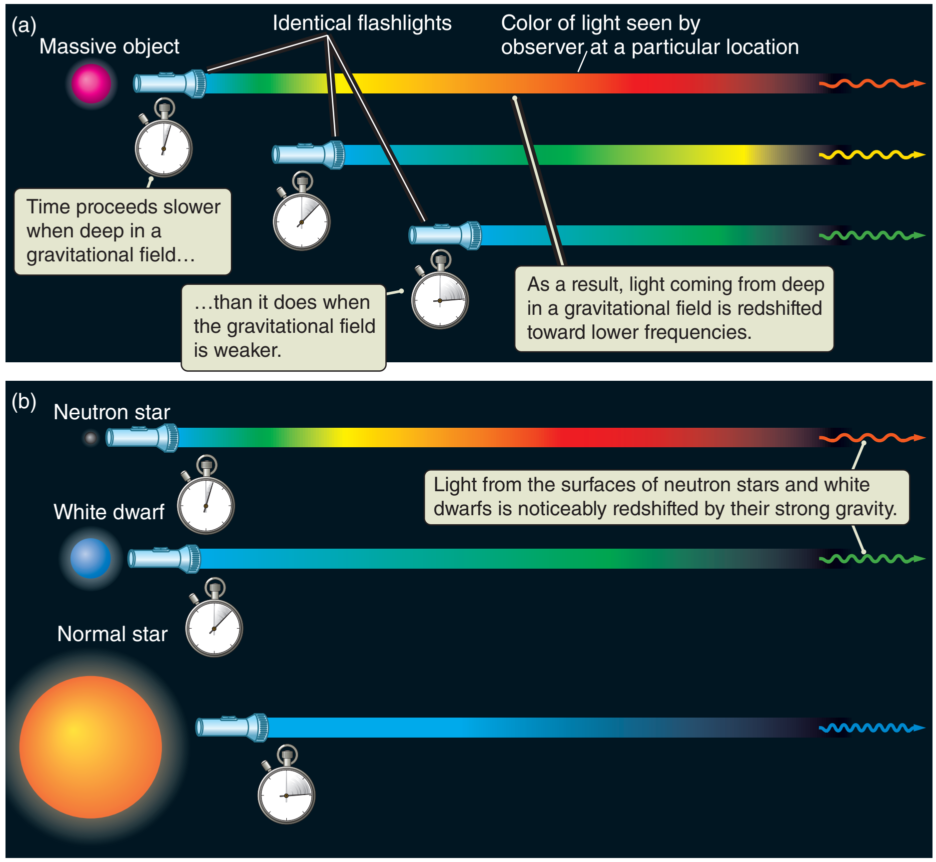


Figure 18.15 Time passes more slowly near massive objects because of the curvature of spacetime. As a result, to a distant observer light from near a massive object will have a lower frequency and longer wavelength. (a) The closer the source of radiation is to the object or (b) the more massive and compact the object is, the greater the gravitational redshift will be.

gravitational redshift, as shown in **Figure 18.15**. The effect of gravitational redshift is similar to the Doppler redshift discussed in Chapter 5. In fact, there is no way to tell the difference between light that has been redshifted by gravity and light from an object moving away from you that has been Doppler shifted.

Bringing the phenomenon of general relativistic time dilation a bit closer to home, a clock on the top of Mount Everest runs faster, gaining about 80 ns a day compared with a clock at sea level. The difference between an object on the surface of Earth and an object in orbit is even greater. A GPS receiver uses the results of sophisticated calculations of the effects of general relativistic gravitational redshift to help you accurately find your position on the surface of Earth. Satellites in orbit travel quickly enough that special relativistic effects are measurable. Even after allowing for slowing due to special relativity, the clocks on the satellites that make up the GPS run faster than clocks on the surface of Earth. If the satellite clocks *and* your GPS receiver did not correct for this and other effects of general relativity, then the position your GPS receiver reported would be in error by up to half a kilometer. The fact that the GPS can be accurate to a few meters provides strong experimental confirmation of two predictions of general relativity—gravitational redshift and general relativistic time dilation.

Gravitational Waves

General relativity also makes predictions that have not yet been confirmed. If you exert a force on the surface of a rubber sheet, accelerating it downward, waves will move away from where you struck it, like ripples spreading out over the surface of a pond. Similarly, the equations of general relativity predict that if you accelerate the fabric of spacetime (for example, with the catastrophic asymmetrical collapse of a high-mass star), then ripples in spacetime, or **gravitational waves**, will move outward at the speed of light. These gravitational waves are like electromagnetic waves in some respects. Accelerating an electrically charged particle gives rise to an electromagnetic wave. Deforming a massive object gives rise to gravitational waves.

Gravitational waves have not yet been directly observed, but there is strong circumstantial evidence for their existence. General relativity predicts that the orbit of binary neutron stars should lose energy, which will be carried away as gravitational waves. In 1974, astronomers discovered a binary system of two neutron stars, one of which is an observable pulsar. Using the pulsar as a precise clock, astronomers accurately measured the orbits of both stars. The orbits are gradually losing energy at the rate predicted by general relativity. Other similar binary pairs have been found with an orbital energy loss consistent with the radiation of gravitational waves. These systems provide indirect evidence for the existence of gravitational waves. As we discussed in Chapter 6, astrophysicists have a new kind of "telescope," called the Laser Interferometer Gravitational-Wave Observatory (LIGO), that might be able to detect the predicted gravitational waves emanating from such events.

CHECK YOUR UNDERSTANDING 18.3

What does gravity mean in relativity? (a) It is a result of mass and energy being the same thing. (b) It is a consequence of length contraction. (c) It is the result of masses acting larger when they move at high speeds. (d) It is the result of the distortion in spacetime around a massive object.

18.4 Black Holes

General relativity also predicts the existence of black holes. There is no well-formulated theory of black holes in the Newtonian world. When a mass is placed on the surface of a rubber sheet, it causes a funnel-shaped distortion that is analogous to the distortion of spacetime by a mass. Now imagine the limit in which the funnel is *infinitely* deep—it gets narrower as it goes deeper but has no bottom. This is the rubber-sheet analog to a black hole. The mathematics describing a black hole approaches infinity in the same way that the mathematical expression $1/x$ does when x approaches zero. Such a mathematical anomaly is called a **singularity**. Black holes contain singularities in spacetime, and this mathematical complication indicates that extreme conditions exist near (and inside) a black hole.

How Black Holes Form

We have seen how stellar evolution can lead to the formation of compact stellar remnants such as white dwarfs and neutron stars. However, there is an even more extreme fate that awaits some massive stars at the end of their evolution. Recall from Chapter 16 that a white dwarf can have a mass of no more than the Chandrasekhar limit, about 1.4 solar masses. If the mass of the object exceeds this limit, then gravity is able to overcome electron degeneracy pressure, and the white dwarf will begin collapsing again.

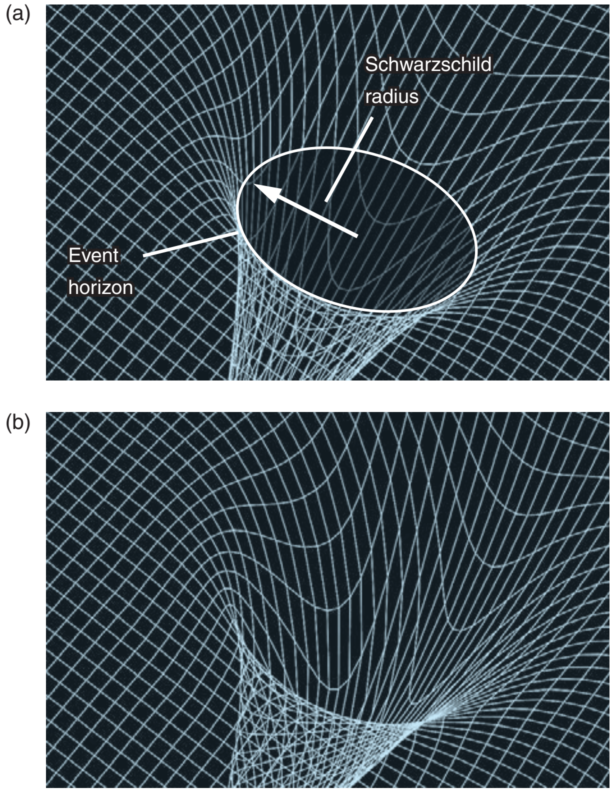


Figure 18.16 (a) A black hole's size is determined by the Schwarzschild radius and the corresponding event horizon. This image is a two-dimensional analogy for a black hole. In reality, the event horizon is a sphere. (b) If the object that formed the black hole was spinning, its angular momentum is conserved, and the black hole twists the spacetime around it.

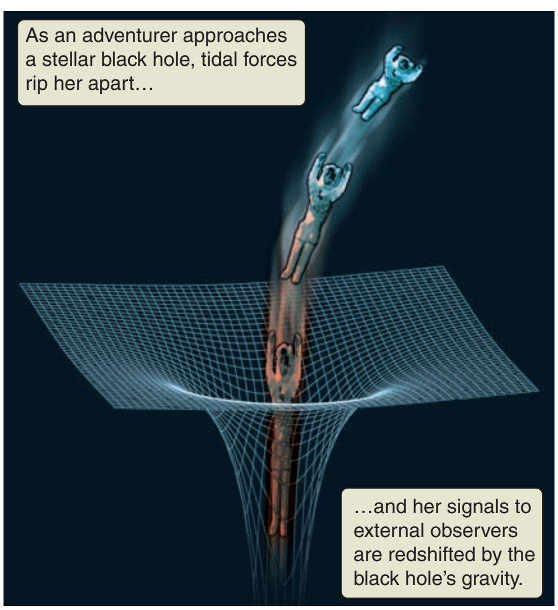


Figure 18.17 An adventurer falling into a black hole would be torn apart by the extreme tidal forces.

The physics of a neutron star is much like the physics of a white dwarf, except that neutrons rather than electrons are what cause a neutron star to be degenerate. Analogous to a white dwarf, if the mass of a neutron star exceeds about 3 M_{Sun}, gravity begins to win out over pressure once again. The neutron star grows smaller, and gravity at the star's surface becomes stronger and stronger at an ever-accelerating pace. Recall from Chapter 4 that the escape velocity from a planet or moon depends on its surface gravity. Now imagine surface gravity so strong that the escape velocity approaches the speed of light. But nothing, not even light, can travel faster than c. So when the neutron star crosses the threshold where the escape velocity from its surface exceeds the speed of light, not even light can escape its gravity. A region of space where neither matter nor radiation can escape the pull of gravity is called a **black hole**.

A neutron star in a binary system can collapse to become a black hole if it accretes enough matter from its companion to push it over the 3-M_{Sun} limit, similar to the way some Type Ia supernovae accrete matter from a companion and approach the Chandrasekhar limit of 1.4 M_{Sun}. Regardless of how it forms, any stellar remnant with a mass greater than 3 M_{Sun} must be a black hole.

Properties of Black Holes

From the outside, you can never actually "see" a black hole. The closer an object is to a black hole, the greater is its escape velocity (the speed that it would need to escape from the gravity of the black hole). There is a distance from the black hole at which the escape velocity reaches the speed of light. The radius where the escape velocity equals the speed of light is called the **Schwarzschild radius**, named for the physicist Karl Schwarzschild (1873–1916), and it is proportional to the mass of the black hole:

$$R_S = \frac{2GM_{\text{BH}}}{c^2} = 3\,\text{km} \times \frac{M_{\text{BH}}}{M_{\text{Sun}}}$$

where R_S is the Schwarzschild radius, G is the universal gravitational constant, M_{BH} is the mass of the black hole, and c is the speed of light. The sphere around the black hole at this distance is called its **event horizon**, a surface through which nothing, not even light, can escape. A black hole with a mass of 1 M_{Sun} has a Schwarzschild radius of about 3 km. A black hole with a mass of 5 M_{Sun} has a Schwarzschild radius 5 times that, or about 15 km. If Earth were squeezed into a black hole, it would have a Schwarzschild radius of only about a centimeter. All the mass of a black hole is concentrated at its very center, but this fact is unobservable from outside the black hole. **Figure 18.16a** shows a rubber-sheet analog to a black hole, with the Schwarzschild radius and the event horizon.

A black hole has only three observable properties: mass, electric charge, and angular momentum. The mass of a black hole determines the Schwarzschild radius. The electric charge of a black hole is the net electric charge of the matter that fell into it. The angular momentum of a rotating black hole twists the spacetime around it, as shown in Figure 18.16b. Apart from these three properties, all information about the material that fell into the black hole is lost. Nothing of its former composition, structure, or history survives.

Imagine that an adventurer falls into a black hole, as illustrated in **Figure 18.17**. From our perspective outside the black hole, the adventurer would appear to fall toward the event horizon. As she fell, her watch would appear to run more and more slowly, and her progress toward the event horizon would slow as well. At the event horizon, the gravitational redshift becomes infinite, and clocks appear to stop altogether. She would approach the event horizon, but from our perspective she would never quite make it. Yet the adventurer's experience would be very different. From her perspective, there would be nothing special about the event horizon. She would fall past the event horizon and continue deeper into the black hole's gravitational well. She would now have entered a region of spacetime cut off from the rest of the universe. The event horizon is like a one-way door: after the adventurer has passed through, she can never again pass back into the larger universe she once belonged to.

Actually, we have overlooked a rather important detail—the adventurer would have been torn to shreds long before she reached the black hole. Near the event horizon of a 3-M_{Sun} black hole, the difference in gravitational acceleration between the adventurer's feet and her head would be about a billion times her gravitational acceleration on the surface of Earth. In other words, her feet would be accelerating a billion times faster than her head. This is not an adventure that anyone would ever find appealing. Although scientific theories must produce testable predictions, not all individual predictions have to be tested directly.

"Seeing" Black Holes

In 1974, the physicist Stephen Hawking realized that black holes should be sources of radiation. In the ordinary vacuum of empty space, particles and their antiparticles spontaneously appear and then, within about 10^{-21} second, annihilate each other and disappear. Such particles are called **virtual particles** because they exist for only a very short time. If a pair of virtual particles comes into existence near the event horizon of a very small black hole, one of the particles might fall into the black hole while the other particle escapes, as illustrated in **Figure 18.18**. Some of the gravitational energy of the black hole will have been used up in making one of the pair of virtual particles real. Hawking showed that through this process, a black hole should emit a Planck blackbody spectrum and that the effective temperature of this spectrum would increase as the black hole became smaller through this "evaporation" process. After a very, very long time (of the order 10^{61} years for a black hole with a mass of the Sun), the black hole would become small enough that it would become unstable and explode. Although the light that emerges, called **Hawking radiation**, is of considerable interest to physicists, in a practical sense the low intensity of Hawking radiation means it is not a likely way to see a black hole.

The strongest direct evidence for black holes that result from supernovae comes from X-ray binary stars. The radio emission from Cygnus X-1 (an object originally identified in X-rays) flickers rapidly, changing in as little as 0.01 second. This means that the source of the X-rays must be smaller than the distance that light travels in 0.01 second, or 3,000 km—smaller than Earth. Cygnus X-1 was also identified both with a radio source and with an already cataloged star called HD 226868. The spectrum of HD 226868 shows that it is a normal O9.7 I supergiant star with a mass of about 19 M_{Sun}, far too cool to produce X-ray emission in the quantity observed. The wavelengths of absorption lines in the spectrum of HD 226868 are Doppler-shifted back and forth with a period of 5.6 days, indicating that HD 226868 is part of a binary system. Using the same techniques we discussed to measure the masses of binary stars in Chapter 13, astronomers

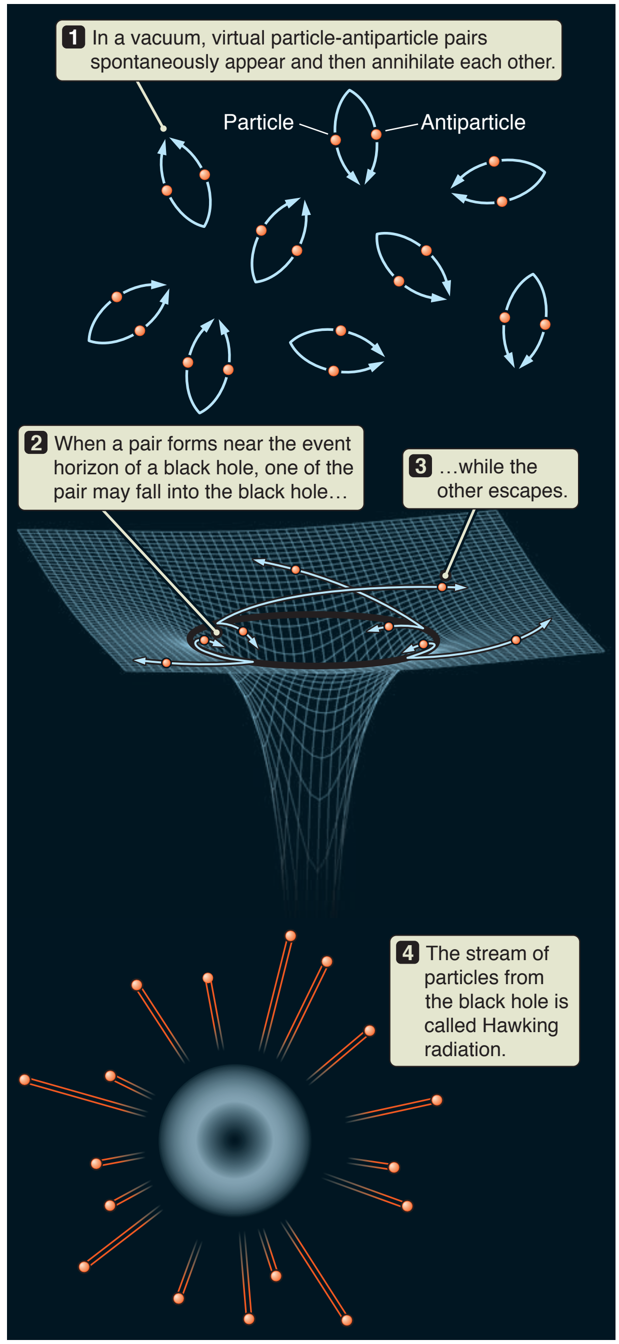


Figure 18.18 In the vacuum of empty space, particles and antiparticles are constantly being created and then annihilating each other. Near the event horizon of a black hole, however, one particle may cross the horizon and fall into the black hole before it recombines with its partner. The remaining particle leaves the vicinity of the event horizon, leading to the production of Hawking radiation.

18.2 Working It Out Masses in X-ray Binaries

Cygnus X-1 is part of a binary system with a blue supergiant star O9.7 I (very close to B0 I) and an unseen compact object located about 0.2 astronomical unit away from it. The blue supergiant and the compact object orbit a common center of mass every 5.6 days. We can use the simple formula from Working It Out 13.4 to calculate the sum of the masses:

$$\frac{M_{\text{blue}}}{M_{\text{Sun}}} + \frac{M_{\text{compact}}}{M_{\text{Sun}}} = \frac{A^3}{P^2}$$

with A in astronomical units (AU) and P in years. In this case, $A = 0.2$ AU and $P = 5.6/365.24$ years, so

$$\frac{M_{\text{blue}}}{M_{\text{Sun}}} + \frac{M_{\text{compact}}}{M_{\text{Sun}}} = \frac{0.2^3}{\left(\dfrac{5.6}{365.24}\right)^2} = 34$$

The sum of the masses of the two stars is 34 M_{Sun}.

To find the values of the two individual masses from their orbits, we need to know the velocities of the two stars or the distance of each star to the center of mass and the orbital inclination of the system. Obtaining such information is difficult when one star is compact and not observed separately. However, the mass of the blue supergiant star can be estimated from spectroscopic and photometric data at many wavelengths, assuming the distance to the system is known. When this is done, the mass of the supergiant is estimated at 19 M_{Sun}. Subtracting from 34, the mass of the compact object is 15 M_{Sun}—well over the mass limit for a neutron star. Therefore, Cygnus X-1 is assumed to be a black hole. A recent study with data from several X-ray telescopes concluded that the black-hole mass of Cygnus X-1 is 14.8 M_{Sun}.

found that the mass of the unseen compact companion of HD 226868 must be about 15 M_{Sun} (**Working It Out 18.2**). The companion to HD 226868 is too compact to be a normal star, yet it is much more massive than the Chandrasekhar limit for a white dwarf or the upper mass limit of a neutron star. Such an object can only be a black hole. The X-ray emission from Cygnus X-1 arises when material from the O9.7 I supergiant falls onto an accretion disk surrounding the black hole as illustrated in the chapter-opening figure.

In some similar systems, winds have been observed blowing off the disk around the black hole. These winds are probably caused by magnetic fields in the disk. The fastest winds observed are in the binary system IGR J17091, where the winds are as high as 32 million km/h (about 3 percent of the speed of light). This wind is blowing in many directions, and it may be carrying away more mass than is being captured by the black hole, as illustrated in **Figure 18.19**.

Astronomers have modeled the observational data of dozens of good candidates for stellar-mass black holes in X-ray binary systems in the Milky Way. They have found that the black-hole masses are greater than 4.5–5 M_{Sun}; that is, not very close to the limit of 2.5–3 M_{Sun} for a neutron star. This gap in mass between the most massive neutron stars and the least massive stellar black holes is not yet understood, and it is assumed to be a result of the mass transfer processes between the stars.

The black holes we discussed in this chapter came from collapsing massive stars, but this is not the only type of black hole. In Chapters 19 and 20, you will learn that supermassive black holes can be found at the centers of galaxies, including the Milky Way.

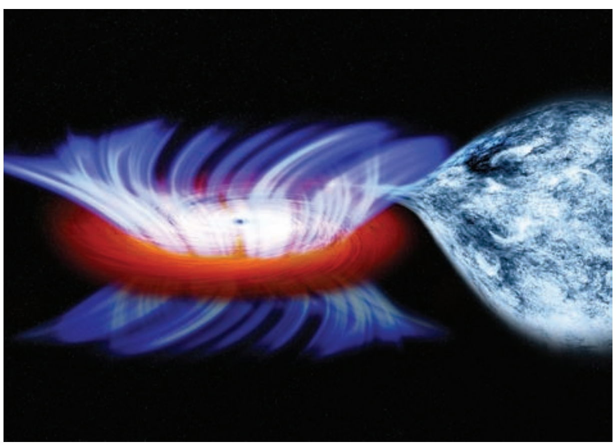


Figure 18.19 This artist's model shows strong winds being emitted from the disk around a stellar black hole. These winds can remove more material than the amount that actually falls into the hole.

CHECK YOUR UNDERSTANDING 18.4

If a black hole suddenly doubled in mass, the event horizon would become ______ its original radius. (a) one-quarter; (b) one-half; (c) twice; (d) 3 times; (e) 4 times

Origins

Gamma-Ray Bursts

The most energetic explosions in the universe are probably related to stellar black holes. **Gamma-ray bursts**, or **GRBs**, are intense bursts of gamma rays. The bursts are followed by a weaker "afterglow" that is observed at many wavelengths. GRBs were first observed in the 1960s by satellites designed to look for radiation from nuclear weapons being tested in space after such tests were banned on Earth. In the 1990s, gamma-ray astronomy satellites discovered that these bursts were coming from all directions in the sky and that they might be associated with supernova explosions in distant galaxies. Short-duration GRBs that last less than 2 seconds probably originate from the merging of two neutron stars or a neutron star and a black hole in a close binary system that collapses into a single black hole. The more common long-duration GRBs are easier to study because they have a longer afterglow. Astronomers think that they originate in the collapse of a very high-mass, rapidly spinning star to a black hole or a neutron star after a supernova explosion. (Supernovae from very high-mass stars are sometimes called **hypernovae**, or Type Ib or Ic supernovae.)

Unlike regular supernovae that radiate equally in all directions, GRBs are beamed events, so most of their enormous energies are concentrated into two opposite jets of emission as illustrated in **Figure 18.20a**. In addition to the electromagnetic radiation, there are relativistic jets of cosmic rays. Astronomers have not observed any GRBs in the Milky Way: there has not been a massive supernova in the Milky Way for at least a century. But the energy of GRBs is so intense that people have wondered what might happen to Earth if one went off nearby with its radiation beamed in Earth's direction, a possibility imagined in Figure 18.20b. A leading candidate for a future GRB in our galaxy is the massive star Eta Carinae, discussed in Chapter 17. This star is 7,500 light-years away—a neighbor, astronomically speaking. However, its rotation axis is such that is unlikely to form a GRB that beams toward Earth.

Some scientists wonder whether past supernova and GRB events could have affected the history of life on Earth. Supernova "archaeologists" may have found evidence on Earth of past supernovae. In one study, rocks deep in the Pacific Ocean were found to have amounts of a radioactive isotope of iron that is too short-lived to be left over from the formation of Earth. This iron-60 could have been deposited 2.8 million years ago on Earth after a supernova explosion. In another study, high concentrations of nitrates were found in some layers in Antarctic ice cores. Gamma radiation from supernovae can produce excess nitrogen oxides in the atmosphere, which then become converted to nitrates that are trapped in snowfall. Nitrate spikes correlate to 1006 and 1054 CE—two years when bright supernovae are known to have appeared in the Milky Way.

What about more drastic effects on Earth from a nearby supernova or a more distant but beamed GRB? Normally, Earth is protected from cosmic radiation and cosmic-ray particles by its ozone layer and its magnetic field. Cosmic-ray particles might not be a major problem if they arose very close to Earth, but the high-energy gamma-ray radiation could have a more serious effect on Earth. The excess nitrogen oxides they produce in the atmosphere can absorb sunlight, which would cool Earth. The gamma radiation could ionize Earth's atmosphere, reducing or destroying the ozone layer that protects life from ultraviolet radiation. Even a short burst of a few seconds could lead to ozone damage lasting for decades. The gamma rays could trigger a burst of solar UV radiation at Earth's surface, which could damage the DNA of phytoplankton a few hundred meters deep in the ocean, affecting their ability to photosynthesize. Phytoplankton are the base of Earth's food chain, so a drastic reduction in phytoplankton could upset the entire biosphere. It has been hypothesized that a GRB may have been responsible for the Ordovician mass extinction event 450 million years ago.

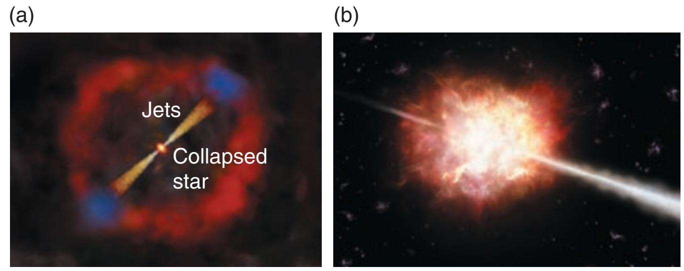


Figure 18.20 An artist's model of a gamma-ray burst. (a) Narrow beams of intense energy are sent in two opposite directions. (b) If one beam is pointed toward the observer, the GRB will appear bright.

Statistically speaking, GRBs that beam to Earth may be quite rare, and some astronomers have argued that they are less likely to be produced in a Milky Way–type galaxy than in other types of galaxies. There is a lot of uncertainty in any estimate of how close the supernova or GRB must be and how often these explosive events must occur to have a serious effect on Earth. According to one estimate, there is a supernova or GRB explosion close enough to alter Earth's biosphere a few times every billion years, possibly leading to mass extinction events. In Chapter 17, we noted that the chemical elements that make up life were created in supernova explosions. The discussion here suggests that supernovae may have had some effect on the *evolution* of life on Earth as well.

ARTICLES QUESTIONS

Astronomers simulate the formation of a black hole.

After Neutron Star Death-Match, a Black Hole Is Born

By **IAN O'NEILL, Discovery News**

What happens when two neutron stars collide? Using a sophisticated computer simulation, NASA scientists have visualized this violent scenario in awesome degenerate-matter-crushing detail.

Neutron stars are the result of supernovae spawned by stars 8–30 times the mass of our Sun. Occasionally, however, two neutron stars may meet, becoming entangled in a deep gravitational embrace. Should this scenario play out, one of the most powerful known explosions in the universe may be sparked—a fast gamma-ray burst (GRB).

But before two neutron stars collide, what happens to their structures? What kind of insanely powerful tidal forces are at play?

In a simulation released today (May 13) by NASA Goddard Space Flight Center scientists, two neutron stars are placed a mere 11 miles apart. Keep in mind that although both neutron stars are 1.5 and 1.7 times the mass of our Sun, all of that matter is packed into a tiny sphere only *12 miles wide*. As a result, their densities and gravitational fields are immense—a teaspoonful of neutron star material would weigh as much as Mount Everest. The crushing gravitational forces ensure that atomic structures cannot be sustained, collapsing the material into a neutron degenerate state—only the structure of the neutrons themselves prevent the neutron star's gravity from collapsing it into a point, forming a black hole.

Should more mass be added to the neutron star, a mass threshold may be reached when the gravitational forces overwhelm even the neutron degenerate pressure, causing it to collapse.

As the simulation unfolds, the neutron stars' savage tidal forces rip each other to shreds, cracking open their thin crusts and shedding huge quantities of material into space. As they are so close to one another, the two neutron stars rapidly spin, merging in a fraction of a second. The progression of this simulated neutron star merger produces a ring doughnut–shaped material that forms about the genesis of the newborn black hole in its center.

Simulating these violent events is important for our understanding of not only how some black holes are created: it can develop the science behind GRBs. As an extension, the mysterious source of the heaviest elements in the universe may be found in these events where the rapid cohesion of neutron stars—and the resulting explosions—could forge rapid-neutron capture (or "r-process") elements.

The majority of r-process elements—heavier than iron—found in the universe are thought to be generated inside core-collapse supernovae. However, there's a growing body of evidence that suggests merging neutron stars may be a fertile environment for the largest atomic nuclei to form.

So neutron star mergers are not only a fascinating field of astrophysical curiosity, they may also be the driving production mechanism responsible for the heaviest elements and complex chemistry throughout the cosmos.

ARTICLES QUESTIONS

1. Explain why the merger of two neutron stars will make a black hole.
2. How do the two stars get so close together?
3. What is space like near the merging neutron stars?
4. R-process elements form from nucleosynthesis after rapid neutron capture. How would capturing a neutron lead to nucleosynthesis?
5. Watch the video at http://svs.gsfc.nasa.gov/vis/a010000/a011500/a011530/index.html. Why do such mergers have to be studied on a computer?

Summary

The highest-mass stars leave behind black holes. In the environment surrounding black holes, relativistic effects become important. A black hole's mass determines its Schwarzschild radius: the boundary from which light cannot escape. Gamma-ray bursts are beamed high-energy explosions that result from the merger of two compact objects or the rapid collapse of a high-mass star to a black hole. The radiation from these bursts could affect life on Earth.

LG 1 **Describe how the motion of the observer affects the observed velocity of objects.** Even at relatively low relative speeds, the motion of the observer can affect the direction of the measured velocity, as in the aberration of starlight. At higher speeds, the magnitude of the measured velocity is also affected. The speed of light in a vacuum, c, is the ultimate speed limit. Observers in all inertial reference frames will measure the same speed of light.

LG 2 **Discuss the observable consequences of the relationship between space and time.** Special relativity connects space and time into four-dimensional spacetime. How the spacetime distance between events is divided between space and time depends on the observer's motion. As observers approach the speed of light, they observe moving clocks appearing to run more slowly than their own clocks, moving objects to be contracted in length, and moving objects to behave as if they were more massive.

LG 3 **Recognize that gravity is a consequence of the way mass distorts the very shape of spacetime.** Inertial mass and gravitational mass are the same, leading to the principle of equivalence, in which acceleration cannot be distinguished from gravity if the acceleration is small enough. In general relativity, mass warps the fabric of spacetime so that objects move on the shortest path in this warped geometry. Gravity is a consequence of the spacetime warping. Time runs more slowly near massive objects, and radiation from any light source near a black hole's event horizon is redshifted.

LG 4 **Explain why the most massive stars end as black holes, and describe the key properties of these stellar black holes.** There is an upper limit to the mass for both white dwarfs and neutron stars. Dense stellar remnants more massive than about 3 M_{Sun} collapse to form black holes. The supernova explosion that ends the life of a massive star leaves behind a neutron star or a black hole. The mathematical singularity at the center of a black hole is still a mystery to science. However, observational evidence for black holes is very strong: scientists have identified many objects that have strong gravity but are too small to be normal matter. Black holes may, after a very, very long time, be destroyed by evaporation through Hawking radiation.

? UNANSWERED QUESTIONS

- What happens to the information that falls into a black hole? We said earlier that a black hole is characterized by only three properties: mass, angular momentum, and electric charge. Where did all the other information go? To a distant observer, it takes an infinitely long time for material to fall into a black hole, so although the observer sees less and less radiation from the material, the properties of the material seem to be the same for all time. But from the perspective of the infalling material, it takes a finite time to cross the event horizon, so information other than mass, angular momentum, and charge can no longer be shared with the outside world.
- Do wormholes exist in spacetime, connecting one region with another, perhaps through black holes? Wormholes are a mathematical solution to the equations of general relativity. The idea is that when something goes into a black hole, it travels through a wormhole and emerges in a different part of the universe. In this way a wormhole acts as a shortcut through spacetime. In science fiction, wormholes are a popular means of traveling large distances by exploiting the strange geometry of spacetime. But many scientists doubt that wormholes can exist in nature, and even if they do exist, strong tidal forces would pull apart anything that falls into a black hole before it emerged from a wormhole.

Questions and Problems

Test Your Understanding

1. Rank the following in terms of the mass of the star that produces each from least to most massive.
   a. neutron star
   b. black hole
   c. white dwarf

2. A car approaches you at 50 km/h. A fly inside the car is flying toward the back of the car at 7 km/h. From your point of view by the side of the road, the fly is moving at __________ km/h.
   a. 7
   b. 28.5
   c. 43
   d. 57

3. A car approaches you at 50 km/h. The driver turns on the headlights. From your point of view, the light from the headlights is moving at
   a. c + 50 km/h.
   b. c − 50 km/h.
   c. (c + 50 km/h)/2.
   d. c.

4. Imagine that you are on a spaceship. A second spaceship rockets past yours at 0.5c. You start a stopwatch and stop it 10 seconds later. For an astronaut in the other spaceship, the number of seconds that have ticked by during the 10 seconds on your stopwatch is
   a. more than 10 seconds.
   b. equal to 10 seconds.
   c. less than 10 seconds.

5. The International Space Station flies overhead. Using a telescope, you take a picture and measure its length to be __________ than its length as it would be measured if it were sitting on the ground.
   a. much greater
   b. slightly greater
   c. slightly less
   d. much less

6. Astronauts in the International Space Station
   a. have no mass.
   b. have no energy.
   c. are outside of Earth's gravitational field.
   d. are in free fall.

7. Einstein's formulation of gravity
   a. is approximately equal to Newton's universal law of gravitation for small gravitation fields.
   b. is always used to calculate gravitational effects in modern times.
   c. explained why Newton's universal law of gravitation describes the motions of masses.
   d. both a and c

8. As the mass of a black hole increases, its Schwarzschild radius
   a. increases as the square of the mass.
   b. increases proportionately.
   c. stays the same.
   d. decreases proportionately.
   e. decreases as the square of the mass.

9. If a neutron star is more than 3 times as massive as the Sun, it collapses because
   a. the force of electron degeneracy is stronger than gravity.
   b. gravity overpowers the force of electron degeneracy.
   c. gravity overpowers the force of neutron degeneracy.
   d. the force of neutron degeneracy is stronger than gravity.

10. Relative motion between two objects is apparent
    a. even at everyday speeds, such as 10 km/h.
    b. only at very large speeds, such as 0.8c.
    c. only near very large masses.
    d. only when both objects are in the same reference frame.

11. If a spaceship approaches you at 0.5c, and a light on the spaceship is turned on pointing in your direction, how fast will the light be traveling when it reaches you?
    a. 1.5c
    b. between 1.0c and 1.5c
    c. exactly c
    d. between 0.5c and 1.0c

12. Imagine two protons traveling past each other at a distance d, with relative speed 0.9c. Compared with two *stationary* protons a distance d apart, the gravitational force between these two protons will be
    a. smaller, because they interact for less time.
    b. smaller, because the moving proton acts as if it has less mass.
    c. the same, because the particles have the same mass.
    d. larger, because the moving proton acts as if it has more mass.

13. If two spaceships approach each other, each traveling at 0.5c relative to an outside observer, spaceship 1 will measure spaceship 2 to be traveling
    a. much faster than c.
    b. slightly faster than c.
    c. at c.
    d. more slowly than c.

14. The strongest evidence for black holes comes from rapidly flickering X-ray sources. This observation indicates that
    a. X-rays are not light, because even light cannot escape from a black hole.
    b. the X-rays are coming from a very small source—so small that it could be a black hole.
    c. black holes are (at least sometimes) surrounded by hot gas.
    d. black holes have a very high temperature.

15. The current model of long-duration GRBs includes jets from the collapsed star. You have seen jets like this before, when studying (choose all that apply)
    a. giant planets.
    b. white dwarfs.
    c. pulsars.
    d. star formation.
    e. supernovae.
    f. planetary nebulae.

Thinking about the Concepts

16. An astronomer sees a redshift in the spectrum of an object. Without any other information, can she determine whether this is an extremely dense object (exhibiting gravitational redshift) or one that is receding from her (exhibiting Doppler redshift)? Explain your answer.

17. Imagine you are traveling in a spacecraft at 0.9999999*c*. You point your laser pointer out the back window of the spacecraft. At what speed does the light from the laser pointer travel away from the spacecraft? What speed would be observed by someone on a planet traveling at 0.000001*c*?

18. Einstein's special theory of relativity says that no object can travel faster than, or even at, the speed of light. Recall that light is both an electromagnetic wave and a particle called a photon. If it acts as a particle, how can a photon travel at the speed of light?

19. Twin A takes a long trip in a spacecraft and returns younger than twin B, who stayed behind. Could twin A ever return before twin B was born? Explain.

20. In one frame of reference, event A occurs before event B. Is it possible, in another frame of reference, for the two events to be reversed, so that B occurs before A? Explain.

21. Imagine you are watching someone whizzing by at very high speed in a spacecraft. Will that person's pulse rate appear to be extremely fast or extremely slow?

22. Suppose you had a density meter that could instantly measure the density of an object. You point the meter at a person in a spacecraft zipping by at very high speed. Is that person's mass density larger than an average person's or smaller? Explain.

23. Imagine a future astronaut traveling in a spaceship at 0.866 times the speed of light. Special relativity says that the length of the spaceship along the direction of flight is only half of what it was when it was at rest on Earth. The astronaut checks this prediction with a meter stick that he brought with him. Will his measurement confirm the contracted length of his spaceship? Explain your answer.

24. Looking into a speeding spaceship, you observe that the travelers are playing soccer with a perfectly round soccer ball. What is the shape of the ball according to observers on the spacecraft?

25. You observe a meter stick traveling past you at 0.9999*c*. You measure the meter stick to be 1 meter long. How is the meter stick oriented relative to you?

26. Suppose astronomers discover a 3-M_{Sun} black hole located a few light-years from Earth. Should they be concerned that its tremendous gravitational pull will lead to Earth's untimely demise?

27. If you could watch a star falling into a black hole, how would the color of the star change as it approached the event horizon?

28. Why don't people detect the effects of special and general relativity in their everyday lives here on Earth?

29. Many movies and television programs (like *Star Wars*, *Star Trek*, and *Battlestar Galactica*) are premised on faster-than-light travel. How likely is it that such technology will be developed in the near future?

30. How could a gamma-ray burst in our galaxy potentially affect life on Earth?

Applying the Concepts

31. Study Figure 18.1b. Does the angle of the rain falling outside the car depend on the speed of the car? Knowing only this angle and the information on your speedometer, how could you determine the speed of the falling rain?

32. Compare Figure 18.1 with Figure 18.2. If you knew the speed of Earth in its orbit from a prior experiment, how could you determine the speed of light from the angle of the aberration of starlight?

33. According to Einstein, mass and energy are equivalent. So, which weighs more on Earth: a cup of hot coffee or a cup of iced coffee? Why? Do you think the difference is measurable?

34. As explained in the Process of Science Figure, a new theory should contain the old theory within it. Study Figure 18.5, which compares two imaginary rocket ships experiencing the same acceleration, *g*.
    a. Approximately how fast (as a fraction of the speed of light) are the two spaceships going when the effects of relativity begin to be significant?
    b. Convert this speed to kilometers per hour. How does this speed compare to the speeds at which you usually travel? Why do you not usually see relativistic effects in your life?

35. Figure 18.6 shows how the Lorentz factor depends on speed. At about what speed (in terms of c) does the Lorentz factor begin to differ noticeably from 1? What happens to the Lorentz factor as the speed of an object approaches the speed of light?

36. The perihelion of Mercury advances 2 degrees per century. How many arcseconds does the perihelion advance in a year? (Recall that there are 60 arcseconds in an arcminute and 60 arcminutes in a degree.) Is it possible to measure Mercury's position well enough to measure the advance of perihelion in 1 year?

37. Study Figure 18.12. If the Sun were twice as massive, would the distance between the apparent positions of stars 1 and 2 increase or decrease during an eclipse?

38. Follow Working It Out 18.1 to find out how much younger you would be than your twin if you made the journey described there at $0.5c$.

39. Use Working It Out 18.1 to predict how much younger you would be than your twin if you traveled at $0.999c$ instead of $0.5c$ as in question 38. Then calculate the difference in ages and compare the calculated result to your prediction.

40. What is the Schwarzschild radius of a black hole that has a mass equal to the average mass of a person ($\sim$70 kilograms)?

41. What is the mass of a black hole with a Schwarzschild radius of 1.5 km?

42. The Moon has a mass equal to $3.7 \times 10^{-8}\ M_{\text{Sun}}$. Suppose the Moon suddenly collapsed into a black hole.
    a. What would be the Schwarzschild radius of the black-hole Moon?
    b. What effect would this collapse have on tides raised by the Moon on Earth? Explain.
    c. Do you think this event would generate gravitational waves? Explain.

43. If a spaceship approaching Earth at 0.9 times the speed of light shines a laser beam at Earth, how fast will the photons in the beam be moving when they arrive at Earth?

44. Suppose you discover signals from an alien civilization coming from a star that is 25 light-years away, and you go to visit it using the spaceship described in the discussion of the twin paradox in Working It Out 18.1.
    a. How long will it take you to reach that planet, according to your clock? According to a clock on Earth? According to the aliens on the other planet?
    b. How likely is it that someone you know will be here to greet you when you return to Earth?

45. Working It Out 18.2 relates the mass of a binary pair to the period and the size of the orbit. Suppose that a spaceship orbited a black hole at a distance of 1 AU, with a period of 0.5 year. What assumptions could you make that would allow you to calculate the mass of the black hole from this information? Make those assumptions, and calculate the mass.

USING THE WEB

46. Go to the "Through Einstein's Eyes" website (http://anu.edu.au/Physics/Savage/TEE/site/tee/home.html), click on "Start Here," and "Take the tour" and "Movie explained." Take the ride on the "relativistic roller coaster." Why do colors look different on the relativistic roller coaster? Click on "Continue tour" and view the cube, tram, and desert road. How do things look different? Why do you get rainbow when driving down the desert at high gamma? Continue for the Solar System tours (or return to the home page). What do you see when you approach the Sun or a planet at a relativistic speed?

47. NASA missions:
    a. Go to NASA's Swift Gamma-Ray Burst Mission website (http://swift.gsfc.nasa.gov/). Locate a recent result related to supernovae, gamma-ray bursts, or stellar black holes. Why would two merging neutron stars likely form a black hole?
    b. NASA's Fermi Gamma-ray Space Telescope (http://www.nasa.gov/content/fermi-gamma-ray-space-telescope and http://fermi.gsfc.nasa.gov) is exploring the gamma-ray universe. What objectives of this mission relate to the study of black holes? What is a recent news story related to black holes?

48. Go to the LIGO website (http://ligo.org/science.php) and read about gravitational waves. Click on "Sources of Gravitational Waves" and listen to the example. What are the differences among the four listed sources of gravitational waves? Click on "Advanced LIGO." What's new with the project?

49. The newest NASA mission to study black holes, gamma-ray bursts, and neutron stars is named NuSTAR (Nuclear Spectroscopic Telescopic Array). Go to the NuSTAR website (http://www.nustar.caltech.edu/news). What wavelengths and energies does this telescope observe? What has been observed? What new science has been learned?

50. Go to the "Inside Black Holes" website (http://jila.colorado.edu/~ajsh/insidebh), enter, and click on "Schwarzschild." Work your way down the page, watching the videos. What does it look like when you go into a black hole? Why is there gravitational lensing when Earth is in orbit? What happens when you fall through the horizon: is everything black? Click on "Reissner-Nordström" to see an electrically charged black hole. What is a wormhole? Why is there a warning at the top of the page? Click on "4D perspective." What does it look like if you move toward the Sun at the speed of light?

smartwork5

If your instructor assigns homework in Smartwork5, access your assignments at digital.wwnorton.com/astro5.

digital.wwnorton.com/astro5

Because it's not possible to go grab a black hole and bring it into the lab, and because Earth has never actually been close to one, astronomers can only conduct thought experiments to explore the properties of black holes. Following are a few thought experiments to help you think about what's happening near and around a black hole.

Imagine a big rubber sheet. It is very stiff and not easily stretched, but it does have some "give" to it. At the moment, it is perfectly flat. Imagine rolling some golf balls across it.

1 Describe the path of the golf balls across the sheet.

Now imagine putting a bowling ball (very much heavier than a golf ball) in the middle of the sheet, so that it makes a big, slope-sided pit. Roll some more golf balls.

2 What happens to the path of the golf balls when they are very far from the bowling ball?

3 What happens to the path of the golf balls when they come just inside the edge of the dip?

4 What happens to the path of the golf balls when they go directly toward the bowling ball?

5 How do each of the three cases in questions 2–4 change if the golf balls are moving very, very fast? What if they are moving very slowly?

6 What happens to the depth and width of the pit as the golf balls fall into the center near the bowling ball? (Imagine putting *lots* of golf balls in.)

All of the preceding thought experiments relate to ordinary stuff. Stars, people, planets—everything interacts in this way because of gravity. With black holes, things are a bit different. In this case, it is more accurate to think of the bowling ball as a hole in the sheet that pulls it down, rather than as an object that sits on it. But the bowling ball still affects the sheet in the same way. The hole is a good analogy for the event horizon of a black hole. Objects outside the event horizon will know that the black hole is there, because the sheet is sloping, but they won't be captured unless they come within the event horizon. Think about light for a moment as though it were, say, grains of sand rolling across the sheet.

7 What happens to the light as it passes far from the pit? What happens if it reaches the hole?

Now suppose you roll another bowling ball across the sheet.

8 What happens to the sheet when the second bowling ball falls in after the first? Would this change affect your golf balls and grains of sand? How? What happens to the hole? What happens to the size of the pit?

None of these thought experiments take into account relativistic effects (length contraction and time dilation). Imagine for a moment that you are traveling close to the black hole.

9 Look out into the galaxy and describe what you see. Consider the lifetimes of stars, the distances between them, their motions in your sky, and how they die. Add anything else that occurs to you.

19 Galaxies

It has been less than a century since astronomers realized that the universe is filled with huge collections of stars, gas, and dust called galaxies. Just as stars vary in their mass or their stage of evolution, galaxies come in many forms. This chapter begins our discussion of galaxies with a survey of the types of galaxies and their basic properties to understand better the differences among them. The next chapter looks in detail at our galaxy, the Milky Way, and in subsequent chapters we will look at the evolution of galaxies and of the universe itself.

