listed in part (a), explain what the source of the released energy is.

30. Why is the maximum mass of neutron stars not known as accurately as the Chandrasekhar limit for white dwarfs?

Advanced Questions

Problem-solving tips and tools

The volume of a sphere of radius r is $4\pi r^3/3$. The small-angle formula is given in Box 1-1. Section 5-4 describes how to relate the temperature of a blackbody to its wavelength of maximum emission. Section 17-6 gives the formula relating a star's luminosity, surface temperature, and radius (see Box 17-4 for worked examples). Appendix 6 gives the conversion between seconds and years as well as the radius of the Sun.

31. Using a diagram like Figure 21-3, explain why the number of pulsars that we observe in nearby space is probably quite a bit less than the number of rotating, magnetized neutron stars in nearby space.
32. There are many more main-sequence stars of low mass (less than 8 $M_\odot$) than of high mass (8 $M_\odot$ or more). Use this fact to explain why white dwarf stars are far more common than neutron stars.
33. The distance to the Crab Nebula is about 2000 parsecs. In what year did the star actually explode? Explain your answer.
34. How do we know that the Crab pulsar is really embedded in the Crab Nebula and not simply located at a different distance along the same line of sight?
35. The Crab Nebula has an apparent size of about 5 arcmin, and this size is increasing at a rate of 0.23 arcsec per year. (**a**) Assume that the expansion rate has been constant over the entire history of the Crab Nebula. Based on this assumption, in what year would Earth observers have seen the supernova explosion that formed the nebula? (**b**) Does your answer to part (a) agree with the known year of the supernova, 1054 A.D.? If not, can you point to assumptions you made in your computations that led to the discrepancies? Or do you think your calculations suggest additional physical effects are at work in the Crab Nebula, over and above a constant rate of expansion?
36. Emission lines in the spectrum of the Crab Nebula exhibit a Doppler shift, which indicates that gas in the part of the nebula closest to us is moving toward us at 1450 km/s. (**a**) Assume that the expanding gas has been moving at the same speed since the original supernova explosion, observed in 1054 A.D., and calculate what radius and what diameter (in light-years) we should observe the nebula to have today. (**b**) Compare your result in part (a) to the actual size of the nebula, given in the caption to Figure 21-4.
37. The supernova remnant G5.4-1.2 shown in Figure 21-5 lies about 5000 pc (16,000 ly) from Earth in the constellation Sagittarius. (**a**) The green arc in the large left-hand image in Figure 21-5 represents part of the outer edge of this spherical supernova remnant. Estimate the diameter of this remnant in parsecs. (*Hint:* To calculate this, you will need to make measurements on the image.) (**b**) How far (in parsecs) did the neutron star travel from where it was formed (at the position of the supernova's progenitor star, presumably at the center of the present-day remnant) to the position shown in Figure 21-5? Explain your answer.
38. To determine accurately the period of a pulsar, astronomers must take into account the Earth's orbital motion about the Sun. (**a**) Explain why. (**b**) Knowing that the Earth's orbital velocity is 30 km/s, calculate the maximum correction to a pulsar's period because of the Earth's motion. Explain why the size of the correction is greatest for pulsars located near the ecliptic.
39. If a pulsar has period P (in seconds) and its period is increasing at a rate R (in seconds per second), an approximate formula for the age T of the pulsar (in seconds) is $T = P/2R$. For the Crab pulsar, $P = 0.0333$ s and $R = 4.21 \times 10^{-13}$ s/s. (**a**) Calculate the approximate age of the Crab pulsar in years. (**b**) Based on the information given in Section 21-1, is your result in (a) an underestimate or an overestimate? Explain your answer.
40. (See Advanced Question 39.) (**a**) Magnetar SGR 1806-20 was the source of the intense burst observed on December 27, 2004. Based on the information given in Section 21-6, calculate the rate R (in seconds per second) at which the period of SGR 1806-20 is increasing. (**b**) The ordinary pulses from SGR 1806-20 have a period of 7.47 s. Calculate the approximate age of SGR 1806-20 in years. (**c**) Theory predicts that a magnetar becomes inactive after about 10,000 years. Is your result in part (b) consistent with this prediction? Explain why or why not.
41. A neutron has a mass of about 1.7×10^{-27} kg and a radius of about 10^{-15} m. (**a**) Compare the density of matter in a neutron with the average density of a neutron star. (**b**) If the neutron star's density is more than that of a neutron, the neutrons within the star are overlapping; if it is less, the neutrons are not overlapping. Which of these seems to be the case for average neutrons within the star? Which do you think is the case at the center of the neutron star, where densities are higher than average?
42. The total luminosity at all wavelengths of the magnetar burst observed on December 27, 2006, was approximately 10^{14} $L_\odot$. At what distance from the magnetar would the brightness of the burst have been equal to the brightness of the Sun as seen on Earth? Give your answer in AU and in parsecs.
43. X-ray pulsars are speeding up but ordinary (radio) pulsars are slowing down. Propose an explanation for this difference.
44. If the model for Hercules X-1 discussed in the text is correct, at what orientation of the binary system do we see its maximum optical brightness? Explain your answer.
45. Explain why heavy elements that are produced by neutron star collisions can still be regarded as having been processed through a supernova.
46. In an X-ray burster, the surface of a neutron star 10 km in radius is heated to a temperature of 3×10^7 K. (**a**) Determine the wavelength of maximum emission of the heated surface (which you may treat as a blackbody). In what part of the electromagnetic spectrum does this lie? (See Figure 5-7.) (**b**) Find the luminosity of the heated neutron star. Give your answer in watts and in terms of the luminosity of the Sun, given in Table 16-1. How does this compare with the peak luminosity of a nova? Of a Type Ia supernova?

47. The nearest neutron star, called RX J185635-3754, is just 60 pc (200 ly) from Earth. It is thought to be the relic of a star that underwent a supernova explosion about 1 million years ago. The explosion ejected the neutron star at high speed, so it is now moving through nearly empty space. (**a**) RX J185635-3754 is *not* a pulsar, that is, it does not emit pulses of radiation. Suggest why this might be so. (**b**) The neutron star has a surface temperature of 600,000 K. Find the wavelength at which it emits most strongly, and explain why the neutron star appears as a steady, nonpulsating object in an X-ray telescope. (**c**) RX J185635-3754 has a total luminosity at all wavelengths of about $0.046 L_{\odot}$. Calculate its radius, and explain why astronomers conclude that it is a neutron star.

Discussion Questions

48. Imagine that we're somehow able to stand (and survive) at one of the magnetic poles of the Crab pulsar. Describe what you would see. How would the stars appear to move in the sky? What would you see if you looked straight up? What factors make this location a very unhealthy place to visit?
49. Accretion disks in close binary systems are too small to be seen directly with even the highest-resolution telescopes. How, then, can astronomers detect the presence of such accretion disks?
50. When neutrons are very close to one another, they repel one another through the strong nuclear force. If this repulsion were made even stronger, what effect might this have on the maximum mass of a neutron star? Explain your answer.

Web/eBook Questions

51. Search the World Wide Web for information about the latest observations of the stellar remnant at the center of SN 1987A. Has a pulsar been detected? If so, how fast is it spinning? Has the supernova's debris thinned out enough to give a clear view of the neutron star?
52. Search the World Wide Web for the latest information about magnetars and soft gamma repeaters (also known as soft gamma-ray repeaters). How many soft gamma repeaters are known? Have any of these produced an intense burst since December 27, 2004?
53