LEARNING GOALS

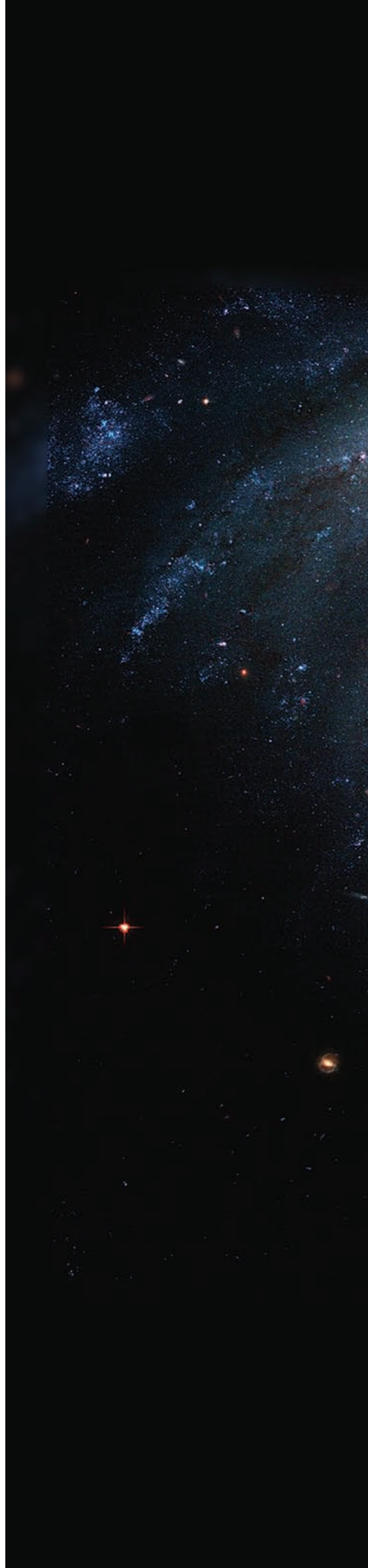
By the conclusion of this chapter, you should be able to:

LG 1 **Determine a galaxy's type from its appearance, and describe the motions of its stars.**

LG 2 **Explain the distance ladder and how distances to galaxies are measured.**

LG 3 **Describe the evidence suggesting that galaxies are composed mostly of dark matter.**

LG 4 **Discuss the evidence indicating that most—perhaps all—large galaxies have supermassive black holes at their centers.**



The large, barred spiral galaxy NGC 1300 is about 20 million parsecs away. ▶▶▶

How do galaxies get their shapes?

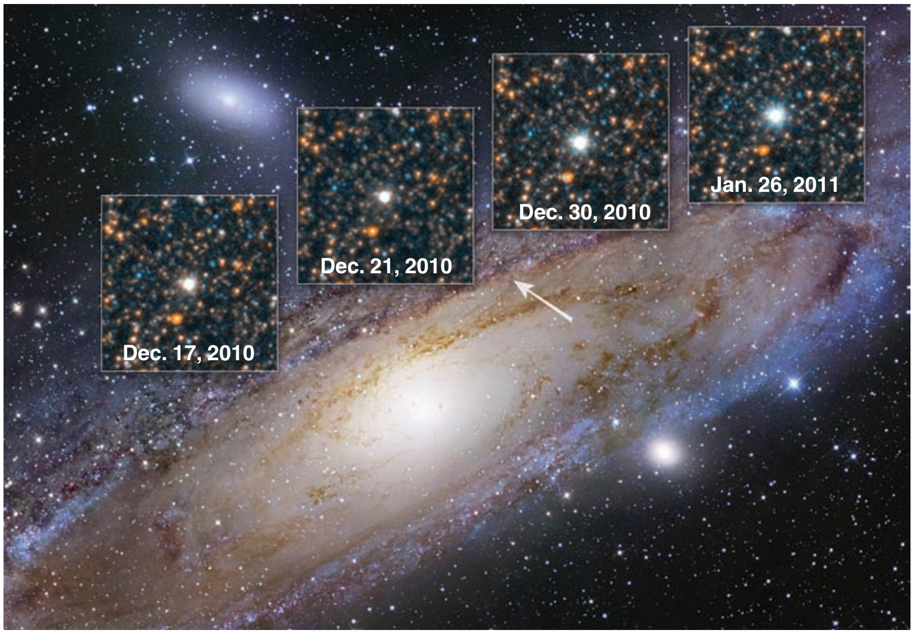


Figure 19.1 The Andromeda Galaxy, the nearest large galactic neighbor to the Milky Way, is about 2.5 million light-years (780,000 pc) away. The arrow points to the Cepheid variable star V1, a standard candle that Hubble used to estimate the distance to Andromeda. The insets show V1's variability. Hubble's measurement provided the first observational evidence of the vastness of the universe.

19.1 Galaxies Come in Different Shapes and Sizes

A **galaxy** is a gravitationally bound collection of dust, gas, and a million to hundreds of billions of stars. The universe contains more galaxies than there are stars in the Milky Way Galaxy. Most of these galaxies are located at such astonishing distances that they appear too small and faint to detect with any but the most powerful telescopes. In this section, we will discuss how astronomers concluded that galaxies were separate from our Milky Way and will differentiate the various types of galaxies.

The Discovery of Galaxies in the 20th Century

Observers have long known that the sky contains faint, misty patches of light. These objects were originally called nebulae (singular: *nebula*) because of their nebulous (fuzzy) appearance. In 1784, astronomer Charles Messier (1730–1817) published a catalog of 103 nebulous objects. Twenty years later, because of the observations by the astronomers William Herschel and his sister Caroline Herschel, that number jumped to 2,500. Although some of the nebulae looked diffuse and amorphous, most were round or elliptical or resembled spiraling whirlpools. These distinctions were the basis for the original three categories—diffuse, elliptical, and spiral—used to classify nebulae.

Over the next 140 years, it was suggested that spiral nebulae might be relatively nearby planetary systems in various stages of formation. Alternatively, the influential 18th century philosopher Immanuel Kant (1724–1804) speculated that spiral nebulae were instead "island universes"—separate from the Milky

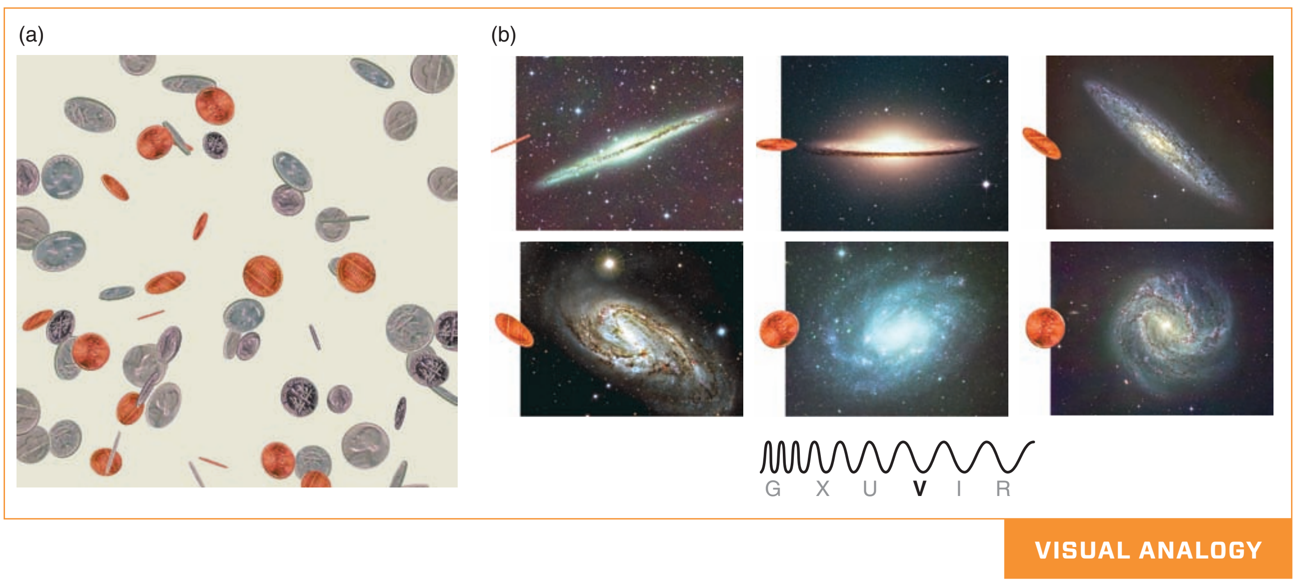


Figure 19.2 (a) A handful of coins thrown in the air provides an analogy for the difficulties in identifying the shapes of certain types of galaxies. Like the coins in this picture, galaxies are seen in various orientations—some face on, some edge on, and most somewhere in between. (b) These disk-shaped galaxies are seen from various perspectives or angles, corresponding to the range of perspectives for coins.

Way Galaxy. The presence of interstellar dust complicated early attempts to understand the size of the Milky Way. Early astronomers did not know of the existence or consequences of the dust that blocks the passage of visible light through the Milky Way. Unable to see past this obscuring shroud, they concluded that the Milky Way is a system of stars some 1,800 parsecs (pc) across (recall from Chapter 13 that 1 parsec = 3.26 light-years). In the beginning of the 20th century, astronomer Harlow Shapley (1885–1972) used observations of globular clusters to estimate that instead, the Milky Way is more than 50 times larger than the earlier estimates—300,000 light-years, or 92,000 pc, in size.

Shapley thought his far larger estimate for the size of the Milky Way meant that it was big enough to encompass everything in the universe, and therefore he thought that the spiral and elliptical nebulae were inside the Milky Way. Astronomer Heber D. Curtis (1872–1942) preferred the earlier, smaller model of the Milky Way. He also favored the idea that the spiral nebulae were in fact galaxies separate from the Milky Way and that therefore the whole universe was larger than the Milky Way. In 1920, Shapley and Curtis met in Washington, D.C., to debate publicly their interpretations of the nature of spiral nebulae. Historians call this meeting astronomy's *Great Debate*. While this debate did not resolve the issue at the time, it set the stage and gave direction to the subsequent work of Edwin P. Hubble (1889–1953), whose discoveries fundamentally changed the modern understanding of the universe.

Using the newly finished 100-inch telescope on Mount Wilson, high above the then small city of Los Angeles, Hubble was able to find some variable stars in the spiral nebula Andromeda, as shown in **Figure 19.1**. He recognized that these stars were very similar to but appeared fainter than other known Cepheid variable stars. Using the period-luminosity relation for Cepheid variable stars discussed in Chapter 17, Hubble turned his observations of these stars into measurements of the distances to these objects. The results showed that the distances to these nebulae are far greater than Shapley's size of the Milky Way. Kant was correct: the Milky Way is one of many island universes. Most diffuse nebulae are clouds of gas and dust near Earth in the Milky Way Galaxy, but what Messier and others thought of as elliptical and spiral nebulae are instead galaxies similar in size to the Milky Way but located at truly immense distances.

Types of Galaxies

Imagine taking a handful of coins and throwing them in the air, as shown in **Figure 19.2a**. You know that all of these objects are flat and circular. When you look at the objects falling through the air, however, they do not appear all the same. Some coins appear face on and look circular. Some coins are seen edge on and look like thin lines. Most coins are seen from an angle between these two extremes and appear with various degrees of ellipticity, or flattening. Even if this image of many coins was the only information you had, you could use it to figure out the three-dimensional shape of a coin—flat and circular.

Astronomers use a similar method to discover the true three-dimensional shapes of galaxies. Figure 19.2b shows a set of galaxies seen from various viewing angles, from face on to edge on. You can infer from these images that, just like the coins in Figure 19.2a, some galaxies are disk-shaped and are randomly oriented on the sky.

The classifications for galaxies used today date back to the 1930s, when Edwin Hubble sorted the different shapes into categories like those shown in **Figure 19.3**. Hubble grouped all galaxies according to appearance and positioned them on a diagram that resembles the tuning fork used in the tuning of a musical in-

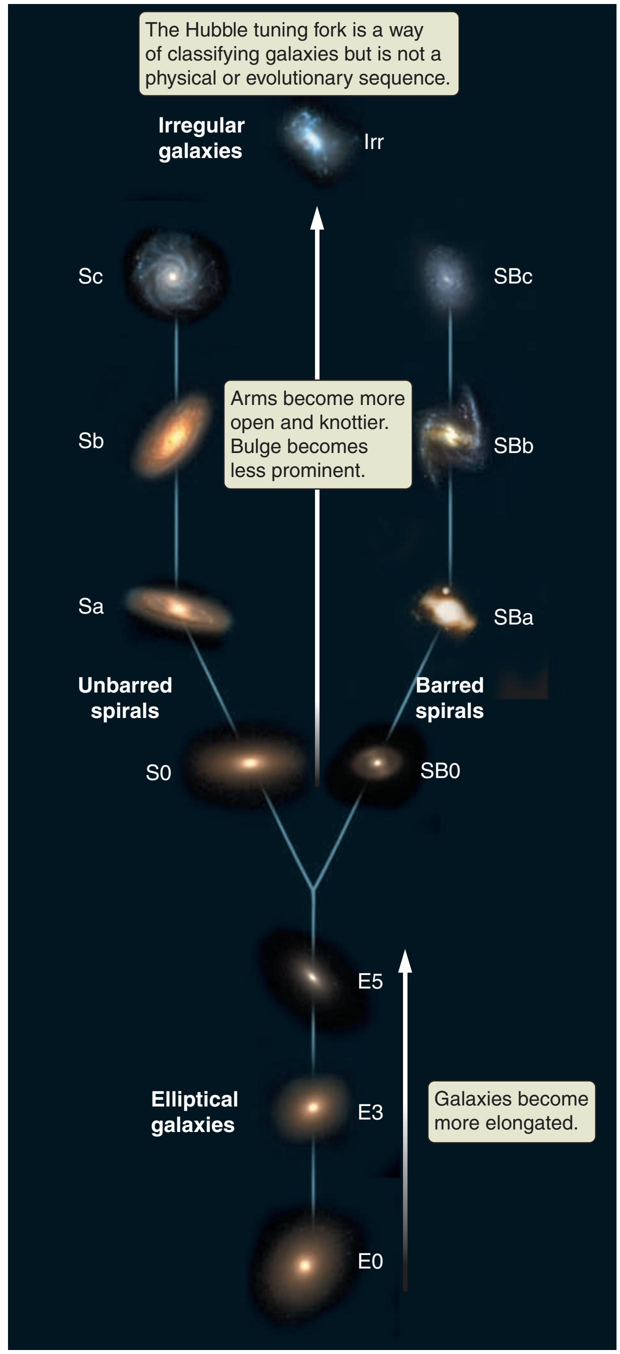


Figure 19.3 This tuning fork diagram illustrates Edwin Hubble's scheme for classifying elliptical (E), spiral (S, SB), and S0 galaxies according to their appearance only. Irregular (Irr) galaxies were not placed on the tuning fork. Hubble's scheme does not indicate an evolutionary sequence of galaxies.

strument. Galaxies come in three basic types: spirals, ellipticals, and irregulars—the latter including all other shapes. Originally, Hubble thought that his tuning fork diagram might indicate an evolutionary sequence for galaxies similar to what the Hertzsprung-Russell (H-R) diagram had done for stars. Even though it is *not* an evolutionary sequence, Hubble's classification scheme is still used for sorting galaxy types by their appearance in visible light.

At the bottom of Figure 19.3 are objects that are generally elliptical in three dimensions. These **elliptical galaxies** (labeled "E" on the diagram) can have either spherical or ellipsoidal shapes, and they show little evidence of the flat, round disk seen in other types of galaxies. They have numbered subtypes ranging from nearly spherical (E0) to quite flattened (E7). As with the coins tossed in the air, the appearance of an elliptical galaxy in the sky does not necessarily tell us its true shape. For example, a galaxy might actually be shaped like a rugby ball (which has rounded ends), but if viewed end on, it looks round like a soccer ball instead.

Spiral galaxies, labeled "S" on the left side of Figure 19.3, are characterized by spiral arms that lie in a flattened, rotating disk. In addition to disks and arms, a spiral galaxy has a central **bulge**, which extends above and below the disk. Hubble noticed that the bulges of about half of the spiral galaxies are bar-shaped: these galaxies are called **barred spirals**, labeled "SB" on the right side of Figure 19.3. Both spirals and barred spirals are subdivided into types a, b, and c according to the prominence of the central bulge and how tightly the spiral arms are wound. For example, Sa and SBa galaxies have the largest bulges and display tightly wound and smooth spiral arms. Sc and SBc galaxies have small central bulges and more loosely woven spiral arms that are often very knotty in appearance. The Milky Way Galaxy is a barred spiral (SBbc).

Astronomy in Action: Galaxy Shapes and Orientation

The distinction between spiral and elliptical galaxies is not always clear. Some galaxies, known as **S0 galaxies**, are a combination of the two types, having stellar disks but no spiral arms, so that the disk is smooth in appearance, like an elliptical galaxy. Hubble differentiated S0 galaxies as either barred (SB0) or unbarred (S0). Modern telescopic observations have revealed that many, if not most, elliptical galaxies contain small rotating disks at their centers, and both elliptical and S0 galaxies have little star formations. Galaxies that fall into none of these classes are **irregular galaxies** (labeled "Irr" in Figure 19.3). As their name implies, irregular galaxies are often without symmetry in shape or structure. About 25 percent of galaxies are irregular, and now astronomers think that most of them once were spirals or ellipticals that became distorted by the gravity of another galaxy. **Table 19.1** summarizes the criteria that Hubble used to classify galaxies.

TABLE 19.1 The Hubble Classification of Galaxies

A Classification Scheme Based on the Properties of Galaxies

Category	Criteria	Abbreviation	Range of Features		
Ellipticals	Mostly bulge Old, red stellar population Smooth-appearing	E0 ↕ E7	More spherical ↕ More elongated		
S0 (unbarred/ barred)	Bulge and disk with no arms and with mostly old, red stars	S0/SB0	Smooth disk and bulge		
Spirals (unbarred/ barred)	Bulge and disk with arms Bulge has old, red stars Disk has both old, red stars and young, blue stars Spirals (S) have roundish bulges Barred spirals (SB) have elongated or barred bulges	Sa/SBa Sb/SBb Sc/SBc	More bulge ↕ Little bulge	Tightly wound arms ↕ Open arms	Smooth arms ↕ Knotty arms
Irregulars	No arms, no bulge Some old stars, but mostly young stars, gas and dust, giving a knotty appearance	Irr			

Stellar Motions and Galaxy Shape

Stellar motions determine galaxy shapes. A galaxy is not a solid object like a coin, but a collection of stars, gas, and dust. In an elliptical galaxy, stars move in all possible directions, following orbits with a wide range

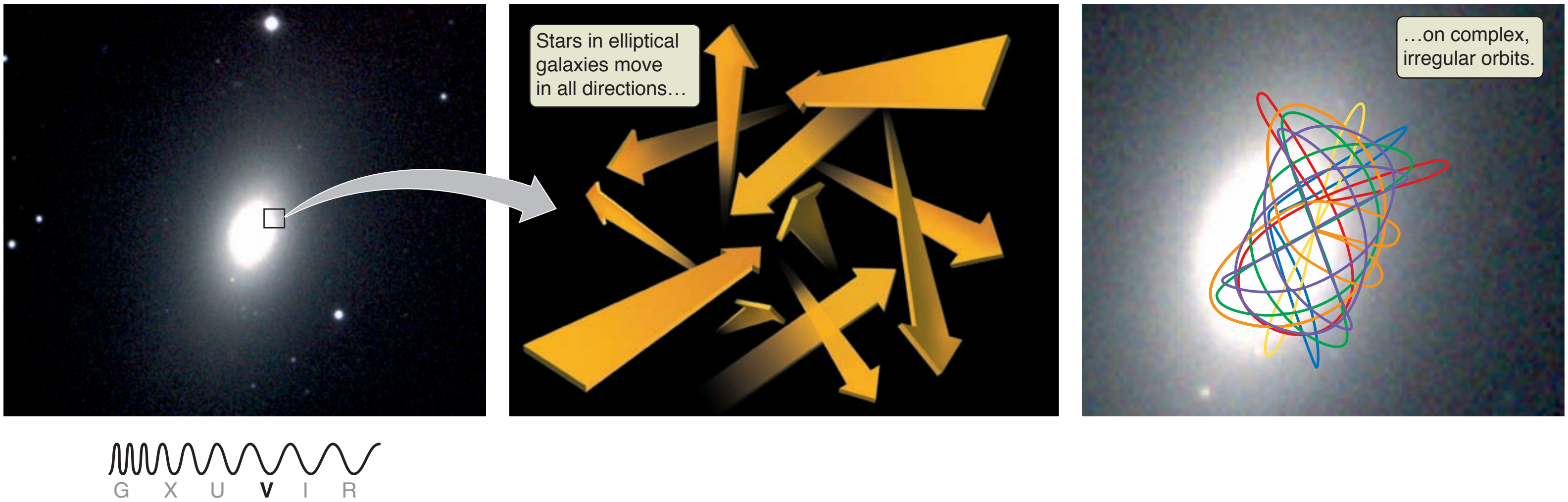


Figure 19.4 Elliptical galaxies take their shape from the orbits of the stars they contain. The colored lines superimposed on the galaxy represent the complex orbits of its stars

of shapes, as shown in **Figure 19.4**. These orbits are more complex than the orbits of planets about a star because the gravitational field within an elliptical galaxy does not come from a single central object. Taken together, all of these stellar orbits give an elliptical galaxy its shape.

Orbital speeds are also a factor. The faster the stars are moving, the more spread out the galaxy is. If the stars in an elliptical galaxy are moving in truly random directions, the galaxy will have a spherical shape. However, if stars are more likely to have certain directions of motion than others at each location, the galaxy will be more spread out in that direction, giving it an elliptical shape.

The orbits of stars in the disks of spiral galaxies are quite different from those of stars in elliptical galaxies. The components of a barred spiral galaxy are shown in **Figure 19.5**. The defining feature of a spiral galaxy is that it has a flattened, rotating disk. Like the planets of the Solar System, most of the stars in the disk of a spiral galaxy follow nearly circular orbits and travel in the same direction around a concentration of mass at the center of the galaxy. But the stellar orbits in a spiral galaxy's central bulge are quite different from those in the galaxy's disk. As with elliptical galaxies, the gravitational field within the bulge does not come from a single object, and the stars therefore follow random orbits. The bulges of unbarred spiral galaxies are thus roughly spherical in shape.

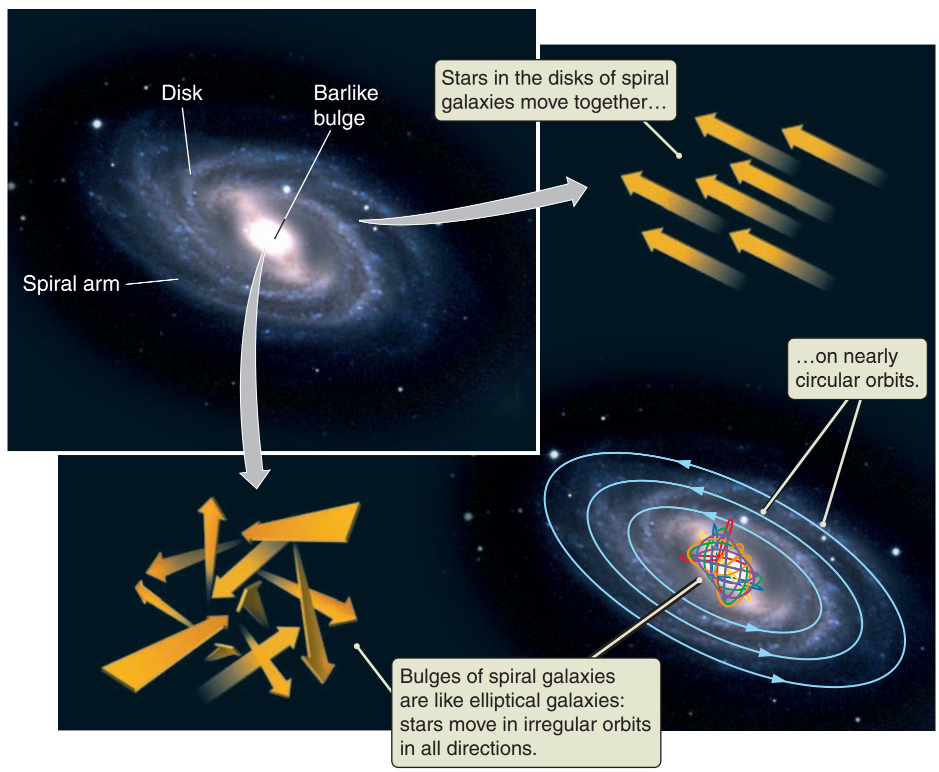


Figure 19.5 The components of a barred spiral galaxy include a barlike bulge, a disk, and spiral arms. The orbits of stars in the rotating disk are different than the orbits of stars in the elliptically shaped bulge.

CHECK YOUR UNDERSTANDING 19.1

Galaxies are classified according to: (a) mass; (b) color; (c) density; (d) shape.

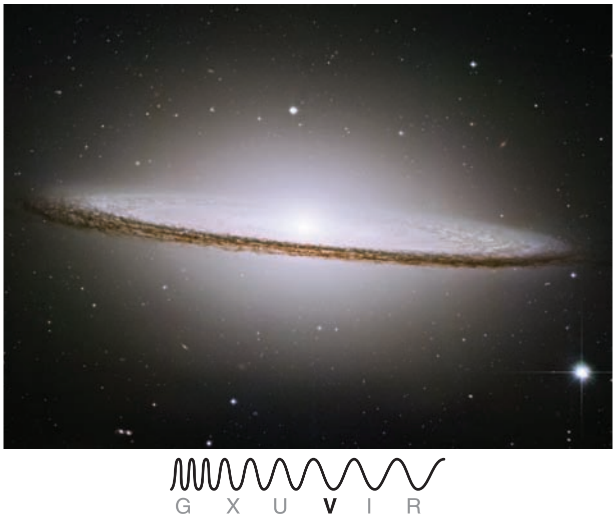


Figure 19.6 This Hubble Space Telescope (HST) image shows the nearly edge-on spiral galaxy M104 (the Sombrero Galaxy; type Sa). The dust in the plane is seen as a dark, obscuring band in the midplane of the galaxy. Note the bright halo made up of stars and globular clusters. Compare this image with Figure 15.1, which shows the dust in the plane of the Milky Way.

Other Differences among Galaxies

In addition to the differences in their stellar orbits, there are other important distinctions between spiral and elliptical galaxies. These distinctions carry information about the way they have evolved in the past and the way they will evolve in the future.

Gas and Dust Most spiral galaxies contain large amounts of dust and cold, dense molecular gas concentrated in the midplanes of their disks. Just as the dust in the disk of the Milky Way can be seen on a clear summer night as a dark band slicing the galaxy in two (see Figure 15.1), the dust in an edge-on spiral galaxy appears as a dark, obscuring band running down the midplane of the disk (**Figure 19.6**). The cold molecular gas that accompanies the dust can also be seen in radio observations of spiral galaxies. In contrast, giant elliptical galaxies contain large amounts of very hot gas that astronomers see primarily by observing the X-rays that the gas emits.

The difference in shape between elliptical and spiral galaxies offers some insight into why the gas in giant ellipticals is hot, while in spirals it is cold and dense. Conservation of angular momentum causes cold gas to settle into the disk of a spiral galaxy, just as gas settles into a disk around a forming star. In contrast, elliptical galaxies do not have a net rotation, so the gas does not settle into a disk. The only place in an elliptical galaxy where cold gas could collect is at the center. However, the density of stars in elliptical galaxies is so high that evolving stars and Type Ia supernovae continually reheat this gas, preventing most of it from cooling off and forming cold interstellar clouds.

Color The colors of spiral and elliptical galaxies reveal a great deal about their star formation histories. Recall from Chapter 15 that stars form from dense clouds of cold molecular gas. Because the gas seen in elliptical galaxies is very hot, astronomers know that active star formation is not taking place in those galaxies today. The reddish colors of elliptical and S0 galaxies confirm that little or no star formation has occurred there for quite some time. The stars in these galaxies are an older population of lower-mass stars. In contrast, the bluish colors of the disks of spiral galaxies indicate that massive, young, hot stars are forming in the cold molecular clouds contained within the disk. Even though most of the stars in a spiral disk are old, the massive, young stars are so luminous that their blue light dominates. Star formation in most irregular galaxies is like that in spiral galaxies. Some irregular galaxies form stars at prodigious rates, given their relatively small sizes. Irregular and disk galaxies that undergo intense bursts of star formation are called starburst galaxies.

Luminosity The relationship between luminosity and radius among the different types of galaxies is not straightforward. Galaxies range in luminosity from tens of thousands up to a trillion (a million million) solar luminosities (10^4 to $10^{12}\ L_{Sun}$) and in size from a few hundred to hundreds of thousands of parsecs. There is no distinct size difference between elliptical and spiral galaxies: about half of both types of galaxies fall within a similar range of sizes. Although it is true that the most luminous elliptical galaxies are more luminous than the most luminous spiral galaxies, there is considerable overlap in the range of luminosities among all Hubble types.

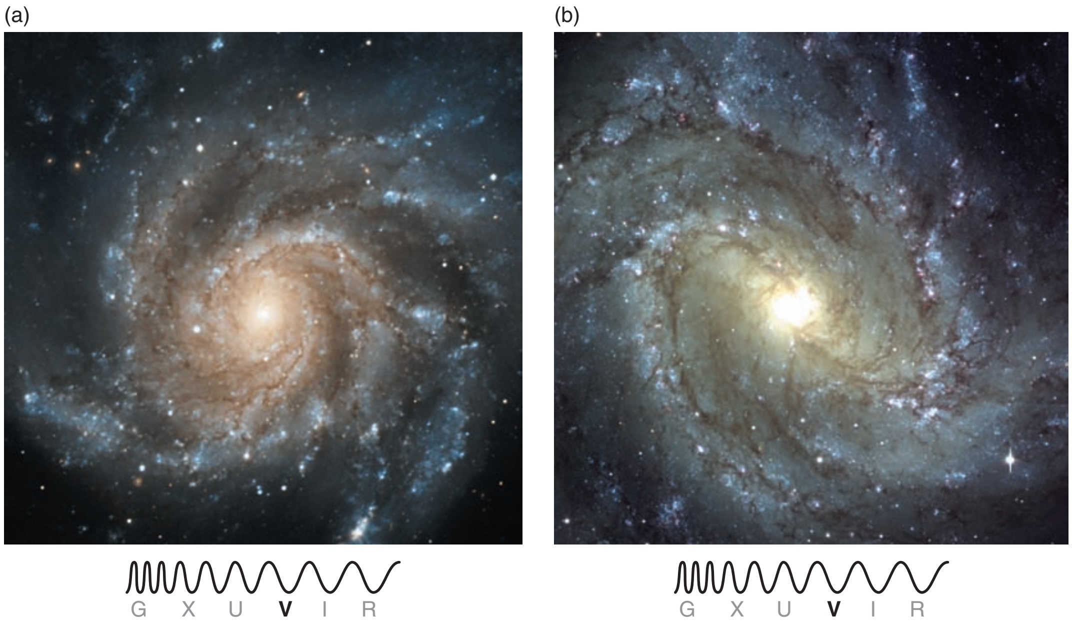


Figure 19.7 The mass or size of a spiral galaxy does not determine its appearance. Even though these galaxies appear to be similar in size and luminosity, the larger galaxy (a) is 4 times more distant and 10 times more luminous than the smaller galaxy (b).

Mass Mass is the single most important parameter in determining the properties and evolution of a star. In contrast, differences in mass and size do not lead to obvious differences among galaxies. Only subtle differences in color and concentration exist between large and small galaxies, making it difficult to distinguish which are large and which are small. Even when a larger, more distant spiral galaxy (**Figure 19.7a**) is seen next to a smaller, nearby spiral galaxy (Figure 19.7b), it can be hard to tell which is which by appearance alone. Galaxies that have relatively low luminosity (less than 1 billion L_{Sun}) are called dwarf galaxies, and those that are more than 1 billion L_{Sun} are called giant galaxies, because the luminosity indicates the number of stars and therefore the total amount of stellar mass. Only elliptical and irregular galaxies come in both types: among spiral and S0 galaxies, there are only giants. It is relatively easy to tell the difference between a dwarf elliptical galaxy and a giant elliptical galaxy (**Figure 19.8**). Giant elliptical galaxies have a much higher density of stars, and these are more centrally concentrated than stars in dwarf ellipticals.

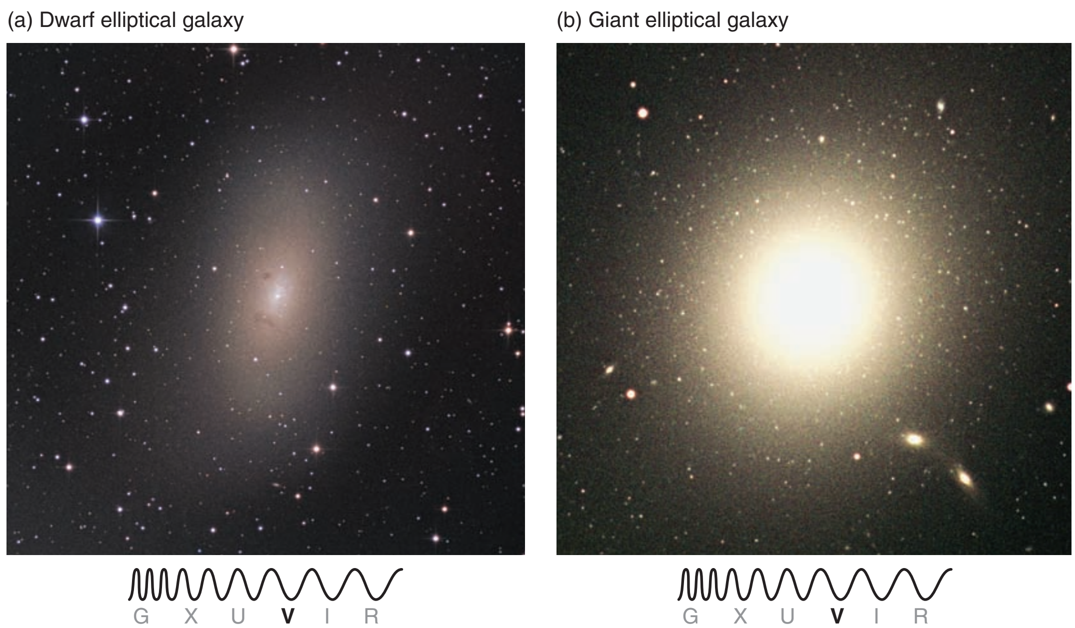


Figure 19.8 Dwarf elliptical galaxies (a) differ in appearance from giant elliptical galaxies (b). Stars in giant elliptical galaxies are more centrally concentrated than those in dwarf elliptical galaxies.

CHECK YOUR UNDERSTANDING 19.2

Currently, star formation rates are highest in (a) elliptical galaxies; (b) S0 galaxies; (c) spiral galaxies; (d) all of the above

19.2 Astronomers Use Several Methods to Find Distances to Galaxies

To determine the distances to galaxies, astronomers started with the closest objects and looked for patterns that would help them to find distances to the farthest objects—just as they did for stars. Distances are measured in a series of different methods called the **distance ladder**, which relates distances on a variety of overlapping scales, each method building on the previous one. In this section, we will discuss these distance methods and how they led to the surprising discovery that the galaxies are moving away from us.

The Distance Ladder

Nebraska Simulation: Spectroscopic Parallax Simulator

The distance ladder is summarized in **Figure 19.9**. In the 18th and 19th centuries, astronomers used several methods to estimate the astronomical unit (AU; the average distance from Earth to the Sun), and thus the distances to the planets. (Since the 1960s, distances within the Solar System are found using radar and signals from space probes.) Once the value of the AU was known, astronomers used trigonometric parallax, as discussed in Chapter 13, to measure distances to nearby stars and thereby to build up the H-R diagram. For more distant stars, astronomers use the spectral and luminosity classification of a star to determine its position on the H-R diagram. That position provides a star's luminosity, which in turn enables astronomers to estimate its distance by comparing its apparent brightness with its luminosity using spectroscopic parallax (as described in Chapter 13 and Appendix 7).

Nebraska Simulation: Supernova Light Curve Fitting Explorer

Astronomers measure the distance to relatively nearby galaxies using *standard candles*, a term borrowed from an old unit of light intensity that was based on actual candles. **Standard candles** are objects that have a known luminosity,

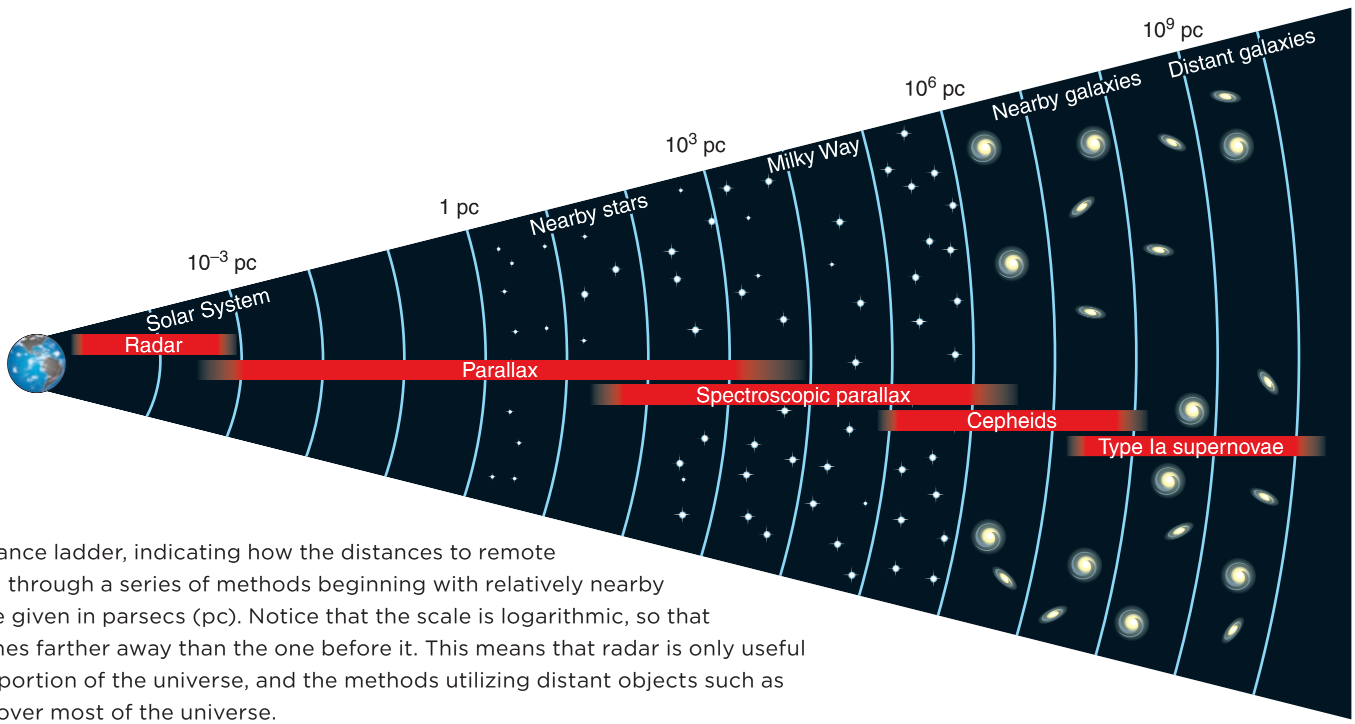


Figure 19.9 The distance ladder, indicating how the distances to remote objects are estimated through a series of methods beginning with relatively nearby objects. Distances are given in parsecs (pc). Notice that the scale is logarithmic, so that each blue arc is 10 times farther away than the one before it. This means that radar is only useful for a vanishingly tiny portion of the universe, and the methods utilizing distant objects such as Type Ia supernovae cover most of the universe.

usually because they have been observed within the Milky Way. The objects must also be bright enough to be recognizable in the distant galaxy. Astronomers assume that the luminosity of each object is the same as that for a similar object of that type in the Milky Way, and then they compare the luminosity and apparent brightness of the standard candle to find its distance.

Objects that can be used as standard candles include main-sequence O stars, globular clusters, planetary nebulae, novae, variable stars such as RR Lyraes and Cepheids, and supernovae. For example, Hubble Space Telescope (HST) observations of Cepheid variables enable astronomers to measure distances accurately to galaxies as far away as 30 million parsecs (also called 30 megaparsecs [Mpc]). Even more luminous than the Cepheids, and thus detectable at greater distances, are Type Ia supernovae.

Recall from Chapter 16 that Type Ia supernovae can occur when gas flows from an evolved star onto its white dwarf companion, pushing the white dwarf up toward the Chandrasekhar limit for the mass of an electron-degenerate object: 1.4 M_{Sun}. When this happens, the white dwarf burns carbon, collapses, and then explodes. At first, astronomers thought that these Type Ia supernovae all occurred in white dwarfs of just below 1.4 M_{Sun}. In that situation, all such explosions would occur at the same mass and have similar luminosity, with some calibration adjustment for the rate at which the brightness declines after it peaks. But now astronomers estimate that 80 percent of Type Ia supernova come from double-degenerate systems; for example, two white dwarfs that merge and then explode. The double white dwarf supernova could have up to twice the mass of a single white dwarf supernova. In addition, there may be other variations in Type Ia supernovae based on their color. To test whether they all have about the same luminosity, astronomers observe nearby Type Ia supernovae in galaxies with distances determined from other methods, such as by Cepheid variables. But more distant supernovae might have different luminosity. With a peak luminosity that can outshine a billion Suns (**Figure 19.10**), Type Ia supernovae can be seen and measured with modern telescopes at very large distances (**Working It Out 19.1**).

The motions of the stars in galaxies give rise to two "secondary" methods for estimating the distance to a galaxy. In rotating spiral galaxies, some of the light is approaching Earth and thus blueshifted, and some light is moving away from Earth and thus redshifted. These redshifts and blueshifts together broaden a spectral line in the galaxy, and the amount of broadening indicates how fast the galaxy is rotating. Astronomers can measure the broadening from the Doppler shifts of the 21-centimeter (21-cm) radio emission line of hydrogen, which tells them the speed of rotation, which then relates to the galaxy's mass by Newton's version of Kepler's third law. The more massive galaxies have more stars and are therefore more luminous. This empirical relation between the measured width of the 21-cm line and the luminosity of the spiral galaxy is called the Tully-Fisher relation. Once the luminosity of the galaxy is known, it can be compared to the galaxy's observed apparent brightness to estimate its distance. This method is thought to work out to about 100 Mpc.

Elliptical galaxies and the bulges of S0 galaxies do not rotate, so instead, astronomers look at the distribution of the surface brightness of a galaxy. Closer galaxies show more variations in the surface brightness because the distribution of stars throughout the galaxy isn't perfectly uniform. For more distant galaxies, these variations are less noticeable, and the surface brightness appears more uniform across the galaxy. This method is less precise than the Tully-Fisher method for spirals, but generally it also is thought to work out to about 100 Mpc.

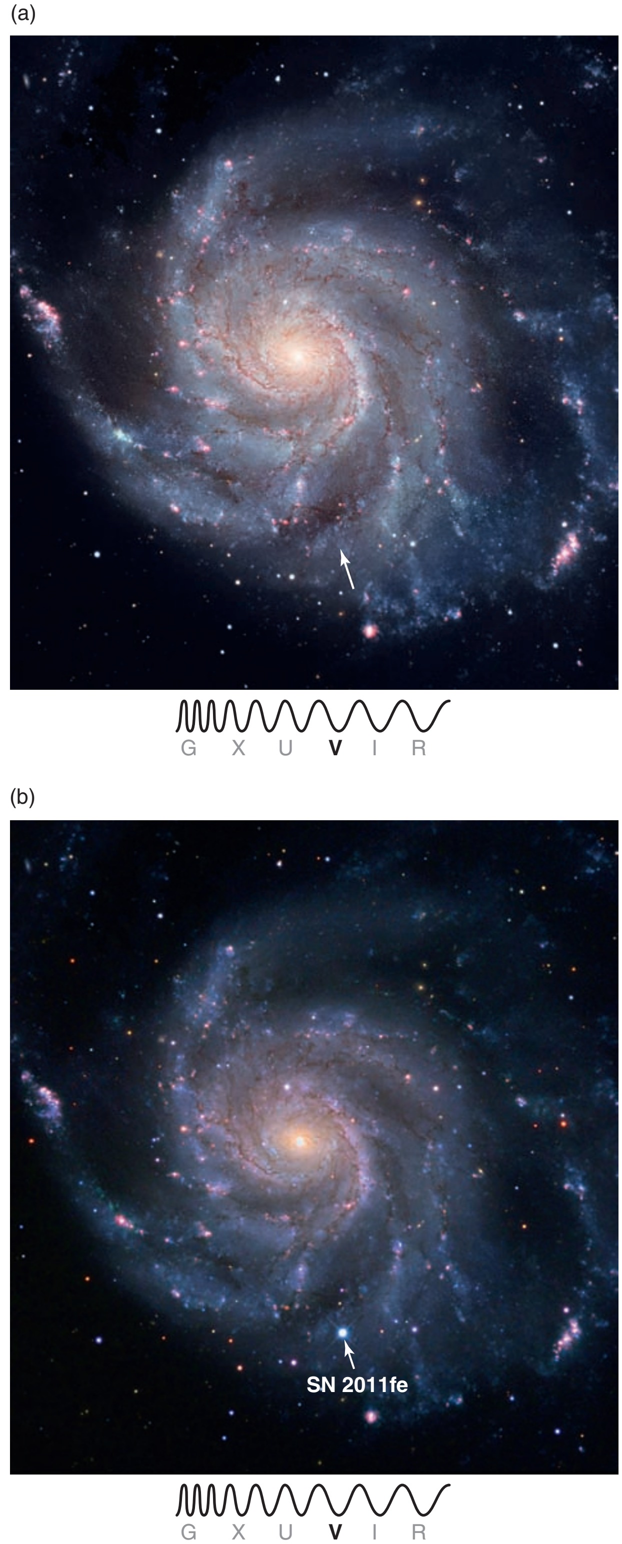


Figure 19.10 Type Ia supernovae are extremely luminous standard candles. The Pinwheel Galaxy is 6.4 Mpc away, shown before (a) and after (b) Supernova 2011fe.

19.1 Working It Out Finding the Distance from a Type Ia Supernova

Let's see how astronomers use a standard candle to estimate distance. Figure 19.10 shows the Pinwheel Galaxy (M101), with a supernova that was observed in 2011. Astronomers can compare the peak observed brightness of this supernova with the peak luminosity for this type of supernova to compute the distance.

In Section 13.1, we gave the equation relating brightness, luminosity, and distance:

$$\text{Brightness} = \frac{\text{Luminosity}}{4\pi d^2}$$

Rearranging to solve for distance gives:

$$d = \sqrt{\frac{\text{Luminosity}}{4\pi \times \text{Brightness}}}$$

The maximum observed brightness of this supernova is 7.5×10^{-12} watts per square meter (W/m^2). The graph in Figure 17.11 shows that the typical maximum luminosity L of a Type Ia supernova is 9.5×10^9 times the luminosity of the Sun:

$$L = 9.5 \times 10^9 \times L_{\text{Sun}}$$

$$L = 9.5 \times 10^9 \times (3.9 \times 10^{26}\,\text{W}) = 3.7 \times 10^{36}\,\text{W}$$

Thus, we can solve the equation:

$$d = \sqrt{\frac{3.7 \times 10^{36}\,\text{W}}{4\pi \times 7.5 \times 10^{-12}\,\text{W/m}^2}} = 2.0 \times 10^{23}\,\text{m}$$

To put this into megaparsecs:

$$d = \frac{2.0 \times 10^{23}\,\text{m}}{3.1 \times 10^{22}\,\text{m/Mpc}} = 6.4\,\text{Mpc}$$

The distance is 6.4 Mpc. This supernova was detected early, while it was in the earliest stages of the explosion. Because the Pinwheel Galaxy is relatively close, other standard candles can be observed in this galaxy to help calibrate the distance. Also note that 6.4 Mpc = 21 million light-years, so this supernova explosion took place 21 million years ago.

The Discovery of Hubble's Law

In the 1920s, Hubble and his coworkers were studying the properties of a large collection of galaxies. Another astronomer, Vesto Slipher (1875–1969), was obtaining spectra of these galaxies at Lowell Observatory in Flagstaff, Arizona. Slipher's galaxy spectra looked like the spectra of ensembles of stars with a bit of glowing interstellar gas mixed in. But he was surprised to find that the emission and absorption lines in the spectra of these galaxies were seldom seen at the same wavelengths as in the spectra of stars observed in the Milky Way Galaxy. The lines were almost always shifted to longer wavelengths as seen in **Figure 19.11**.

Nebraska Simulation: Galactic Redshift Simulator

Slipher characterized most of the observed shifts in galaxy spectra as redshifts because the light from these galaxies is shifted to longer, or redder, wavelengths. Hubble interpreted Slipher's redshifts as Doppler shifts, and he concluded that almost all of the galaxies in the universe are moving away from the Milky Way. Recall from Chapter 5 that objects with larger Doppler redshifts are moving away more quickly than those with smaller redshifts. When Hubble combined the measurements of galaxy velocities with his own estimates of the distances to these galaxies, he found that distant galaxies are moving away from Earth more rapidly than are nearby galaxies. Specifically, *the velocity at which a galaxy is moving away from an observer is proportional to the distance of that galaxy*. This relationship between distance and recession velocity has become known as **Hubble's law**.

AstroTour: Hubble's Law

Figure 19.12 plots the measured recession velocities of galaxies against their measured distances. Because the velocity and distance are proportional to each other, the points lie along a line on the graph with a slope equal to the proportionality constant H_0, called the **Hubble constant**. Notice how well the data line up along the line. This strong correlation indicates that the universe follows Hubble's law; for example, a galaxy at a distance of 30 Mpc from Earth moves away twice as fast as a galaxy that is 15 Mpc distant. The original value for the

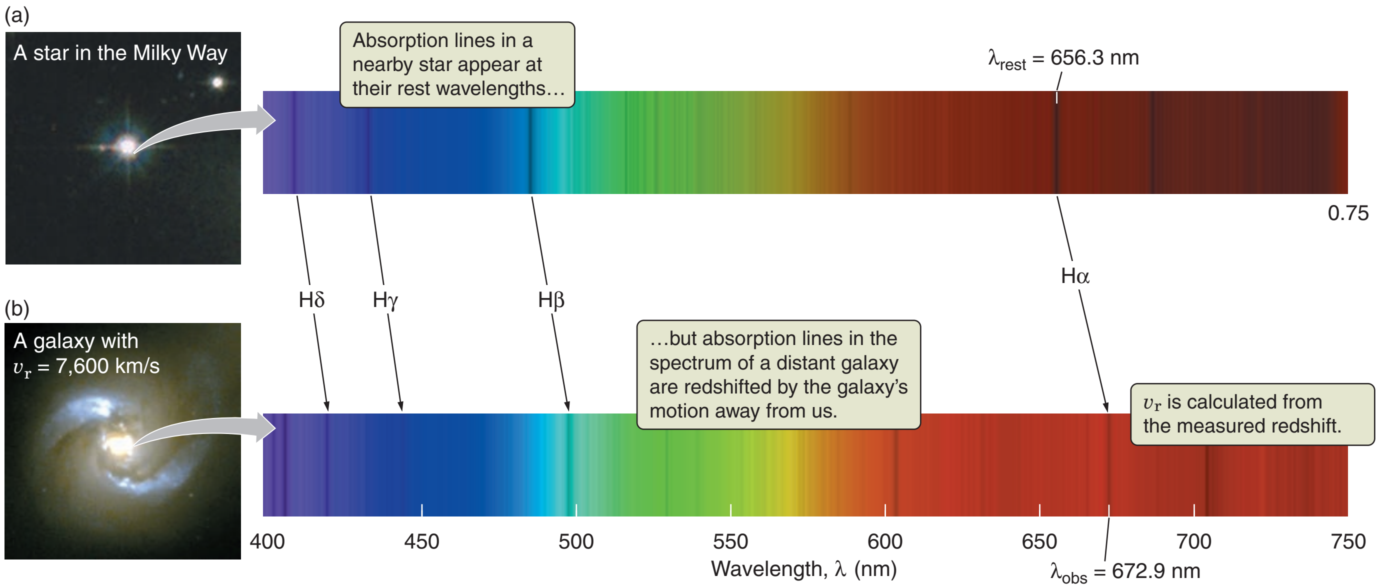


Figure 19.11 (a) The spectrum of a star in our galaxy shows absorption lines, which in this case lie at the rest wavelength. (b) A distant galaxy, shown with its spectrum at the same scale as that of the star, has lines that are redshifted to longer wavelengths. v_r is recession velocity, or radial velocity.

Hubble constant was 8 times too large, which led to inconsistencies, but the problem was resolved when astronomers realized there that there are two types of Cepheids with slightly different period-luminosity relationships. Today, astronomers have measured the Hubble constant to an accuracy of a few percent by several different methods, using observations from the Wilkinson Microwave Anisotropy Probe (*WMAP*), HST, Spitzer Space Telescope, and the Planck space observatory. The value is likely to be further refined in the years to come. In this text, we use a value of 70 km/s/Mpc as an approximation to the best current measured values of 67 to 74 km/s/Mpc, with uncertainties of 1–3 km/s/Mpc depending on the measurement.

Using this constant, Hubble's law is written as $v_r = H_0 \times d_G$, where d_G is the distance to the galaxy, and v_r is the galaxy's recession velocity (**Working It Out 19.2**). Hubble's law gives astronomers a practical tool for measuring distances to remote galaxies. Once they know the value of H_0, they can use a straightforward measurement of the redshift of a galaxy to find its distance. In other words, once H_0 is known, Hubble's law makes the once-difficult task of measuring distances in the universe relatively easy, providing astronomers with a tool to map the structure of the observable universe. We will return to Hubble's law and its implications for understanding the universe as a whole in Chapter 21.

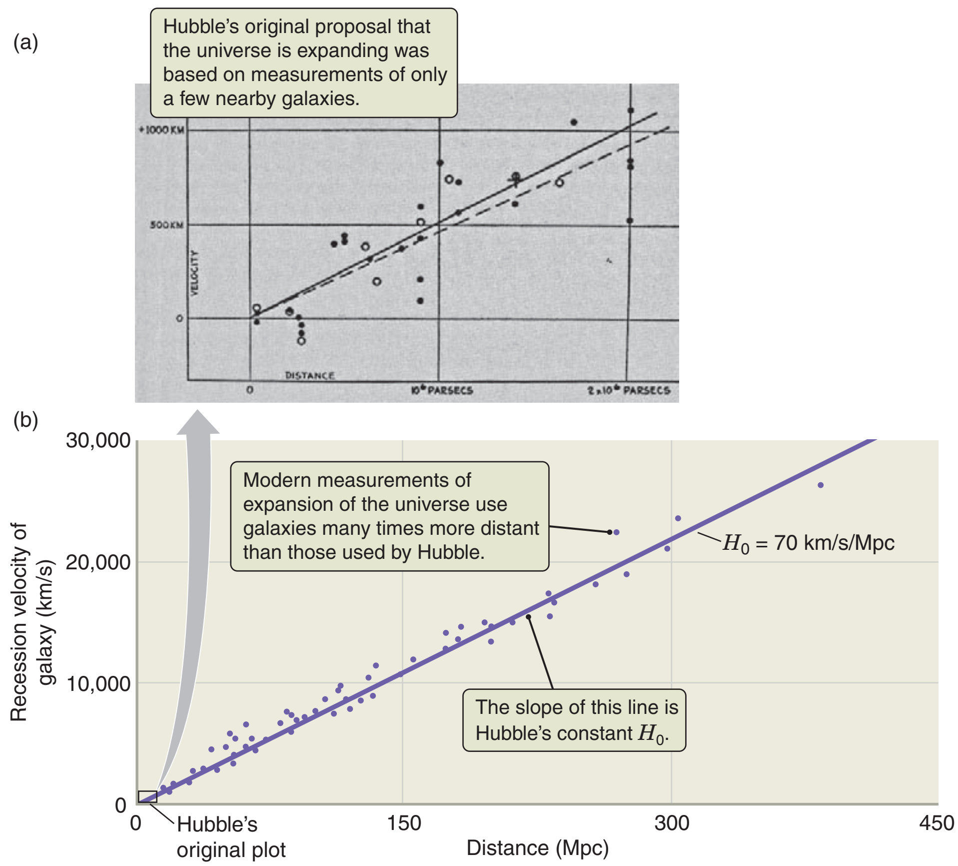


Figure 19.12 (a) Hubble's original graph shows that more distant galaxies are receding faster than less distant galaxies. (b) Modern data on galaxies many times farther away than those studied by Hubble show that recession velocity is proportional to distance.

19.2 Working It Out Redshift—Calculating the Recession Velocity and Distance of Galaxies

Recall from Working It Out 5.2 that the Doppler equation for spectral lines showed that

$$v_r = \frac{\lambda_{obs} - \lambda_{rest}}{\lambda_{rest}} \times c$$

The fraction in front of the c is equal to z, the redshift. Substituting for the fraction, we get

$$v_r = z \times c$$

(Note: This correspondence requires a correction as velocities approach the speed of light.)

Because spectral lines from distant galaxies have wavelengths shifted to the red, the galaxies must be moving away from Earth. Suppose astronomers observe a spectral line with a rest wavelength of 373 nanometers (nm) in the spectrum of a distant galaxy. If the observed wavelength of the spectral line is 379 nm, then its redshift (z) is

$$z = \frac{\lambda_{obs} - \lambda_{rest}}{\lambda_{rest}}$$

$$z = \frac{379\,\text{nm} - 373\,\text{nm}}{373\,\text{nm}} = 0.0161$$

Note that the value of the redshift of a galaxy is independent of the wavelength of the line used to measure it: the same result would have been calculated if a different line had been observed.

We can now calculate the recession velocity from this redshift as follows:

$$v_r = z \times c = 0.0161 \times 300{,}000\ \text{km/s} = 4{,}830\ \text{km/s}$$

How far away is the distant galaxy? This is where *Hubble's law* and the *Hubble constant* apply. Hubble's law relates a galaxy's recession velocity to its distance as

$$v_r = H_0 \times d_G$$

where d_G is the distance to a galaxy measured in megaparsecs. Dividing through by $H_0 = 70$ km/s/Mpc yields

$$d_G = \frac{v_r}{H_0} = \frac{4{,}830\,\text{km/s}}{70\,\text{km/s/Mpc}} = 69\,\text{Mpc}$$

From a measurement of the wavelength of a spectral line, we see that the distant galaxy is approximately 69 Mpc away.

CHECK YOUR UNDERSTANDING 19.3

Hubble's law was discovered using measurements of two properties of a galaxy: __________ and __________. (a) size; mass (b) distance; rotation speed (c) distance; recession velocity (d) size; recession velocity

19.3 Galaxies Are Mostly Dark Matter

Efforts to measure the masses of galaxies during the 20th century led to the discovery of dark matter—mass that does not interact with light and cannot be detected via the light they emit. To understand this discovery, we first need to understand how astronomers go about measuring the mass of a galaxy, and then see how they concluded that much of the mass in a galaxy is dark matter.

Finding the Mass of a Galaxy

▶II **AstroTour:** Dark Matter

To measure the mass of a galaxy, astronomers add up the mass of the stars, dust, and gas that they observe. Because a galaxy's spectrum is composed primarily of starlight, once astronomers know what types of stars are in the galaxy, they can use what is known about stellar evolution to estimate the total stellar mass from the galaxy's luminosity. Astronomers then estimate the mass of the dust and gas by using the physics of radiation from interstellar gas at X-ray, infrared, and radio wavelengths. Together, the stars, gas, and dust in a galaxy are called **luminous**

matter (or simply **normal matter**) because this matter emits or scatters electromagnetic radiation.

There is an independent method for determining mass that does not involve luminosity: the effect of gravity on an object's motion can be used to determine its mass. Stars in disks follow orbits that are much like the Keplerian orbits of planets around their parent stars and binary stars around each other (see Working It Out 13.4). To measure the mass of a spiral galaxy, astronomers apply Kepler's laws, just as they do for those other systems.

Detections of Dark Matter

Astronomers originally hypothesized that the mass and the light in a galaxy are distributed in the same way; that is, they assumed that all the mass in a galaxy is luminous matter. They observed that the light of all galaxies, including spiral galaxies, is highly concentrated toward the center (**Figure 19.13a**). On the basis of the observed location of light, astronomers predicted that nearly all mass in a spiral galaxy is concentrated toward its center (Figure 19.13b). This situation is much like the Solar System, where nearly all the mass is at the Solar System's center—in the Sun. Therefore, they predicted faster orbital velocities near the center of the spiral galaxy and slower orbital velocities farther out (Figure 19.13c).

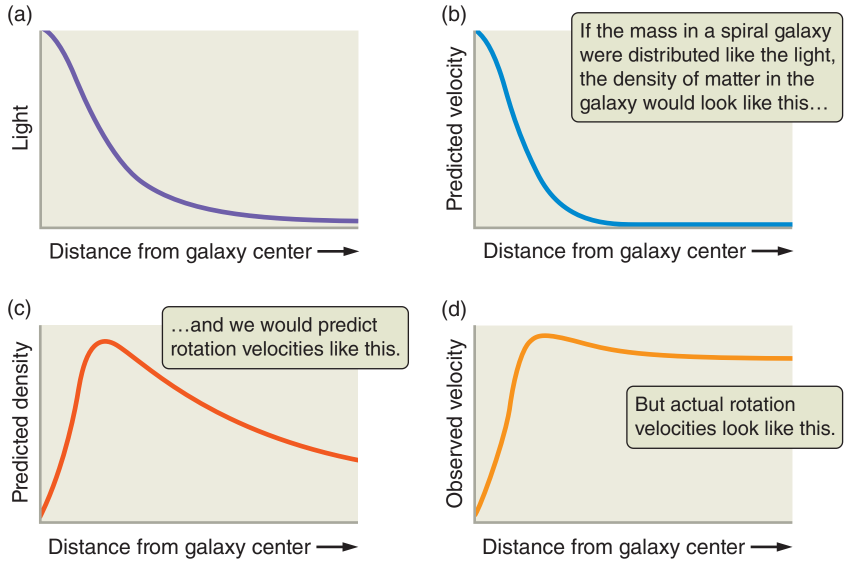


Figure 19.13 (a) The profile of visible light in a typical spiral galaxy drops off with distance from the center. (b) The predicted mass density of stars and gas located at a given distance from the galaxy's center follows the light profile. If stars and gas accounted for all of the mass of the galaxy, then the galaxy's rotation curve would be as shown in (c). However, observed galaxy rotation curves look more like the curve shown in (d).

To test this prediction, astronomers used the Doppler effect to measure orbital motions of stars, gas, and dust at various distances from a galaxy's center. The velocities of stars are obtained from observations of absorption lines in their spectra. The velocities of interstellar gas are obtained using emission lines such as those produced by hydrogen alpha (Hα) emission or 21-cm emission from neutral hydrogen (see Chapter 15). Once the velocities have been found, astronomers create a graph—called a **rotation curve**—that shows how orbital velocity in a galaxy varies with distance from the galaxy's center. The rotation curve of a spiral galaxy enables astronomers to determine directly how the mass in that galaxy is distributed by applying Kepler's laws to the rotation curves.

Vera Rubin pioneered work on galaxy rotation rates in the 1970s. She discovered that, contrary to earlier prediction (Figure 19.13c), the rotation velocities of spiral galaxies remain about the same out to the most distant measured parts of the galaxies (Figure 19.13d). Observations of 21-cm radiation from neutral hydrogen show that the rotation curves appear level, or "flat," in their outer parts even well outside the extent of the visible disks. These observations indicated that mass and light are distributed differently.

What mass distribution would cause this unexpected rotation curve? Recall from Chapter 4 that only the mass inside a given radius contributes to the net gravitational force felt by an orbiting object. From the rotation velocity, you can calculate the mass within the orbit of the object. **Figure 19.14** shows the result of such a calculation for the spiral galaxy NGC 3198. The black line shows the speed of rotation of this galaxy at a particular radius. The red line shows how much luminous mass is inside that radius. To produce a rotation curve like the one shown in black, this galaxy must have a second component consisting of matter that does not show up in the census of stars, gas, and dust. This material, which does not interact with light, and therefore reveals itself only by the influence of its gravity, is called

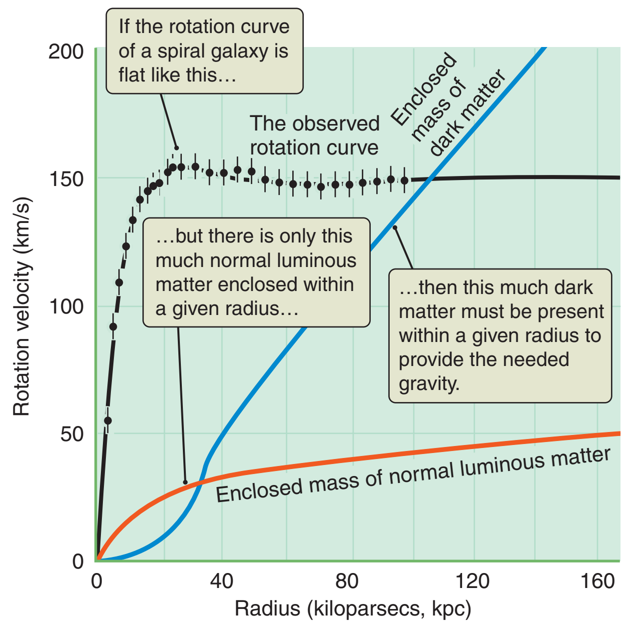


Figure 19.14 The flat rotation curve of the spiral galaxy NGC 3198 can be used to determine the total mass within a given radius. Notice that the normal mass that can be accounted for by stars and gas provides only part of the needed gravity. Extra dark matter is needed to explain the rotation curve. (Note: 1 kiloparsec [kpc] = 1,000 parsecs.)

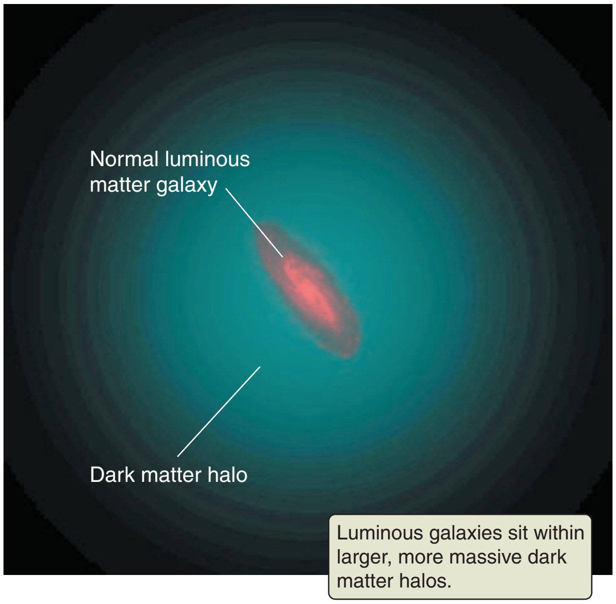


Figure 19.15 In addition to the matter that is visible, galaxies are surrounded by halos containing a large amount of dark matter.

dark matter. The blue line shows how much dark matter must be inside a particular radius in order to provide enough mass to make the galaxy rotate as it does.

The rotation curves of the inner parts of spiral galaxies match predictions based on their luminous matter, indicating that the inner parts of spiral galaxies are mostly luminous matter. Within the entire *visual* image of a galaxy, the mix of dark and luminous matter is about half and half. However, rotation curves measured with 21-cm radiation from neutral hydrogen indicate that the outer parts of spiral galaxies are mostly dark matter. Astronomers currently estimate that as much as 95 percent of the total mass in some spiral galaxies consists of a **dark matter halo**, shown in **Figure 19.15**, which can extend up to 10 times farther than the visible spiral portion of the galaxy located at the galaxy's center. This is a startling statement. The luminous part of a spiral galaxy is only a part of a much larger distribution of mass that is dominated by some type of invisible dark matter.

What about elliptical galaxies? Again, astronomers need to compare the luminous mass measured from the light they can see with the gravitational mass measured from the effects of gravity. Because elliptical galaxies do not rotate, astronomers cannot use Kepler's laws to measure the gravitational mass from the rotation of a disk. Instead, they noticed that an elliptical galaxy's ability to hold on to its hot, X-ray-emitting gas depends on its mass. If the galaxy is not massive enough, the hot atoms and molecules will escape into intergalactic space. To find the mass of an elliptical galaxy, astronomers first infer the total amount of gas from X-ray images, such as the (false-color) blue and purple halo seen in **Figure 19.16**. Then they calculate the mass that is needed to hold on to the gas and compare that gravitational mass with the luminous mass. The amount of dark matter is the difference between what is needed to hold on to the inferred amount of gas and the observed amount of luminous matter.

Some elliptical galaxies contain up to 20 times as much mass as can be accounted for by their stars and gas alone, so they must be dominated by dark matter, just like spiral galaxies. As with spirals, the luminous matter in ellipticals is more centrally concentrated than the dark matter. The transition from the inner parts of galaxies (where luminous matter dominates) to the outer parts (where dark matter dominates) is remarkably smooth. Some galaxies may contain less dark matter than others, but about 90–95 percent of the total mass in a typical galaxy is in the form of dark matter. The high percentage of dark matter distinguishes smaller dwarf galaxies from globular clusters, which do not have dark matter. This difference is an important observation that will need to be explained in the context of the evolution of galaxies.

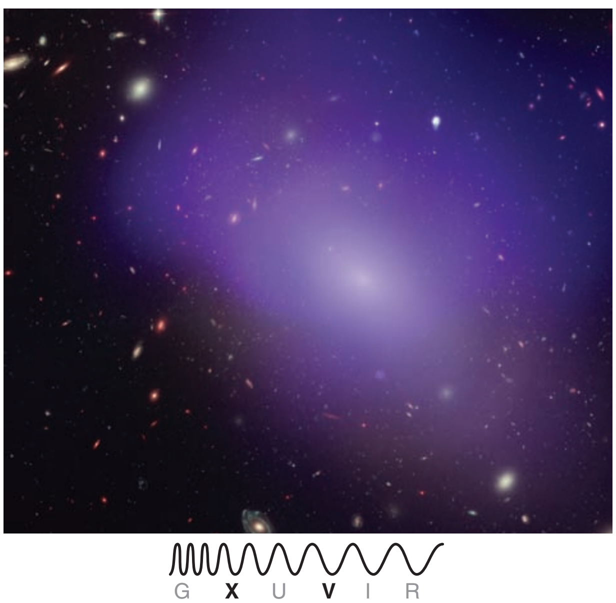


Figure 19.16 In this combined visible-light and X-ray image of elliptical galaxy NGC 1132, the false-color blue and purple halo is X-ray emission from hot gas surrounding the galaxy. The hot gas extends well beyond the visible light from stars.

The Composition of Dark Matter?

What is the dark matter that makes up most of a galaxy? A number of suggestions are under investigation. Some candidates are objects such as large planets, compact stars, black holes, and exotic unknown elementary particles. These candidates have been lumped into two groups: MaCHOs and WIMPs.

Dark matter candidates such as small main-sequence M stars, Jupiter-sized planets, white dwarfs, neutron stars, or black holes are collectively referred to as **MaCHOs**, which stands for *massive compact halo objects*. If the dark matter in a galaxy's halo consists of MaCHOs, there must be a lot of these objects, and they must each exert gravitational force but not emit much light. Because they have mass, MaCHOs gravitationally deflect light according to Einstein's general theory of relativity—a phenomenon called gravitational lensing (see Chapter 18). If astronomers were observing a distant star and a MaCHO passed between Earth and the star, the star's light would be deflected and, if the geometry were just right, focused by the intervening MaCHO as it passed across their line of sight. Because gravity affects all wavelengths equally, such lensing events should look the same in all colors, ruling out other causes of variability.

Astronomers would be remarkably lucky if such an event occurred just as they were observing a single distant star. When they monitored the stars in two small companion galaxies of the Milky Way, observing tens of millions of stars for several years, they saw some lensing events but not nearly enough to account for the amount of dark matter in the halo of our galaxy. The implication of this result is that dark matter in the Milky Way (and therefore other galaxies too) must be composed of something other than MaCHOs.

The other dark matter candidates are the exotic unknown elementary particles commonly known as **WIMPs**, which stands for *weakly interacting massive particles*. WIMPs are predicted to be similar to neutrinos in that they would barely interact with ordinary matter, yet would be more massive and would move more slowly. WIMPs are currently the favored candidates for dark matter because there are not enough MaCHOs to account for the observed effects. Experiments are under way at the Large Hadron Collider and on the International Space Station to detect the existence of such particles, and additional experiments are being done to detect such particles from the dark matter halo of the Milky Way as they pass through Earth.

CHECK YOUR UNDERSTANDING 19.4

Astronomers detect dark matter: (a) by comparing luminous mass to gravitational mass; (b) because it blocks background light; (c) because more distant galaxies move away faster; (d) because it emits lots of X-rays.

19.4 Most Galaxies Have a Supermassive Black Hole at the Center

Studying the centers of galaxies is difficult because there are so many stars and so much dust and gas in the way that astronomers cannot get a clear picture of the center, even for nearby galaxies. Instead, it was observations of the most distant objects in the universe that provided the clues to understanding what lies in the centers of massive galaxies.

The Discovery of Quasars

In the late 1950s, radio surveys detected a number of bright, compact objects that at first seemed to have no optical counterparts. Improved radio positions revealed that the radio sources coincided with faint, very blue, starlike objects. Unaware of the true nature of these objects, astronomers called them "radio stars." Obtaining spectra of the first two radio stars was a laborious task, requiring 10-hour exposures. Astronomers were greatly puzzled by the results because these spectra did not display the expected absorption lines characteristic of blue stars. Instead, the spectra showed only a single pair of emission lines that were broad—indicating very rapid motions within these objects—and that did not seem to correspond to the lines of any known substances.

For several years, astronomers believed they had discovered a new type of star, until astronomer Maarten Schmidt realized that these broad spectral lines, shown in **Figure 19.17**, were the highly redshifted lines of ordinary hydrogen.

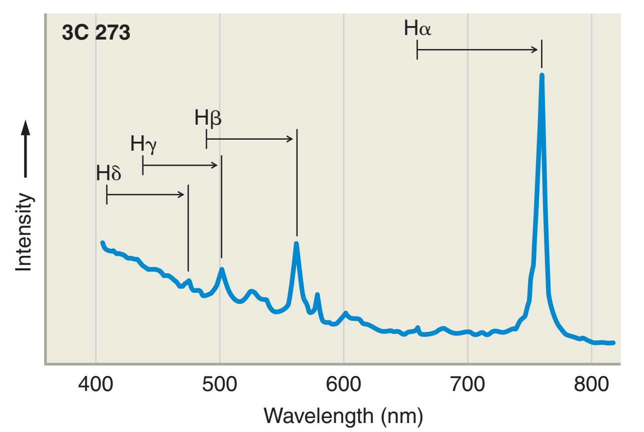


Figure 19.17 This figure shows the spectrum of quasar 3C 273, one of the closest and most luminous known quasars. The emission lines are redshifted by $z = 0.16$, from marked rest wavelengths, indicating that the quasar is at a distance of about 750 Mpc.

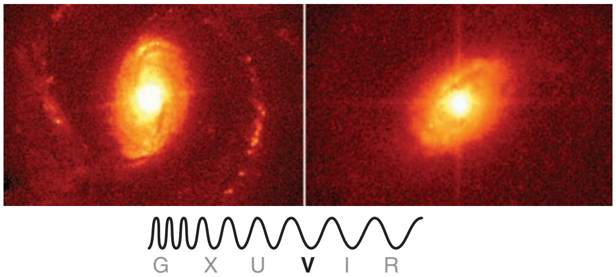


Figure 19.18 These HST images show quasars embedded in the centers of galaxies.

The implications were surprising: these "stars" were not stars. They were extraordinarily luminous objects at enormous distances. These "quasi-stellar radio sources" were named **quasars**. Other quasars were soon found by the same techniques. As still more were found, astronomers began cataloging them.

Quasars are phenomenally powerful, shining with the luminosity of a trillion to a thousand trillion (10^{12} to 10^{15}) Suns. They are also very distant from Earth: hundreds or thousands of megaparsecs. Billions of galaxies are closer to Earth than is the nearest quasar. Recall that the distance to an object also indicates the amount of time that has passed since the light from that object left its source. The fact that quasars are seen only at great distances implies that they are quite rare in the universe at this time but were once much more common. The discovery that quasars existed in the distant and therefore earlier universe provided one of the first pieces of evidence demonstrating that the universe has evolved over time.

Quasars are now recognized as the result of the most extreme form of activity that can occur in the nuclei of galaxies (**Figure 19.18**), often resulting from interactions with other galaxies. Quasars are a type of **active galactic nuclei**, or simply **AGNs**. The distinct types of active nuclei are identified from the spectrum of the galaxy. A "normal" galaxy has an absorption spectrum that is a composite of the light from its billions of stars. A galaxy with an AGN exhibits emission lines in addition to the stellar absorption spectrum. Active galactic nuclei are identified by the emission lines in their spectra, which distinguishes them from normal galaxies, which mostly show only absorption lines.

AGNs come in several types and can occur in spiral or elliptical galaxies. **Seyfert galaxies**, named after Carl Seyfert (1911–1960), who discovered them in 1943, are spiral galaxies whose centers contain AGNs. The luminosity of a typical Seyfert nucleus can be 10 billion to 100 billion L_{Sun}, comparable to the luminosity of the rest of the host galaxy as a whole. Similarly, **radio galaxies** are elliptical galaxies whose centers contain AGNs; their emission is usually most prominent in radio wavelengths. Radio galaxies and the more distant and luminous quasars are often the sources of slender jets that extend outward millions of light-years from the galaxy, powering twin lobes of radio emission (**Figure 19.19**).

Much of the light from AGNs is synchrotron radiation. This is the same type of radiation that comes from extreme environments such as the Crab Nebula supernova remnant. Synchrotron radiation comes from relativistic charged particles spiraling around the direction of a magnetic field. The fact that AGNs accelerate large amounts of material to nearly the speed of light indicates that they are very violent objects. In addition to the continuous spectrum of synchrotron emission, the spectra of many quasars and Seyfert nuclei also show emission lines that are smeared out by the Doppler effect across a wide range of wavelengths. This observation implies that gas in AGNs is swirling around the centers of these galaxies at speeds of thousands or even tens of thousands of kilometers per second.

Figure 19.19 Radio galaxy Centaurus A is the closest AGN to Earth, at a distance of 3.4 Mpc. In this composite image, the visible-light image shows the galaxy, the X-ray image (pink) shows the hot gas and an energetic jet blasting from the AGN, and the radio image (purple) shows the jets and lobes.

AGNs Are the Size of the Solar System

The enormous radiated power and mechanical energy of active galactic nuclei are made even more spectacular by the fact that all of this power emerges from a region that can be no larger than a light-day or so across—comparable in size to the Solar System. Although the HST and large, ground-based telescopes show faint fuzz—light from the surrounding galaxy—around the images of some quasars and other AGNs, the objects themselves remain as unresolved points of light.

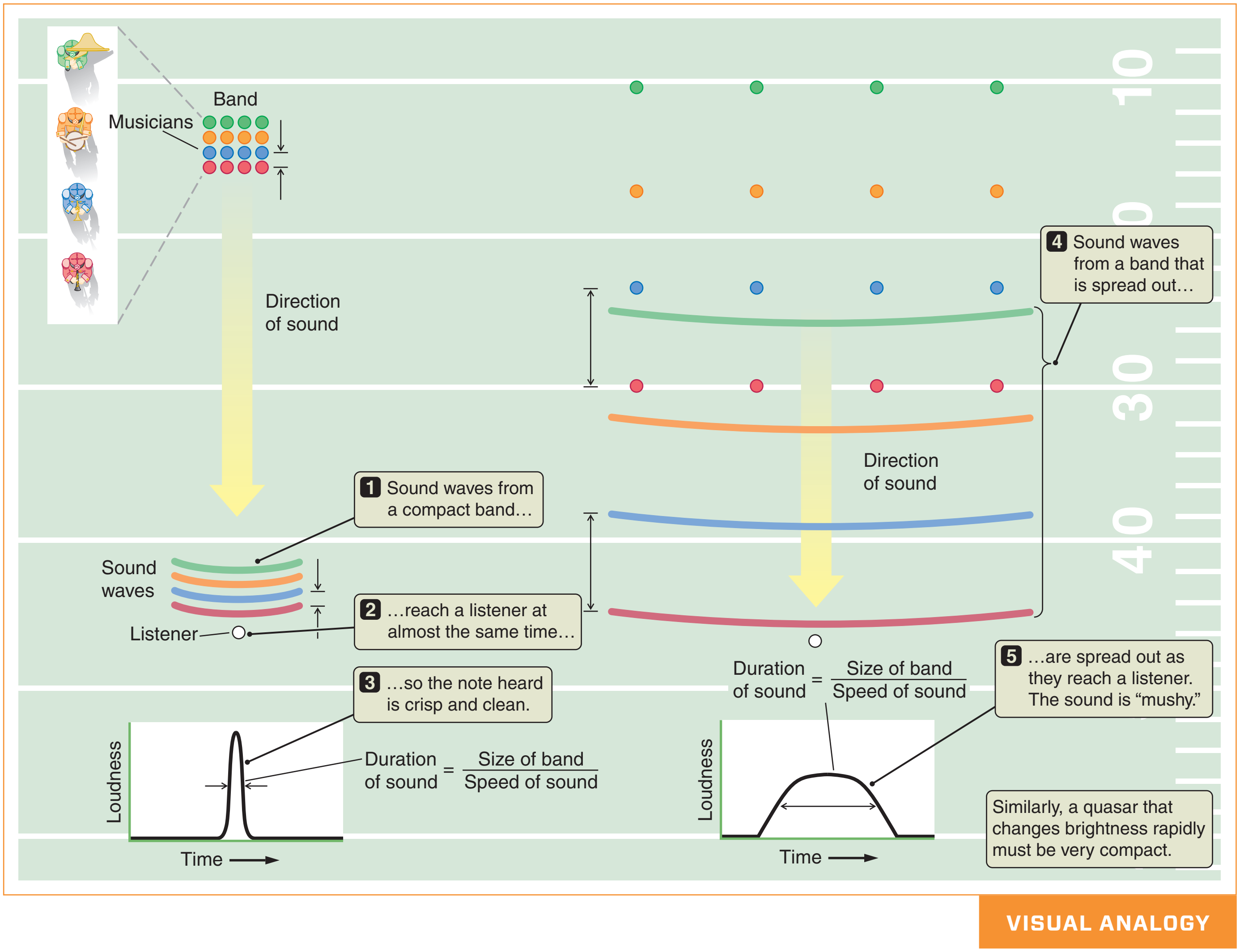


Figure 19.20 A marching band spread out across a field cannot play a clean note. Similarly, AGNs must be very compact to explain their rapid variability.

To understand why astronomers think that AGNs are compact objects, think about the halftime show at a local football game. **Figure 19.20** illustrates a problem faced by every director of a marching band. When a band is all together in a tight formation at the center of the field, the notes you hear in the stands are clear and crisp; the band plays together beautifully. But as the band spreads out across the field, its sound begins to get mushy. This is not because the marchers are poor musicians. Instead, it is because sound travels at a finite speed. On a cold, dry December day, sound travels at a speed of about 330 meters per second (m/s). At this speed, it takes sound approximately $\frac{1}{3}$ of a second to travel from one end of the football field to the other. Even if every musician on the field plays a note at exactly the same instant in response to the director’s cue, in the stands you hear the instruments close to you first but have to wait longer for the sound from the far end of the field to arrive.

If the band is spread from one end of the field to the other, then the beginning of a note will be smeared out over about $\frac{1}{3}$ of a second, which is the time it takes for sound to travel from one end of the field to the other. If the band were spread out

▶II **AstroTour:** Active Galactic Nuclei over two football fields, it would take about $\frac{2}{3}$ of a second for the sound from the most distant musicians to arrive at your ear. If the marching band were spread out over a kilometer, then it would take roughly 3 seconds—the time it takes sound to travel a kilometer—for you to hear a crisply played note start and stop. Even with your eyes closed, it would be easy to tell whether the band was in a tight group or spread out across the field.

Astronomy in Action: Size of Active Galactic Nuclei

Exactly the same principle applies to the light observed from active galactic nuclei. Quasars and other AGNs change their brightness dramatically over the course of only a day or two—and in some cases as quickly as in a few hours. This rapid variability sets an upper limit on the size of the AGN, just as hearing clear music from a marching band indicates that the band musicians are close together. The AGN powerhouse must therefore be no more than a light-day or so across because if it were larger, the light astronomers see could not possibly change in a day or two. An AGN has the light of up to 10,000 galaxies pouring out of a region of space that would come close to fitting within the orbit of Neptune.

Supermassive Black Holes and Accretion Disks

The conclusion that AGNs are small compared to their entire host galaxy and have incredible energy outputs had to be explained. In thinking about what type of object could be very small yet very energetic, astronomers hypothesized that galaxies with an AGN contain **supermassive black holes**—black holes with masses from thousands to tens of billions of solar masses. Violent accretion disks surround these supermassive black holes. As matter in the disk falls inward, gravitational energy is converted into heat, providing the luminous energy of the AGN. Recall that you have already encountered accretion disks several times in this book. Accretion disks surround young stars, providing the raw material for planetary systems. Accretion disks around white dwarfs, fueled by material torn from their bloated evolving companions, lead to novae and some Type Ia supernovae. Accretion disks around neutron stars and stellar-mass black holes a few kilometers across are seen as X-ray binary stars. Now take the neutron star or stellar black hole examples and scale them up to a black hole with a mass of a billion solar masses and a radius comparable in size to the orbit of Neptune. To attain such a high luminosity, an AGN has an accretion disk that is fed by several solar masses every year rather than by the small amounts of material siphoned off a star (**Working It Out 19.3**).

In our discussion of star formation in Chapter 15, we showed that gravitational energy is converted to thermal energy as material moves inward toward the growing protostar. Here, as material moves inward toward the supermassive black hole, conversion of gravitational energy heats the accretion disk to hundreds of thousands of kelvins, causing it to glow brightly in visible, ultraviolet, and X-ray light. Conversion of gravitational energy to thermal energy as material falls onto the accretion disk is also a source of energetic emission. As much as 20 percent of the mass of infalling material around a supermassive black hole is converted to luminous energy. The rest of that mass is pulled into the black hole itself, causing it to grow even more massive.

The interaction of the accretion disk with the black hole creates powerful radio jets that emerge perpendicular to the disk (as in the jet in the upper left of Figure 19.19). Throughout, twisted magnetic fields accelerate charged particles such as electrons and protons to relativistic speeds, producing synchrotron emission. Gas in the accretion disk or in nearby clouds orbiting the central black hole

19.3 Working It Out The Size, Density, and Power of a Supermassive Black Hole

Size. What are the sizes of supermassive black holes? Recall from Chapter 18 our discussion of the Schwarzschild radius, where you saw that stellar-mass black holes are kilometers in size. The formula for the Schwarzschild radius is given by

$$R_S = \frac{2GM_{BH}}{c^2}$$

where G is the gravitational constant, and c is the speed of light.

The largest supermassive black holes observed have about 10 billion solar masses (M_{Sun}). For example, the black hole at the center of the galaxy M87 is 6.6 billion M_{Sun}. To compute its size, recall that $M_{Sun} = 1.99 \times 10^{30}$ kilograms (kg), $c = 3 \times 10^5$ km/s, and $G = 6.67 \times 10^{-20}$ km^3/(kg s^2). Then, a 6.6-billion-M_{Sun} black hole has a Schwarzschild radius of

$$R_S = \frac{2 \times [6.67 \times 10^{-20}\,\text{km}^3/(\text{kg}\,\text{s}^2)] \times (6.6 \times 10^9 \times 1.99 \times 10^{30}\,\text{kg})}{(3 \times 10^5\,\text{km/s})^2}$$

$$R_S = 2.0 \times 10^{10}\,\text{km}$$

We can convert this value into astronomical units. Recall that 1 astronomical unit (AU) = 1.5×10^8 km. Therefore, this supermassive black hole has a radius of 130 AU—somewhat larger than the Solar System. We know that light takes $8\frac{1}{3}$ minutes to reach Earth from the Sun at a distance of 1 AU, so this 130 AU corresponds to a distance of 1,080 light-minutes, or 18 light-hours.

Density. What is the average density of this object inside of the event horizon? The mass of the black hole divided by the volume within the Schwarzschild radius is

$$\text{Density} = \frac{\text{Mass}}{\text{Volume}} = \frac{(6.6 \times 10^9) \times (1.99 \times 10^{30}\,\text{kg})}{\frac{4}{3} \times \pi \times (2.0 \times 10^{10}\,\text{km})^3}$$

$$\text{Density} = 3.9 \times 10^8\,\text{kg/km}^3 = 0.39\,\text{kg/m}^3$$

This is 2,500 times less dense than water. Supermassive black holes do not have the extremely high mean densities of stellar-mass black holes.

Feeding an AGN. Power for an AGN is produced when matter falls onto the accretion disk around the central supermassive black hole. Some of this high-velocity mass is radiated away according to Einstein's mass-energy equation: $E = mc^2$. About how much material has to be accreted to produce the observed luminosities? Astronomers estimate the efficiency of the accretion to be about 10–20 percent. Here, we'll assume that 15 percent of the infalling matter is radiated away as energy, or

$$E = 0.15mc^2$$

Astronomers can measure how much energy is produced by the infalling material and radiated to space. For a relatively weak AGN like that of the earlier example of the galaxy M87, $L = 5 \times 10^{35}$ joules per second (J/s), or 5×10^{35} kg m^2/s^2 each second. Here, we'll use $c = 3 \times 10^8$ m/s. Dividing both sides of Einstein's equation by $0.15c^2$ gives us the mass consumed each second:

$$m = \frac{E}{0.15c^2} = \frac{5 \times 10^{35}\,\text{kg}\,\text{m}^2/\text{s}^2}{0.15 \times (3 \times 10^8\,\text{m/s})^2}$$

$$m = 3.7 \times 10^{19}\,\text{kg}$$

Multiplying this result (the mass consumed each second) by 3.2×10^7 seconds per year shows that this AGN accretes 10^{27} kg, or about half the mass of Jupiter, each year, which is then radiated away as energy.

If we consider a quasar with a luminosity (L) of 10^{39} J/s = 10^{39} kg m^2/s^2 each second (= 2.5 trillion L_{Sun}), then the mass accreted each second is given by

$$m = \frac{10^{39}\,\text{kg}\,\text{m}^2/\text{s}^2}{0.15 \times (3 \times 10^8\,\text{m/s})^2}$$

$$m = 7.4 \times 10^{22}\,\text{kg}$$

Multiplying by 3.2×10^7 seconds per year yields a mass of 2.4×10^{30} kg per year. Recall that the mass of the Sun is 1.99×10^{30} kg. Therefore, this quasar supermassive black hole is accreting about 1.2 M_{Sun} each year to radiate this much energy. A quasar with 10 times this luminosity would be accreting 10 times the mass.

at high speeds produces emission lines that are smeared out by the Doppler effect into the broad lines seen in AGN spectra. This accretion disk surrounding a supermassive black hole is the "central engine" that powers AGNs.

The Unified Model of AGN

Astronomers have developed the basic picture of a supermassive black hole surrounded by an accretion disk into a more complete AGN model. The **unified model of AGN** attempts to explain all types of AGNs—quasars, Seyfert galaxies,

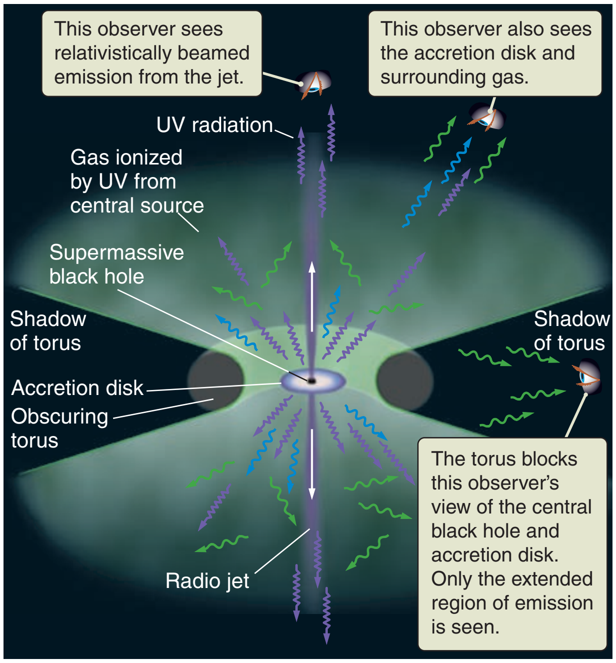


Figure 19.21 This figure illustrates the basic model of an active galactic nucleus, with a supermassive black hole surrounded by an accretion disk at the center. A larger, dusty torus sometimes blocks the view of the black hole. The mass of the central black hole, the rate at which it is being fed, and the viewing angle determine the observational properties of an AGN.

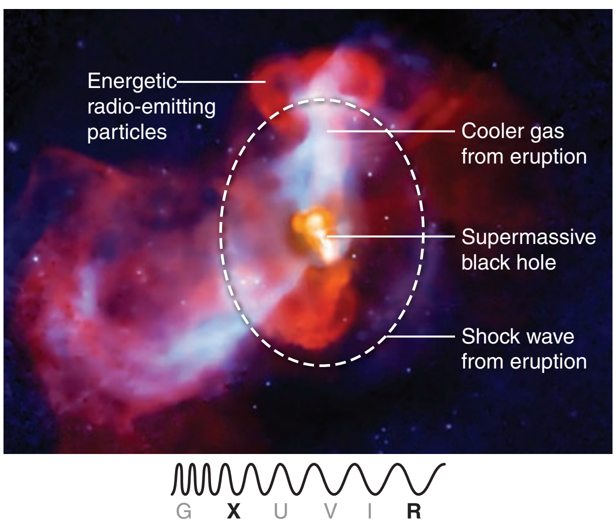


Figure 19.22 This image of M87 in radio and X-rays shows the location of the supermassive black hole.

and radio galaxies. **Figure 19.21** shows the various components of this AGN model, in which an accretion disk surrounds a supermassive black hole. Much farther out from the accretion disk lies a large torus, or "doughnut" of gas and dust consisting of material that is feeding the central engine. Located far from the inner turmoil of the accretion disk, and far larger than the central engine, some of this torus is ionized by UV light from the AGN (**Process of Science Figure**).

In the unified model of AGN, the different types of AGNs observed from Earth are partly explained by astronomers' view of the central engine. The torus of gas and dust obscures this view in different ways, depending on the viewing angle. Variation in this angle, in the mass of the black hole, and in the rate at which it is being fed accounts for a wide range of AGN properties. When the AGN is viewed edge on, astronomers see emission lines from the surrounding torus and other surrounding gas. They can also sometimes see the torus in absorption against the background of the galaxy. From this nearly edge-on orientation, they cannot see the accretion disk itself, so they do not see the Doppler-smeared lines that originate closer to the supermassive black hole. If jets are present in the AGN, however, these should be visible emerging from the center of the galaxy.

If astronomers observe the accretion disk somewhat more face on, they can see over the edge of the torus and thus get a more direct look at the accretion disk and the location of the black hole. In this case, they see more of the synchrotron emission from the region around the black hole and the Doppler-broadened lines produced in and around the accretion disk. **Figure 19.22** shows an image of one such object, the galaxy M87, at an intermediate inclination. M87 is a source of powerful jets that continue outward for 100,000 light-years but originate in the tiny engine at the heart of the galaxy. Spectra of the disk at the center of this galaxy show the rapid rotation of material around a central black hole that has a mass of 3 billion (3×10^9) M_{Sun}.

The material in an AGN jet travels very close to the speed of light. As a result, what astronomers see is strongly influenced by relativistic effects. One of these is called **relativistic beaming**: matter traveling at close to the speed of light concentrates any radiation it emits into a tight beam pointed in the direction in which it is moving. So astronomers often observe only one side of the jets from AGNs, even though the radio lobes of radio galaxies are usually two-sided. The jet moving away is just too faint to observe.

In rare instances when the accretion disk in a quasar or radio galaxy is viewed almost directly face on, relativistic beaming dominates the observations. In these *blazars*, emission lines and other light coming from hot gas in the accretion disk are overwhelmed by the bright glare of jet emission beamed directly at Earth.

Normal Galaxies and AGNs

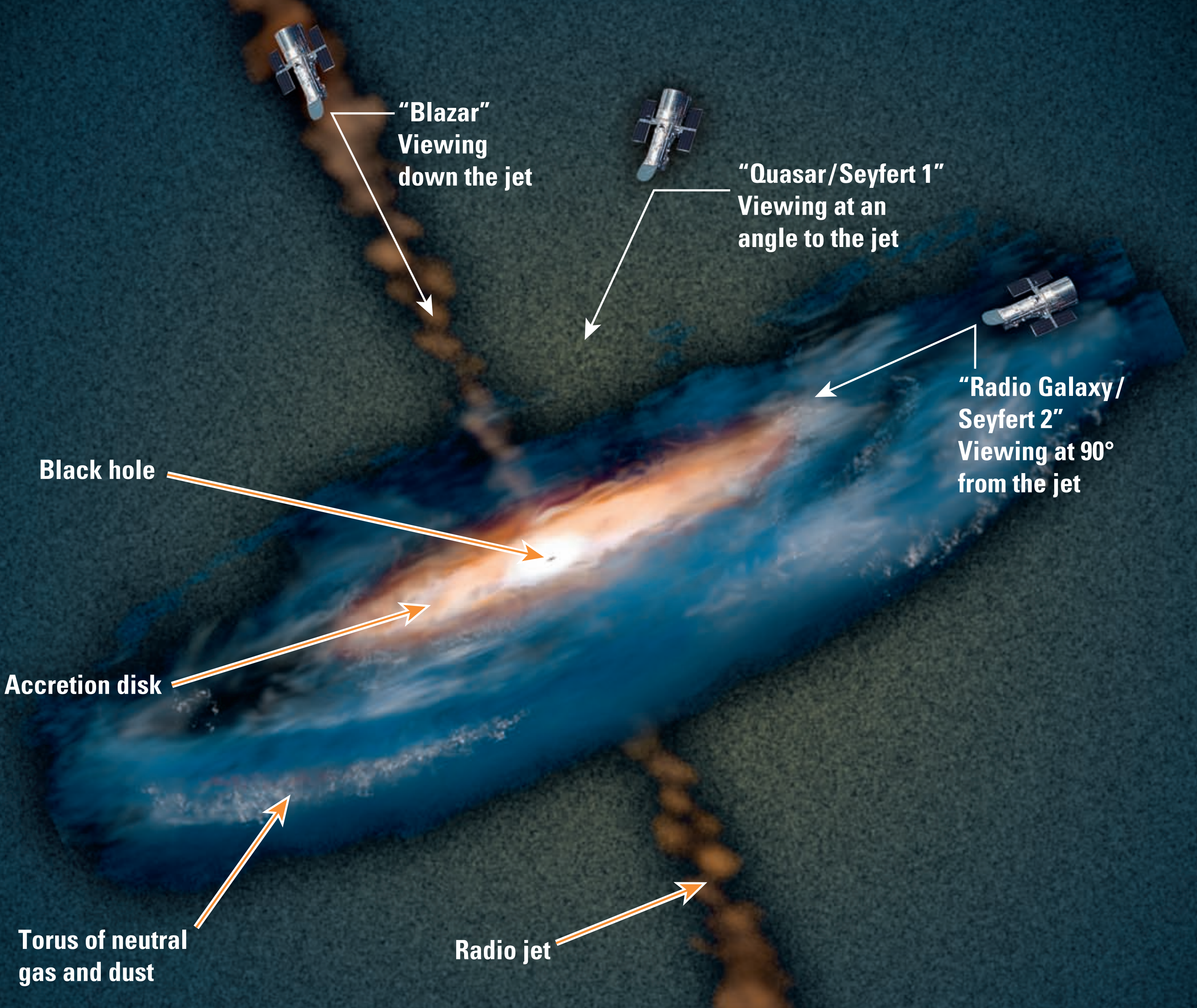
The essential elements of an AGN are a central engine (an accretion disk surrounding a supermassive black hole) and a source of fuel (gas and stars flowing onto the accretion disk). Without a source of matter falling onto the black hole, an AGN would no longer be an active nucleus. Astronomers looking at such an object would observe a normal galaxy with a supermassive black hole sitting in its center.

Only a few percent of present-day galaxies contain AGNs as luminous as the host galaxy. But when astronomers look at more distant galaxies (and therefore look further back in time), the percentage of galaxies with AGNs is much larger. These observations show that when the universe was younger, there were many more AGNs than there are today. If astronomers' understanding of AGNs is correct, then all the supermassive black holes that powered those dead AGNs should

Process of Science

FINDING THE COMMON THREAD

When first discovered, active galactic nuclei seemed to come in many different types, with dramatically different spectra. Later, it was realized that this could be an orientation effect: the many different types of objects could all be explained with one type of object, viewed from different angles.



Scientists seek underlying principles that explain more than one phenomenon. In the case of AGNs, scientists developed one model to explain many types of objects. This model encompasses the idea that our classification can be affected by our viewing angle, as in the picture of the coins at the beginning of the chapter. A unified model is much simpler than having a different model for every type of object.

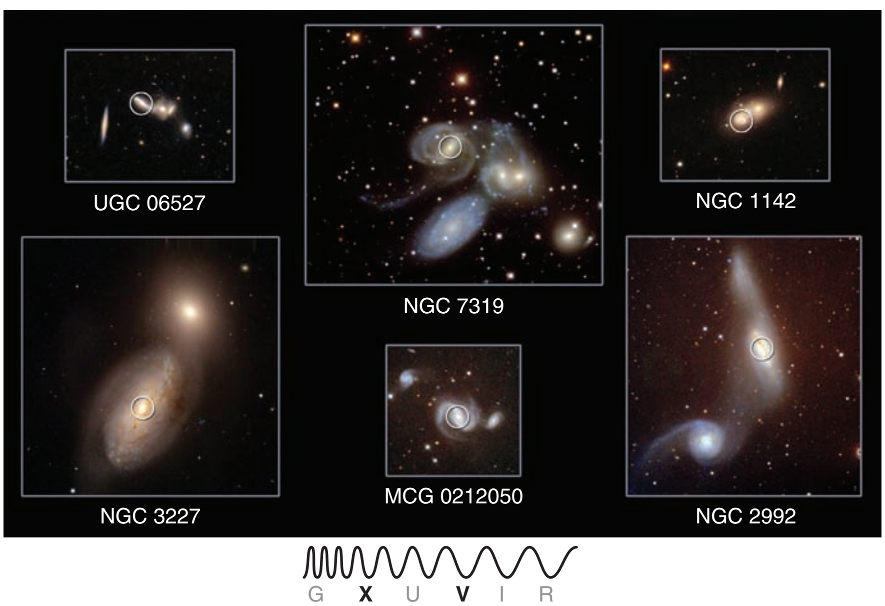
still be around. If astronomers combine what they know of the number of AGNs in the past with ideas about how long a given galaxy remains in an active AGN phase, they are led to predict that many—perhaps even *most*—normal galaxies today contain supermassive black holes.

If supermassive black holes are present in the centers of normal galaxies, these black holes should reveal themselves in a number of ways. For one thing, such a concentration of mass at the center of a galaxy should draw surrounding stars close to it. The central region of such a galaxy should be much brighter than could be explained if stars alone were responsible for the gravitational field in the inner part of the galaxy. Stars feeling the gravitational pull of a supermassive black hole in the center of a galaxy should also orbit at very high velocities and therefore show large Doppler shifts. Astronomers have found these large Doppler shifts in every normal galaxy with a substantial bulge in which a careful search has been conducted. The masses inferred for these black holes range from 10,000 M_{Sun} to 20 billion M_{Sun}. The mass of the supermassive black hole seems to be related to the mass of the bulge in which it is found. Most large galaxies, whether elliptical or spiral, probably contain supermassive black holes. These observations reveal something remarkable about the structure and history of normal galaxies.

Apparently the only difference between a normal galaxy and an active galaxy is whether the supermassive black hole at its center is being fed at the time we see that galaxy. The rarity of present-day galaxies with very luminous AGNs does not indicate which galaxies have the potential for AGN activity. Rather, it indicates which galaxy centers are being lit up at the moment. If a large amount of gas and dust were dropped directly into the center of any large galaxy, this material would fall inward toward the central black hole, forming an accretion disk and a surrounding torus. This process would change the nucleus of this galaxy into an AGN.

In Chapter 23, we will discuss galaxy evolution and note that many of the observed properties of galaxies discussed in this chapter, including the formation of galaxy type, spiral structure, star formation, and AGN, depend on the interactions and mergers between galaxies. To account for the many large galaxies visible today, interactions and mergers must have been much more prevalent in the past when the universe was younger; this is one explanation for the larger number of AGNs that existed in the past. Computer models show that galaxy-galaxy interactions can cause gas located thousands of parsecs from the center of a galaxy to fall inward toward the galaxy's center, where it can provide fuel for an AGN. During mergers, a significant fraction of a galaxy might wind up being cannibalized. HST images of quasars often show that quasar host galaxies are tidally distorted or are surrounded by other visible matter that is probably still falling into the galaxies. Galaxies that show evidence of recent interactions with other galaxies are more likely to house AGNs in their centers (**Figure 19.23**). Any large galaxy might be only an encounter away from becoming an AGN.

Figure 19.23 The Swift Gamma-Ray observatory has detected active black holes (circles) in these merging galaxies.



CHECK YOUR UNDERSTANDING 19.5

Supermassive black holes: (a) are extremely rare—there are only a handful in the universe; (b) are completely hypothetical; (c) occur in most, perhaps all, large galaxies; (d) occur only in the space between galaxies.

Origins

Habitability in Galaxies

In this chapter, we discussed the different types of galaxies that have been observed. Can we say anything about their potential for life? The short answer is that there is no solid information. These galaxies are too far away for astronomers to have detected any planets around their stars. So all we can do is speculate about the habitability of other galaxies. Two key requirements are the presence of heavy elements to form planets (and life) and an environment without too much radiation that might be damaging to life.

A study of the host stars of exoplanet candidates discovered by the Kepler telescope suggests that stars with a higher percentage of heavier elements may be more likely to have planets. (This finding fits with the core accretion models of planet formation discussed in Chapter 7.) The first generation of stars made from the hydrogen and helium of Big Bang nucleosynthesis do not have heavy elements. Recall from the chapters on stellar evolution that elements heavier than helium are created in the cores of dying stars and then are scattered into the galactic environment through planetary nebulae, stellar winds, and supernova explosions. So the amount of heavier elements in a star depends on the cosmic history of the material from which the star formed. Therefore, astronomers must consider the galactic environment of the star, which varies among different types of galaxies and different locations within the galaxies.

Spiral galaxies have had more continual star formation in their disks throughout their history. They contain more stars born from recycled material, and therefore more stars with a higher fraction of heavy elements. Elliptical (and S0) galaxies have older, redder populations of stars and very little current star formation. Old, massive ellipticals have a larger percentage of lower-mass stars than that of smaller ellipticals or spirals. Astronomers had previously thought this difference meant that large elliptical galaxies would not be good environments for planet formation. But the Kepler telescope has found many planets around small, red, main-sequence stars like the ones that populate elliptical galaxies. One study of two elliptical galaxies showed that both had some fraction of stars with a heavy-element fraction similar to that of the stars hosting Kepler exoplanets in the Milky Way.

Another issue is the presence of radiation that might be hazardous to life. This radiation is most likely to come from the center of the galaxy. Galaxies that are in an active AGN state might have too much radiation in regions close to their centers to be conducive to life. Stars whose orbits cross spiral arms many times might also be exposed to higher-than-average levels of radiation, but this is not the case for the majority of stars in a galaxy.

The conditions in these galaxies may also change as the galaxies evolve. Galaxy mergers can shake up stellar orbits and relocate stars and their planets to different locations. Mergers may also affect the growth and activity level of supermassive black holes and thus the presence of radiation. Some galactic environments just may not remain habitable for the length of time—billions of years—that it took life to evolve from bacteria to intelligence on Earth.

READING ASTRONOMY **NEWS**

ARTICLES QUESTIONS

In this news release story, astronomers used the Hubble and Gemini telescopes to find a massive black hole in a small galaxy.

Hubble Helps Find Smallest Known Galaxy with a Supermassive Black Hole

Hubblesite.org

Astronomers have found an unlikely object in an improbable place: a monster black hole lurking inside one of the tiniest galaxies known.

Though the black hole is five times the mass of the black hole at the center of our Milky Way, it is inside a galaxy that crams 140 million stars within a diameter of about 300 light-years, only 1/500th of our galaxy's diameter.

The dwarf galaxy containing the black hole, called M60-UCD1, is the densest galaxy ever seen (**Figure 19.24**). If you lived inside of it, the night sky would dazzle with at least 1 million stars visible to the naked eye (as opposed to 4,000 stars in our nighttime sky, as seen from Earth's surface).

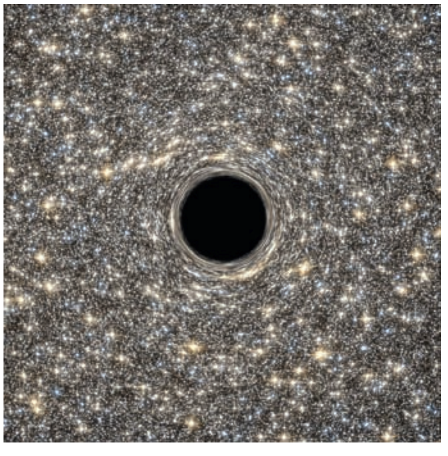


Figure 19.24 An artist's concept of a giant black hole in the center of the ultracompact galaxy M60-UCD1.

The finding implies that there are many other very compact galaxies in the universe that contain supermassive black holes. The observation also suggests that dwarf galaxies may actually be the stripped remnants of larger galaxies that were torn apart during collisions with yet other galaxies—rather than small islands of stars born in isolation.

"We don't know of any other way you could make a black hole so big in an object this small," said University of Utah astronomer Anil Seth, lead author of an international study of the dwarf galaxy published in the journal *Nature*.

His team of astronomers used the Hubble Space Telescope and the Gemini North 8-meter optical and infrared telescope on Hawaii's Mauna Kea to observe M60-UCD1 and measure the black hole's mass. The sharp Hubble images provide information about the galaxy's diameter and stellar density. Spectroscopy with Gemini measures the stellar motions as affected by the black hole's pull. These data are used to calculate the mass of the unseen black hole.

Black holes are gravitationally collapsed, ultracompact objects that have a gravitational pull so strong that even light cannot escape. Supermassive black holes—those with the mass of at least 1 million stars like our Sun—are thought to be at the centers of many galaxies.

The black hole at the center of our Milky Way galaxy has the mass of 4 million Suns, but as heavy as that is, it is less than 0.01 percent of the Milky Way's total mass. By comparison, the supermassive black hole at the center of M60-UCD1 is a stunning 15 percent of the small galaxy's total mass. "That is pretty amazing, given that the Milky Way is 500 times larger and more than 1,000 times heavier than the dwarf galaxy M60-UCD1," Seth said.

One explanation is that M60-UCD1 was once a large galaxy containing 10 billion stars, but then it passed very close to the center of an even larger galaxy, M60, and in that process all the stars and dark matter in the outer part of the galaxy got torn away and became part of M60.

The team believes that M60-UCD1 may eventually be pulled back to merge with the center of M60, which has its own monster black hole, weighing a whopping 4.5 billion solar masses (more than 1,000 times bigger than the black hole in our galaxy). When that happens, the black hole in M60-UCD1 will merge with the far more massive black hole in M60. The galaxies are 50 million light-years away.

ARTICLES QUESTIONS

1. How far away is M60-UCD1 in parsecs and kiloparsecs?
2. On average, how close are stars to each other in this galaxy?
3. Explain how the black hole's mass was estimated.
4. How is this merger different from those shown in Figure 19.23?
5. Watch the short video of a computer simulation of M60-UCD1 (http://www.spacetelescope.org/videos/heic1419a/). How do astronomers think this galaxy formed?

Summary

Galaxies are classified on the basis of their shape and the types of orbits of their stars. Most of the mass of a galaxy is dark matter, which interacts with light very weakly, if at all. The form of this matter is not yet known. Most galaxies have a supermassive black hole at the center, which may become an AGN if gas accretes onto it. In thinking about the potential habitability of galaxies, astronomers consider the activity state of the galaxy, including the presence or absence of an AGN and mergers, and the amount of heavy elements in the stars in the galaxy, which is related to the star formation rate and galaxy type.

LG 1 **Determine a galaxy's type from its appearance, and describe the motions of its stars.** Spiral galaxies are distinguished by their flat disk and spiral arms. The stars in this disk all orbit the center of the galaxy in the same direction. Elliptical galaxies are roughly egg-shaped, and the stars orbit in all directions. Irregular galaxies are galaxies that fit neither of these classifications, usually because they are interacting with another galaxy.

LG 2 **Explain the distance ladder and how distances to galaxies are estimated.** Astronomers build a distance ladder to galaxies by observing objects of known luminosity, such as Cepheid variable stars and Type Ia supernovae, in distant galaxies. Observations of the distance and the velocity of galaxies show that the two factors are related: more distant galaxies move away from us faster. Hubble's law $v = H_0d$ provides a method to find the distances of the most remote objects.

LG 3 **Describe the evidence suggesting that galaxies are composed mostly of dark matter.** Most of the mass in galaxies does not reside in gas, dust, or stars; instead, galaxy rotation curves indicate that about 90 percent of a galaxy's mass is in the form of dark matter, which does not emit or absorb light to any significant degree. Dark matter is identified by its gravitational interaction with ordinary matter. The two main groups of candidates for the composition of dark matter are MaCHOs (astronomical objects such as planets, dead stars, and black holes) and WIMPs (massive, weakly interacting elementary particles).

LG 4 **Discuss the evidence indicating that most—perhaps all—large galaxies have supermassive black holes at their centers.** Observations of distant quasars reveal extremely luminous, compact sources near the centers of galaxies. These active galactic nuclei (AGNs) are best explained as supermassive black holes, surrounded by an accretion disk and a torus of dust and gas. AGNs can emit as much as 1,000 times the light of the whole galaxy, all coming from a region the size of the Solar System. Surveys of the centers of galaxies show that many have AGNs at their centers.

? UNANSWERED QUESTIONS

- How standard are "standard candles"? For example, Cepheid variable light curves are slightly different, depending on the amount of heavy elements in the stars, so this variation must be calibrated. A majority of Type Ia supernovae may originate from the merging of two compact objects rather than from one white dwarf accreting mass and exploding when it gets close to 1.4 M_{Sun}. Astronomers try to address these difficulties by using multiple methods to find distances to galaxies—for example, observing numerous Cepheids, bright O stars, and Type Ia supernovae in the same galaxy—to check that their calculated distances agree. In one recent study, astronomers used HST to observe more than 600 Cepheid variable stars in eight galaxies in which there were Type Ia supernova detections, and they were able to reduce the uncertainty in their distances (and thus their calculated uncertainty in their value of the Hubble constant). The Type Ia supernovae calibrated in this way are consistent with earlier-calibrated maximum luminosities. But it is possible that those supernovae in galaxies far enough away that there are no other standard candles may have slightly different maximum luminosities, leading to a somewhat incorrect distance estimate (and value of H_0).
- Where do supermassive black holes come from? To explore this question, astronomers are studying computer models in which the supermassive black hole grew from a large number of stellar-mass black holes, or grew along with the galaxy by swallowing large amounts of central gas, or increased after the merger of two or more galaxies. We will return to this question when we discuss galaxy evolution in Chapter 23.

Questions and Problems

Test Your Understanding

1. Which of the following contributes the largest percentage to the total mass of a spiral galaxy?
   a. dark matter
   b. central black hole
   c. stars
   d. dust and gas

2. In the context of spiral galaxies, Kepler's laws could be used to estimate
   a. P, the period.
   b. A, the semimajor axis.
   c. M, the mass of the galaxy.
   d. v, the rotation speed of the galaxy.

3. Astronomers determine the radius of an AGN by measuring
   a. how much light comes from it.
   b. how hard it pulls on stars nearby.
   c. how quickly its light varies.
   d. how quickly it rotates.

4. If you observed a galaxy with an Hα emission line that had a wavelength of 756.3 nm, what would be the galaxy's redshift? Note that the rest wavelength of the Hα emission line is 656.3 nm.
   a. 0.01
   b. 0.05
   c. 0.10
   d. 0.15

5. As astronomers extend their distance ladder beyond 30 Mpc, they change their measuring standard from Cepheid variable stars to Type Ia supernovae. Why is this change necessary?
   a. Type Ia supernovae are more luminous than Cepheid variables.
   b. Type Ia supernovae are less luminous than Cepheid variables.
   c. Type Ia supernovae vary more slowly than do Cepheid variables.
   d. Type Ia supernovae vary more quickly than do Cepheid variables.

6. Which galaxy type has a spherical bulge and a well-defined disk?
   a. spiral
   b. barred spiral
   c. elliptical
   d. irregular

7. Which galaxy type is shaped like a rugby ball?
   a. Sb
   b. SBb
   c. E0
   d. E5

8. For a galaxy, the term *morphology* refers to
   a. its shape.
   b. its evolution over time.
   c. the motion of its stars.
   d. its overall density.

9. If all the stars in an elliptical galaxy traveled in random directions in their orbits, the elliptical galaxy would be type
   a. E0.
   b. E2.
   c. E5.
   d. E7.

10. The flat rotation curves of spiral galaxies imply that the distribution of mass resembles
    a. the Solar System; most mass is concentrated in the center.
    b. a wheel; the density remains the same as the radius increases.
    c. the light distribution of the galaxy; a large concentration occurs in the middle, but significant mass exists quite far out.
    d. an invisible sphere much larger than the visible galaxy.

11. Astronomers observe two galaxies, A and B. Galaxy A has a recession velocity of 2,500 km/s, while galaxy B has a recession velocity of 5,000 km/s. According to these data,
    a. galaxy A is 4 times as far away as galaxy B.
    b. galaxy A is twice as far away as galaxy B.
    c. galaxy B is twice as far away as galaxy A.
    d. galaxy B is 4 times as far away as galaxy A.

12. The Hubble constant is found from the
    a. slope of the line fit to the data in Hubble's law.
    b. y-intercept of the line fit to the data in Hubble's law.
    c. spread in the data in Hubble's law.
    d. inverse of the slope of the line fit to the data in Hubble's law.

13. Astronomers know that dark matter is present in galactic halos because the speeds of orbiting stars __________ far from the center of the galaxy.
    a. decrease
    b. increase
    c. remain about constant
    d. fluctuate dramatically

14. What accounts for the differences among various types of AGNs?
    a. the type of the host galaxy
    b. the size of the central black hole
    c. the amount of dark matter in the galaxy's halo
    d. our viewing angle

15. If a Seyfert galaxy's nucleus varies in brightness on the timescale of 10 hours, then approximately what is the size of the emitting region?
    a. 20 AU
    b. 70 AU
    c. 90 AU
    d. 140 AU

Thinking about the Concepts

16. What was the subject of the Great Debate, and why was it important to astronomers' understanding of the scale of the universe?

17. How did observations of Cepheid variable stars finally settle the Great Debate?

18. Why is it better to observe more than one type of standard candle in a distant galaxy?

19. Explain what astronomers mean by *distance ladder*.

20. Why is it important to know the type of progenitor of a Type Ia supernova in a distant galaxy?

21. Some galaxies have regions that are relatively blue; other regions appear redder. What does this variation indicate about the differences between these regions?

22. Describe how elliptical galaxies and spiral bulges are similar.

23. Which is more luminous: a quasar or a galaxy with 100 billion solar-type stars? Explain your answer.

24. The nearest observed quasar is about 750 Mpc away. Why don't astronomers observe any that are closer?

25. What distinguishes a normal galaxy from one that contains an AGN? It is a principle of science, often attributed to Einstein, that one should make things as simple as possible, but not simpler. Explain how this principle is at work in the distinction between normal galaxies and AGNs.

26. Contrast the size of a typical AGN with the size of the Solar System. How do astronomers know the size of an AGN?

27. Describe what astronomers think is happening at the center of a galaxy that contains an AGN.

28. It is likely that most galaxies contain supermassive black holes, yet many galaxies display no obvious evidence for their existence. Why do some black holes reveal their presence while others do not?

29. Study Figure 19.9.
    a. Why is it important that the different "rungs" of the distance ladder overlap in the distances that they measure?
    b. Does the figure end at the right edge because there are no more ways to measure distance or because there is no more universe to measure? How do you know?

30. Why do astronomers think that planets around stars near the centers of galaxies would not be good locations for the formation of life?

Applying the Concepts

31. Study Figure 19.14. The small vertical bars (known as error bars) on the data points indicate the size of the measurement error.
    a. At a radius of 25,000 parsecs (pc), what is the approximate measurement error in the rotation velocity?
    b. What is this value as a percentage of the measured velocity?
    c. Error bars are important because they show how wrong the measurement could possibly be. One way to think about this is that the black line could be as high as the top of the error bars or as low as the bottom of the error bars. In either case, would shifting the black line change the overall conclusion about redshift and distance? Why or why not?

32. Suppose the number density of galaxies in the universe is, on average, 3×10^{-68} galaxy/m^3. If astronomers could observe all galaxies out to a distance of 10^{10} pc, how many galaxies would they find?

33. The spectrum of a distant galaxy shows the Hα line of hydrogen ($\lambda_{\text{rest}} = 656.28$ nm) at a wavelength of 750 nm. Assume that $H_0 = 70$ km/s/Mpc.
    a. What is the redshift (z) of this galaxy?
    b. What is its recession velocity (v_r) in kilometers per second?
    c. What is the distance of the galaxy in megaparsecs?

34. The nearest known quasar is 3C 273. It is located in the constellation Virgo and is bright enough to be seen in a medium-sized amateur telescope. With a redshift of 0.158, what is the distance to 3C 273 in parsecs?

35. The quasar 3C 273 has a luminosity of $10^{12} L_{\text{Sun}}$. Assuming that the total luminosity of a large galaxy, such as the Andromeda Galaxy, is 10 billion times that of the Sun, compare the luminosity of 3C 273 with that of the entire Andromeda Galaxy.

36. A quasar has the same brightness as a galaxy that is seen in the foreground 2 Mpc distant. If the quasar is 1 million times more luminous than the galaxy, what is the distance of the quasar?

37. Estimate the Schwarzschild radius for a supermassive black hole with a mass of 26 billion M_{Sun}.

38. You read in the newspaper that astronomers have discovered a "new" cosmological object that appears to be flickering with a period of 83 minutes. Because you have read *21st Century Astronomy*, you are able to estimate quickly the maximum size of this object. How large can it be?

39. A quasar has a luminosity of 10^{41} W, or J/s, and $10^8\ M_{Sun}$ to feed it. Assuming constant luminosity and 20 percent conversion efficiency, what is your estimate of the quasar's lifetime?

40. A solar-type star ($M = 2 \times 10^{30}$ kg) approaches a supermassive black hole. As it crosses the event horizon, half of its mass falls into the black hole while the other half is completely converted to energy in the form of light. How much energy does this dying star send out to the rest of the universe?

41. Suppose a Type Ia supernova is found in a distant galaxy. The measured supernova brightness is 10^{-17} W/m^2. What is the distance of the galaxy?

42. Suppose that an object with the mass of Earth ($M_{Earth} = 5.97 \times 10^{24}$ kg) fell into a supermassive black hole with a 10 percent energy conversion.
    a. How much energy (in joules) would be radiated by the black hole?
    b. Compare your answer with the energy radiated by the Sun each second: 3.85×10^{26} J.

43. If a luminous quasar has a luminosity of 2×10^{41} W, or J/s, how many solar masses ($M_{Sun} = 2 \times 10^{30}$ kg) per year does this quasar consume to maintain its average energy output?

44. Material ejected from the supermassive black hole at the center of galaxy M87 extends outward from the galaxy to a distance of approximately 30,000 pc. M87 is approximately 17 Mpc away.
    a. If this material were visible to the naked eye, how large would it appear in the nighttime sky? Give your answer in degrees (1 radian = 57.3°).
    b. Compare this size with the angular size of the Moon.

45. A lobe in a visible jet from galaxy M87 is observed at a distance of 1,530 pc (5,000 light-years) from the galaxy's center moving outward at a speed of 0.99 times the speed of light (0.99*c*). Assuming constant speed, how long ago was the lobe expelled from the supermassive black hole at the galaxy's center?

USING THE WEB

46. Go to the Goddard Media Studios website and view the animation of a Cepheid variable star in a spiral galaxy (http://svs.gsfc.nasa.gov/goto?10145). Explain how astronomers use data like these to estimate the distance to the galaxy. What is actually observed, what is assumed, and what is calculated? (Review the discussion in Chapter 16 as needed.)

47. Go to the Astronomy Picture of the Day app or website (http://apod.nasa.gov/apod) and look at some recent pictures of galaxies. In each case, consider the following questions: Was the picture taken from a large or small telescope; from the ground or from space? Are galaxies in the image face on, edge on, or at an angle? What wavelengths were used for making the image? Are any of the colors "false colors"? If the picture is a combination of images from several telescopes, what do the different colors indicate?

48. Citizen science:
    a. Go to the website for Galaxy Zoo (http://galaxyzoo.org), the original Zooniverse citizen science project. (Log in with your Zooniverse password.) The specific project in action at any given time depends on the real data that need to be examined. One of the projects is likely a classification project. Click on and read, "Story," "Science," and "Classify," and then classify some galaxies. Save a copy of your classifications for your homework if necessary.
    b. Go to the website for Radio Galaxy Zoo (http://radio.galaxyzoo.org/). This project asks people to help locate supermassive black holes and their associated jets. Click on "Science" to read about the project, "Begin Hunting" to see some data, and then "Classify." Work through the given example, and then classify a few more.

49. a. Go to the website for the Fermi Gamma-ray Space Telescope (http://fermi.gsfc.nasa.gov). Scroll down to click on "Full News Archive" and look for a story about dark matter. What has this telescope discovered about dark matter?
    b. Go to the website for the Alpha Magnetic Spectrometer (http://ams02.org), a particle physics detector located on the International Space Station to search for dark matter, including WIMPs. What are the latest results?

50. a. Go to the website for the NASA Swift Gamma-Ray observatory (swift.gsfc.nasa.gov), which studies gamma-ray bursts. Click on "Latest Swift News" and look for a story about supermassive black holes. What has been discovered?
    b. Go to the website for NuSTAR (Nuclear Spectroscopic Telescope Array—http://www.nustar.caltech.edu), a space telescope launched by NASA in 2012. This mission is studying active galaxies hosting supermassive black holes. What type of telescope is this (wavelengths observed, general design)? What has been discovered?

smartwork5

If your instructor assigns homework in Smartwork5, access your assignments at digital.wwnorton.com/astro5.

EXPLORATION

Galaxy Classification

digital.wwnorton.com/astro5

Galaxy classification sounds simple, but it can become complicated when you actually attempt it. **Figure 19.25**, taken by the Hubble Space Telescope, shows a small portion of the Coma Cluster of galaxies. The Coma Cluster contains thousands of galaxies, each containing billions of stars. Some of the objects in this image (the ones with a bright cross) are foreground stars in the Milky Way. Some of the galaxies in this image are far behind the Coma Cluster. Working with a partner, in this Exploration you will classify the 20 or so brightest galaxies in this cluster.

First, make a map by laying a piece of paper over the image and numbering the 20 or so brightest (or largest) galaxies in the image (label them "galaxy 1," "galaxy 2," and so on). Copy this map so that you and your partner each have a list of the same galaxies.

Separately, classify each galaxy by type. If it is a spiral galaxy, what is its subtype: a, b, or c? If it is an elliptical, how elliptical is it? Make a table that contains the galaxy number, the type you have assigned it, and any comments that will help you remember why you made that choice. When you are done classifying, compare your list with your partner's. Now comes the fun part! Argue about the classifications until you agree—or until you agree to disagree.

Figure 19.25 This Hubble Space Telescope image of the Coma Cluster shows a diversity of shapes.

1 Which galaxy type was easiest to classify?

2 Which galaxy type was hardest to classify?

3 What makes it hard to classify some of the galaxies?

4 Which galaxy type did you and your partner agree about most often?

5 Which galaxy type did you and your partner disagree about most often?

6 How might you improve your classification technique?

If you found this activity interesting and rewarding, astronomers can use your help: go to http://galaxyzoo.org to get involved in a citizen science project to classify galaxies, some of which have never been viewed before by human eyes.

20 The Milky Way—A Normal Spiral Galaxy

Of the hundreds of billions of galaxies in the universe, the Milky Way is the only one that astronomers can study at close range. In this chapter, we focus our attention on the Milky Way and how it offers clues to understanding all galaxies.

LEARNING GOALS

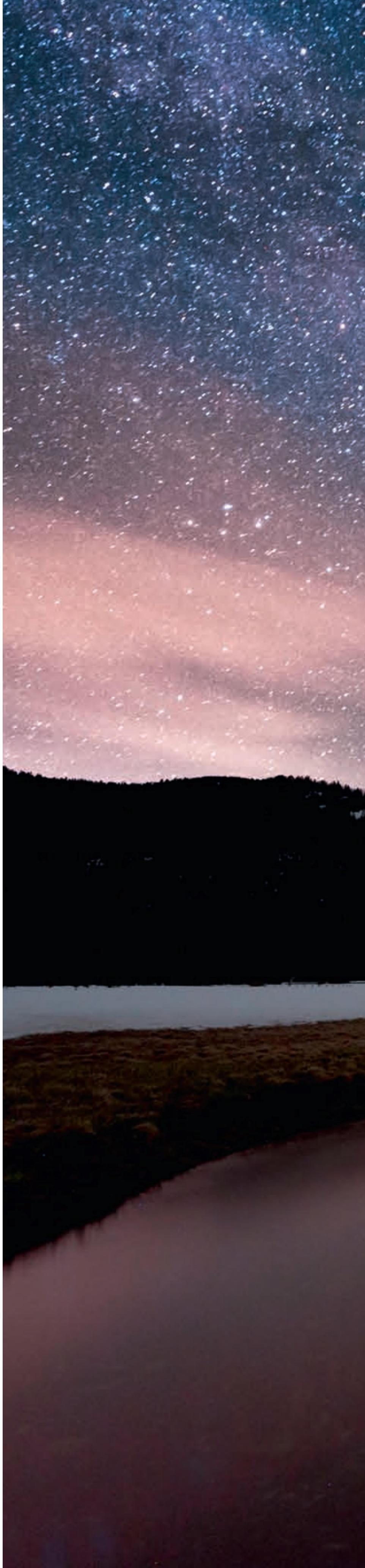
By the conclusion of this chapter, you should be able to:

LG 1 Explain how astronomers discovered the size and spiral structure of the Milky Way.

LG 2 List the clues of galaxy formation that can be found from the components of the Milky Way.

LG 3 Explain the evidence for a dark matter halo and for the supermassive black hole at the center of the Milky Way.

LG 4 Describe the Local Group of galaxies and how it provides clues about the evolution of the Milky Way.



The Milky Way in Earth's night sky has both bright patches of starlight and dark patches where stars are hidden by dust and gas. ▶▶▶

How do we know what the Milky Way Galaxy looks like from the outside?

20.1 Astronomers Have Measured the Size and Structure of the Milky Way

As you saw in Chapter 19, the universe is full of galaxies of many sizes and types. Because Earth is embedded within the Milky Way, the details of the shape and structure of the Milky Way are actually more difficult to determine than for other galaxies. Comparing observations of the Milky Way with observations of more distant galaxies improves our understanding of the Milky Way. In this section, we explore how astronomers infer the size and structure of the Milky Way from observations both within and without it.

Spiral Structure in the Milky Way

Figure 20.1a shows the Milky Way Galaxy in Earth's night sky. From a dark location at night, you can see dark bands of interstellar gas and dust that obscure much of the central plane of the Milky Way. This view of the Milky Way from inside offers a different and much closer perspective of a galaxy than can be obtained by viewing external galaxies. Compare this image of the Milky Way to the image in Figure 20.1b, which shows an edge-on spiral galaxy. The similarities between these two images suggest that the Milky Way is a spiral galaxy and that we are viewing it edge on, from a location inside the disk.

Finding further information about the size and shape of the Milky Way requires more extensive observations in the visible, infrared, and radio regions of the electromagnetic spectrum. Recall from Chapter 15 that neutral hydrogen

(a)

(b)

G X U V I R

Figure 20.1 (a) We see the Milky Way as a luminous band across the night sky. Prominent dark lanes caused by interstellar dust obscure the light from more distant stars. (b) The edge-on spiral galaxy NGC 891, whose disk greatly resembles the Milky Way.

emits radiation at a wavelength of 21 centimeters (cm), in the radio region of the spectrum. Maps of this radiation show spiral structure in other galaxies and suggest spiral structure in the Milky Way. In addition, observations of ionized hydrogen gas in visible light show two spiral arms with concentrations of young, hot O and B stars. These observations confirm that the Milky Way is a spiral galaxy.

In 2005, Spitzer Space Telescope observations of the distribution and motions of stars toward the center of the galaxy confirmed that the Milky Way has a substantial bar with a modest bulge at its center. **Figure 20.2** shows an artist's rendering of the major features of the Milky Way. Two major spiral arms—Scutum-Centaurus and Perseus—connect to the ends of the central bar and sweep through the galaxy's disk, just like the arms observed in external spiral galaxies. There are several smaller arm segments, including the Orion Spur, which contains the Sun and Solar System. Astronomers conclude that the Milky Way is a giant barred spiral that is more luminous than an average spiral. If viewed from the outside, the Milky Way would look much like the barred spiral galaxy M109, shown in **Figure 20.3**.

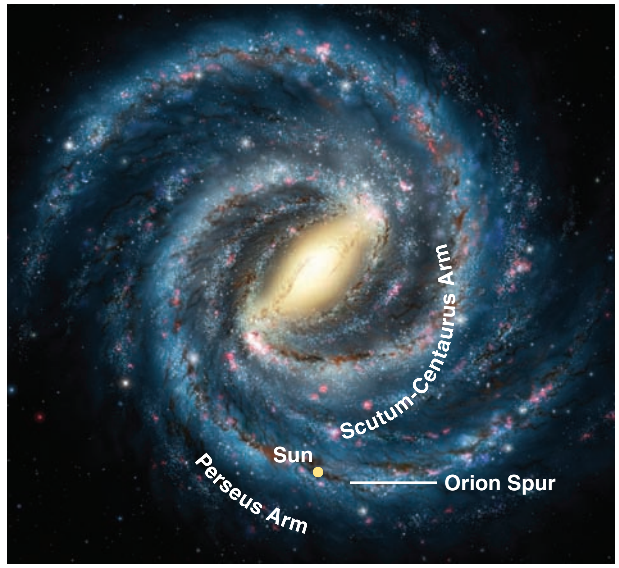


Figure 20.2 Infrared and radio observations contribute to an artist's model of the Milky Way Galaxy. The galaxy's two major arms (Scutum-Centaurus and Perseus) are seen attached to the ends of a thick central bar.

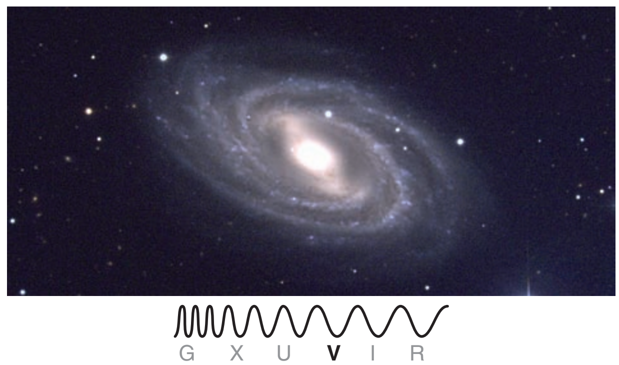


Figure 20.3 From the outside, the Milky Way would look much like this barred spiral galaxy, M109.

Spiral Arms and Star Formation

In pictures of other spiral galaxies, the arms are often the most prominent feature, as in the Andromeda Galaxy, shown in **Figure 20.4**. The spiral arms are quite prominent in the ultraviolet image (Figure 20.4a), and while they are less prominent in visible light (Figure 20.4b), they are still clearly defined. From this, you might conclude that most stars in the disk of a spiral galaxy are concentrated in the spiral arms. This turns out not to be the case: although stars are slightly concentrated in spiral arms, this concentration is not strong enough to account for their prominence. However, structures associated with star formation, such as molecular clouds and associations of luminous O and B stars, are all concentrated in spiral arms. Spiral arms are prominent because star formation is occurring there, so the arms contain significant concentrations of young, massive, luminous stars.

Recall from Chapter 15 that stars form when dense interstellar clouds become so dense that they begin to collapse under the force of their own gravity. If stars form in spiral arms, then spiral arms must be places where clouds of interstellar gas and dust pile up and are compressed. These clouds can be observed where the dust

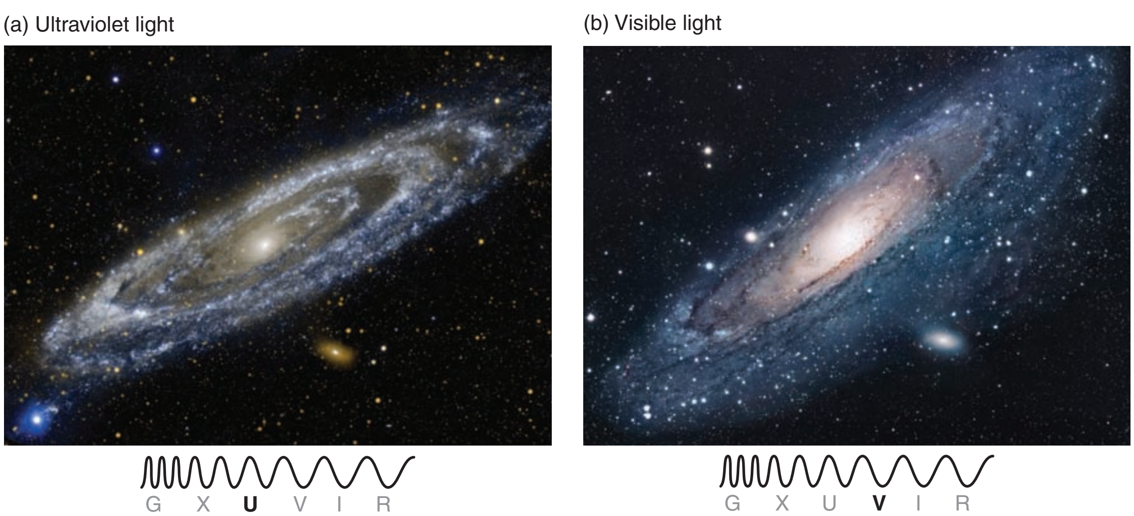


Figure 20.4 These photos show the Andromeda Galaxy in ultraviolet light (a) and visible light (b). Note that the spiral arms, which are dominated by hot young stars, are most prominent in ultraviolet light. The spiral arms are less prominent in visible light.

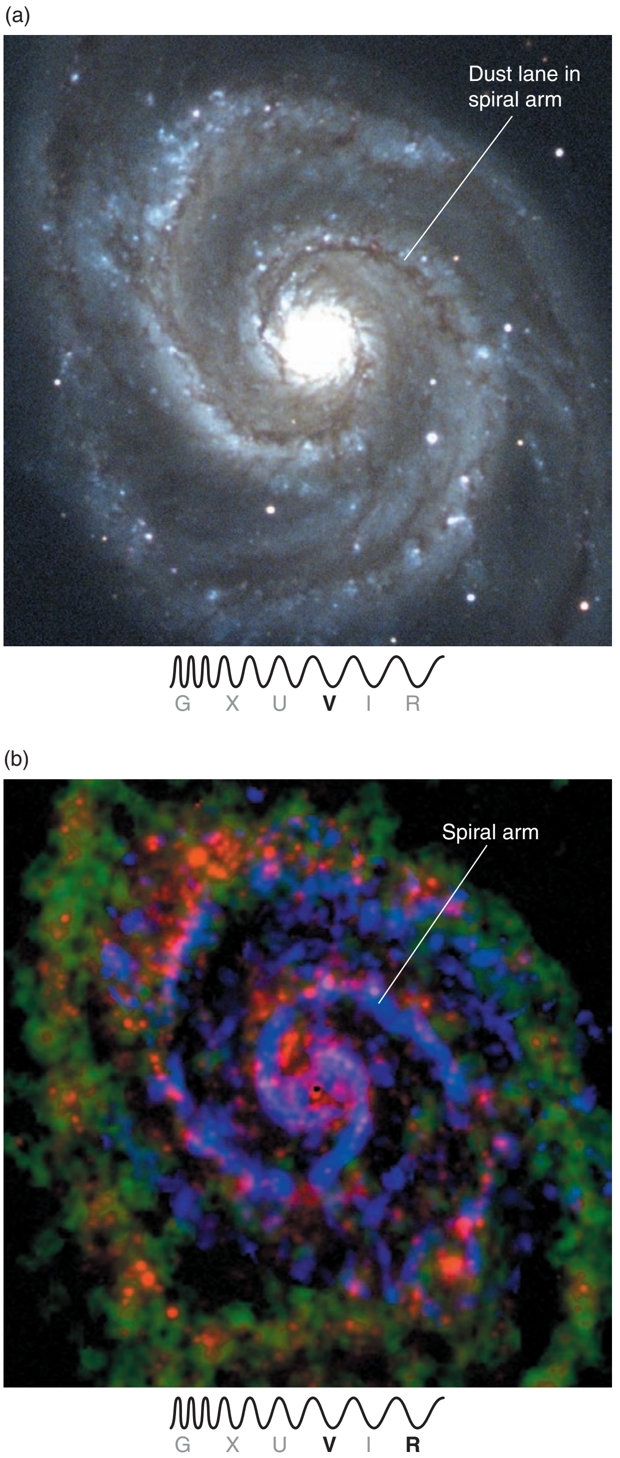


Figure 20.5 These two images of a face-on spiral galaxy show the spiral arms. (a) This visible-light image also shows dust absorption. (b) This image shows the distribution of neutral interstellar hydrogen (green), carbon monoxide (CO) emission from cold molecular clouds (blue), and hydrogen alpha (Hα) emission from ionized gas (red).

and gas block starlight, as in **Figure 20.5a**, or where gases such as neutral hydrogen or carbon monoxide produce emission at various wavelengths (Figure 20.5b).

Why do spiral arms exist at all? Part of the answer is that disks do not rotate like a solid body, such as a wheel. Instead, material that is closer to the center takes less time to complete a revolution around the galaxy than material farther out in the galaxy, and so the inner part of the disk gets ahead of the outer part. This means that any disturbance in the disk of a spiral galaxy will produce a spiral pattern as the disk rotates. **Figure 20.6** illustrates the point. The first frame shows a single linear disturbance through the center of a model galaxy. In the second frame, the outer part of the line is trailing behind the inner part. As the galaxy rotates, a straight line through the center becomes a spiral. In the time it takes for objects in the inner part of the galaxy to complete several rotations, objects in the outer part of the galaxy may not have completed even a single revolution. In the process, the originally straight arms are slowly made into the spiral structure shown.

A spiral galaxy can be disturbed, for example, by gravitational interactions with other galaxies or by a burst of star formation. However, a single disturbance will not produce a stable spiral-arm pattern. Spiral arms from a single disturbance will wind tightly around the center in two or three rotations and disappear. Disturbances that are repetitive, however, can sustain spiral structure indefinitely. For example, a bar in the center of a spiral galaxy gravitationally disturbs the disk. As the disk rotates through this disturbance, the disturbance is repeated. Repeated episodes of star formation occur and maintain stable spiral arms.

Many galaxies show clear evidence of a relationship between the shapes of their bulges and the structure of their spiral arms. Barred spirals, for example, have a characteristic two-armed spiral pattern that is connected to the elongated bulge, as seen in Figure 20.3. Even the bulges of galaxies that are not obviously barred may be elongated enough to contribute to the formation of a two-armed spiral structure. Smaller galaxies in orbit about larger galaxies can also give rise to a periodic gravitational disturbance, triggering the same sort of two-armed structure.

The process of star formation itself can also create spiral structure. Regions of star formation release considerable energy into their surroundings through UV radiation, stellar winds, and supernova explosions. This energy compresses clouds of gas and triggers more star formation. Typically, many massive stars form in the same region at about the same time, and their combined mass outflows and supernova explosions occur one after another in the same region of space over the course of only a few million years. The result can be large, expanding bubbles of hot gas that sweep out cavities in the interstellar medium and concentrate the swept-up gas into dense, star-forming clouds, much like the snow that piles up in front of a snowplow. In this way, star formation can propagate through the disk of a galaxy. Rotation bends the resulting strings of star-forming regions into spiral structures.

Stars move in and out of arms as they orbit the center of a galaxy. Consequently, the stars in an arm today are not the same stars that were in the arm 20 million years ago. This is roughly analogous to a traffic jam on a busy highway. The cars in the jam are changing all the time, yet the traffic jam persists as a place of higher density—where there are more cars than usual. The traffic jam itself moves slowly backward, even as the cars move forward and pass through it. Just like the traffic jam, the disturbance of the spiral arm also moves at a different speed than the individual stars. These disturbances in the disks of spiral galaxies are called **spiral density waves** because they are waves of greater mass density and increased pressure in the galaxy's interstellar medium. These waves move around a disk in the pattern of a two-armed spiral that does not rotate at the same rate as the stars,

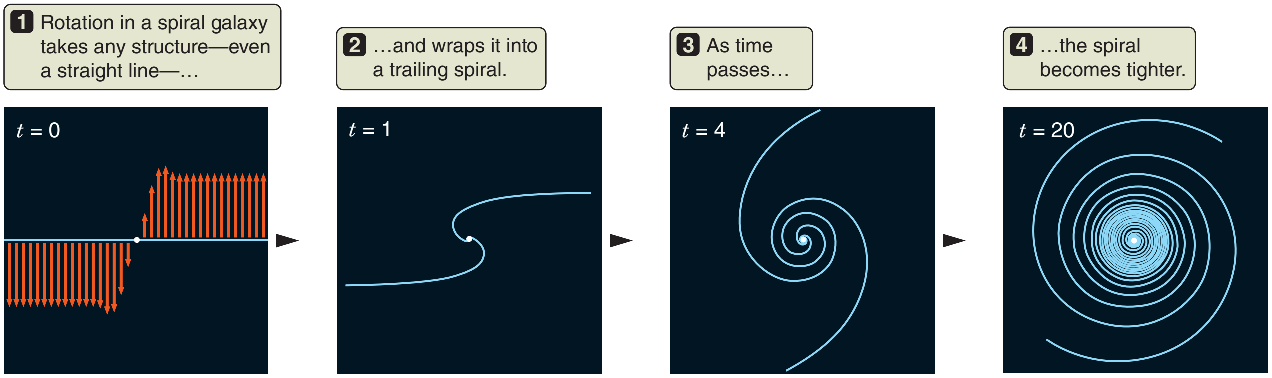


Figure 20.6 The rotation of a spiral galaxy will naturally take even an originally linear structure (t = 0) and wrap it into a progressively tighter spiral as time (t) goes by.

gas, or dust. As material in the disk orbits the center of the galaxy, it passes through these spiral density waves.

A spiral density wave has very little effect on the motions of stars as they pass through it, but it does compress the gas that flows through it. As an analogy for this process, consider what happens when you turn on the tap in your kitchen sink (**Figure 20.7**). The water hits the bottom of the sink and spreads out in a thin, rapidly moving layer. A few centimeters out, depending on the rate at which water is flowing, there is a sudden increase in the depth of the water. Spiral arms in galaxies work in much the same way. Gas flows into the spiral density wave and piles up like the water in the sink. Stars form in the resulting compressed gas. Massive stars are concentrated in the arms because they have such short lives (typically 10 million years or so) that they never have the chance to drift far from the spiral arms where they were born. Less massive stars, however, have plenty of time to move away from their places of birth, so they form a smooth underlying disk.

Nebraska Simulation: Traffic Density Analogy

The Size of the Milky Way Galaxy

Because the Solar System is inside the dusty disk of the galaxy itself, the visible-light view of the Milky Way is badly obscured. If you go out on a dark night, away from any streetlights, and look in the direction of the center of the Milky Way—located in the constellation Sagittarius—you will see the dark lane of dusty clouds shown in Figure 20.1a. To probe the structure of the Milky Way, modern astronomers use long-wavelength infrared and radio radiation that can penetrate the dust in the disk. The most powerful tool for this work is the same 21-cm line from neutral interstellar hydrogen, which (as we described in Chapter 19) is used to measure the rotation of other galaxies. The distance to the center of the galaxy, however, still cannot be measured directly. In the 1920s, Harlow Shapley made a three-dimensional map of globular clusters in the Milky Way, which led to the first determination of the size of the Milky Way and the Sun's offset from the center.

Recall from Chapter 17 that globular clusters are large, spheroidal groups of stars held together by gravity. The Milky Way contains more than 150 cataloged globular clusters, and many more are hidden by dust in the disk. Globular clusters are very luminous (as much as 1 million L_{Sun}), so the ones that lie outside of the dusty disk can be easily seen as round, fuzzy blobs even through small telescopes, and even at great distances.

In a Hertzsprung-Russell (H-R) diagram of an old cluster, the horizontal branch crosses the instability strip, which contains pulsating stars such as RR Lyrae stars and Cepheid variables. RR Lyrae stars are easy to spot in globular clusters because they are relatively luminous and have a distinctive light curve.

Figure 20.7 Water from a tap flows in a thin layer along the bottom of a kitchen sink. There is a sudden increase in the depth of the water away from the drain.

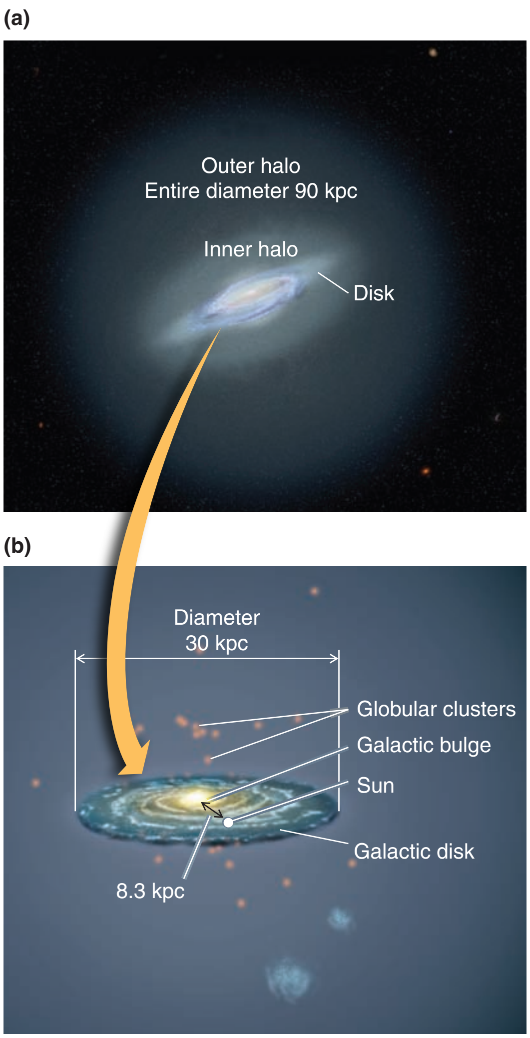


Figure 20.8 Parts of the Milky Way include (a) the disk and the inner and outer halos; (b) the galactic bulge and disk. Globular clusters are located in the halo, and the Sun is located in the disk, about 8.3 kpc from the center. (1 kpc = 1,000 pc.)

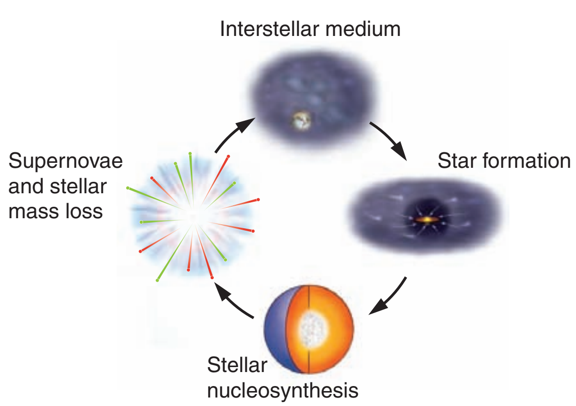


Figure 20.9 Matter moves from the interstellar medium into stars and back again in a progressive cycle that has enriched today's universe with massive elements.

As with Cepheid variables, the time it takes for an RR Lyrae star to undergo one pulsation is related to the star's luminosity. Harlow Shapley used this period-luminosity relationship to find the luminosities of RR Lyrae stars in globular clusters. He then combined these luminosities with measured brightnesses to determine the distances to globular clusters. Finally, Shapley cross-checked his results by noting that more distant clusters (as measured by the RR Lyrae stars) also tended to appear smaller in the sky, as expected.

These globular clusters trace out the luminous part of the galactic halo of the Milky Way Galaxy, a large spherical volume of space surrounding the disk and bulge. The center of the distribution of globular clusters coincides with the gravitational center of the galaxy. Shapley realized that because he could determine the distance to the center of this distribution, he had actually determined the Sun's distance from the center of the Milky Way, as well as the size of the galaxy itself. His map showed that globular clusters occupy a roughly spherical region of space with a diameter of about 90 kiloparsecs (kpc), or 90,000 parsecs (pc). Shapley, however, did not know about gas and dust, so he overestimated the distance to the globular clusters (**Process of Science Figure**). A modern determination indicates that the Sun is located about 8,300 pc (27,000 light-years) from the center of the galaxy, or roughly halfway out toward the edge of the disk.

CHECK YOUR UNDERSTANDING 20.1

Name three pieces of evidence in support of the theory that the Milky Way is a spiral galaxy.

20.2 The Components of the Milky Way Provide Clues about the Formation of Spiral Galaxies

Figure 20.8 illustrates the Sun's position and relationship to the rest of the Milky Way. The Sun is a middle-aged disk star located among other middle-aged stars that orbit around the galaxy within the galactic disk, as do the gas and dust in the disk. The stars in the halo move in random orbits similar to those of stars in elliptical galaxies, sometimes at high velocities. Some of these stars can be observed near the Sun, as their orbits carry them swiftly through the disk. Most of these stars are much older than the Sun. The bar in the galactic bulge of the Milky Way is shaped primarily by stars and gas moving both in highly elongated orbits up and down the long axis of the bar and in short orbits aligned perpendicular to the bar. All these stellar orbits determine the shapes of the different parts of the galaxy, and it is much easier to measure stellar orbits in the Milky Way than in other galaxies. Using the ages, chemical abundances, and motions of nearby stars, astronomers can differentiate between disk and halo stars to learn more about the galaxy's structure. In this section, you will learn how astronomers study the constituents of the Milky Way to find direct clues about how spiral galaxies form.

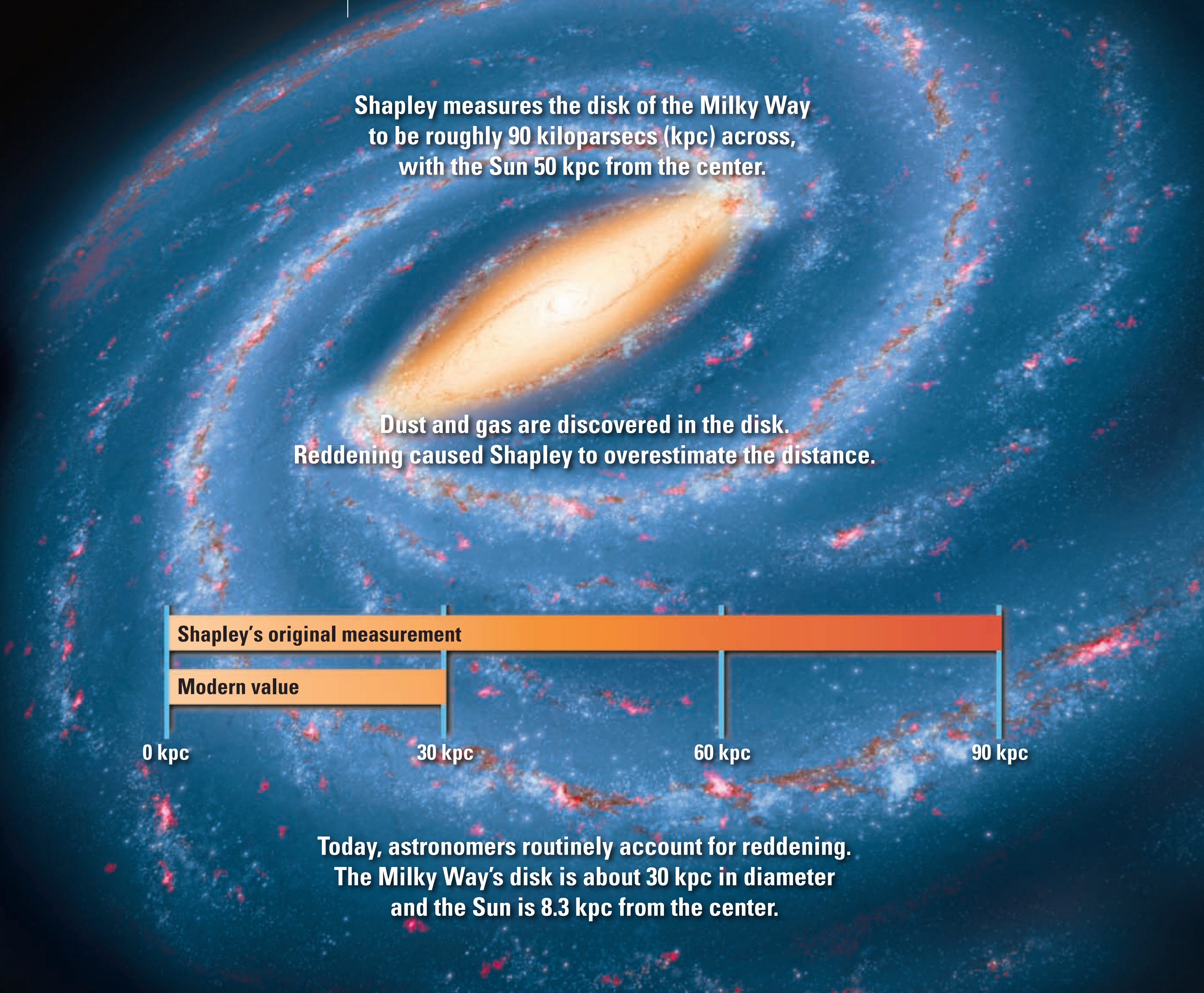
Age and Chemical Compositions of Stars

Over time, the chemical content of a galaxy changes as stars are born, live and die, and progressively enrich the interstellar medium, as shown in **Figure 20.9**. The interstellar medium therefore reflects all the stellar evolution that has taken place

Process of Science

UNKNOWN UNKNOWNS

Shapley's initial efforts to measure the Milky Way did not include dust and gas, because he did not know about them.



Often, scientists are very aware of what they don't know—for example, the composition of dark matter. Other times, an "unknown unknown" is later discovered, and prior results must be modified to incorporate the new knowledge.

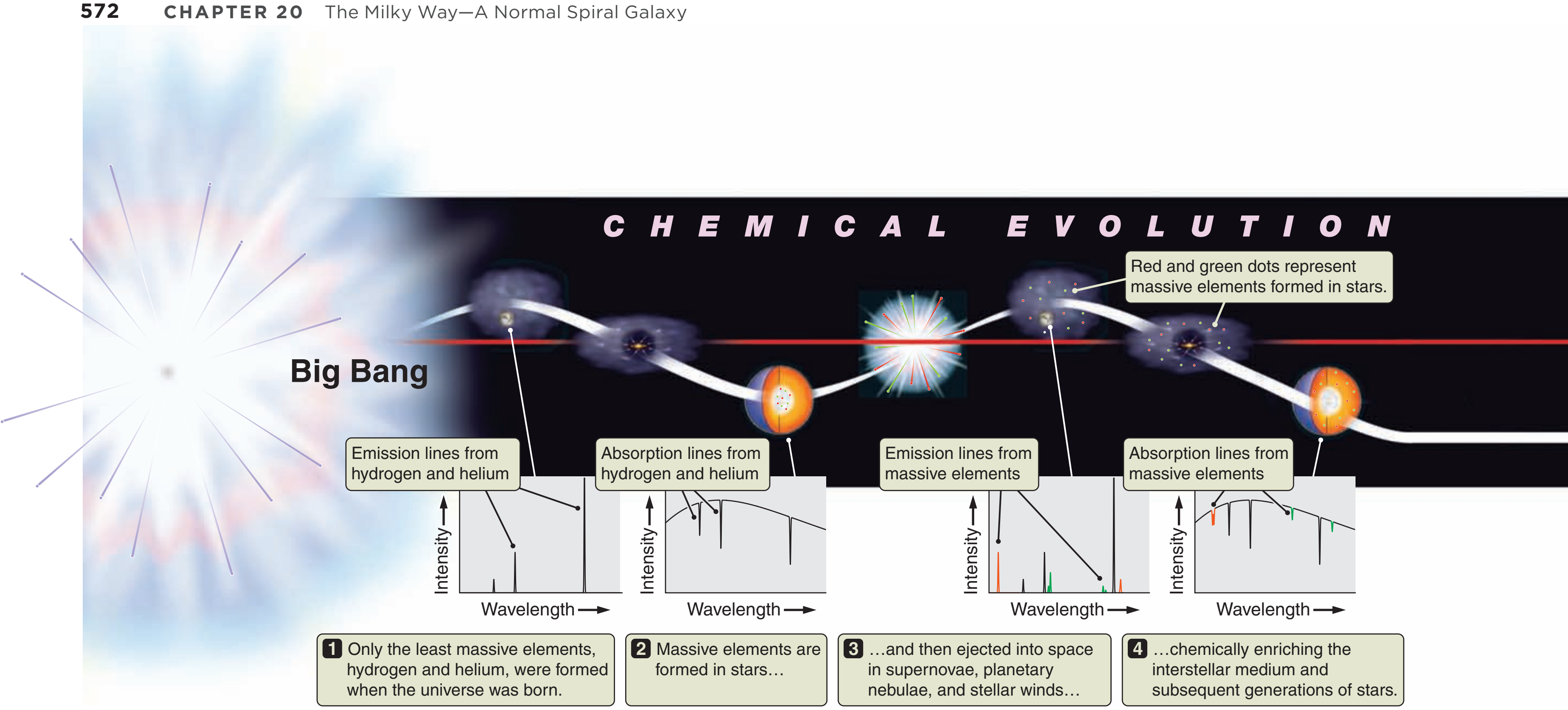
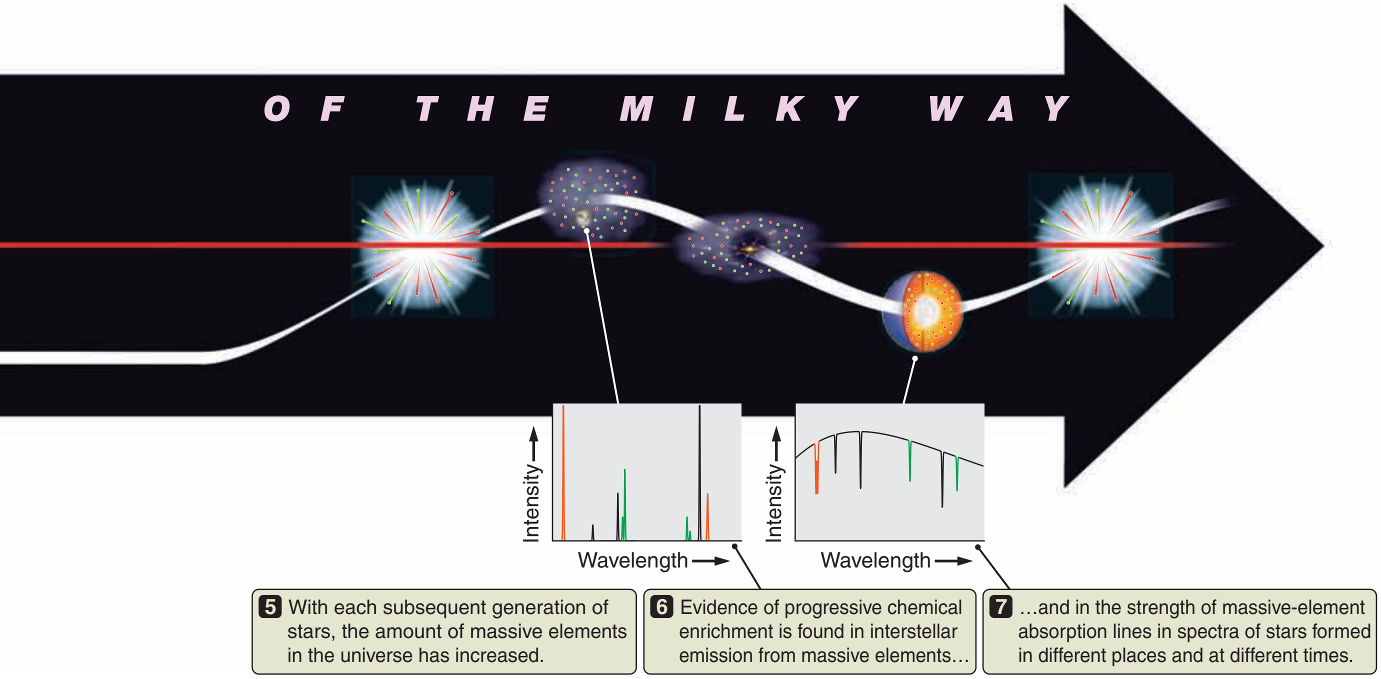


Figure 20.10 As subsequent generations of stars form, live, and die, they enrich the interstellar medium with massive elements—the products of stellar nucleosynthesis. The chemical evolution of the Milky Way and other galaxies can be traced in many ways, including by the strength of interstellar emission lines and stellar absorption lines.

up to the present time. Gas that is rich in massive elements has gone through a great deal of stellar processing. This evolution is similarly reflected in the composition of stars within the galaxy. As illustrated in **Figure 20.10**, the chemical composition of a star's atmosphere reflects the cumulative amount of star formation that occurred before that star formed. Although the details are complex, several clear and important lessons can be learned from the observed patterns in the amounts of massive elements in the galaxy.

Stars in globular clusters, being among the earliest stars to form, should contain only very small amounts of massive elements. Some globular-cluster stars contain only 0.5 percent as much of these massive elements as the Sun. This relationship between age and abundances of massive elements is evident throughout much of the galaxy. Within the disk, younger stars typically have higher abundances of massive elements than those of older stars. Similarly, older stars in the outer parts of the galaxy's bulge have lower massive-element abundances than those of young stars in the disk. Such lower amounts of heavy elements characterize not only globular-cluster stars but also all of the stars in the galaxy's halo, where globular-cluster stars constitute only a minority among the total number of stars in the galactic halo.

In addition to the globular clusters in the halo, younger open clusters orbit in the disk of the Milky Way. As with globular clusters, the stars in an open cluster all formed in the same region at about the same time. Open clusters have a wider range of ages. Some open clusters contain the very youngest stars known; others contain stars that are somewhat older than the Sun. The oldest open clusters in the disk are several billion years younger than the youngest globular clusters in the halo. The differences in ages between globular and open clusters indicate that stars in the halo formed first, but this epoch of star formation did not last long. No young globular clusters are seen. Star formation in the disk started later and has been continuing ever since.



Within the galaxy's disk, astronomers observe differences in abundances of massive elements from place to place, related to the rate of star formation in different regions. Star formation is generally more active in the denser, inner part of the Milky Way than in the outer parts. Observations of chemical abundances in the interstellar medium, based both on interstellar absorption lines in the spectra of stars and on emission lines in glowing clouds of gas known as H II regions, confirm this prediction by showing that there is a smooth decline in abundances of massive elements from the inner to the outer parts of the disk. Astronomers have observed similar trends in other galaxies. Within a galactic disk, relatively old stars near the center of a galaxy often have greater massive-element abundances than those of young stars in the outer parts of the disk.

The basic idea that higher massive-element abundances should follow from the more prodigious star formation in the inner galaxy seems correct, but the full picture is not this simple. New material falling into the galaxy might affect the amounts of heavy elements in the interstellar medium. Chemical elements produced in the inner disk might be blasted into the halo in great "fountains" powered by the energy of massive stars, only to fall back onto the disk elsewhere. Past interactions with other galaxies might have stirred the Milky Way's interstellar medium, mixing gas from those other galaxies with gas already there. The variations of chemical abundances within the Milky Way and other galaxies—and what these variations tell us about the history of star formation and the formation of elements—remain active topics of research.

Even the very oldest globular-cluster stars contain some chemical elements fused in previous generations of more massive stars. This observation implies that globular-cluster stars and other halo stars were not the first stars in the Milky Way to form. At least one generation of massive stars lived and died, ejecting newly synthesized massive elements into space, before even the oldest globular clusters formed. (We will return to these first stars in Chapter 23.) Every star less

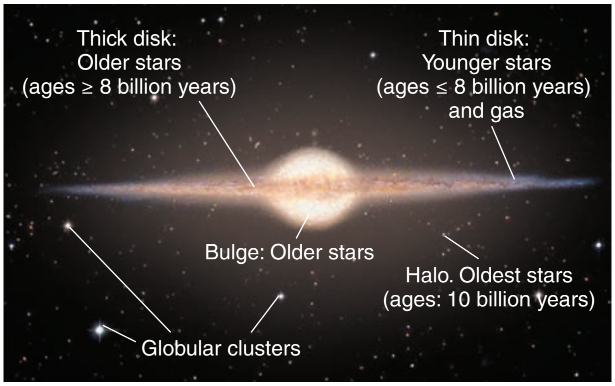


Figure 20.11 This illustration shows the disks, bulge, and inner halo of the Milky Way and the location of globular clusters.

massive than about 0.8 M_{Sun} that ever formed is still around as a main-sequence star today. Even so, astronomers find no disk stars with exceptionally low massive-element abundances. The gas that wound up in the plane of the Milky Way must have seen a significant amount of star formation before it settled into the disk of the galaxy and made stars. Still, even a chemically "rich" star like the Sun, which is made of gas processed through approximately 9 billion years of previous generations of stars, is composed of less than 2 percent massive elements. Luminous matter in the universe is still dominated by hydrogen and helium formed just after the Big Bang, long before the first stars.

Components of the Disk

The disk of the Milky Way has a thin component and a thick component, as illustrated in **Figure 20.11**. The youngest stars in the galaxy are most strongly concentrated in the galactic plane, defining a disk about 300 pc (1,000 light-years) thick but more than 30,000 pc (100,000 light-years) across. This ratio of the diameter to the thickness of the disk is similar to that of a DVD. The older population of disk stars, distinguishable by lower abundances of massive elements, has a much thicker distribution, about 3,700 pc (12,000 light-years) thick. The youngest stars are concentrated closest to the plane of the galaxy because this is where the molecular clouds are. These stars show a decrease in heavy-element content as the distance from the galactic center increases. Older stars make up the thicker parts of the disk. Among these older stars, the stars farther from the galactic plane have similar amounts of heavier elements at all distances from the galactic center.

There are two hypotheses for the origin of this thicker disk. One suggests that these stars formed in the midplane of the disk long ago but were affected by gravitational interactions with massive molecular clouds in the spiral arms that kicked them up out of the plane of the galaxy. The other hypothesis suggests that these stars were acquired from the merging processes that formed the Milky Way Galaxy. As galaxies merge, their clouds of gas cannot pass through each other, so the colliding gas clouds instead settle into a disk at the midplane of the newly formed spiral galaxy. These clouds appear as the concentrated dust lanes that slice the disks of spiral galaxies (as shown in Figure 20.1). This thin lane slicing through the middle of the disk is the place where new stars form and where new stars are found. However, stars are free to move back and forth from one side of the disk to the other, so the older stars can be found above and below this thin disk.

The interstellar medium is a dynamic place—energy from star-forming regions can shape it into impressively large structures. As we mentioned earlier, energy from regions of star formation can form interesting structures in the interstellar medium, clearing out large regions of gas in the disk of a galaxy. Many massive stars forming in the same region can blow "fountains" of hot gas out through the disk of the galaxy via a combination of supernova explosions and strong stellar winds. In the process, dense interstellar gas can be thrown high above the plane of the galaxy as illustrated in the artist's depiction of **Figure 20.12**. Once the gas is a few kiloparsecs above or below the disk, it radiates and cools, falling back to the disk. Maps of the 21-cm emission from neutral hydrogen in the galaxy, X-ray observations, and visible-light images of hydrogen emission from some edge-on external galaxies show numerous vertical structures in the interstellar medium of disk galaxies. These vertical structures are often interpreted as the "walls" of the fountains. If enough massive stars are formed together, sufficient energy may be deposited to blast holes all the way through the plane of the galaxy.

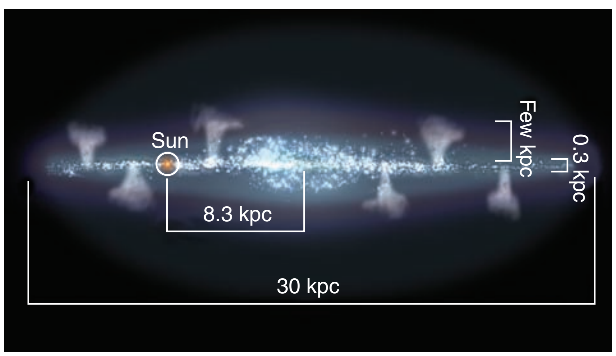


Figure 20.12 This artist's sketch illustrates the concept of a "galactic fountain" in which gas is pushed away from the plane of the galaxy by energy released by young stars and supernovae and then falls back onto the disk. The distance of the Sun from the center of the galaxy and the diameter of the disk are shown here for scale.

Components of the Halo

The globular clusters in the galactic halo indicate that star formation there was early and brief. Yet globular clusters account for only about 1 percent of the total mass of stars in the halo. As halo stars fall through the disk of the Milky Way, some pass close to the Sun, providing a sample of the halo that can be studied at closer range. Most of the stars near the Sun are disk stars like the Sun, but astronomers can distinguish nearby halo stars in two ways. First, most halo stars have much lower amounts of massive elements than those of disk stars. Second, halo stars appear to be moving at higher relative velocities than disk stars. The disk stars near the Sun orbit the center of the galaxy at nearly the same speed, in roughly the same direction. In contrast, halo stars orbit the center of the galaxy in random directions. So the relative velocity between the halo stars and the Sun tends to be high. These stars are known as high-velocity stars.

By studying the orbits of high-velocity stars, astronomers have determined that the halo has two separate components: an inner halo that includes stars up to about 15 kpc (50,000 light-years) from the center, and an outer halo that extends far beyond that (see Figure 20.8a). The stars in the outer halo have lower fractions of heavier elements, suggesting that they formed very early. Many of them are moving in a direction opposite to the rotation of the galaxy. This suggests that the outer halo may have its origins in a merger with a small dwarf galaxy long ago. The orbits of halo stars suggest that these stars fill a volume of space similar to that occupied by the globular clusters in the halo.

X-ray observations indicate that there may be a halo of hot gas surrounding the Milky Way, as shown in the artist's concept in **Figure 20.13**. This gas halo may extend for about 100 kpc (300,000 light-years) from the galactic center, encompassing two nearby small galaxies, and containing as much mass as that of all the stars in the galaxy. Its temperature is estimated to be about 2 million kelvins (K), so the gas particles are moving very quickly. But the gas is extremely diffuse, so the particles are not colliding with each other and transferring energy. The gas wouldn't feel hot, much like the solar corona that we discussed in Chapter 14.

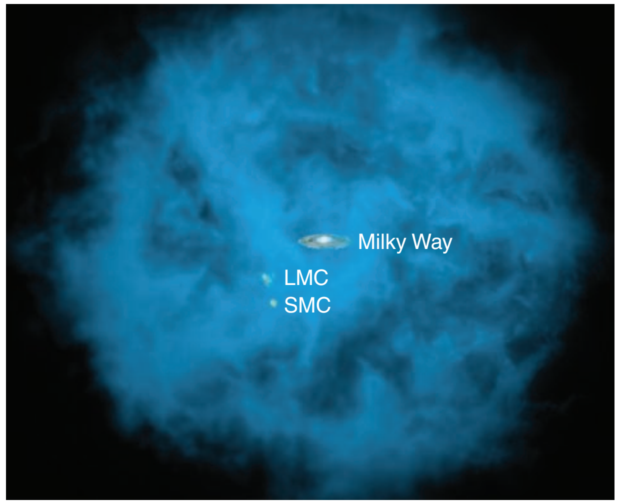


Figure 20.13 This artist's concept shows a hot gas halo surrounding the Milky Way, which may contain as much mass as that of all of the stars in the galaxy combined. The Large Magellanic Cloud (LMC) and Small Magellanic Cloud (SMC) are nearby dwarf galaxies.

Magnetic Fields and Cosmic Rays Fill the Galaxy

The interstellar medium of the Milky Way is laced with magnetic fields that are wound up and compressed by the rotation of the galaxy's disk. The total interstellar magnetic field, however, is about a hundred thousand times weaker than Earth's magnetic field. Charged particles and magnetic fields interact strongly; the particles spiral around magnetic fields, moving along the field rather than across it. Conversely, magnetic fields cannot freely escape from a cloud of gas containing even a small amount of charged particles. The dense clouds of interstellar gas in the midplane of the Milky Way (**Figure 20.14**), anchor the galaxy's magnetic field to the disk, in turn anchoring high-energy charged particles, known as cosmic rays, to the galaxy.

Cosmic rays are charged particles that originate in space and travel close to the speed of light. Despite their name, cosmic rays are not a form of electromagnetic radiation: they were named before their true nature was known. Most cosmic-ray particles are protons, but some are nuclei of helium, carbon, and other elements produced by nucleosynthesis. A few are high-energy electrons and other subatomic particles. Cosmic rays span an enormous range in particle

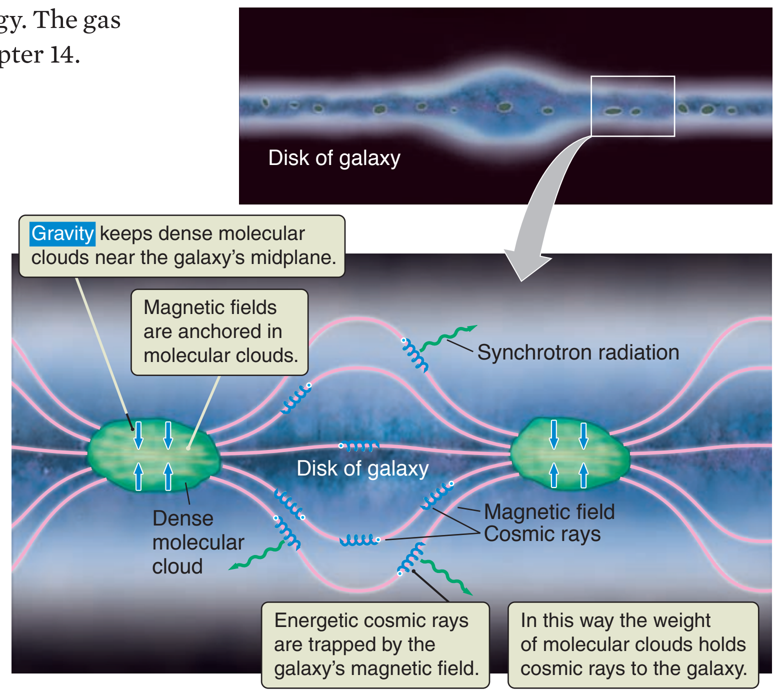


Figure 20.14 The weight of interstellar clouds anchors the magnetic field of the Milky Way to the disk of the galaxy. The magnetic field, in turn, traps the galaxy's cosmic rays.

energy. Astronomers can observe the lowest-energy cosmic rays with interplanetary spacecraft. These cosmic rays have energies as low as about 10^{-11} joule (J), which corresponds to the energy of a proton moving at a velocity of a few tenths the speed of light.

In contrast, the most energetic cosmic rays are 10 trillion (10^{13}) times as energetic as the lowest-energy cosmic rays, and they move very close to the speed of light, 0.999999*c*. These high-energy cosmic rays are continually hitting Earth and are detected from the showers of elementary particles that they cause when crashing through Earth's atmosphere. Astronomers hypothesize that most of these cosmic rays are accelerated to these incredible energies by the shock waves produced in supernova explosions. The very highest-energy cosmic rays are as much as a hundred million times more energetic than any particle ever produced in a particle accelerator on Earth. These extremely high energies make these cosmic rays much more difficult to explain than those with lower energies.

The disk of the galaxy glows from synchrotron radiation (see Chapter 17) produced by cosmic rays, mostly electrons, spiraling around the galaxy's magnetic field. Such synchrotron emission is seen in the disks of other spiral galaxies as well, indicating that they, too, have magnetic fields and populations of energetic cosmic rays. The very highest-energy cosmic rays are moving much too fast to be confined by the gravitational force of their originating galaxy. Any such cosmic rays that formed in the Milky Way would soon stream away from the galaxy into intergalactic space. Thus, it is likely that some of the energetic cosmic rays reaching Earth originated in energetic events outside of the Milky Way Galaxy.

The total energy of all the cosmic rays in the galactic disk can be estimated from the energy of the cosmic rays reaching Earth. The strength of the interstellar magnetic field can be measured by observing the effect that it has on the properties of radio waves passing through the interstellar medium. These measurements indicate that in the Milky Way Galaxy, the magnetic-field energy and the cosmic-ray energy are about equal to each other. Both are comparable to the energy present in other energetic components of the galaxy, including the motions of interstellar gas and the total energy of electromagnetic radiation within the galaxy.

CHECK YOUR UNDERSTANDING 20.2

What parts of the Milky Way contain old stars, and what parts contain young stars?

20.3 Most of the Milky Way Is Unseen

Nebraska Simulation: Milky Way Rotational Velocity

As in other galaxies, the most interesting parts of the Milky Way may be the parts that can't be seen directly but only detected by their gravitational influence on the stars around them. Dark matter accounts for the vast majority of the mass in a galaxy and extends far beyond a galaxy's visible boundary. As in all other spiral galaxies, there is compelling evidence that dark matter dominates the Milky Way. From radio and infrared observations, astronomers can figure out how the disk of the Milky Way moves, and from that motion, in turn, they can determine its mass. The supermassive black hole at the center of the Milky Way provides a different kind of observation problem. This object is also detected by its gravitational effects on the stars nearby but cannot be seen directly. In this section, we will explore what can be inferred about these two components of the Milky Way from the effect of their gravity on other objects.

Dark Matter in the Milky Way

The rotation of the disk of the Milky Way can be determined from observations of the relative velocities of interstellar hydrogen measured from 21-cm radiation. **Figure 20.15** shows how these velocities vary with viewing direction from the Sun. Looking toward the center of the galaxy, the gas is stationary relative to the Sun. Looking in the direction of the Sun's motion around the galactic center, hydrogen clouds appear to be moving toward Earth, while in the opposite direction, clouds are moving away from Earth. In other directions, the measured velocities are complicated by Earth's moving vantage point within the disk and so are more difficult to interpret at a glance. This is a pattern of the rotation velocity of gas in a disk like those you learned about in the past chapter. The only difference is that instead of looking at it from outside, we see the Milky Way Galaxy's rotation curve from a vantage point located within—and rotating with—the galaxy. Even so, observed velocities of neutral hydrogen enable astronomers to measure the Milky Way Galaxy's rotation curve and even determine the structure present throughout its disk.

Recall from the previous chapter that observations of rotation curves led astronomers to conclude that the masses of spiral galaxies consist mostly of dark matter. **Figure 20.16** shows the rotation curve of the Milky Way as inferred primarily from 21-cm observations. The orbital motion of the nearby dwarf galaxy called the Large Magellanic Cloud provides data for the outermost point in the rotation curve, at a distance of roughly 50,000 pc (160,000 light-years) from the center of the galaxy. Like other spiral galaxies, the Milky Way has a fairly flat rotation curve. The mass outside of the Sun's orbit does not greatly affect the Sun's orbit.

The total mass of the Milky Way Galaxy is currently estimated to be about 1.0 trillion to 1.5 trillion times the mass of the Sun. However, the luminous mass, estimated by adding the masses of stars, dust, and gas, is only about one-tenth

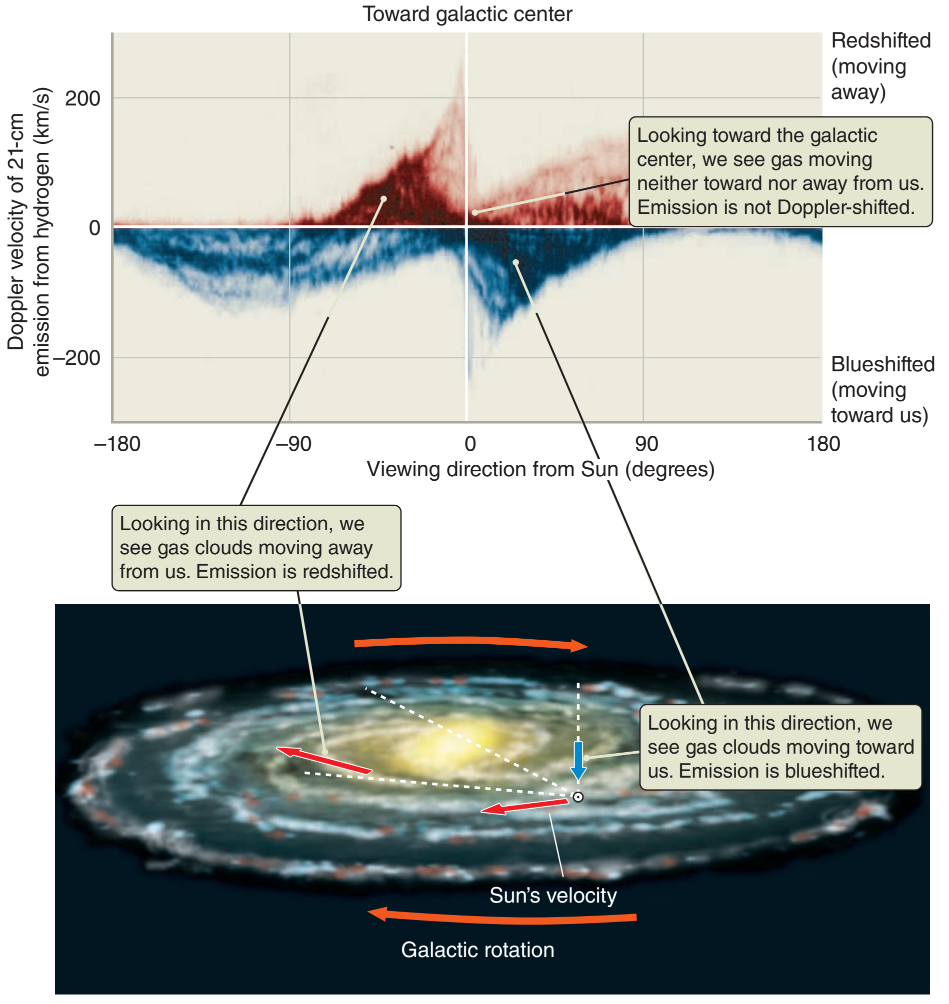


Figure 20.15 The graph shows Doppler velocities measured from observations of 21-cm emission from interstellar clouds of neutral hydrogen. These velocities vary between redshift and blueshift as an observer looks around in the plane of the Milky Way, along the sight lines shown as dashed white lines. From Earth's perspective within the Solar System, the signature of a rotating disk is clear from views of either side of the galactic center.

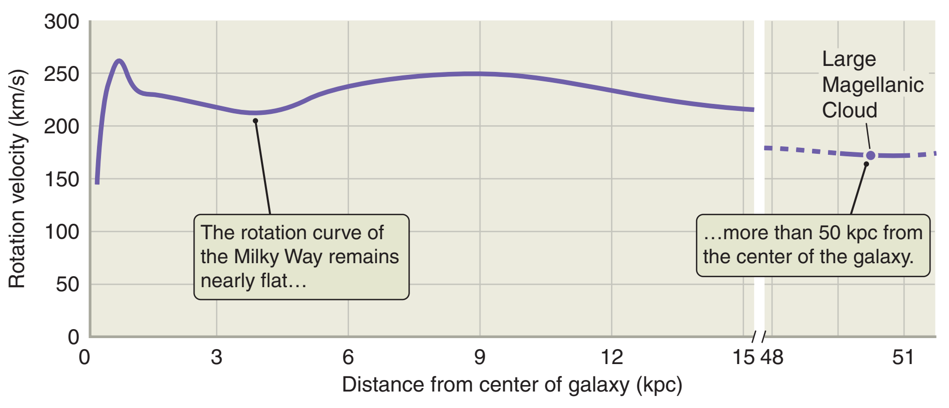


Figure 20.16 Rotation velocity is plotted against distance from the center of the Milky Way. The most distant point comes from measurements of the orbit of the Large Magellanic Cloud. The nearly flat rotation curve indicates that dark matter dominates the outer parts of the Milky Way. Part of this graph has been hidden in order to fit the Large Magellanic Cloud into the figure. This is indicated by the broken x-axis, marked by the diagonal hatch marks.

20.1 Working It Out The Mass of the Milky Way inside the Sun's Orbit

The Sun orbits about the center of the Milky Way Galaxy. Recall from Chapter 4 that even though the gravitational pull on the Sun comes from all of the material inside this orbit, Newton showed that we could treat the system as if all the mass were concentrated at the center. Then we can apply Newton's and Kepler's laws to calculate the mass of the Milky Way inside of the Sun's orbit. Newton's version of Kepler's third law relates the period of the orbit to the orbital radius and the masses of the objects. But in this case, the mass of the galaxy is much larger than the mass of the Sun, so the Sun's mass is negligible by comparison. In addition, what astronomers can *measure* is the orbital speed of the Sun or other stars about the galactic center, rather than the orbital period. Thus, we can use a rearranged form of the same equation from Working It Out 4.2 that we used to estimate the mass of the Sun from the orbit of Earth:

$$M = \frac{rv^2_{\text{circ}}}{G}$$

The Sun orbits the center of the galaxy at 220 kilometers per second (km/s). The distance of the Sun from the center of the galaxy is 8,300 pc, which converts to kilometers as

$$8{,}300 \text{ pc} \times (3.09 \times 10^{13} \text{ km/pc}) = 2.56 \times 10^{17} \text{ km}$$

Because we know that the gravitational constant $G = 6.67 \times 10^{-20}$ km^3/(kg s^2), we can calculate the mass of the portion of the Milky Way that is inside of the Sun's orbit:

$$M = \frac{(2.56 \times 10^{17} \text{ km}) \times (220 \text{ km/s})^2}{6.67 \times 10^{-20} \text{ km}^3/(\text{kg s}^2)}$$

$$M = 1.86 \times 10^{41} \text{ kg}$$

To put this in units of the Sun's mass, we divide this answer by $M_{\text{Sun}} = 1.99 \times 10^{30}$ kg, yielding

$$M = \frac{1.86 \times 10^{41} \text{ kg}}{1.99 \times 10^{30} \text{ kg}/M_{\text{Sun}}} = 9.35 \times 10^{10} M_{\text{Sun}}$$

The mass of the Milky Way inside of the Sun's orbit is about 94 billion times the mass of the Sun.

as much. Like other spiral galaxies, the Milky Way's mass consists mainly of dark matter. The spatial distribution of dark and normal matter within the Milky Way is also much like that of other galaxies, with dark matter dominating its outer parts (**Working it Out 20.1**).

The Supermassive Black Hole

Figure 20.17 shows images of the Milky Way's center taken with the Chandra X-ray Observatory and the Spitzer Space Telescope. The X-ray view (Figure 20.17a) shows the location of a strong radio source called Sagittarius A* (abbreviated

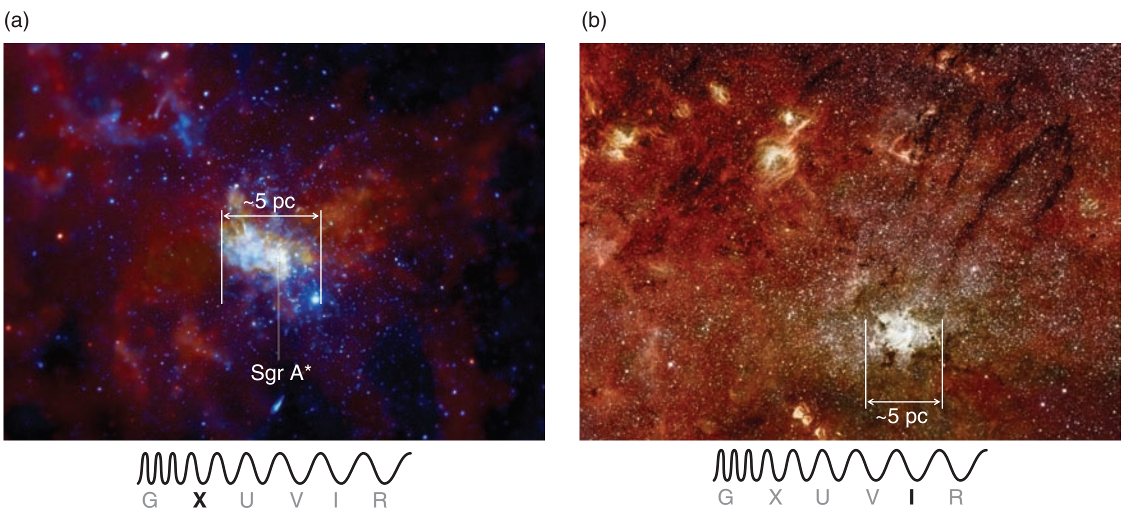


Figure 20.17 (a) This X-ray view of the Milky Way's central region shows the active source, Sagittarius A* (Sgr A*), as the brightest spot at the middle of the image. Lobes of superheated gas (shown in red) are evidence of recent, violent explosions happening near Sgr A*. (b) This infrared view of the central core of the Milky Way shows hundreds of thousands of stars. The bright white spot at the lower right marks Sgr A*, the location of the supermassive black hole.

Sgr A*), which lies at the center of the Milky Way. The infrared image (Figure 20.17b) cuts through the dust to reveal the galaxy's crowded, dense core containing hundreds of thousands of stars.

Studies of the motions of stars closest to the Sgr A* source suggest a central mass very much greater than that of the few hundred stars orbiting there. Furthermore, observations of the galaxy's rotation curve show rapid rotation velocities very close to the galactic center. Stars closer than 0.1 light-year from the galactic center follow Kepler's laws, indicating that their motion is dominated by mass within their orbit. The closest stars studied are only about 0.01 light-year from the center of the galaxy—so close that their orbital periods are only about a dozen years. The positions of these stars change noticeably over time, and astronomers can see them speed up as they whip around what can only be a supermassive black hole at the focus of their elliptical orbits, as shown in **Figure 20.18**. Using Newton's version of Kepler's third law, we can then estimate that the black hole at the center of the Milky Way Galaxy is a relative lightweight, having a mass of "only" 4 million times the mass of the Sun (**Working It Out 20.2**).

Clouds of interstellar gas at the galaxy's center are heated to millions of degrees by shock waves from supernova explosions and colliding stellar winds blown outward by young, massive stars. Superheated gas produces X-rays, and the Chandra X-ray Observatory has detected more than 9,000 X-ray sources within the central region of the galaxy. These include frequent, short-lived X-ray flares near Sgr A* (see Figure 20.17a), which provide direct evidence that matter falling toward the supermassive black hole fuels the energetic activity at the galaxy's center.

The Fermi Gamma-ray Space Telescope has observed gamma-ray-emitting bubbles that extend 8 kpc (25,000 light-years) above and below the galactic plane. The bubbles may have formed after a burst of star formation a few million years ago produced massive star clusters near the center of the galaxy. If some of the gas formed stars and about 2,000 M_{Sun} of material fell into the supermassive black hole, enough energy could have been released to power the bubbles. More recently, faint gamma-ray signals were observed that look like jets coming from the

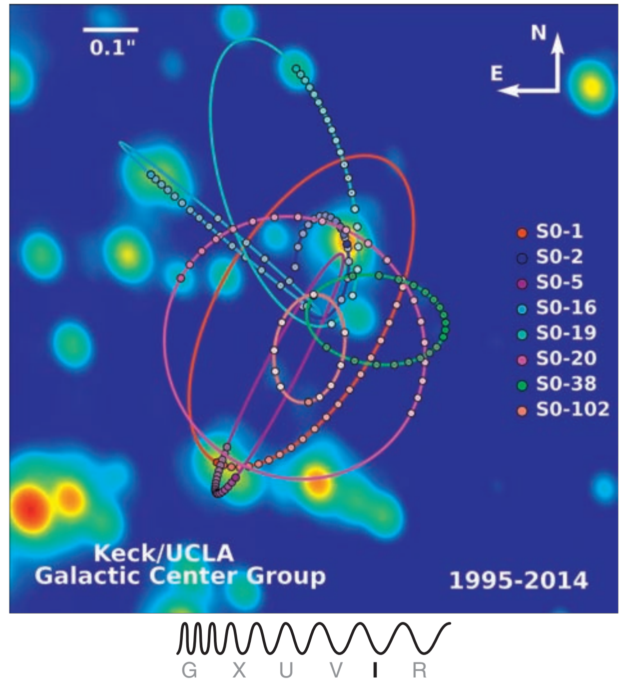


Figure 20.18 This figure shows orbits of seven stars within 0.03 pc (0.1 light-year, or about 6,000 astronomical units [AU]) of the Milky Way's center. The Keplerian motions of these stars reveal the presence of a 4-million-M_{Sun} supermassive black hole at the galaxy's center. Colored dots show the measured positions of each of the stars over many years: the dots progress from lighter in 1995 to darker in 2014.

20.2 Working It Out The Mass of the Milky Way's Central Black Hole

Figure 20.18 illustrates data points for the stars in the central region orbiting closely to the central black hole of the Milky Way Galaxy. These stars have highly elliptical orbits with changing speeds, but the orbital periods are short enough that they can be observed and measured. Star S0-2 in the figure has a measured orbital period of 15.8 years. The semimajor axis of its orbit is estimated to be 1.5×10^{11} km $= 1{,}000$ AU. With this information, we can use Newton's version of Kepler's third law to estimate the mass inside of S0-2's orbit. Setting up the equation as we did in Working It Out 13.4:

$$\frac{m_{\text{BH}}}{M_{\text{Sun}}} + \frac{m_{\text{S0-2}}}{M_{\text{Sun}}} = \frac{A_{\text{AU}}^3}{P_{\text{years}}^2}$$

The mass of star S0-2 is much less than the mass of the black hole, so the sum of the two is very close to the mass of the black hole. Therefore, we can write

$$\frac{m_{\text{BH}}}{M_{\text{Sun}}} = \frac{A_{\text{AU}}^3}{P_{\text{years}}^2} = \frac{1{,}000^3}{15.8^2} = 4.0 \times 10^6$$

$$m_{\text{BH}} = 4.0 \times 10^6\, M_{\text{Sun}}$$

The supermassive black hole at the center of the Milky Way has a mass 4 million times that of the Sun. This is quite a bit less than the billion-solar-mass black holes in some active galactic nuclei (AGNs) that we discussed in Chapter 19.

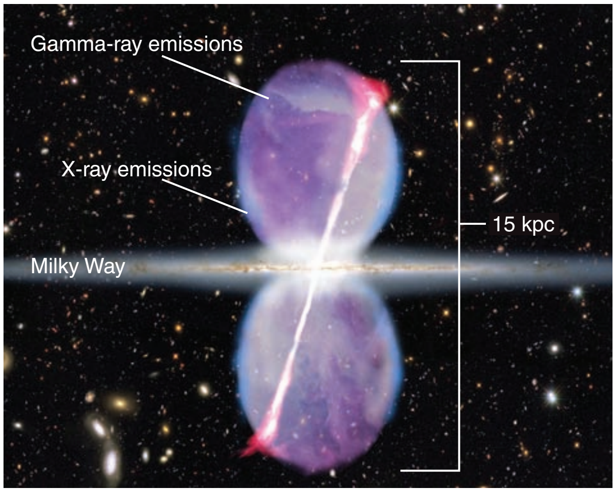


Figure 20.19 The Fermi Gamma-ray Space Telescope observed gamma-ray bubbles (purple) extending 8 kpc above and below the galactic plane. Hints of the edges of the bubbles were first observed in X-rays (blue) in the 1990s. In this artist's conceptual view from outside of the galaxy, the gamma-ray jets (pink) are tilted with respect to the bubbles, which might suggest that the accretion disk around the black hole is tilted as well.

center, within the bubbles, as shown in the artist's depiction in **Figure 20.19**. If these jets are originating from material falling into the supermassive black hole, activity might be even more recent—perhaps 20,000 years ago. Some astronomers predict that gas clouds are heading toward the center and will soon be accreted by the black hole. Currently, the observed activity is not as intense as that seen in active galaxies with central, supermassive black holes. The inner Milky Way is a reminder that it was almost certainly "active" in the past and could become active once again.

CHECK YOUR UNDERSTANDING 20.3

Which property is detectable for both dark matter and the supermassive black hole at the center of the Milky Way? (a) their luminosities; (b) their temperatures; (c) their gravity; (d) their composition

20.4 The History and Future of the Milky Way

One of the fundamental goals of stellar astronomy is to understand the life cycle of stars, including how stars form from clouds of interstellar gas. In Chapter 15, we told a fairly complete story of this process, at least as it occurs today, and tied this story strongly to observations of Earth's galactic neighborhood. Galactic astronomy has a similar basic goal. Astronomers would like to have a complete and well-tested theory of how the Milky Way formed and to be able to make predictions about its future. The distribution of stars of different ages with different amounts of heavy elements is one clue. Additional clues come from studying other galaxies at different distances (and therefore of different ages), their supermassive black holes, and their merger history. In this section, we explore the history and the future of the Milky Way.

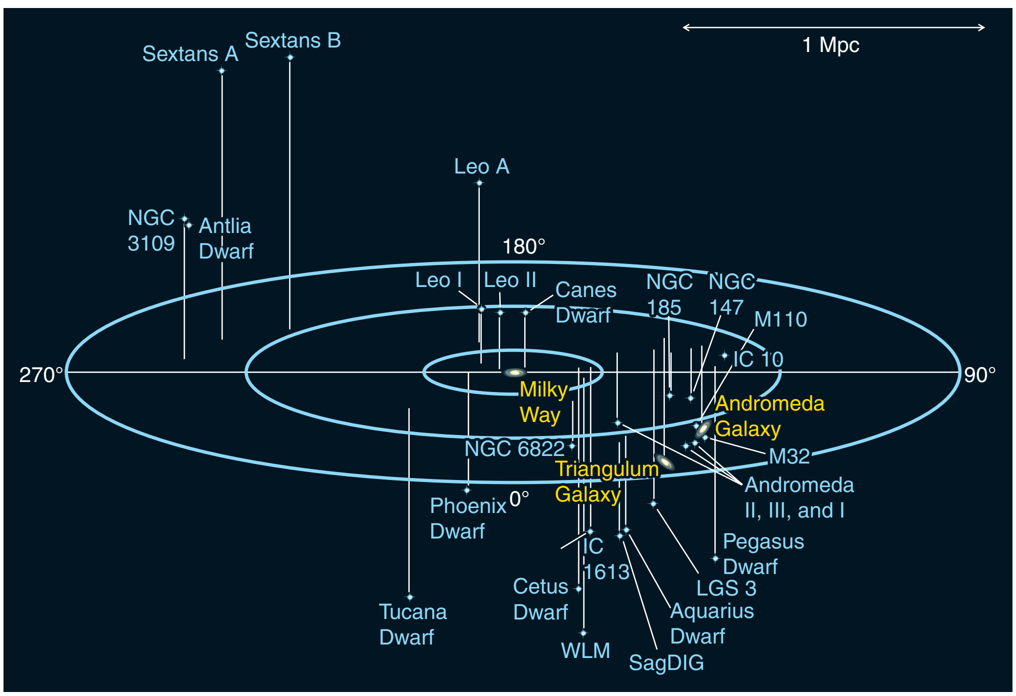


Figure 20.20 This graphical map shows some of the members of the Local Group of galaxies. Most are dwarf galaxies. Spiral galaxies are shown in yellow. The closest galaxies to the Milky Way are not seen on this scale. (1 Mpc = 1 megaparsec = 1 million parsecs.)

The Local Group

Galaxies do not exist in isolation. The vast majority of galaxies are parts of gravitationally bound collections of galaxies. The smallest and most common of these are called **galaxy groups**. A galaxy group contains as many as several dozen galaxies, most of them dwarf galaxies. As we saw in Chapter 1, the Milky Way is a member of the Local Group, first identified by Edwin Hubble in 1936. Hubble labeled 12 galaxies as part of the Local Group, but now astronomers count at least 50. As shown in **Figure 20.20**, the Local Group includes the two giant barred spirals—the Milky Way Galaxy and the Andromeda Galaxy—along with a few ellipticals and irregulars and at least 30 smaller dwarf galaxies in a volume of space about 3 million pc (10 million light-years) in diameter. Almost 98 percent of all the galaxy mass in the Local Group resides in just its two giant galaxies. The third largest galaxy, Triangulum, is an unbarred spiral with about one-tenth the mass of the Milky Way or Andromeda. Most but not all of the dwarf elliptical and dwarf spheroidal galaxies in the group are satellites of the Milky Way or Andromeda. The Local Group interacts with a few nearby groups, which will be discussed further in Chapter 23.

There are more than 20 of these satellite dwarf galaxies, although it's not certain that all are gravitationally bound to the Milky Way. Some of the fainter dwarf galaxies were discovered only very recently because of their low luminosity. The dwarf galaxies are the lowest-mass galaxies observed, and they are dominated by an even greater percentage of invisible dark matter than are other known galaxies. They also contain stars very low in elements more massive than helium. These ultrafaint dwarf galaxies offer clues to the formation of the Local Group. In addition, observations of the motions and speeds of the dwarf galaxies about the Milky Way may lead to new estimates of the dark matter mass within the Milky Way itself.

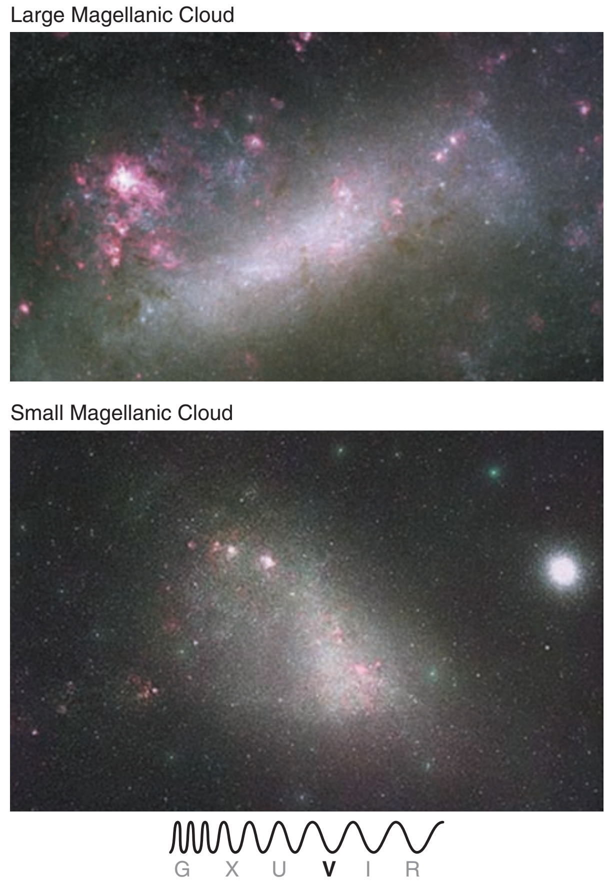


Figure 20.21 The Milky Way is surrounded by more than 20 dwarf companion galaxies: the largest among them are the Large Magellanic Cloud (LMC) and Small Magellanic Cloud (SMC) (see Figure 20.13).

The Formation of the Milky Way

We have seen that globular clusters and high-velocity stars must have been among the first stars formed in the Milky Way that still exist. The fact that they are not concentrated in the disk or bulge of the galaxy indicates that they formed from clouds of gas well before those clouds settled into the galaxy's disk. This hypothesis is supported by observations that globular clusters are very old and that the youngest globular cluster is older than the oldest disk stars. The presence of extremely small amounts of massive elements in the atmospheres of halo stars also indicates that at least one generation of stars must have lived and died *before* the formation of the halo stars visible today. This line of reasoning implies that the Milky Way formed from the merger of a number of smaller clumps of matter, which included both stars and clouds of dust and gas.

Combining that conclusion with the presence of the central, supermassive black hole and the number of nearby dwarf galaxies, astronomers conclude that the Milky Way must have formed when the gas within a huge "clump" of dark matter collapsed into a large number of small protogalaxies. Some of these protogalaxies then merged to form the large, barred spiral galaxies in the Local Group, but some of these smaller protogalaxies are still around today in the form of the small, satellite dwarf galaxies near the Milky Way. The largest among them are the Large Magellanic Cloud and the Small Magellanic Cloud (**Figure 20.21**), which are easily seen by the naked eye in the Southern Hemisphere and appear much like detached pieces of the Milky Way. The Magellanic Clouds were named for Ferdinand Magellan (circa 1480–1521), who headed the first European expedition that ventured far enough into the Southern Hemisphere to see them.

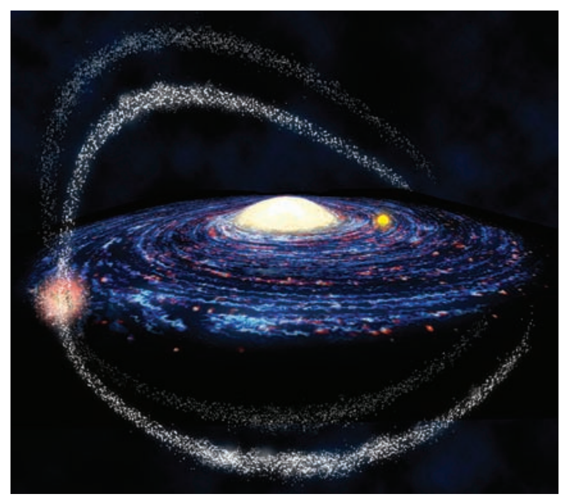


Figure 20.22 This artist's impression shows tidal tails of stars from the Sagittarius Dwarf elliptical galaxy (reddish-orange). These stars have been stripped from the dwarf galaxy by the much more massive Milky Way, and the two galaxies will eventually merge in billions of years.

The Future of the Milky Way

Mergers and collisions of Local Group galaxies continue today. Among the closest companions to the Milky Way is the Sagittarius Dwarf Galaxy, which is plowing through the disk of the Milky Way on the other side of the bulge. Astronomers have observed streams of stars, as sketched in **Figure 20.22**, from Sagittarius

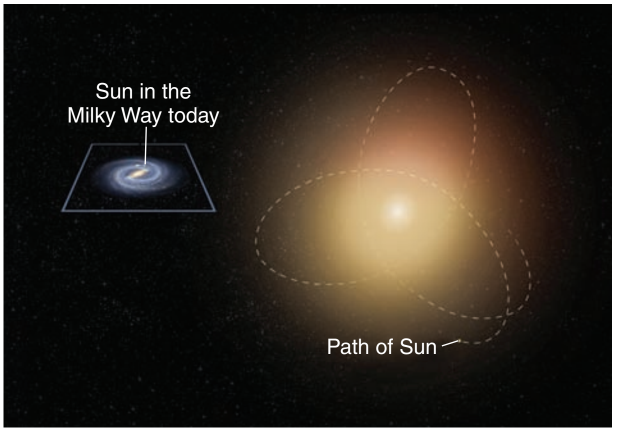


Figure 20.23 A computer simulation of the orbit of the Sun within the Milky Way–Andromeda merger remnant 10 billion years from now.

Dwarf and some of the other dwarf galaxies that are being tidally disrupted by the Milky Way. These dwarf galaxies will become incorporated into the Milky Way—an indication that the galaxy is still growing. Computer simulations suggest that such mergers could cause the spiral-arm structure.

The Andromeda Galaxy appears to violate Hubble's law because its spectrum shows blueshifts, indicating that the galaxy is moving toward, not away from, the Milky Way at 110 km/s (400,000 km/h). Andromeda and the Milky Way are the two largest galaxies in the Local Group, about 770 kpc (2.5 million light-years) apart. (If each galaxy were the size of a quarter, they would be about an arm's length apart.) Since Edwin Hubble noticed nearly 100 years ago that Andromeda's spectrum was blueshifted, indicating that it was approaching the Milky Way, astronomers have wondered if and when Andromeda and the Milky Way would collide and form a merged galaxy.

Astronomers used the Hubble Space Telescope to measure the perpendicular motion of Andromeda and determine whether a collision will be head-on, partial, or a total miss. They concluded that the two galaxies will "collide" head-on in about 4 billion years, although because most of a galaxy is space between the stars, actual collisions between the stars themselves are quite unlikely: while diffuse gas in the interstellar medium will collide and cause hot shocks, most of the material of one galaxy will pass through the other. It will take another 2 billion years for the two galaxies to merge completely and form one giant elliptical galaxy. The third spiral in the Local Group, the Triangulum Galaxy, may merge first with the Milky Way, or with Andromeda, or be ejected from the Local Group.

Note the timing here. This first close encounter with Andromeda in 4 billion years will occur *before* the Sun runs out of hydrogen in its core, although it is likely that the Sun will have increased its luminosity enough by then that Earth's habitability will have been affected. In 6 billion years, the Sun and its planets could end up near the center of the merged galaxy, or more likely they will have a new location farther from the center with a different orbit and in a different stellar neighborhood (**Figure 20.23**). If by chance the encounter leads to a star passing close to the Sun, the orbits of the Solar System planets could be disrupted, causing them to be at different distances from the Sun. **Figure 20.24**

Figure 20.24 This computer simulation depicts the view in the sky on Earth when the Andromeda and Milky Way galaxies collide. The spiral Andromeda appears larger in the sky as it gets closer, and then the sky becomes increasingly bright as the two collide.

shows what this merged galaxy might look like in Earth's sky—if anyone is still around to see it!

CHECK YOUR UNDERSTANDING 20.4

Why is Andromeda now moving toward us?

Origins

The Galactic Habitable Zone

In Chapter 19, we discussed the concept of galactic habitable zones in general; in this Origins, we will focus more specifically on ideas about the habitable zone of the Milky Way. Stars that are situated too far from the galactic center may have protoplanetary disks with insufficient quantities of heavy elements—such as oxygen, silicon, iron, and nickel (to make up rocky planets like Earth) or carbon, nitrogen, and oxygen (to make up the molecules of life). Stars that are too close to the galactic center may have planets that are too strongly affected by its high-energy radiation environment (X-rays and gamma rays from the supermassive black hole) and by supernova explosions and gamma-ray bursts (GRBs). The bulge has a higher density of stars creating a strong radiation field, and the halo and thick disk have stars with lower amounts of heavy elements, so perhaps only stars in the thin disk of the galaxy are candidates for residence in a galactic habitable zone.

Astronomers must also consider stellar lifetimes versus the 4 billion years after the formation of Earth that it took for life to evolve into land animals—so only stars with masses low enough that they will live at least 4 billion years on the main sequence are considered potential hosts of more complex life. In this simple model based on the evolution times for life on Earth, the galactic habitable zone could be a doughnut-shaped region around the galactic center. In one version of the model, this zone is estimated to contain stars born 4 billion to 8 billion years ago located between 7 and 9 kpc from the galactic center, with the Sun exactly in the middle of the doughnut. The doughnut would grow larger over time as heavier-element formation spread outward from the galactic center.

Some have conducted additional studies and proposed more complex models. For example, one group initiated a search for the molecule formaldehyde (H_2CO)—a key "prebiotic" molecule—by observing molecular clouds in the outer parts of the Milky Way, 12–23.5 kpc from the galactic center. Formaldehyde was detected in two-thirds of the group's sample of 69 molecular clouds, suggesting that at least one important prebiotic molecule is available far from the galactic center.

Another computer model took into account details of the evolution of individual stars within the Milky Way, including birth rates, locations, distribution within the galaxy, abundances of heavy elements, stellar masses, main-sequence lifetimes, and the likelihood that stars became or will become supernovae. The model also assumed 4 billion years to the development of complex life. One result of this model is that stars in the inner part of the Milky Way are more likely to be affected by supernova explosions, but these stars are even more likely to have the heavier elements for the formation of planets. So in this model, the inner part of the galaxy, about 2.5–4 kpc from the center, in and near the midplane of the thin disk, is the most likely place for habitable planets. In this case, the Sun is *not* in the middle of the most probable zone for habitable planets. As more extrasolar planets are discovered, astronomers will have a better idea of their distribution throughout the Milky Way.

Nebraska Simulation: Milky Way Habitability Explorer

As noted, mergers with other galaxies could cause stars to migrate into or out of the galactic habitable zone. The uncertainties increase as the assumptions in these models move from the astronomical to the biological. For example, maybe life evolved faster on other planets, so stars of higher mass and shorter lifetimes should be included. Or maybe life evolved slower elsewhere, in which case the older stars would be the best candidates. It is unknown whether intense radiation from a supernova or a gamma-ray burst would permanently sterilize a planet or only affect evolution for a while. For example, if Earth's ozone layer was temporarily destroyed, life on land might die out, but life in the oceans would continue. Habitability in the Milky Way Galaxy is complicated, with many unanswered questions.

READING ASTRONOMY NEWS

ARTICLES QUESTIONS

In this article, astronomers report on their estimate of the amount of dark matter in the Milky Way Galaxy.

Dark Matter Half What We Thought, Say Scientists

International Centre for Radio Astronomy Research

A new measurement of dark matter in the Milky Way has revealed there is half as much of the mysterious substance as previously thought.

Australian astronomers used a method developed almost 100 years ago to discover that the weight of dark matter in our own galaxy is 800,000,000,000 (or 8×10^{11}) times the mass of the Sun.

They probed the edge of the Milky Way, looking closely, for the first time, at the fringes of the galaxy about 5 million trillion kilometers from Earth.

Astrophysicist Dr. Prajwal Kafle, from The University of Western Australia node of the International Centre for Radio Astronomy Research, said we have known for a while that most of the universe is hidden.

"Stars, dust, you and me, all the things that we see, only make up about 4 percent of the entire universe," he said.

"About 25 percent is dark matter and the rest is dark energy."

Dr. Kafle, who is originally from Nepal, was able to measure the mass of the dark matter in the Milky Way by studying the speed of stars throughout the galaxy, including the edges, which had never been studied to this detail before.

He used a robust technique developed by British astronomer James Jeans in 1915—decades before the discovery of dark matter.

Dr. Kafle's measurement helps to solve a mystery that has been haunting theorists for almost two decades.

"The current idea of galaxy formation and evolution, called the Lambda Cold Dark Matter theory, predicts that there should be a handful of big satellite galaxies around the Milky Way that are visible with the naked eye, but we don't see that," Dr. Kafle said.

"When you use our measurement of the mass of the dark matter the theory predicts that there should only be three satellite galaxies out there, which is exactly what we see; the Large Magellanic Cloud, the Small Magellanic Cloud, and the Sagittarius Dwarf Galaxy."

University of Sydney astrophysicist Professor Geraint Lewis, who was also involved in the research, said the missing satellite problem had been "a thorn in the cosmological side for almost 15 years."

"Dr. Kafle's work has shown that it might not be as bad as everyone thought, although there are still problems to overcome," he said.

The study also presented a holistic model of the Milky Way, which allowed the scientists to measure several interesting things such as the speed required to leave the galaxy.

"Be prepared to hit 550 kilometers per second if you want to escape the gravitational clutches of our galaxy," Dr. Kafle said.

"A rocket launched from Earth needs just 11 kilometers per second to leave its surface, which is already about 300 times faster than the maximum Australian speed limit in a car!"

ARTICLES QUESTIONS

1. How far from the center of the galaxy in parsecs were the measured stars?
2. Where would this measurement fit on Figure 20.16?
3. How does this mass of dark matter compare with the estimate of mass inside of the Sun's orbit?
4. Where is this additional mass located?
5. Explain how a new estimate of the mass of the Milky Way affects the calculation of the escape velocity from the Milky Way.

Summary

Astronomers compare images and spectra of other galaxies to observations of the Milky Way. From this, they determine that the Milky Way is a barred spiral of type SBbc. The Milky Way formed from a collection of smaller protogalaxies that collapsed out of a halo of dark matter. The idea of a galactic habitable zone is that certain parts of the Milky Way may be more suitable for the existence of habitable planets. This zone would have enough heavy elements for the formation of rocky planets and organic molecules, but not too much radiation that it would damage any life.

LG 1 **Explain how astronomers discovered the size and spiral structure of the Milky Way.** The distances to globular clusters can be found from the luminosity of variable stars within them. Because these globular clusters are symmetrically distributed around the center of the Milky Way, the center of the distribution is located at the center of the galaxy. The Sun is located about 8,300 pc (27,000 light-years) from the Milky Way's center, and the Milky Way is 30,000 pc (100,000 light-years) across.

LG 2 **List the clues of galaxy formation that can be found from the components of the Milky Way.** The chemical composition of the Milky Way has evolved with time, as material cycles between stars and the interstellar medium. There must have been a generation of stars before the oldest halo and globular-cluster stars we see today formed. The Milky Way has a disk consisting of two parts, the thick disk of old stars and the thin disk of young stars, implying that gas and dust from merging galaxies settled onto the disk, while the stars passed through. The galactic halo consists of an inner halo and outer halo of stars and globular clusters, as well as a large, hot gas halo. The abundance of heavy elements in the Milky Way has increased with time as each generation of stars has produced more of these elements during the final phases of the stars' lives.

LG 3 **Explain the evidence for the dark matter halo and for the supermassive black hole at the center of the Milky Way.** The Doppler velocities of radio spectral lines show that the rotation curve of the Milky Way is flat, like those of other galaxies. But the inferred mass cannot be accounted for by the mass that is observed directly. This indicates that the Milky Way's mass is mostly in the form of dark matter. Evidence for the black hole at the center of the galaxy includes rapid orbital velocities of nearby stars and symmetric X-ray and gamma-ray outflows of material.

LG 4 **Describe the Local Group of galaxies and how it provides clues about the evolution of the Milky Way.** The Milky Way is part of the Local Group of galaxies, which consists of two large, barred spirals and several dozen smaller galaxies. Collisions and mergers between these galaxies likely happened in the past, and a merger with the Andromeda Galaxy may be part of the Milky Way's future. The dwarf satellites and other neighbors in the Local Group are evidence that the Milky Way is growing through accretion.

? UNANSWERED QUESTIONS

- Are there many more ultrafaint dwarf galaxies than have so far been detected? These types of dwarf galaxies are so faint that they are hard to detect even when close to the Milky Way. As we will discuss in Chapter 23, the giant spirals such as the Milky Way were built up by mergers of these small, faint galaxies, so models predict that there should be hundreds or even thousands of them. They could be so dominated by dark matter that they are not at all visible, or too small ever to have formed stars, or they might have merged with the Milky Way or other Local Group members long ago.
- Will the merger of the Milky Way and Andromeda form a quasar at the center of the new elliptical galaxy? If the dust and gas at the center of the galaxy did not completely block the view from Earth, such a quasar could be brighter than the full Moon. The supermassive black hole at the center of Andromeda is thought to be 25–50 times larger than the one in the Milky Way, but both black holes combined still would be on the lower end of the masses of the black holes in AGNs discussed in Chapter 19.

Questions and Problems

Test Your Understanding

1. The size of the Milky Way is determined from studying __________ stars in globular clusters.
   a. Cepheid variable
   b. blue supergiant
   c. RR Lyrae
   d. Sun-like
2. Detailed observations of the structure of the Milky Way are difficult because
   a. the Solar System is embedded in the dust and gas of the disk.
   b. the Milky Way is mostly dark matter.
   c. there are too many stars in the way.
   d. the galaxy is rotating too fast (about 200 km/s).
3. Older stars are found farther from the midplane of a galactic disk because
   a. the disk used to be thicker.
   b. the stars have lived long enough to move there.
   c. the younger stars in the thick disk were more massive and have already died.
   d. they passed through the disk when the galaxy formed.
4. The magnetic field of the Milky Way has been detected by
   a. synchrotron radiation from cosmic rays.
   b. direct observation of the field.
   c. its interaction with Earth's magnetic field.
   d. studying molecular clouds.
5. Evidence of a supermassive black hole at the center of the Milky Way comes from
   a. direct observations of stars that orbit it.
   b. visible light from material that is falling in.
   c. strong radio emission from the black hole itself.
   d. streams of cosmic rays from the center of the galaxy.
6. The Large Magellanic Cloud and Small Magellanic Cloud will likely
   a. become part of the Milky Way.
   b. remain orbiting forever.
   c. become attached to another passing galaxy.
   d. escape from the gravity of the Milky Way.
7. Globular clusters are important to understanding the Milky Way because
   a. they are so young that they provide information about current star formation.
   b. they provide information about dwarf ellipticals from which the Milky Way formed.
   c. they reveal the size of the Milky Way and Earth's location in it.
   d. the stars in them are highly enhanced in metals.
8. A globular cluster with no variable stars would have been left out of Shapley's study because
   a. the cluster would be too far away.
   b. the distance to the cluster could not be determined.
   c. the cluster would be too faint to see.
   d. the cluster would be too young to determine its evolutionary state.
9. The best evidence for the presence of dark matter in the Milky Way comes from the observation that the rotation curve
   a. is flat at great distances from the center.
   b. rises swiftly in the interior.
   c. falls off and then rises again.
   d. has a peak at about 2,000 light-years from the center.
10. Cosmic rays are
    a. a form of electromagnetic radiation.
    b. high-energy particles.
    c. high-energy dark matter.
    d. high-energy photons.
11. What kind of galaxy is the Milky Way?
    a. elliptical
    b. spiral
    c. barred spiral
    d. irregular
12. Where are the youngest stars in the Milky Way Galaxy?
    a. in the core
    b. in the bulge
    c. in the disk
    d. in the halo
13. Halo stars are found in the vicinity of the Sun. What observational evidence does not distinguish them from disk stars?
    a. the direction of their motion
    b. their speed
    c. their composition
    d. their temperature
14. Why are most of the Milky Way's satellite galaxies so difficult to detect?
    a. They are very small.
    b. They are very far away.
    c. The halo of the Milky Way blocks the view.
    d. They are very faint.
15. The concept of a galactic habitable zone does not consider
    a. the radiation field.
    b. the ages of stars.
    c. the amount of heavy elements.
    d. the distance of a planet from its central star.

Thinking about the Concepts

16. Scientists can never know everything, especially at the beginning of a set of research programs. For example, Shapley did not know about dust in the Milky Way. Yet, scientists must often set aside this problem of the unknowns they don't know and do the best they can with the knowledge they have. Explain why this is a necessary step toward scientific understanding.

17. Describe the distribution of globular clusters within the Milky Way, and explain what that implies about the size of the galaxy and our distance from its center.

18. In which parts of the Milky Way do astronomers find open clusters? In which parts do they find globular clusters?

19. Old stars in the inner disk of the Milky Way have higher abundances of massive elements than those of young stars in the outer disk. Explain how this difference might have developed.

20. How do 21-cm radio observations reveal the rotation of the Milky Way Galaxy?

21. What does the rotation curve of the Milky Way indicate about the presence of dark matter in the galaxy?

22. Halo stars are found in the vicinity of the Sun. What observational evidence distinguishes them from disk stars?

23. What is one source of synchrotron radiation in the Milky Way, and where is it found?

24. Why must astronomers use X-ray, infrared, and 21-cm radio observations to probe the center of the galaxy?

25. What is Sgr A* and how was it detected?

26. Explain the evidence for a supermassive black hole at the center of the Milky Way. How does the mass of the supermassive black hole at the center of our galaxy compare with that found in most other spiral galaxies?

27. To observers in Earth's Southern Hemisphere, the Large Magellanic Cloud and Small Magellanic Cloud look like detached pieces of the Milky Way. What are these "clouds," and why is it not surprising that they look so much like pieces of the Milky Way?

28. What is the origin of the Milky Way's satellite galaxies? What has been the fate of most of the Milky Way's satellite galaxies? Why are most of the Milky Way's satellite galaxies so difficult to detect?

29. Use your imagination to describe how Earth's skies might appear if the Sun and Solar System were located (a) near the center of the galaxy; (b) near the center of a large globular cluster; (c) near the center of a large, dense molecular cloud.

30. What factors do astronomers consider when thinking about a galactic habitable zone in the Milky Way?

Applying the Concepts

31. From Figure 20.8, estimate the ratio between the radius of the Milky Way's outer halo and the radius of the disk.

32. Study Figure 20.16.
    a. What is the rotation velocity of a disk star located 6,000 pc from the center of the Milky Way?
    b. Assuming a circular orbit, how long does it take that star to orbit once?

33. From the data in Figure 20.16, estimate the time it would take the Large Magellanic Cloud to orbit the Milky Way if the Large Magellanic Cloud were on a circular orbit.

34. The Sun completes one trip around the center of the galaxy in approximately 230 million years. How many times has the Solar System made the circuit since its formation 4.6 billion years ago?

35. The Sun is located about 8,300 pc from the center of the galaxy, and the galaxy's disk probably extends another 9,000 pc farther out from the center. Assume that the Sun's orbit takes 230 million years to complete.
    a. With a truly flat rotation curve, how long would it take a globular cluster located near the edge of the disk to complete one trip around the center of the galaxy?
    b. How many times has that globular cluster made the circuit since its formation about 13 billion years ago?

36. Parallax measurements of the variable star RR Lyrae indicate that it is located 230 pc from the Sun. A similar star observed in a globular cluster located far above the galactic plane appears 160,000 times fainter than RR Lyrae.
    a. How far from the Sun is this globular cluster?
    b. What does your answer to part (a) tell you about the size of the galaxy's halo compared to the size of its disk?

37. How do astronomers know that the center of the Milky Way contains a black hole and not just dark matter?

38. Given what you have learned about the distribution of massive elements in the Milky Way and what you know about the terrestrial planets, where do you think such planets are most likely and least likely to form?

39. A cosmic-ray proton is traveling at nearly the speed of light (3×10^8 m/s).
    a. Using Einstein's familiar relationship between mass and energy ($E = mc^2$), show how much energy (in joules) the cosmic-ray proton would have if m were based only on the proton's rest mass (1.7×10^{-27} kg).
    b. The actual measured energy of the cosmic-ray proton is 100 J. What, then, is the relativistic mass of the cosmic-ray proton?
    c. How much greater is the relativistic mass of this cosmic-ray proton than the mass of a proton at rest?

40. One of the fastest cosmic rays ever observed had a speed of $(1.0 - [1.0 \times 10^{-24}]) \times c$ (very, very close to c). Assume that the cosmic ray and a photon left a source at the same instant. To a stationary observer, how far behind the photon would the cosmic ray be after traveling for 100 million years?

41. Consider a black hole with a mass of 5 million M_{Sun} ($M_{Sun} = 2 \times 10^{30}$ kg). A star's orbit about the black hole has a semimajor axis of 0.02 light-year (1.9×10^{14} meters). Calculate the star's orbital period. (Hint: You may want to refer back to Chapter 4.)

42. A star in a circular orbit about the black hole at the center of the Milky Way (whose mass $M_{BH} = 8 \times 10^{36}$ kg) has an orbital radius of 0.0131 light-year (1.24×10^{14} meters). What is the average speed of this star in its orbit? (Hint: You may want to refer back to Chapter 4.)

43. What is the Schwarzschild radius of the black hole at the center of the Milky Way? What is its density? How does this compare with the density of a stellar black hole?

44. A star is observed in a circular orbit about a black hole with an orbital radius of 1.5×10^{11} km and an average speed of 2,000 km/s. What is the mass of this black hole in solar masses?

45. One model of the galactic habitable zone contains stars in a doughnut-shaped region between 7 and 9 kpc from the center of the galaxy, in the disk. Assuming that this doughnut is as thick as the disk itself, what fraction of the disk of the Milky Way lies in this habitable zone?

USING THE WEB

46. Go to the "Night Sky" Web page of the National Park Service (NPS; http://nature.nps.gov/night). What is a "natural lightscape"? Click on and read "Light Pollution" and "Measuring Lightscapes." Why is it becoming more rare for people to see the Milky Way? Why does the NPS consider viewing the Milky Way an important part of the experience for people visiting the park?

47. a. Go to the Astronomy Picture of the Day (APOD) app or website for July 2, 2012, and watch the video clip "Zooming into the Center of the Milky Way" (http://apod.nasa.gov/apod/ap120702.html). Why is there a shift in wavelengths of the selected pictures? On APOD, run a search for "Milky Way" to look at some of the best photographs of it. Where were the pictures taken from? Can you see the Milky Way from *your* location on a clear night?

    b. Why does the galactic center have to be observed in infrared and X-ray wavelengths? Go to the "Milky Way Galaxy" page of the Chandra X-ray telescope website (http://chandra.harvard.edu/photo/category/milkyway.html) and to the Spitzer infrared telescope website (http://spitzer.caltech.edu). Are there any new images of the galactic center? What has been learned?

48. Go to the websites of the two main groups studying the supermassive black hole at the galactic center: the UCLA Galactic Center Group (http://astro.ucla.edu/~ghezgroup/gc) and the Galactic Center Research group at the Max Planck Institute for Extraterrestrial Physics (http://mpe.mpg.de/ir/GC). Watch some of the time-lapse animations of the stars orbiting something unseen. Why is it assumed that the unseen object is a black hole? What new results are these groups reporting?

49. Citizen science: Go to the Milky Way Project website (http://milkywayproject.org). This project examines Spitzer telescope observations of the dusty material in the galaxy. Log in with your Zooniverse account name, and read the information under Menu: "Science," "FAQ," and "Classify." Participants in this project have already discovered some of these bubbles (http://spitzer.caltech.edu/images/4938-sig12-002-Finding-Bubbles-in-the-Milky-Way). Classify some images.

50. Go to the Hubble Space Telescope website and watch the videos about the possible collision of the Milky Way and Andromeda (http://hubblesite.org/newscenter/archive/releases/galaxy/2012/20/video). Read the report under "The Full Story." A newer simulation is here: www.icrar.org/multimedia/videos/video-pages/andromeda-and-the-milky-way-collide! Why will Andromeda "eat" the Milky Way instead of the other way around?

smartwork5

If your instructor assigns homework in Smartwork5, access your assignments at digital.wwnorton.com/astro5.

digital.wwnorton.com/astro5

Adapted from *Learning Astronomy by Doing Astronomy* by Ana Larson.

Astronomers once thought that the Sun was at the center of the Milky Way. In this Exploration, you will repeat Harlow Shapley's globular-cluster experiment that led to a more accurate picture of the size and shape of the Milky Way.

Imagine that the disk of the Milky Way is a flat, round plane, like a pizza. Globular clusters are arranged in a rough sphere around this plane. To map globular clusters on **Figure 20.25**, imagine that a line is drawn straight "down" from a globular cluster to the plane of the Milky Way. The "projected distance" in kiloparsecs is the distance from the Sun to the place where the line hits the plane. The galactic longitude indicates the direction toward that point; it is marked around the outside of the graph, along with the several constellations.

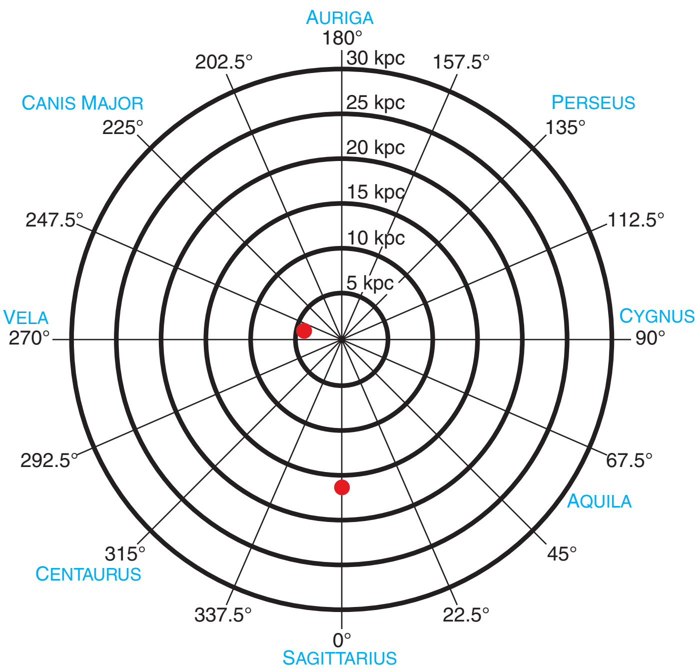


Figure 20.25 This polar graph can be used to plot distance and direction.

1 What is the approximate distance from the Sun to the center of the Milky Way?

---

2 What is the galactic longitude of the center of the Milky Way?

---

3 How do astronomers know that the Sun is not at the center of the Milky Way?

Make a dot at the location of each globular cluster by finding the galactic longitude indicated outside the circle and then coming in toward the center to the projected distance. The two globular clusters in boldface in **Table 20.1** have been plotted for you as examples. After plotting all of the globular clusters, estimate the center of their distribution and mark it with an X. This is the center of the Milky Way.

TABLE 20.1 Globular Cluster Data

Cluster	Galactic Longitude	Projected Distance (kpc)	Cluster	Galactic Longitude	Projected Distance (kpc)
104	306	3.5	6273	357	7
362	302	6.6	**6287**	**0**	**16.6**
2808	283	8.9	6333	5	12.6
4147	**251**	**4.2**	6356	7	18.8
5024	333	3.4	6397	339	2.8
5139	309	5	6535	27	15.3
5634	342	17.6	6712	27	5.7
Pal 5	1	24.8	6723	0	7
5904	4	5.5	6760	36	8.4
6121	351	4.1	Pal 10	53	8.3
O 1276	22	25	Pal 11	32	27.2
6638	8	15.1	6864	20	31.5
6171	3	15.7	6981	35	17.7
6218	15	6.7	7089	54	9.9
6235	359	18.9	Pal 12	31	25.4
6266	353	11.6	288	147	0.3
6284	358	16.1	1904	228	14.4
6293	357	9.7	Pal 4	202	30.9
6341	68	6.5	4590	299	11.2
6366	18	16.7	5053	335	3.1
6402	21	14.1	5272	42	2.2
6656	9	3	5694	331	27.4
6717	13	14.4	5897	343	12.6
6752	337	4.8	6093	353	11.9
6779	62	10.4	6541	349	3.9
6809	9	5.5	6626	7	4.8
6838	56	2.6	6144	352	16.3
6934	52	17.3	6205	59	4.8
7078	65	9.4	6229	73	18.9
7099	27	9.1	6254	15	5.7

21 The Expanding Universe

The cosmological principle has been at the center of astronomers' conceptual understanding of the universe. An important prediction of the cosmological principle is that the conclusions we reach about our universe should be more or less the same, regardless of whether we live in the Milky Way or in a galaxy billions of light-years away. In this chapter, we look at the evidence for the cosmological principle and for the Big Bang.

LEARNING GOALS

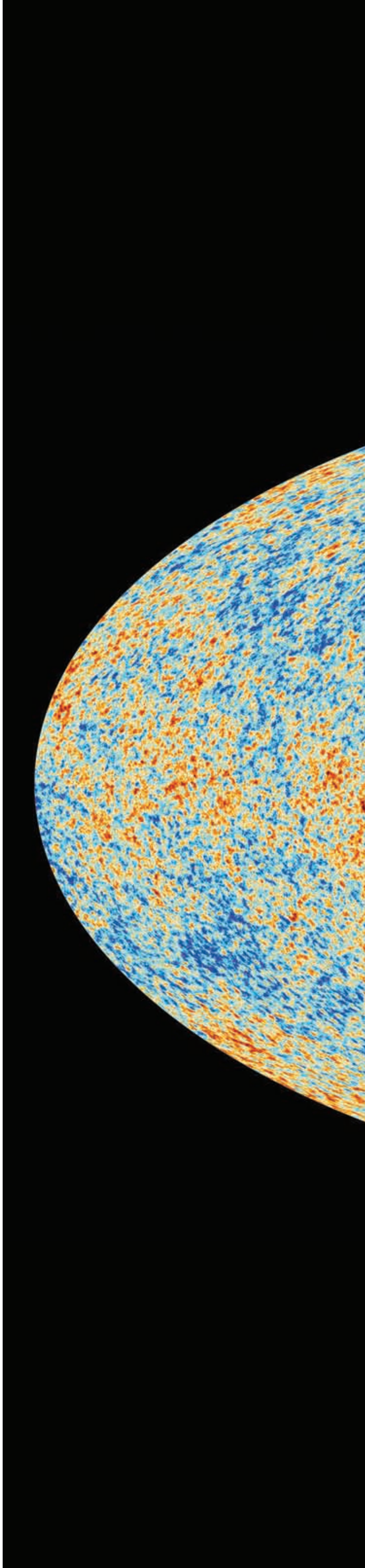
By the conclusion of this chapter, you should be able to:

LG 1 Explain in detail the cosmological principle.

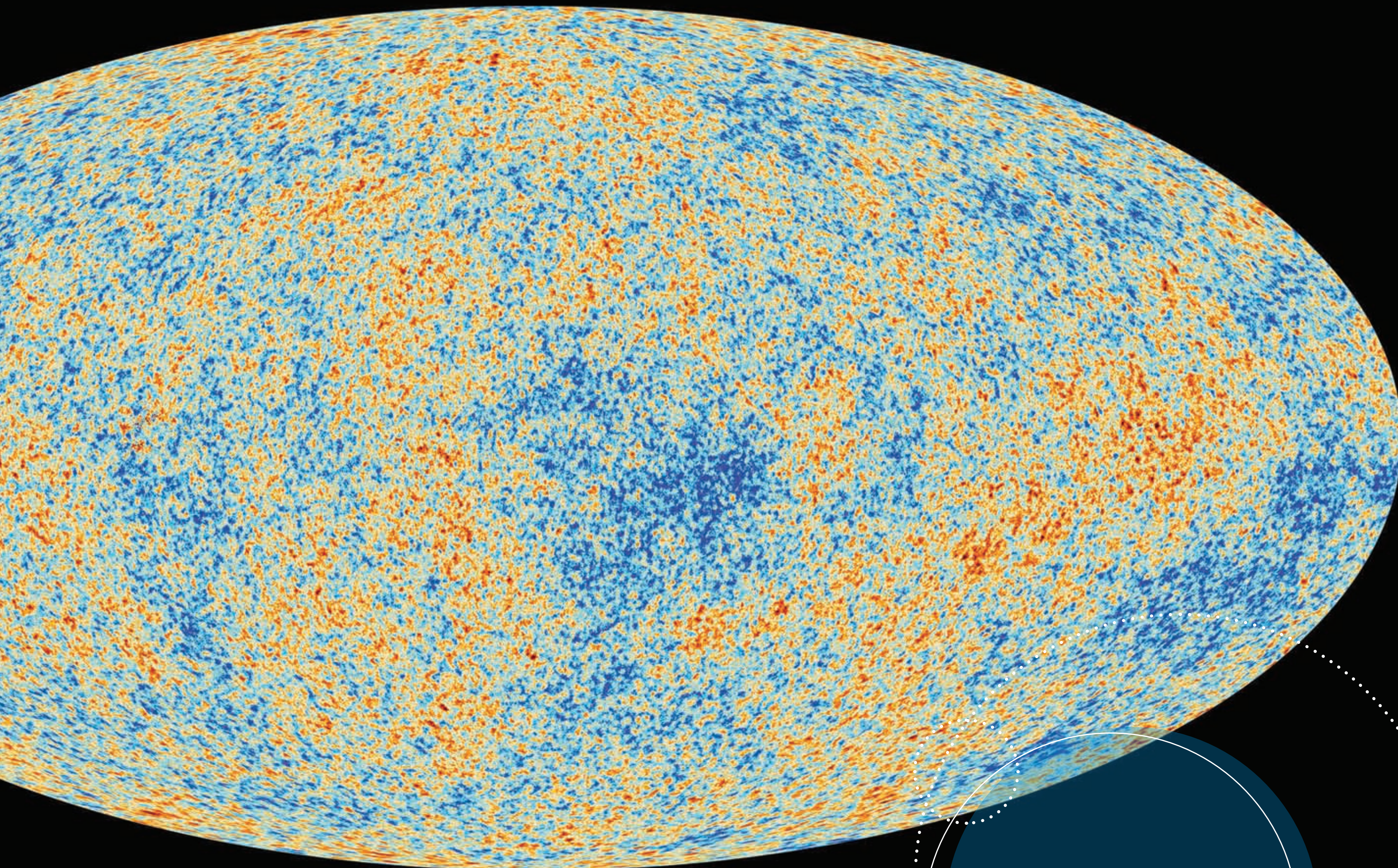
LG 2 Describe how the Hubble constant can be used to estimate the age of the universe.

LG 3 Describe the observational evidence for the Big Bang.

LG 4 Explain which chemical elements were created in the early hot universe.



This image shows the oldest light in the universe, the cosmic microwave background radiation, detected by the Planck space observatory. ▶▶▶



What is the evidence for the Big Bang?

21.1 The Cosmological Principle

In Chapter 19, we described observations of galaxy motions and distances. From these observations, Edwin Hubble created a graph that showed the galaxies moving apart, with the more distant ones moving faster—a result known as Hubble's law. In this section, we will further explore the implications of Hubble's law.

The Homogeneous and Isotropic Universe

Cosmology is the study of space and time and the dynamics of the universe as a whole. In the 1920s, around the same time that astronomers were first measuring distances to galaxies, theoretical physicists were applying Einstein's general theory of relativity to cosmology. The cosmologist Alexander Friedmann (1888–1925) produced mathematical models of the universe that assumed the **cosmological principle**, which is stated as follows: *the physical laws that apply to one part of the universe apply everywhere*. The cosmological principle forms the basis of our study of distant galaxies and of the universe itself and is now a fundamental tenet of modern cosmology.

We have used this principle throughout this book when applying laws of physics and chemistry to objects near and far in the universe. For example, according to the cosmological principle, gravity works the same way in distant galaxies as it does here on Earth. The cosmological principle is a testable scientific theory. An important prediction of the principle is that the conclusions that scientists reach about the universe should be the same whether we observe the universe from the Milky Way or from a galaxy billions of parsecs away. In other words, if the cosmological principle is correct, then the universe is **homogeneous**, having the same composition and properties at all places.

Clearly, the universe is not homogeneous in an absolute sense, as the conditions on Earth are very different from those in deep space or in the center of the Sun. In cosmology, homogeneity of the universe means that stars and galaxies in Earth's part of the universe are much the same, and behave in the same manner, as stars and galaxies in remote corners of the universe. It also means that stars and galaxies everywhere are distributed in space in much the same way that they are distributed in Earth's cosmic neighborhood, and that observers in those galaxies would see the same properties for the universe that astronomers see from here.

It is not easy to verify the prediction of homogeneity directly. Scientists cannot travel from the Milky Way to a galaxy in the remote universe to see whether conditions are the same. However, they can compare light arriving from closer and farther locations in the distant universe and see the ways in which features look the same or different. For example, they can look at the way galaxies are distributed in distant space and see whether that distribution is similar (that is, homogeneous) to the distribution nearby.

In addition to predicting that the universe is homogeneous, the cosmological principle requires that all observers measure the same properties of the universe, regardless of the direction in which they are looking. If something is the same in all directions, then it is **isotropic**. This prediction of the cosmological principle is much easier to test directly than is homogeneity. For example, if galaxies were lined up in rows, astronomers would measure very different properties of the universe, depending on the direction they looked. The universe would still be homogeneous, but not isotropic, so it would not satisfy the cosmological principle.

In most instances, isotropy goes together with homogeneity, and the cosmological principle requires both. **Figure 21.1** shows examples of how the universe could have violated the cosmological principle by not being homogeneous or isotropic, as well as examples of how the universe might satisfy the cosmological principle. All observations show that the properties of the universe are basically the same, regardless of the direction in which observers are looking. When averaged over very large scales, thousands of millions of parsecs, the universe appears homogeneous as well.

The Hubble Expansion

Recall from Chapter 19 our discussion of Hubble's law—that the distance of a galaxy is proportional to its recession velocity. Hubble's law helps astronomers investigate whether the universe is homogeneous and isotropic. They can confirm its isotropy by observing that galaxies in one direction in the sky obey the same Hubble law as galaxies in other directions in the sky. Hubble's law says that Earth is located in an expanding universe and that the expansion looks the same,

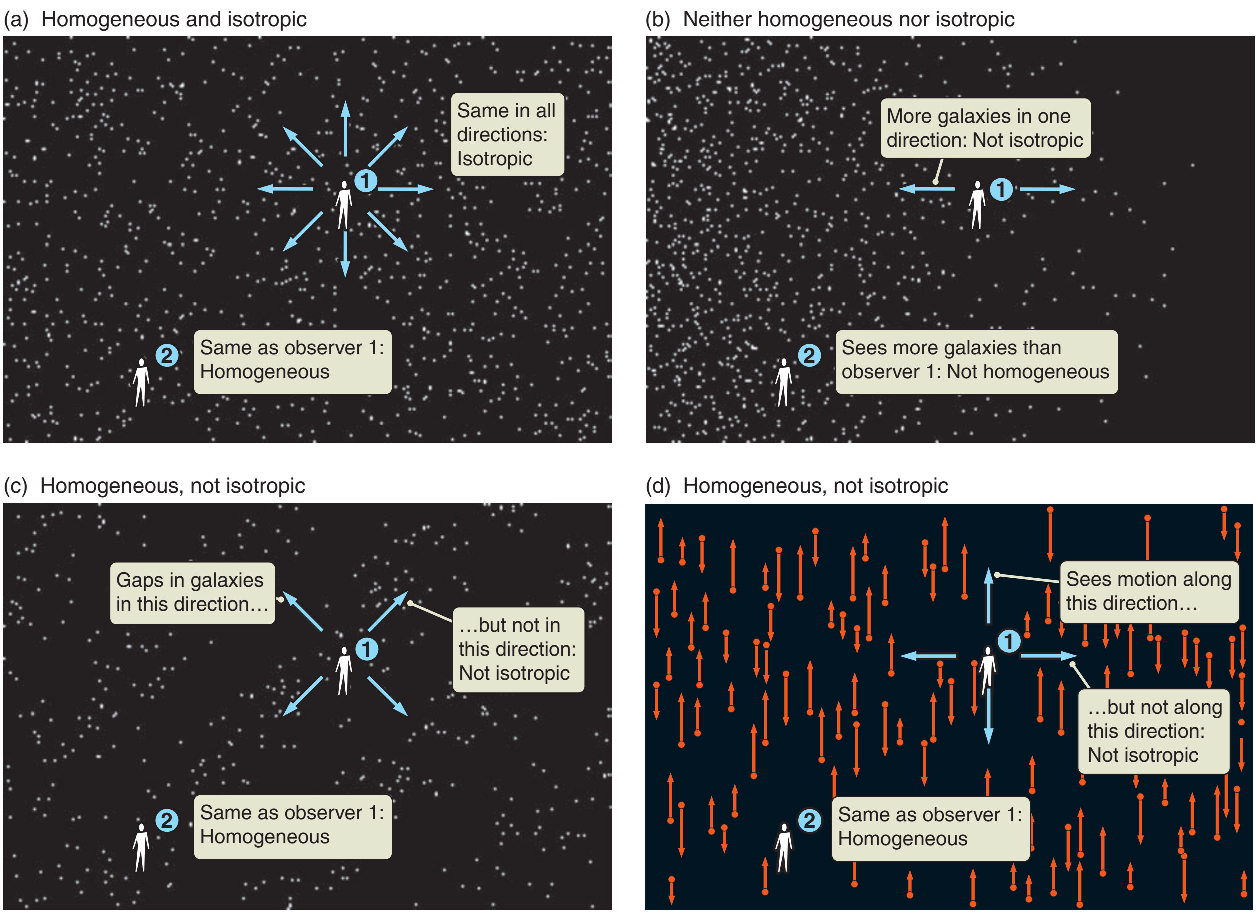


Figure 21.1 Homogeneity and isotropy in four different theoretical models of a universe. Blue arrows indicate the direction of view. (a) The distribution of galaxies is uniform, so this universe is both homogeneous and isotropic. (b) The density of galaxies is decreasing in one direction, so this universe is neither homogeneous nor isotropic. (c) The bands of galaxies lie along a unique axis, making this universe not isotropic. (d) The distribution of galaxies is uniform, but galaxies move along only one direction, so this universe also is not isotropic.

regardless of the location of the observer. To help you visualize this, we now turn to a useful model that you can build for yourself with materials you probably have in your desk.

Figure 21.2 shows a long rubber band with paper clips attached along its length. If you stretch the rubber band, the paper clips—which represent galaxies in an expanding universe—get farther and farther apart. To get a sense of what this expansion would look like up close, imagine yourself an ant riding on paper clip A. As the rubber band is stretched, you notice that all of the paper clips are moving away from you. Clip B, the closest one, is moving away slowly. Clip C,

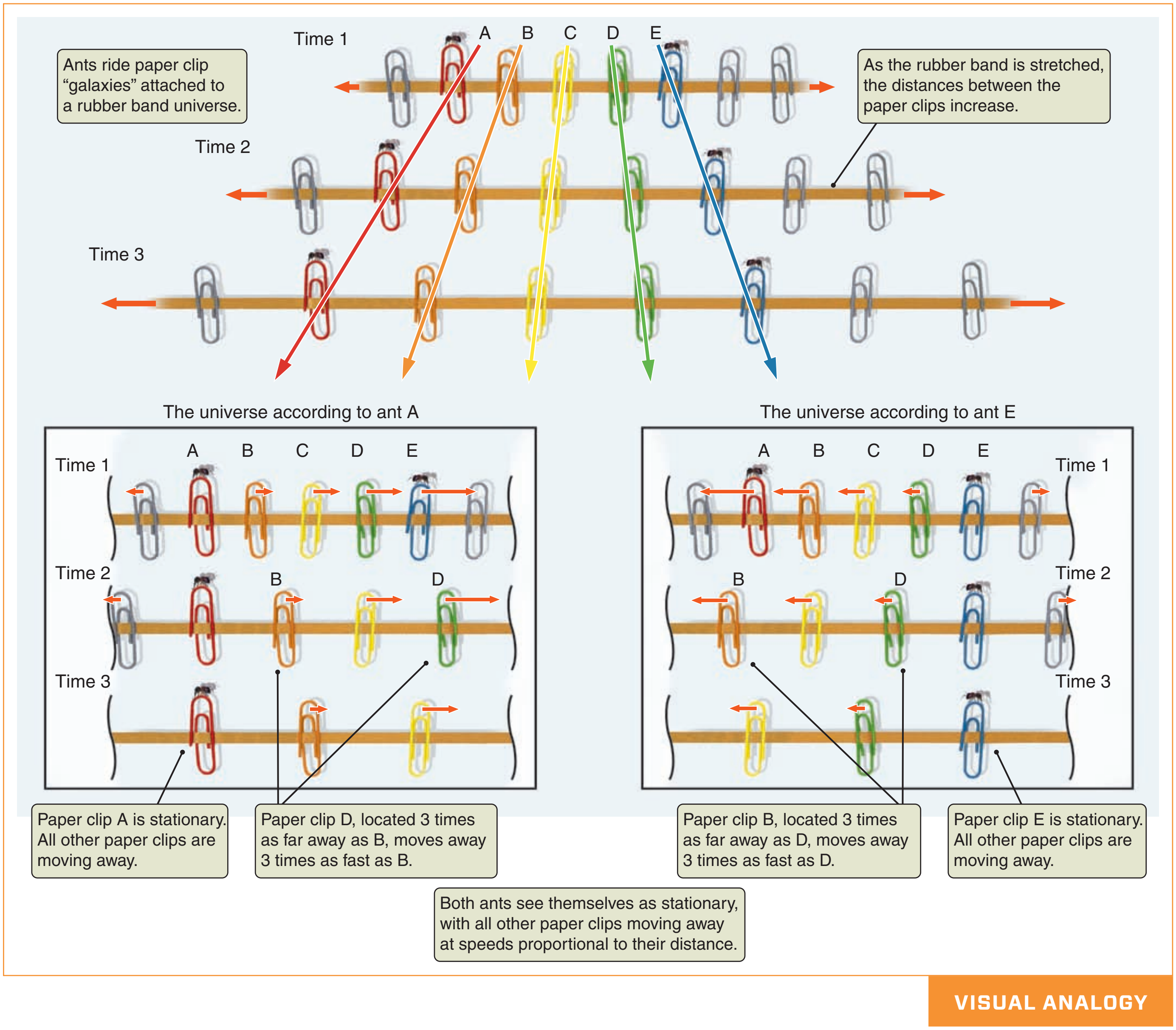


Figure 21.2 In this one-dimensional analogy of Hubble's law, a rubber band with paper clips evenly spaced along its length is stretched. As the rubber band stretches, an ant riding on clip A observes clip C moving away twice as fast as clip B. Similarly, an ant riding on clip E sees clip C moving away twice as fast as clip D. Any ant sees itself as stationary, regardless of which paper clip it is riding, and it sees the other clips moving away with speed proportional to distance.

located twice as far from you as clip B, is moving away twice as fast as clip B. Clip E, located 4 times as far away as clip B, is moving away 4 times as fast as clip B. From your perspective as an ant riding on clip A, all of the other paper clips on the rubber band are moving away with a velocity proportional to their distance. The paper clips located along the rubber band obey a Hubble-like law.

The key insight to the analogy comes from realizing that there is nothing special about the perspective of paper clip A. If, instead, the ant were riding on clip E, clip D would be the one moving away slowly and clip A would be moving away 4 times as fast. Repeat this experiment for any paper clip along the rubber band, and you will arrive at the same result. For an ant on any paper clip along the rubber band, the speed at which other clips are moving away from the ant is proportional to their distance. The stretching rubber band, like the universe, is "homogeneous." The same Hubble-like law applies, regardless of where the observer ant is located.

The observation that nearby paper clips move away slowly and distant paper clips move away more rapidly does not say that the paper clip selected as a vantage point is at the center of anything. Instead, it says that the rubber band is being stretched uniformly along its length. Similarly, Hubble's law for galaxies means that nearby galaxies are carried away slowly by expanding space, and distant galaxies are carried away more rapidly (**Process of Science Figure**). Our galaxy is not at the center of an expanding universe; the universe is expanding uniformly. Any observer in any galaxy sees nearby galaxies moving away slowly and more distant galaxies moving away more rapidly. The same Hubble law applies from their vantage point as applies from our vantage point on Earth. The expansion of the universe is homogeneous.

The only exception to this rule is the case of galaxies that are close together, in which case gravitational attraction dominates over the expansion of space. As you saw in the previous chapter, the Andromeda Galaxy and the Milky Way are being pulled together by their gravity. The Andromeda Galaxy is approaching the Milky Way at about 110 kilometers per second (km/s), so the light from the Andromeda Galaxy is blueshifted, not redshifted. This velocity is caused by the gravitational interaction between the two nearby galaxies, rather than the expansion of space. The fact that gravitational or electromagnetic forces can overwhelm the expansion of space also explains why the Solar System is not expanding, and neither are you.

The Universe in Space and Time

Hubble's law gives astronomers a practical tool for measuring distances to remote objects. Once they know the value of the Hubble constant, H_0, they can use a straightforward measurement of the redshift of a galaxy to find its distance. In other words, once H_0 is known, Hubble's law makes the once-difficult task of measuring distances in the universe relatively easy, providing astronomers with a tool to map the structure of the observable universe. This may seem like a logical impossibility, because they are using redshifts and distances to find H_0 and then using H_0 to find the distances. But astronomers find H_0 from nearby galaxies using standard candles, and then use that value to find the distance to a different, more distant set of galaxies.

Hubble's law does more than place galaxies in space. It also places galaxies in time. Light travels at a huge but finite speed. Recall from Chapter 1 that when you look

Process of Science

DATA ARE THE ULTIMATE AUTHORITY

Einstein is one of the most famous scientists of all time. His genius is recognized by everyone. But even Einstein had to change his mind in the face of new data.

Einstein develops general relativity, which predicts the universe is expanding or contracting.

He inserts an extra constant into his equations to make the universe stationary because he believes it should be.

Hubble presents his data to Einstein. The universe is expanding!

Einstein changes his mind, and describes his reluctance to accept the prediction of his original theory as "his biggest blunder."

Even the most brilliant scientists must yield when the data contradict their conclusions.

at the Sun, you see it as it existed $8\frac{1}{3}$ minutes ago. When you look at Alpha Centauri, the nearest stellar system beyond the Sun, you see it as it existed 4.3 years ago. If you look at the center of the Milky Way, the light you see is 27,000 years old. The **look-back time** of a distant object is the time it has taken for the light from that object to reach a telescope on Earth. As astronomers look into the distant universe, look-back times become very large. The distance to a galaxy whose redshift $z = 0.1$ is 1.4 billion light-years (assuming $H_0 = 70$ km/s/Mpc), so the look-back time to that galaxy is 1.4 billion years. The look-back time to a galaxy where $z = 0.2$ is 2.7 billion years. The look-back time of the most distant galaxies observed, at about $z = 10$, is 13.2 billion years. As astronomers observe objects with greater and greater redshifts, they are seeing increasingly younger stages of the universe.

CHECK YOUR UNDERSTANDING 21.1

In astronomy, *isotropy* means that the universe is the same __________, and *homogeneity* means that the universe is the same __________. (a) in all locations; (b) in all directions; (c) at all times; (d) at all size scales.

21.2 The Universe Began in the Big Bang

Imagine watching a video of the universe, with the galaxies moving apart. Now, reverse the video, and run it backward in time. The galaxies become closer and closer together as the universe becomes younger and younger. In this way, the observation of the expansion of the universe leads to the idea of a beginning to the universe at a time that can be estimated from Hubble's law. In this section, we explore this implication of Hubble's law.

Expansion and the Age of the Universe

Hubble's law gives an estimate of the age of the universe. If we assume the speed of the expansion has always been constant, then the age of the universe can be estimated from the slope of the line in a graph of the velocity of galaxies plotted against their distance. This slope has units that reduce to 1/time, so its inverse has units of time. This slope is the Hubble constant, H_0, and its inverse (1 divided by H_0) is the **Hubble time**. The Hubble time is an estimate of the universe's age: 13.8 billion years (**Working It Out 21.1**). If the expansion were faster, the Hubble constant would be larger, and the universe would be younger; similarly, a slower expansion would yield a smaller Hubble constant and an older universe.

AstroTour: Hubble's Law

Astronomy in Action: Expanding Balloon Universe

If the universe expanded uniformly, then when it was about half its current age, about 6.9 billion years ago, all of the galaxies in the universe were half as far apart as they are now, and 12.4 billion years ago, all of the galaxies in the universe were about a tenth as far apart. Assuming that galaxies have been moving apart all that time at the same speed as they do today, then a little less than one Hubble time ago—13.8 billion years ago—there was almost no space between the particles that constitute today's universe. All such matter as well as energy in the universe then must have been unimaginably dense. Because

21.1 Working It Out Expansion and the Age of the Universe

We can use Hubble's law to estimate the age of the universe. Consider two galaxies located 30 Mpc ($d_G = 9.3 \times 10^{20}$ km) away from each other, as shown in **Figure 21.3**. If these two galaxies are moving apart, then at some time in the past they must have been together in the same place at the same time. According to Hubble's law, and assuming that $H_0 = 70$ km/s/Mpc, the distance (d_G) between these two galaxies is increasing at the following rate:

$$v_r = H_0 \times d_G$$
$$v_r = 70 \text{ km/s/Mpc} \times 30 \text{ Mpc}$$
$$v_r = 2{,}100 \text{ km/s}$$

Knowing the velocity (v_r) at which they are traveling, we can calculate the time it took for the two galaxies to become separated by 30 Mpc:

$$\text{Time} = \frac{\text{Distance}}{\text{Velocity}} = \frac{9.3 \times 10^{20}\text{ km}}{2{,}100\text{ km/s}} = 4.4 \times 10^{17}\text{ s}$$

Dividing by the number of seconds in a year (about 3.16×10^7 s/yr) gives

$$\text{Time} = 1.4 \times 10^{10}\text{ yr} = 14 \text{ billion yr}$$

In other words, *if* expansion of the universe has been constant, two galaxies that today are 30 Mpc apart started out at the same place 14 billion years ago.

Now let's do the same calculation with two galaxies that are 60 Mpc, or 18.6×10^{20} km, apart. These two galaxies are twice as far apart, but the distance between them is increasing twice as rapidly:

$$v_r = H_0 \times d_G = 70 \text{ km/s/Mpc} \times 60 \text{ Mpc} = 4{,}200 \text{ km/s}$$

Therefore,

$$\text{Time} = \frac{18.6 \times 10^{20}\text{ km}}{4{,}200\text{ km/s}} = 4.4 \times 10^{17}\text{ s} = 1.4 \times 10^{10}\text{ yr}$$

Again, we calculate time as distance divided by velocity (twice the distance divided by twice the velocity) to find that these galaxies also took about 14 billion years to reach their current locations. We can do this calculation again and again for any pair of galaxies in the universe today. The farther apart the two galaxies are, the faster they are moving. But all galaxies took the same amount of time to get to where they are today.

Working out the example using words instead of numbers makes it clear why the answer is always the same. Because the velocity we are calculating comes from Hubble's law, velocity equals the Hubble constant multiplied by distance. Writing this out as an equation, we get

$$\text{Time} = \frac{\text{Distance}}{\text{Velocity}} = \frac{\text{Distance}}{H_0 \times \text{Distance}}$$

Distance divides out to give

$$\text{Time} = \frac{1}{H_0}$$

where $1/H_0$ is the Hubble time. This is one way of estimating the age of the universe.

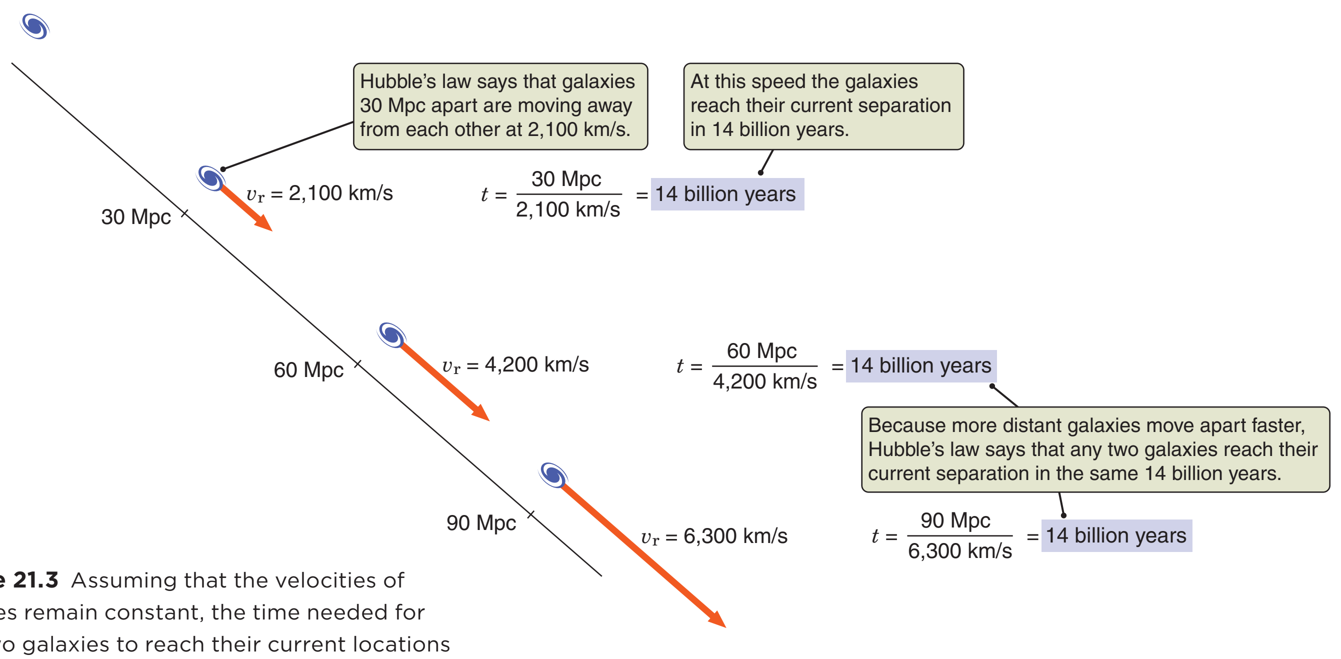


Figure 21.3 Assuming that the velocities of galaxies remain constant, the time needed for any two galaxies to reach their current locations is the same, regardless of their separation.

expanding gases cool down, the universe then must have been much hotter than it is today in its expanded state. This hot, dense beginning, 13.8 billion years ago, is called the **Big Bang** (**Figure 21.4**).

Georges Lemaître (1894–1966) was the first to propose the theory of the Big Bang. This idea greatly troubled many astronomers in the early and middle years of the 20th century. Several different suggestions were put forward to explain the observed fact of Hubble expansion without resorting to the idea that the universe came into existence in an extraordinarily dense "fireball" billions of years ago. However, as more and more distant galaxies have been observed, and as more discoveries about the structure of the universe have been made, the Big Bang theory has grown stronger. The major predictions of the Big Bang theory have proven to be correct.

The implications of Hubble's law forever changed the scientific concepts of the origin, history, and possible future of the universe. At the same time, Hubble's law has pointed to many new questions about the universe. To address them, we next need to consider exactly what is meant by the term *expanding universe*.

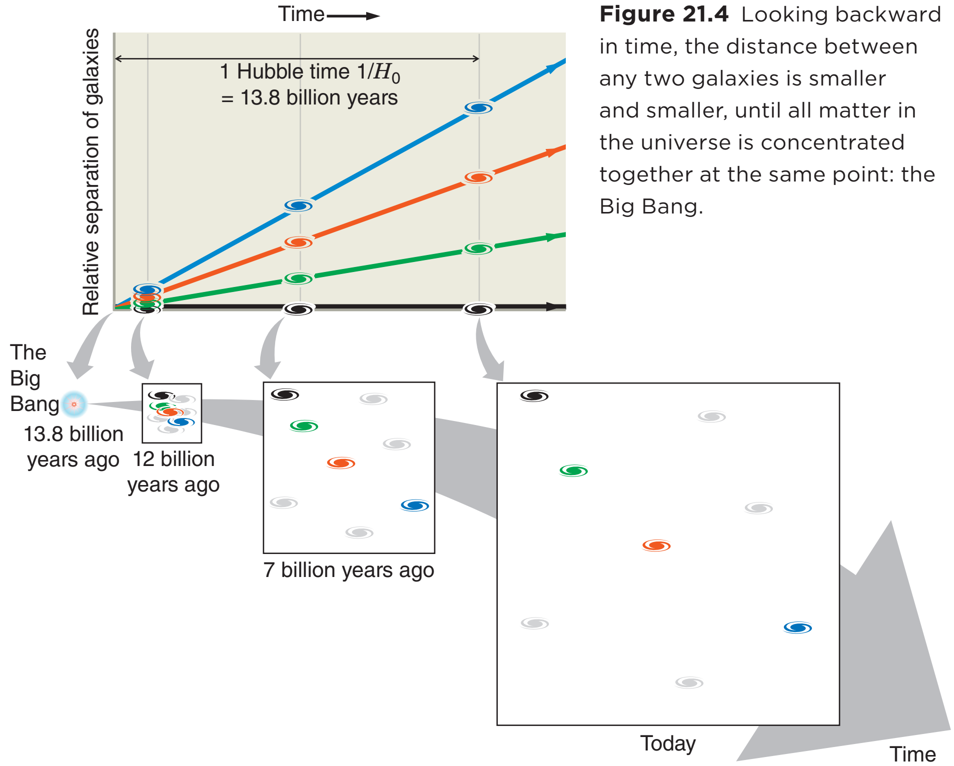


Figure 21.4 Looking backward in time, the distance between any two galaxies is smaller and smaller, until all matter in the universe is concentrated together at the same point: the Big Bang.

Galaxies and the Expansion

At this point in our discussion, you may be picturing the expanding universe as a cloud of debris from an explosion flying outward through surrounding space. This is a common depiction of the Big Bang in movies and television shows, where they portray a tiny bright spot that explodes to fill the screen. However, the Big Bang is not an explosion in the usual sense of the word, and in fact there is no surrounding space into which the universe expands.

We must draw a distinction between the universe and the observable universe. The observable universe is the part of the universe that we can see. The observable universe extends 13.8 billion light-years in every direction. This limit exists because that is the length of time the universe has been around. The light from more distant regions has not yet had time to travel to us, and so we cannot see it yet.

A common question about the Big Bang is, Where did it take place? The answer is that the Big Bang took place *everywhere*. Wherever anything is in the universe today, it is at the site of the Big Bang. The reason is that galaxies are not flying apart through space at all. Rather, space itself is expanding, carrying the stars and galaxies that populate the universe along with it.

We have already dealt with the basic ideas that explain the expansion of space. In our discussion of black holes in Chapter 18, you encountered Einstein's general theory of relativity. General relativity says that space is distorted by the presence of mass, and that the consequence of this distortion is gravity. For example, the mass of the Sun, like any other object, distorts the geometry of spacetime around it; so Earth, coasting along in its inertial frame of reference, follows a curved path around the Sun. We illustrated this phenomenon in Figure 18.10 with the analogy of a ball placed on a stretched rubber sheet, showing how the ball distorted the surface of the sheet.

Astronomy in Action: Observable vs. Actual Universe

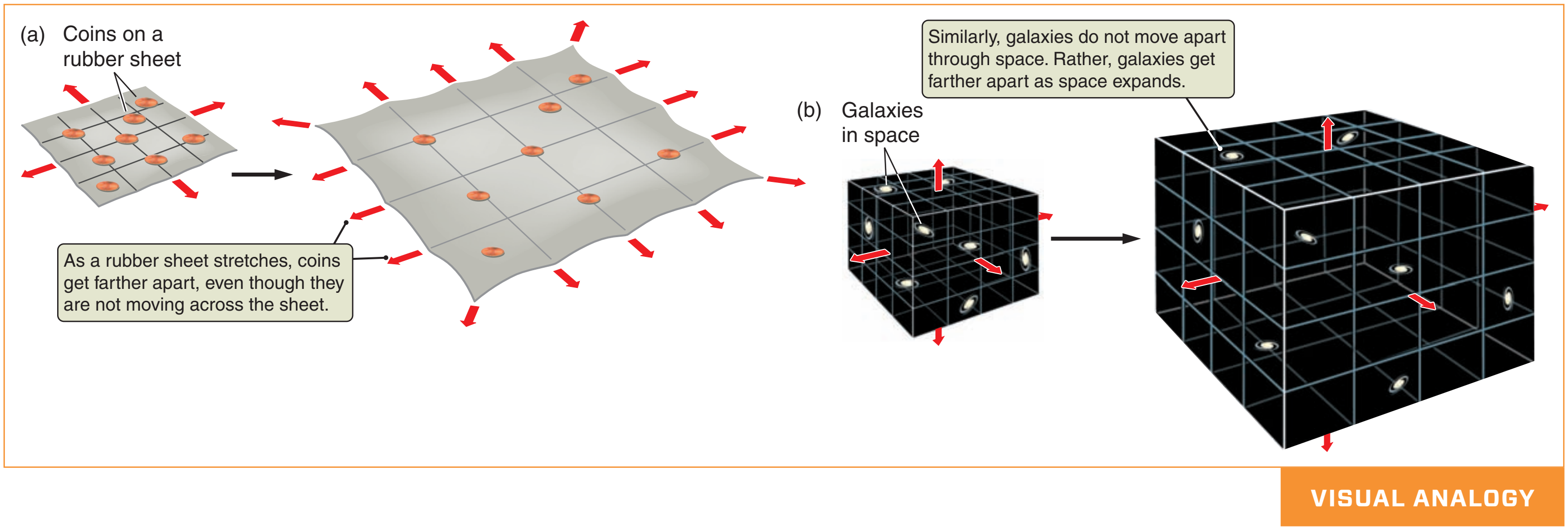


Figure 21.5 (a) As a rubber sheet is stretched, coins on its surface move farther apart, even though they are not moving with respect to the sheet itself. Any coin on the surface of the sheet observes a Hubble-like law in every direction. (b) Galaxies in an expanding universe are not flying apart through space. Rather, space itself is stretching.

The surface of a rubber sheet can be distorted in other ways as well. Imagine a number of coins placed on a rubber sheet, illustrated in **Figure 21.5**. Then imagine grabbing the edges of the sheet and beginning to pull them outward. As the rubber sheet stretches, each coin remains at the same location on the surface of the sheet, but the distances between the coins increase. Two coins sitting close to each other move apart only slowly, while coins farther apart move away from each other more rapidly. The distances and relative motions of the coins on the surface of the rubber sheet obey a Hubble-like relationship as the sheet is stretched.

This movement is analogous to what is happening in the universe, with galaxies taking the place of the coins and space itself taking the place of the rubber sheet. In the case of coins on a rubber sheet, there is a limit to how far the sheet can be stretched before it breaks. With space and the real universe, there is no such limit. The fabric of space can, in principle, go on expanding forever. Hubble's law is the observational consequence of the fact that the space making up the universe is expanding.

How will this expansion of the universe behave in the future? You have learned that most of the mass in galaxies consists of dark matter, and the gravity of the dark matter has the effect of slowing down the expansion of the universe. In the next chapter, you will also learn that there is another unseen constituent of the universe, called *dark energy*, and this constituent causes the expansion of the universe to *accelerate*. At the current stage in the expansion of the universe, the accelerating effect of dark energy dominates over the slowing effect of dark matter, and therefore the universe will continue to expand at an ever-faster rate.

CHECK YOUR UNDERSTANDING 21.2

Where in the universe did the Big Bang take place? (a) near the Milky Way Galaxy; (b) near the center of the universe; (c) near some unknown location on the other side of the universe; (d) everywhere in the universe

21.3 Expansion Is Described with a Scale Factor

As the universe expands, the distance between any two objects increases because of the stretching of space. Astronomers find it useful to discuss this expansion in terms of the *scale factor* of the universe.

Scale Factor

Let's return to the analogy of the rubber sheet. Suppose you place a ruler on the surface of the rubber sheet and draw a tick mark every centimeter, as illustrated in **Figure 21.6a**. To measure the distance between two points on the sheet, you can count the marks between the two points and multiply by 1 centimeter (cm) per tick mark.

As the sheet is stretched, however, the distance between the tick marks does not remain 1 cm. When the sheet is stretched to 150 percent of the size it had when the tick marks were drawn, each tick mark is separated from its neighbors by $1\frac{1}{2}$ times the original distance, or 1.5 cm. The distance between two points can still come from counting the marks, but you need to scale up the distance between tick marks by 1.5 to find the distance in centimeters. If the sheet were twice the size it was when the tick marks were drawn (Figure 21.6b), each mark would correspond to 2 cm of actual distance. Astronomers use the term **scale factor** to indicate the size of the sheet relative to its size at the time when the tick marks were drawn. The scale factor also indicates how much the distance between points on the sheet has changed. In the first example, the scale factor of the sheet is 1.5; in the second, the scale factor of the sheet is 2.

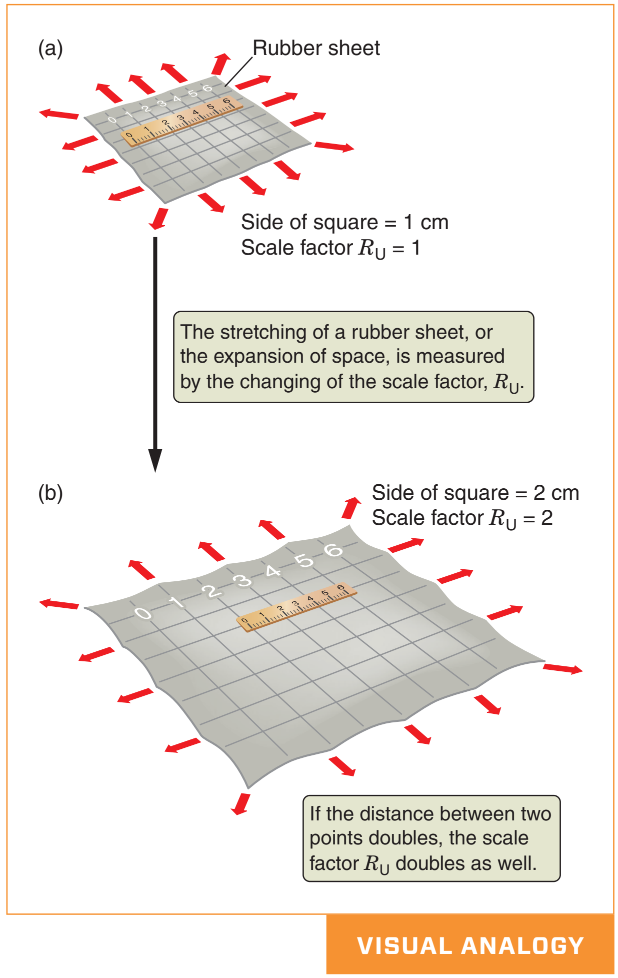


Figure 21.6 (a) On a rubber sheet, tick marks are drawn 1 cm apart. As the sheet is stretched, the tick marks move farther apart. (b) When the spacing between the tick marks is 2 cm, or twice the original value, the scale factor of the sheet, R_U, is said to have doubled. A similar scale factor, R_U, is used to describe the expansion of the universe.

Suppose astronomers choose today to lay out a "cosmic ruler" on the fabric of space, placing an imaginary tick mark every 10 Mpc. The scale factor of the universe at this time is defined to be 1. In the past, when the universe was smaller, distances between the points in space marked by this cosmic ruler would have been less than 10 Mpc. The scale factor of that younger, smaller universe would have been less than 1 compared to the scale factor today. In the future, as the universe continues to expand, the distances between the tick marks on this cosmic ruler will grow to more than 10 Mpc, and the scale factor of the universe will be greater than 1. Astronomers use the scale factor, usually written as R_U, to keep track of the changing scale of the universe.

It is important to remember that the laws of physics are themselves unchanged by the expansion of the universe, just as stretching a rubber sheet does not change the properties of the coins on its surface. At noncosmological scales, the nuclear and electromagnetic forces within and between atoms, as well as the gravitational forces between relatively nearby objects, dominate over the expansion. As the universe expands, the sizes and other physical properties of atoms, stars, and galaxies also remain unchanged.

Looking back in time, the scale factor of the universe gets smaller and smaller, approaching zero as it comes closer and closer to the Big Bang. The fabric of space that today spans billions of parsecs spanned much smaller distances when the universe was young. When the universe was only a day old, all of the space visible today amounted to a region only a few times the size of the Solar System. When the universe was 1/50 of a second old, the vast expanse of space that makes up today's observable universe (and all the matter in it) occupied a volume only the size of today's Earth. Going backward in time approaching the Big Bang itself, the space that makes up today's observable universe becomes smaller and smaller—the size of a grapefruit, a marble, an atom, a proton. Every point in the fabric of space that makes up today's universe was right there at the beginning, a part of that unimaginably tiny, dense universe that emerged from the Big Bang.

The Big Bang did not occur at a specific point in space, because space itself came into existence with the Big Bang. The Big Bang happened everywhere; there is no particular point in today's universe that marked the site of the Big Bang. A Big Bang universe is homogeneous and isotropic, consistent with the cosmological principle.

CHECK YOUR UNDERSTANDING 21.3

The scale factor keeps track of: (a) the movement of galaxies through space; (b) the current distances between many galaxies; (c) the changing distance between any two galaxies; (d) the location of the center of the universe.

Redshift Is Due to the Changing Scale Factor of the Universe

The ideas of general relativity discussed in Chapter 18 are powerful tools for interpreting Hubble's great discovery of the relationship between the velocity and distance of galaxies and of the expanding universe. These velocities were determined from the redshift of the galaxies using the Doppler effect. Although it is true that the distance between galaxies is increasing as a result of the expansion of the universe, and that we can use the equation for Doppler shifts to measure the redshifts of galaxies, these redshifts are not due to Doppler shifts in the same way that we described for a moving star. Light that comes from very distant objects was emitted at a time when the universe was younger and therefore smaller. As this light comes toward Earth from distant galaxies, the scale factor of the space through which the light travels is constantly increasing, and as it does, the distance between adjacent light-wave crests increases as well. The light is "stretched out" as the space it travels through expands.

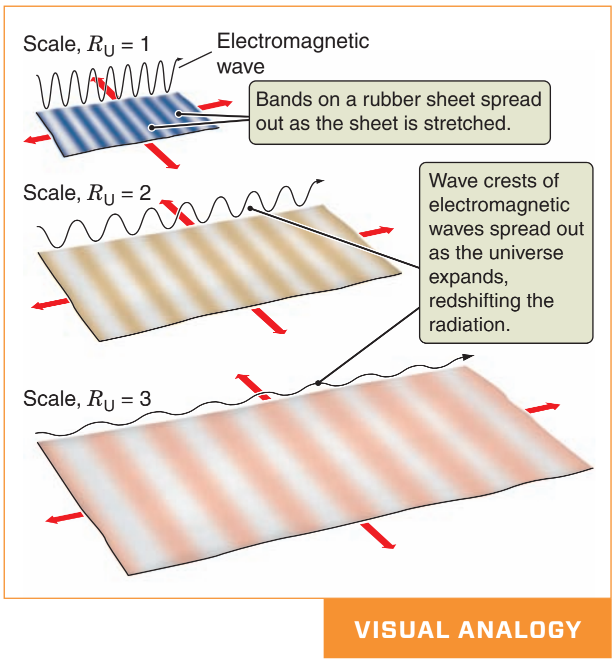


Figure 21.7 Bands drawn on a rubber sheet represent the positions of the crests of an electromagnetic wave in space. As the rubber sheet is stretched—that is, as the universe expands—the wave crests get farther apart. The light is redshifted.

Recall the rubber sheet analogy that we used when discussing black holes in Chapter 18. In that case, we imagined space as a two-dimensional rubber sheet. **Figure 21.7** uses this rubber-sheet analogy to explain why the increasing distance between galaxies causes the light to be redshifted. If you draw a series of bands on the rubber sheet to represent the crests of an electromagnetic wave, you can watch what happens to the wave as the sheet is stretched out. By the time the sheet is stretched to twice its original size—that is, by the time the scale factor of the sheet is 2—the distance between wave crests has doubled. When the sheet has been stretched to 3 times its original size (a scale factor of 3), the wavelength of the wave will be 3 times what it was originally.

Consider the light coming from a distant galaxy. When the light left the distant galaxy, the scale factor of the universe was smaller than it is today. As light comes toward us from distant galaxies, the space through which the light travels is stretching, and the light is also "stretched out" as the space through which it travels expands. The universe expanded while the light was in transit, and as it did so, the wavelength of the light grew longer in proportion to the increasing scale factor of the universe. The redshift of light from distant galaxies is therefore a direct measure of how much the universe has expanded since the time when the radiation left its source. Redshift measures how much the scale factor of the universe, R_U, has changed since the light was emitted. This redshift can be greater than 1 if the galaxy is far enough away that the recession velocity is relativistic, as described in **Working It Out 21.2**.

21.2 Working It Out When Redshift Exceeds One

In our discussion of the Doppler shift in Chapter 5, we noted that $(\lambda_{obs} - \lambda_{rest})/\lambda_{rest}$ is equal to the velocity of an object moving away divided by the speed of light. Edwin Hubble used this result to interpret the observed redshifts of galaxies as evidence that galaxies throughout the universe are moving away from the Milky Way. Einstein's special theory of relativity says that nothing can move faster than the speed of light. Hubble initially assumed that redshifts are due to the Doppler effect. The resulting relation, $z = v_r/c$, would then seem to imply that no object can have a redshift (z) greater than 1. Yet that is not the case. Astronomers routinely observe redshifts significantly in excess of 1. As of this writing, the most distant objects known have redshifts as large as 9 or 10. How can redshifts exceed 1?

To arrive at the expression for the Doppler effect, $v_r/c = (\lambda_{obs} - \lambda_{rest})/\lambda_{rest}$, we have to *assume* that v_r is much less than c. If v_r was close to c, we would have to consider more than just the fact that the waves from an object are stretched out by the object's motion away. We would also have to consider relativistic effects, including the fact that moving clocks run slowly (see Working It Out 18.1). When combining these effects, we would find that as the speed of an object approaches the speed of light, its redshift approaches infinity, as shown in **Figure 21.8**.

Doppler's original formula is essentially correct—for objects near enough that their measured velocities are far less than the speed of light. When astronomers look at the motions of orbiting binary stars or the peculiar velocities of galaxies relative to the Hubble flow, this equation works just fine. But anytime there is a redshift of 0.4 or greater, relativity must be taken into account.

Another source of redshift is the gravitational redshift discussed in Chapter 18. As light escapes from deep within a gravitational well, it loses energy, so photons are shifted to longer and longer wavelengths. If the gravitational well is deep enough, then the observed redshift of this radiation can be boundlessly large. In fact, the event horizon of a black hole—that is, the surface around the black hole from which not even light can escape—is where the gravitational redshift becomes infinite.

Cosmological redshift, which is most relevant to this chapter, results from the amount of "stretching" that space has undergone during the time the light from its original source has been en route to Earth. The amount of stretching is given by the factor $1 + z$. When astronomers observe light from a distant galaxy whose redshift $z = 1$, then the wavelength of this light is twice as long as when it left the galaxy. When the light left its source, the universe was half the size that it is today. When they see light from a galaxy with $z = 2$, the wavelength of the radiation is 3 times its original wavelength, and they are seeing the universe when it was one-third its current size. This direct relationship enables astronomers to use the observed redshift of the galaxy to calculate the size of the universe at the look-back time to that galaxy. Nearby, this means that distance and look-back time are proportional to z. As they look back closer and closer to the Big Bang, however, redshift climbs more and more rapidly.

Written as an equation, the scale factor of the universe that astronomers see when looking at a distant galaxy is equal to 1 divided by 1 plus the redshift of the galaxy:

$$R_U = \frac{1}{1 + z}$$

For example, when astronomers report they have observed a galaxy with a redshift of 9, the scale factor at the time the light was emitted was $R_U = 1/(1 + 9)$, or 1/10.

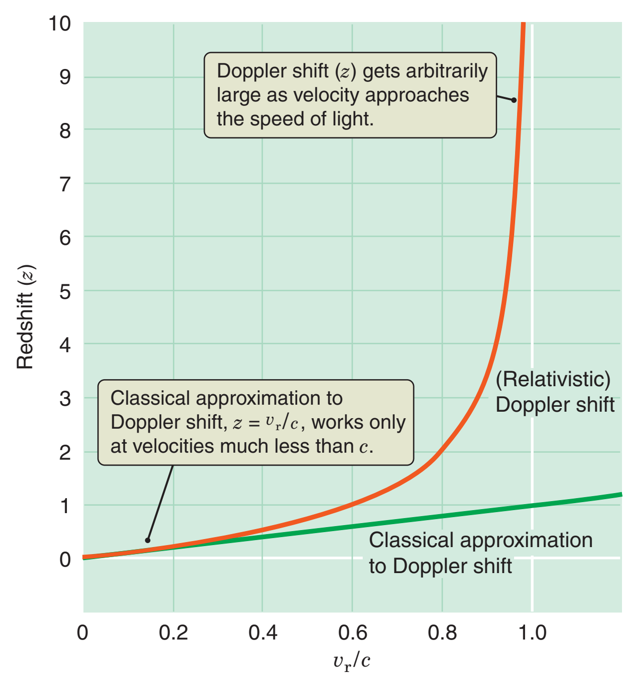


Figure 21.8 This graph shows the plot of the redshift (z) of an object versus its recession velocity (v_r) as a fraction of the speed of light. According to special relativity, as v_r approaches c, the redshift becomes large without limit.

CHECK YOUR UNDERSTANDING 21.4

What is the interpretation of a redshift larger than 1? (a) The object is moving faster than the speed of light. (b) The universe has more than doubled in size since the light from that object was emitted. (c) The light was shifted to longer wavelengths from gravitational radiation. (d) The rate of expansion of the universe is increasing.

21.4 Astronomers Observe Cosmic Microwave Background Radiation

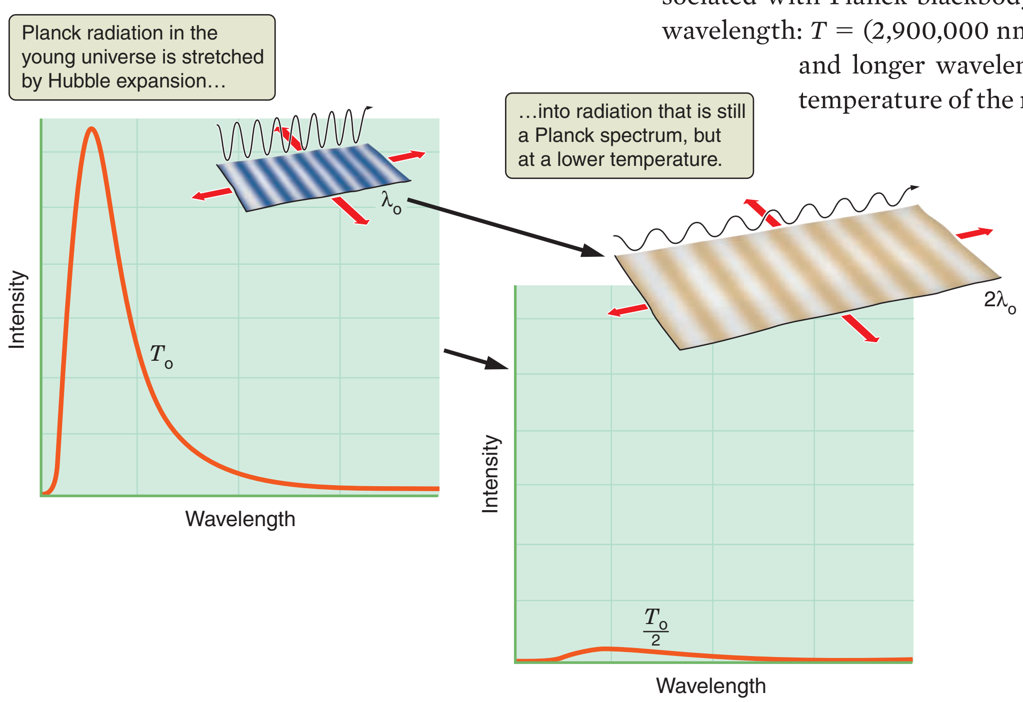
What is the evidence that the Big Bang actually took place? One piece of evidence comes from observations of the early universe across the entire sky. In this section, you will learn about the cosmic microwave background radiation, one of the major confirming observations of the theory that the universe had a beginning.

Radiation from the Big Bang

In the late 1940s, cosmologists Ralph Alpher (1921–2007), Robert Herman (1914–1997), and George Gamow (1904–1968) reasoned that because a compressed gas cools as it expands, the universe should also be cooling as it expands. When the universe was very young and small, it must have consisted of an extraordinarily hot, dense gas. As with any hot, dense gas, this early universe would have been awash in blackbody radiation that exhibits a Planck blackbody spectrum (see Chapter 5).

Gamow and Alpher took this idea a step further, noting that as the universe expanded, this radiation would have been redshifted to longer and longer wavelengths. Recall Wien's law from Chapter 5, which states that the temperature associated with Planck blackbody radiation is inversely proportional to the peak wavelength: $T = (2{,}900{,}000 \text{ nm K})/\lambda_{\text{peak}}$. Shifting the Planck radiation to longer and longer wavelengths is therefore the equivalent of shifting the temperature of the radiation to lower and lower values. As illustrated in **Figure 21.9**, doubling the wavelength of the photons in a Planck blackbody spectrum by stretching space and doubling the scale factor of the universe is equivalent to cutting the temperature of the Planck spectrum in half. The conclusion was that this radiation should still be detectable today and should have a Planck blackbody spectrum with a temperature of about 5–50 kelvins (K). Alpher searched for the signal, but the technology of the late 1940s and early 1950s was not sufficiently advanced. A decade later, physicist Robert Dicke (1916–1997) and his colleagues at Princeton University also predicted a hot early universe, arriving independently at the same basic conclusions that Alpher and Gamow had reached earlier.

Figure 21.9 As the universe expanded, Planck radiation left over from the hot young universe was redshifted to longer wavelengths. Redshifting a Planck spectrum is equivalent to lowering its temperature.



Measuring the Temperature of the Cosmic Microwave Background Radiation

As we noted at the end of Chapter 6, in the early 1960s two physicists at Bell Laboratories—Arno Penzias and Robert Wilson (**Figure 21.10**)—detected a faint microwave signal in all parts of the sky. When Penzias and Wilson had found the signal, they interpreted it as the radiation left behind by the hot early universe. The strength of the detected signal was consistent with the glow from a blackbody with a temperature of about 3 K, very close to the predicted value. Their results, published in 1965, reported the discovery of the "glow" left behind by the Big Bang.

This radiation left over from the early universe is called the **cosmic microwave background radiation (CMB)**. When the universe was young, it was hot enough that all atoms were ionized, so that the electrons were separate from the atomic nuclei. Recall from our discussion of the structure of the Sun and stars in Chapter 14 that free electrons in such conditions interact strongly with radiation, blocking its progress so radiation does not travel well through ionized plasma.

As illustrated in **Figure 21.11a**, the conditions within the early universe were much like the conditions within a star: hot, dense, and opaque. As the universe

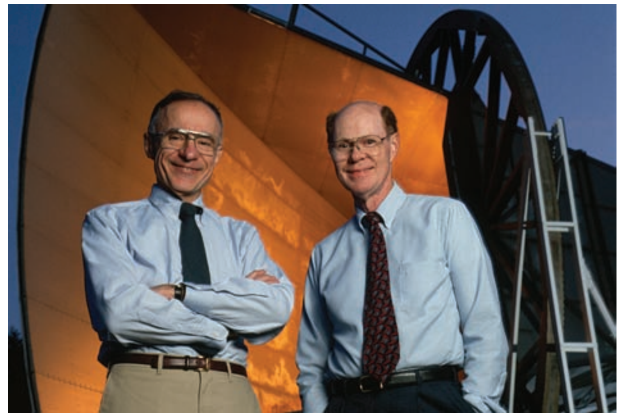


Figure 21.10 Arno Penzias (left) and Robert Wilson next to the Bell Labs radio telescope antenna with which they discovered the cosmic background radiation. This antenna is now a U.S. National Historic Landmark.

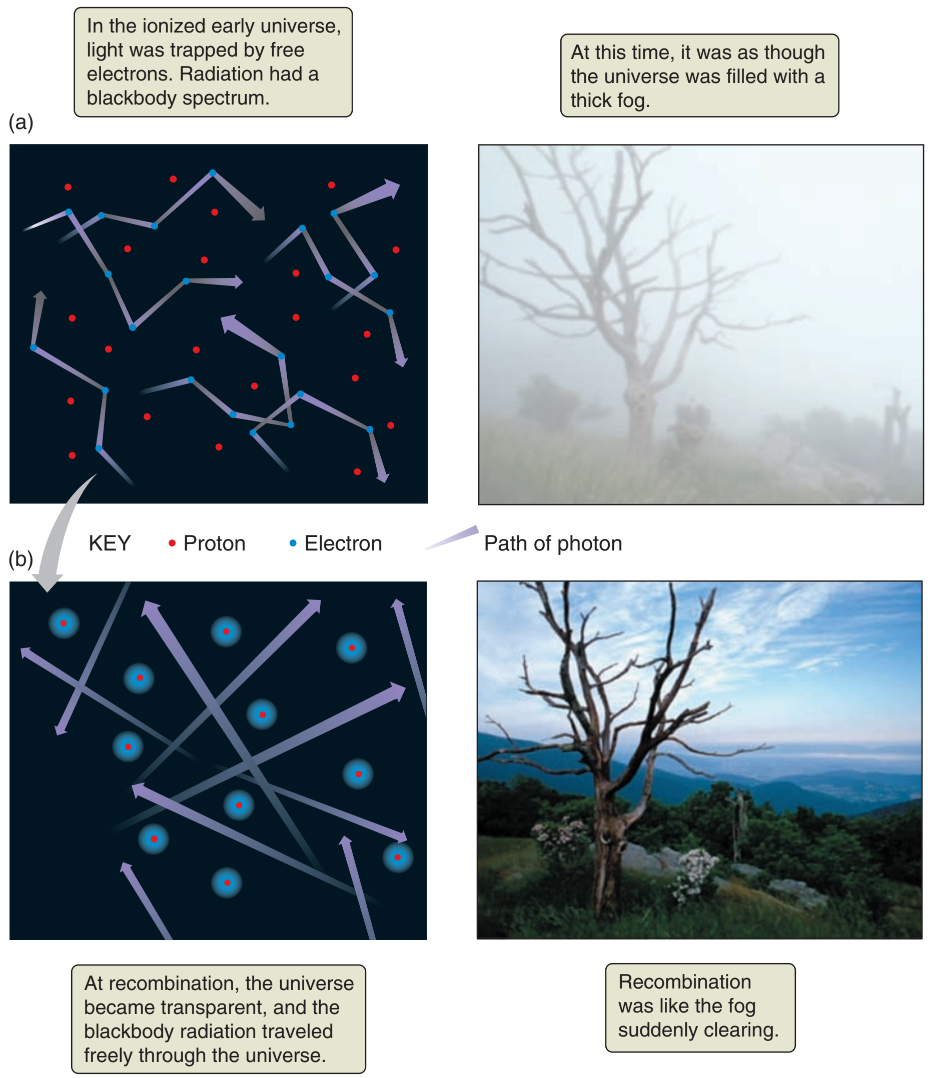


Figure 21.11 The cosmic microwave background radiation originated at the moment the universe became transparent. (a) Before recombination, the universe was like a foggy day, except that the "fog" was a sea of electrons and protons. Radiation interacted strongly with free electrons and so could not travel far. The trapped radiation had a Planck blackbody spectrum. (b) When the constituents of the universe recombined to form neutral hydrogen atoms, the fog cleared and this radiation was free to travel unimpeded.

expanded, the gas filling it cooled. By the time the universe was several hundred thousand years old, and about a thousandth of its current size, the temperature had dropped to a few thousand kelvins. Hydrogen and helium nuclei combined with electrons to form neutral atoms, an event called the **recombination** of the universe, illustrated in Figure 21.11b.

Hydrogen atoms block radiation much less effectively than do free electrons, so when recombination occurred, the universe suddenly became transparent to radiation. Since that time, the radiation left behind from the Big Bang has been able to travel largely unimpeded throughout the universe. At the time of recombination, when the temperature of the universe was about 3000 K, the wavelength of this radiation peaked at about 1 micron (μm), according to Wien's law (see Chapter 5). As the universe expanded, this radiation was redshifted to longer and longer wavelengths—and therefore cooler temperatures as in Figure 21.9. Today, the scale of the universe has increased a thousandfold since recombination, and the peak wavelength of the cosmic background radiation has increased by a thousandfold as well, to a value close to 1 millimeter (mm). The spectrum of the CMB still has the shape of a Planck blackbody spectrum, but with a characteristic temperature a thousandth what it was at the time of recombination.

Variations in the CMB

The presence of cosmic background radiation with a Planck blackbody spectrum is a strong prediction of the Big Bang theory. Penzias and Wilson had confirmed that a signal with the correct strength was there, but they could not say for certain whether the signal they saw had the spectral shape of a Planck blackbody spectrum. From the late 1960s to the 1980s, experiments at different wavelengths supported these same conclusions. The Cosmic Background Explorer (COBE) satellite made extremely precise measurements of the CMB at many wavelengths, from a few microns out to 1 cm. As you can see in **Figure 21.12**, the COBE spectrum of the CMB is a Planck blackbody spectrum with a temperature of 2.73 K. The observed spectrum so perfectly matches the one predicted by Big Bang cosmology that there can be no real doubt that this is the residual radiation left behind from the primordial fireball of the early universe.

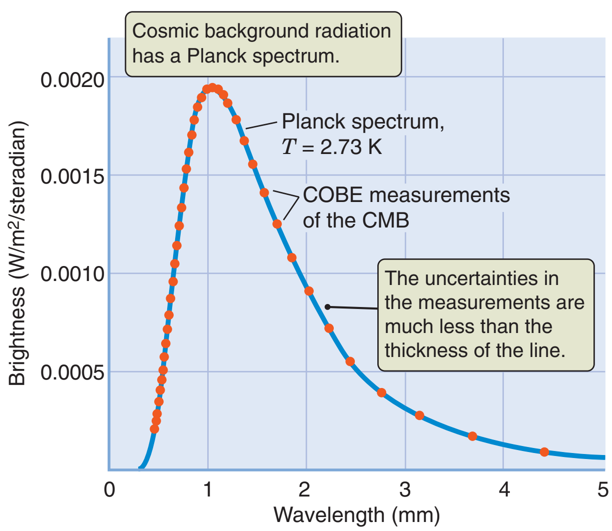


Figure 21.12 The spectrum of the CMB as measured by the Cosmic Background Explorer (COBE) satellite is shown by the red dots. The uncertainty in the measurement at each wavelength is much less than the size of a dot. The line running through the data is a blackbody spectrum with a temperature of 2.73 K.

COBE data included much more than a measurement of the spectrum of the cosmic background radiation. **Figure 21.13a** shows a map obtained by COBE of the CMB from the entire sky. The different colors in the map correspond to variations of about 0.1 percent in the temperature of the CMB. Most of this range of temperature is present because one side of the sky looks slightly warmer than the opposite side of the sky. This difference has nothing to do with the large-scale structure of the universe itself, but rather is the result of the motion of Earth with respect to the CMB.

We have emphasized that there is no preferred frame of reference. The laws of physics are the same in any inertial reference frame, so none is better than any other. Yet in a certain sense there is a preferred frame of reference at every point in the universe. This is the frame of reference that is at rest with respect to the expansion of the universe and in which the CMB is isotropic, or the same in all directions. The COBE map shows that one side of the sky is slightly hotter than the other because Earth and the Sun are moving at a velocity of 368 km/s in the direction of the constellation Crater relative to this cosmic reference frame.

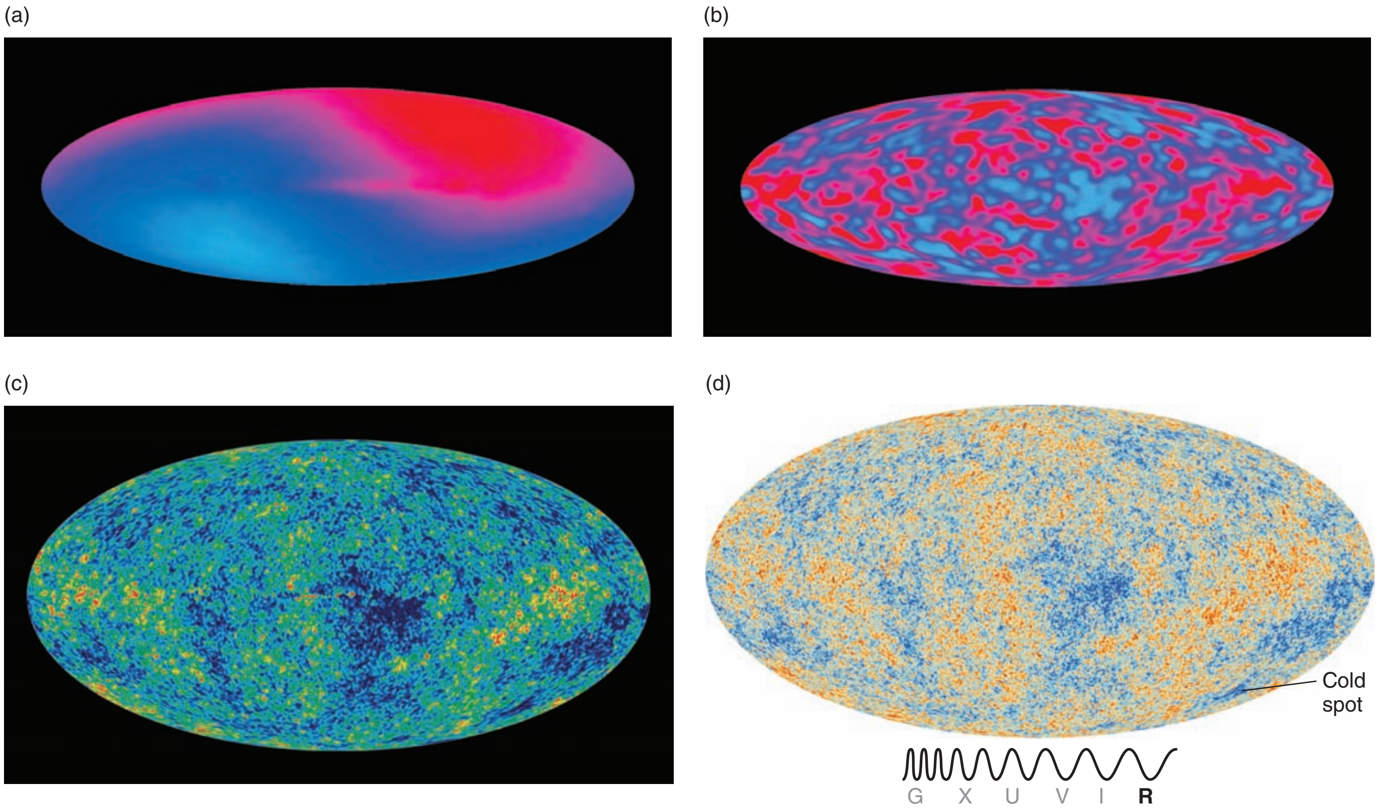


Figure 21.13 (a) The COBE satellite mapped the temperature of the CMB. The CMB is slightly hotter (by about 0.003 K) in one direction in the sky than in the other direction. This difference is due to Earth's motion relative to the CMB. (b) The COBE map with Earth's motion removed, showing tiny ripples remaining in the CMB. (c) *WMAP* (Wilkinson Microwave Anisotropy Probe) confirmed the fundamentals of cosmological theory at small and intermediate scales. (d) The Planck space observatory has provided the highest resolution yet of the CMB and has detected some surprises, such as the "cold spot." The radiation seen here was emitted less than 400,000 years after the Big Bang. Blue spots are cooler, red spots are warmer.

Radiation coming from the direction in which Earth is moving is slightly blueshifted (and thus shifted to a higher temperature) by this motion, whereas radiation coming from the opposite direction is Doppler-shifted toward the red (or cooler temperatures). Earth's motion is due to a combination of factors, including the motion of the Sun around the center of the Milky Way Galaxy and the motion of the Milky Way relative to the CMB.

When this asymmetry in the CMB caused by the motion of Earth is subtracted from the COBE map, only slight variations in the CMB remain, as shown in Figure 21.13b. The brighter parts of this image are only about 0.001 percent brighter than the fainter parts. These slight variations might not seem like much, but they are actually of crucial importance in the history of the universe. Recall from Chapter 18 that gravity itself can create a redshift. These tiny fluctuations in the cosmic background radiation are the result of gravitational redshifts caused by concentrations of mass that existed in the early universe. These concentrations later gave rise to galaxies and the rest of the structure that is evident in the universe today.

Beginning in 2001, NASA's *WMAP* (Wilkinson Microwave Anisotropy Probe) satellite made more precise measurements of the variations of the CMB. Figure 21.13c shows the ripples measured by *WMAP* with much higher resolution than could be detected by COBE. The European Space Agency's Planck space observatory collected data of even higher resolution from 2009 to 2012. A "cold spot" first detected in the *WMAP* data was confirmed to be real (Figure 21.13d). The much higher-resolution maps obtained by *WMAP* and Planck enable astronomers to

Origins

Big Bang Nucleosynthesis

The expansion of the universe and the cosmic microwave background radiation are two of the key pieces of observational evidence supporting the Big Bang theory. The third major piece of supporting evidence comes from observations of the number and types of chemical elements in the universe. For a short time after the Big Bang, the temperature and density of the universe was high enough for nuclear reactions to take place. Collisions between protons in the early universe built up low-mass nuclei, including deuterium (heavy hydrogen) and isotopes of helium, lithium, and beryllium. This process of element creation, called **Big Bang nucleosynthesis**, determined the final chemical composition of the matter that emerged from the hot phase of the Big Bang. Because the universe was rapidly expanding and cooling, the density and temperature of the universe fell too low for fusion to heavier elements such as carbon to occur. Therefore, all elements more massive than beryllium, including most of the atoms that make up Earth and its life, must have formed in subsequent generations of stars.

▶II **AstroTour:** Big Bang Nucleosynthesis

Figure 21.14 shows the observed and calculated predictions of the amounts of deuterium, helium, and lithium from Big Bang nucleosynthesis, plotted as a function of the observed present-day density of normal (luminous) matter in the universe. Observations of current abundances are shown as horizontal bands. Theoretical predictions, which depend on the density of the universe, are shown as darker, thick lines. Big Bang nucleosynthesis predicts that about 24 percent of the mass of the normal matter formed in the early universe should have ended up in the form of the very stable isotope ^{4}He, regardless of the total density of matter in the universe. Indeed, this is what is observed—about 24 percent of the mass of normal matter in the universe today is in the form of ^{4}He, in complete agreement with the prediction of Big Bang nucleosynthesis. This agreement between theoretical predictions and observation provides powerful evidence that the universe began in a Big Bang.

Unlike helium, most other isotope abundances depend on the density of normal matter in the universe, so comparing current abundances with models of isotope formation in the Big Bang helps pin down the density of the early universe. Beginning with the abundances of isotopes such as ^{2}H (deuterium) and ^{3}He found in the universe today (shown as the roughly horizontal, light-colored bands in Figure 21.14) and comparing them to predictions in different models of how abundant isotopes *should* be when formed at different densities (dark-colored curves in Figure 21.14), cosmologists can find the density of normal matter (the vertical yellow band in Figure 21.14). The best current measurements give a value of about 3.9×10^{-28} kg/m^3 for the average density of normal matter in the universe today (at sea level, the density of air is about 1.2 kg/m^3). This value lies refine their ideas about the development of structure in the early universe, which we will discuss in Chapters 22 and 23.

CHECK YOUR UNDERSTANDING 21.5

The existence of the cosmic microwave background radiation tells us that the early universe was: (a) much hotter than it is today; (b) much colder than it is today; (c) about the same temperature as today but was much more dense.

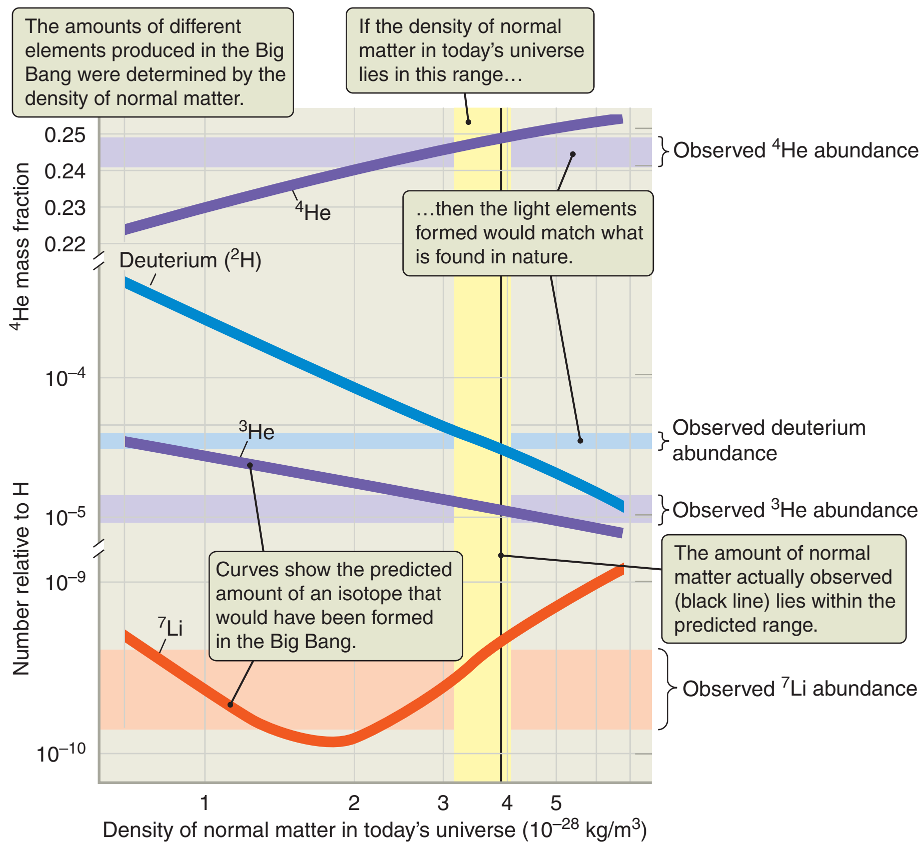


Figure 21.14 Observed and calculated abundances of the products of Big Bang nucleosynthesis, plotted against the density of normal matter in today's universe. Big Bang nucleosynthesis correctly predicts the amounts of these isotopes found in the universe today. (Note the two scale breaks on the *y*-axis.)

well within the range predicted by the observations shown in Figure 21.14. The agreement is remarkable, and it holds for many different isotopes.

Turning this around, cosmologists can begin with an observation of the amount of normal matter in and around galaxies and then compare that with calculations of what the chemical composition emerging from the Big Bang should have been. The observations agree remarkably well with the amounts of these elements actually found in nature. The idea that light elements originated in the Big Bang is resoundingly confirmed.

This agreement also provides a powerful constraint on the nature of dark matter, which dominates the mass in the universe. Dark matter cannot consist of normal matter made up of neutrons and protons; if it did, the density of neutrons and protons in the early universe would have been much higher, and the resulting abundances of light elements in the universe would have been much different from what is actually observed.

ARTICLES QUESTIONS

Scientists celebrate 50 years since the detection of the cosmic microwave background radiation.

50th Anniversary of the Big Bang Discovery

By **JOANNE COLELLA, The Journal NJ**

A unique gathering of some of the most brilliant minds in the world was held last month on Holmdel's Crawfords Hill on a beautiful spring day, celebrating the 50th anniversary of a truly stellar discovery: the detection of cosmic microwave background radiation (CMB), the thermal echo of the universe's explosive birth, and the evidence that proved the famed Big Bang theory, which would have taken place about 13.8 billion years ago. On May 20, 1964, American radio astronomers Robert Wilson and Arno Penzias confirmed that discovery, admittedly by accident but only after an exhaustive amount of investigative research to rule out every possible explanation for an odd buzzing sound that came from all parts of the sky at all times of day and night. The hum was detected by the enormous Horn Antenna at the Bell Labs site, now a national landmark. Puzzled by the noise, but initially not suspecting its significance, the pair went to great lengths to determine, or rule out, any possible source—including some pigeons that had nested in the antenna and were determined to return, even after being shipped to a distant location.

In 1978, Dr. Wilson and Dr. Penzias won the esteemed Nobel Prize in Physics for their work. Now 78 and 81 years old, respectively, the two came together again at the Horn Antenna site to celebrate the momentous anniversary with current and former Bell Labs colleagues. The event was headed up by Bell Labs President and Corporate CTO Marcus Weldon, who lauded their achievement and spoke about the company initiative to return "back to the future"—back to the classic model of Bell Labs, the research arm of Alcatel-Lucent, working to invent the future. Regarding the Big Bang theory, Mr. Weldon brought chuckles by stating, "In the beginning there was nothing and then it exploded . . . and then there was Arno and Bob." Subtle humor was a recurring theme by some speakers, who even poked fun at myths about Bell Labs, which has produced 12 Nobel Prize laureates, harboring extraterrestrial staff and facilities. Dr. Wilson and Dr. Penzias each spoke personally about their backgrounds, their work together, and the meaning of their discovery. The pair's distinct styles and personalities complemented each other perfectly. "It is very satisfying to look back and see we did our job right," said Dr. Wilson quietly and humbly. "This was pretty heady stuff," stated the talkative Dr. Penzias. "This is as close to being religious as I can be."

Throughout the presentations and celebratory luncheon, held under oversized tents on the expansive grounds, the air of excitement, pride, and mutual admiration among the attendees was palpable, as they looked back—and looked ahead—to the incredible legacy and pool of talent shared by Bell Labs personnel over the years. Robert Wilson and his wife still reside in Holmdel, as do many other Bell Labs employees, and the company has been an integral component of Holmdel and surrounding communities.

During the event, details were also announced about the establishment of the Bell Labs Prize, an annual competition to give scientists around the globe the chance to introduce their ideas in the fields of information and communications technology. The challenge offers a grand prize of $100,000, second prize of $50,000, and third prize of $25,000. Winners may also get the chance to develop their ideas at Bell Labs. The program is intended to inspire world-changing discoveries and innovations by young researchers.

ARTICLES QUESTIONS

1. What were the other main pieces of evidence supporting the Big Bang that were known before this discovery?
2. Why was this discovery seen as the final confirmation of the Big Bang?
3. Is "thermal echo of the universe's explosive birth" a good way to describe the CMB?
4. Why might Bell Labs have been supporting the research that led to this discovery in 1964?

"50th anniversary of the Big Bang discovery," by Joanne Colella. The Journal, May 29, 2014. Reprinted by permission of The Holmdel Journal, Shrewsbury, NJ.

Summary

The universe has been expanding since the Big Bang, which occurred nearly 14 billion years ago. The Big Bang happened everywhere: it is not an explosion spreading out from a single point. There are three major pieces of evidence for the Big Bang. Hubble's law states that more distant galaxies have higher redshifts, and these observed redshifts result from the increasing space between them. Observations of the cosmic microwave background radiation independently confirm that the universe had a hot, dense beginning. Lastly, the amounts of helium and trace amounts of other light elements measured today agree with what would be expected from nuclear reactions of normal matter in the hot early universe.

LG 1 **Explain in detail the cosmological principle.** Observations suggest that on the largest scales, the universe is homogeneous (looks the same to all observers) and isotropic (looks the same in all direction), in agreement with the cosmological principle.

LG 2 **Describe how the Hubble constant can be used to estimate the age of the universe.** Hubble found that all galaxies are moving away, with a recession velocity proportional to distance, indicating that the universe is expanding. The local physics and structure of objects are not affected by the expansion. Running the Hubble expansion backward leads to the Hubble time, and the age of the universe, found from the inverse of the Hubble constant.

LG 3 **Describe the observational evidence for the Big Bang.** Hubble's law states that light from distant galaxies is redshifted; this occurs because space itself is expanding. Hubble's law suggests that the universe was once very hot and very dense, a beginning known as the Big Bang. Big Bang theory predicts that we should be able to observe the radiation from a few hundred thousand years after the Big Bang. This radiation, called the cosmic microwave background radiation, has the same spectrum in every direction.

LG 4 **Explain which chemical elements were created in the early hot universe.** During the first few minutes after the Big Bang, the universe was hot enough for nucleosynthesis to take place through the process of fusion. Deuterium, helium, lithium, and beryllium, but not the heavier elements, were created before the universe cooled too much for this fusion to take place.

? UNANSWERED QUESTIONS

- What existed before the Big Bang? The usual (and somewhat unsatisfactory) answer is that the Big Bang was the beginning of space and *time*, so there could be no time before it happened. A more recent answer is that this universe may be just one of many universes, and the Big Bang was the beginning of this universe only. We will return to this topic in Chapter 22.
- If there were evidence that the cosmological principle were not true, then could we deduce anything about the evolution of the universe?

Questions and Problems

Test Your Understanding

1. What do astronomers mean when they say that the universe is homogeneous?
   a. The universe looks exactly the same from every perspective.
   b. Galaxies are generally distributed evenly throughout the universe.
   c. All stars in all galaxies have planetary systems just like ours.
   d. The universe has looked the same at all times in its history.
2. What do astronomers mean when they say that the universe is isotropic?
   a. More distant parts of the universe look just like nearby parts.
   b. Intergalactic gas has the same density everywhere in the universe.
   c. The laws of physics apply everywhere in the universe.
   d. The universe looks the same in every direction.

3. Cosmological redshifts are calculated from observations of spectral lines from
   a. individual stars in distant galaxies.
   b. clouds of dust and gas in distant galaxies.
   c. spectra of entire galaxies.
   d. rotations of the disks of distant galaxies.

4. When they look into the universe, astronomers observe that nearly all galaxies are moving away from the Milky Way. This observation suggests that
   a. the Milky Way is at the center of the universe.
   b. the Milky Way must be at the center of the expansion.
   c. the Big Bang occurred at the location of the Milky Way.
   d. an observer in a distant galaxy would make the same observation.

5. Some galaxies have redshifts z that if equated to v_r/c correspond to velocities greater than the speed of light. Special relativity is not violated in this case
   a. because of relativistic beaming.
   b. because of superluminal motion.
   c. because redshifts carry no information.
   d. because these velocities do not measure motion through space.

6. The Big Bang theory predicted (select all that apply)
   a. the Hubble law.
   b. the cosmic microwave background radiation.
   c. the cosmological principle.
   d. the abundance of helium.
   e. the period-luminosity relationship of Cepheid variables.

7. The simplest way to estimate the age of the universe is from
   a. using the slope of Hubble's law.
   b. the age of Moon rocks.
   c. models of stellar evolution.
   d. measurements of the abundances of elements.

8. The CMB includes information about (select all that apply)
   a. the age of the universe.
   b. the temperature of the early universe.
   c. the density of the early universe.
   d. density fluctuations in the early universe.
   e. the motion of Earth around the center of the Milky Way.

9. Repeated measurements showing that the current helium abundance is much less than the value predicted by the Big Bang would imply that
   a. some part of the Big Bang theory is incorrect or incomplete.
   b. the current helium abundance is wrong.
   c. scientists don't know how to measure helium abundances.

10. The cosmological principle says that
    a. the universe is expanding.
    b. the universe began in the Big Bang.
    c. the rules that govern the universe are the same everywhere.
    d. the early universe was 1,000 times hotter than the characteristic temperature of the CMB.

11. Why is the Milky Way Galaxy not expanding together with the rest of the universe?
    a. It is not expanding because it is at the center of the expansion.
    b. It is expanding, but the expansion is too small to measure.
    c. The Milky Way is a special location in the universe.
    d. Local gravity dominates over the expansion of the universe.

12. The scale factor keeps track of
    a. the movement of galaxies through space.
    b. the current distances between many galaxies.
    c. the changing distance between any two galaxies.
    d. the location of the center of the universe.

13. The Big Bang is
    a. the giant supernova explosion that triggered the formation of the Solar System.
    b. the explosion of a supermassive black hole.
    c. the eventual demise of the Sun.
    d. the beginning of space and time.

14. The CMB is essentially uniform in all directions in the sky. This is an example of
    a. anisotropy.
    b. isotropy.
    c. thermal fluctuations.
    d. Wien's law.

15. Which of the following was *not* created as a result of Big Bang nucleosynthesis?
    a. helium
    b. lithium
    c. hydrogen
    d. deuterium
    e. carbon

Thinking about the Concepts

16. Imagine that you are standing in the middle of a dense fog.
    a. Would you describe your environment as isotropic? Why or why not?
    b. Would you describe it as homogeneous? Why or why not?

17. As the universe expands from the Big Bang, galaxies are not actually flying apart from one another. What is really happening?

18. We see the universe around us expanding, which gives distant galaxies an apparent velocity of 70 km/s/Mpc. If you were an astronomer living today in a galaxy that was located 1 billion light-years away from us, at what rate would you see the galaxies moving away from you?

19. Does Hubble's law imply that our galaxy is sitting at the center of the universe? Explain.

20. What does the value of R_U, the scale factor of the universe, tell us?

21. Does the expansion of the universe make the Sun bigger? What about the Milky Way? Why or why not?

22. Science's greatest strength is the self-correcting nature of scientific inquiry: minds can be changed in the face of new data, as described in the Process of Science Figure. Consider a modern paradigm of science such as the Big Bang, climate change, or the power source of stars. In a short paragraph, describe your current viewpoint, and give a piece of evidence that would make you change your mind if it were true. For example, Einstein believed the universe was static. Measurements of the movement of galaxies caused him to change his mind.

23. Name two predictions of the standard Big Bang theory that have been verified by observations.

24. Knowing that you are studying astronomy, a curious friend asks where the center of the universe is located. You answer, "Right here and everywhere." Explain in detail why you give this answer.

25. The general relationship between recession velocity (v_r) and redshift (z) is $v_r = cz$. This simple relationship fails, however, for very distant galaxies with large redshifts. Explain why.

26. Why is it significant that the CMB displays a Planck blackbody spectrum?

27. What is the significance of the tiny brightness variations that are observed in the CMB?

28. What important characteristics of the early universe are revealed by today's observed abundances of various isotopes, such as ^{2}H and ^{3}He?

29. Why were only a few of the chemical elements created at the Big Bang?

30. How do astronomers know that some of the observed helium is left over from the Big Bang?

Applying the Concepts

31. In Figure 21.6, a rubber sheet is shown as an analogy to help you think about the scale factor. Between the moments shown in parts (a) and (b), each square doubles in size on every edge. How does the area of a square change? Imagine that the sheet is now a block of rubber, expanding in three dimensions instead of two. How would the volume of a cube change between the moment shown in part (a) and the moment shown in part (b)?

32. Study Figure 21.12. Error bars have not been plotted in this figure. Why not? Was this a very precise measurement or a very imprecise measurement? How does the precision of the measurement affect your confidence in the conclusions drawn from it?

33. Study Figure 21.14. This figure includes both predictions and observations.
    a. What do the vertical yellow bar and the slanted lines and curves represent: theory or observation?
    b. What do the pastel horizontal lines and the vertical black line represent: predictions or observations?
    c. Do the predictions and observations match? Choose one example, and explain how you know.

34. The Hubble time ($1/H_0$) represents the age of a universe that has been expanding at a constant rate since the Big Bang. Calculate the age of the universe in years if $H_0 = 80$ km/s/Mpc. (Note: 1 year $= 3.16 \times 10^7$ seconds, and 1 Mpc $= 3.09 \times 10^{19}$ km.)

35. Throughout the latter half of the 20th century, estimates of H_0 ranged from 50 to 100 km/s/Mpc. Calculate the age of the universe in years for each of these estimated values of H_0. Do any answers contradict the ages obtained from stellar evolution or geology on Earth?

36. Suppose a galaxy is observed with a redshift equal to 2. How much has the universe expanded since that light was emitted from these galaxies?

37. How much has the universe expanded since light was emitted from a galaxy with a redshift of $z = 8$?

38. You observe a distant quasar in which a spectral line of hydrogen with rest wavelength $\lambda_{rest} = 121.6$ nm is found at a wavelength of 547.2 nm. What is its redshift? When the light from this quasar was emitted, how large was the universe compared to its current size?

39. A distant galaxy has a redshift $z = 5.82$ and a recession velocity $v_r = 287{,}000$ km/s (about 96 percent of the speed of light).
    a. If $H_0 = 70$ km/s/Mpc and if Hubble's law remains valid out to such a large distance, then how far away is this galaxy?
    b. Assuming a Hubble time of 13.8 billion years, how old was the universe at the look-back time of this galaxy?
    c. What was the scale factor of the universe at that time?

40. The spectrum of the CMB is shown as the red dots in Figure 21.12, along with a blackbody spectrum for a blackbody at temperature of 2.73 K. From the graph, determine the peak wavelength of the CMB spectrum. Use Wien's law to find the temperature of the CMB. How does this rough measurement that you just made compare to the accepted temperature of the CMB?

41. COBE observations show that the Solar System is moving in the direction of the constellation Crater at a speed of 368 km/s relative to the cosmic reference frame. What is the blueshift (negative value of z) associated with this motion?

42. The average density of normal matter in the universe is 4×10^{-28} kg/m^3. The mass of a hydrogen atom is 1.66×10^{-27} kg. On average, how many hydrogen atoms are there in each cubic meter in the universe?

43. To get a feeling for the emptiness of the universe, compare its density (4×10^{-28} kg/m^3) with that of Earth's atmosphere at sea level (1.2 kg/m^3). How much denser is Earth's atmosphere? Write this ratio using standard notation.

44. Assume that the most distant galaxies have a redshift $z = 10$. The average density of normal matter in the universe today is 4×10^{-28} kg/m^3. What was its density when light was leaving those distant galaxies? (Hint: Keep in mind that volume is proportional to the cube of the scale factor.)

45. What was the size of the universe (compared to the present) when the CMB was emitted, at $z = 1{,}000$?

USING THE WEB

46. For more details on the history of the discovery of the expanding universe, go to the American Institute of Physics' "Cosmic Journey: A History of Scientific Cosmology" website (www.aip.org/history/cosmology/). Read through the sections titled "Island Universes," "The Expanding Universe," and "Big Bang or Steady State?" Why was Albert Einstein "irritated" by the idea of an expanding universe? What was the contribution of Belgian astrophysicist (and Catholic priest) Georges Lemaître? What is the steady-state theory, and what was the main piece of evidence against it?

47. Go to the American Institute of Physics page on the age of the universe, http://www.aip.org/history/curie/age-of-earth.htm. How does the age of the universe from Hubble's law compare with the age of the Solar System? How does it compare with the age of the oldest globular clusters? Do a search for "age of the universe"—has the value been updated?

48. A series of public lectures at Harvard commemorating the 50th anniversary of the discovery of the CMB can be found here: http://astronomy.fas.harvard.edu/news/public-talk-50-year-anniversary-discovery-cosmic-micrrowave-background. Robert Wilson's talk starts at the 27-minute mark. Why does he say this is the most important thing he ever measured? How did he identify this radiation?

49. Updated observations of the CMB from the Planck space observatory are reported on the European Space Agency's Planck website (http://esa.int/SPECIALS/Planck). What has been learned from this mission? What is their value of the Hubble constant?

50. Go to the University of Washington Astronomy Department's "Hubble's Law: An Introductory Astronomy Lab" Web page (http://www.astro.washington.edu/courses/labs/clearinghouse/labs/HubbleLaw/hubbletitle) and do the lab exercise, which uses real data from galaxies to calculate Hubble's constant. Your instructor will indicate whether you should use the regular or the shorter version.

smartwork5

If your instructor assigns homework in Smartwork5, access your assignments at digital.wwnorton.com/astro5.

EXPLORATION

Hubble's Law for Balloons

digital.wwnorton.com/astro5

The expansion of the universe is extremely difficult to visualize, even for professional astronomers. In this Exploration, you will use the surface of a balloon to get a feel for how an "expansion" changes distances between objects. Throughout this Exploration, remember to think of the surface of the balloon as a two-dimensional object, much as the surface of Earth is a two-dimensional object for most people. The average person can move east or west, or north or south, but into Earth and out to space are not options. For this Exploration you will need a balloon, 11 small stickers, a piece of string, and a ruler. A partner is helpful as well. **Figure 21.15** shows some of the steps involved.

Blow up the balloon partially and hold it closed, but *do not tie it shut*. Stick the 11 stickers on the balloon (these represent galaxies) and number them. Galaxy 1 is the reference galaxy.

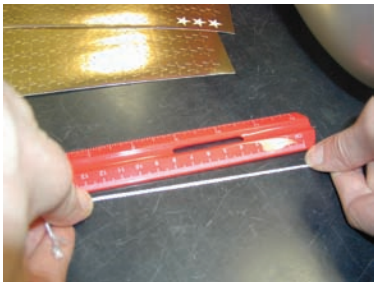
Measure the distance between the reference galaxy and each of the galaxies numbered 2–10. The easiest way to do this is to use your piece of string. Lay it along the balloon between the two galaxies and then measure the length of the string. Record these data in the "Distance 1" column of a table like the one shown later.

Simulate the expansion of your balloon universe by *slowly* blowing up the balloon the rest of the way. Have your partner count the number of seconds it takes you to do this, and record this number in the "Time Elapsed" column of the table (each row has the same time elapsed, because the expansion occurred for the same amount of time for each galaxy). Tie the balloon shut. Measure the distance between the reference galaxy and each numbered galaxy again. Record these data under "Distance 2."

Subtract the first measurement from the second. Record the difference in the table.

Divide this difference, which represents the distance traveled by the galaxy, by the time it took to blow up the balloon. Distance divided by time gives an average speed.

Make a graph with velocity on the *y*-axis and distance 2 on the *x*-axis to get "Hubble's law for balloons." You may wish to roughly fit a line to these data to clarify the trend.



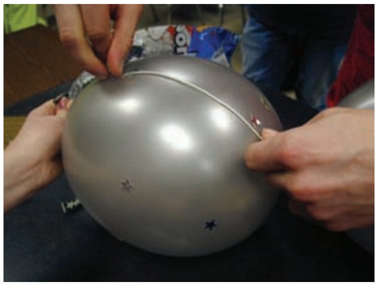


Figure 21.15 Measure the distance around a curved balloon using a string.

Galaxy Number	Distance 1	Distance 2	Difference	Time Elapsed	Velocity
1 (reference)	0	0	0		0
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					

1 Describe your data. If you fit a line to them, is it horizontal or does it trend upward or downward?

2 Is there anything special about your reference galaxy? Is it different in any way from the others?

3 If you had picked a different reference galaxy, would the trend of your line be different? If you are not sure of the answer, get another balloon and try it.

4 The expansion of the universe behaves similarly to the movement of the galaxies on the balloon. We don't want to carry the analogy too far, but there is one more thing to think about. In your balloon, some areas probably expanded less than others because the material was thicker; there was more "balloon stuff" holding it together. How is this similar to some places in the actual universe?

22 Cosmology

Cosmology is the study of the large-scale universe, including its nature, origin, evolution, and ultimate destiny. In this chapter, we take a closer look at the nature of the universe, how it has evolved over time, and its ultimate fate. We also discuss the physics of the smallest particles, which is necessary for describing the earliest moments of the universe.

LEARNING GOALS

By the conclusion of this chapter, you should be able to:

LG 1 Explain how mass within the universe and the gravitational force it produces affect the history, shape, and fate of the universe.

LG 2 Describe the evidence for the accelerating expansion of the universe.

LG 3 Describe the early period of rapid expansion known as inflation.

LG 4 Explain how the events that occurred in the earliest moments of the universe are related to the forces that operate in the modern universe.

The arrows point to a Type Ia supernova observed in the galaxy NGC 1365. ▶▶▶

What will be the fate of the universe?

22.1 Gravity and the Expansion of the Universe

The fate of the universe is a central question of modern cosmology. The simplest answer depends in part on the average mass in the universe within a fixed volume. A number of factors determine the way mass is distributed on large scales across the universe. In this section, we will look at one of these factors: gravity.

Mass Distribution

To see how gravity affects the expansion of the universe, recall gravity's effects on the motion of projectiles and the discussion of escape velocity from Chapter 4. For example, the fate of a projectile fired straight up from the surface of a planet depends on its speed. As long as the projectile speed is less than the planet's escape velocity, gravity will eventually stop the rise of the projectile and pull it back to the planet's surface. But if the speed of the projectile is greater than the planet's escape velocity, gravity will lose. In this case, although the projectile will slow down, it will never stop, and it will escape from the planet entirely.

Recall that the escape velocity from a planet's surface depends on the mass and radius of the planet. If the planet is massive enough, it will slow the projectile, stop it, and pull it back down. However, if the planet is not massive enough, the projectile will slow down but still escape to space. The size of the planet is also important. A smaller planet's mass is more densely packed: if the projectile is fired from the surface, which is closer to the plant's center, the gravitational pull is stronger, and the planet can pull the projectile down. However, if the planet has lower density, then the projectile at the surface starts farther from the center of the planet, the gravitational pull is weaker, and the projectile will escape. Whether the projectile escapes depends on both the planet's mass and the planet's radius.

Just as the mass of the planet gravitationally pulls on a projectile to slow its climb, the mass distributed across the universe, as characterized by its average density, gravitationally slows the expansion of the universe. If there is enough mass in the universe, then gravity will be strong enough to stop the expansion. However, if the radius of the universe is large, even a lot of mass won't lead to a high escape velocity. Thus, the density is critical in determining the fate of the universe. If the mass is packed closely together, so that the average density is high, the expansion will slow, stop, and reverse. If the mass is very spread out, so that the average density is low, the universe will expand forever.

Critical Density

The "escape velocity" of the universe is also determined by its average density. A faster expansion requires a higher density to stop the expansion. Astronomers define a **critical density** for which the mass in the universe would cause it to just barely stop expanding after a very long time. This critical density determines the dividing line between two possible fates of the universe: expanding forever or collapsing. If the universe is less dense than the critical density, gravity will be too weak, and the universe will expand forever. If the universe is denser on average than the critical density, then gravity will be strong enough to stop and reverse the expansion eventually.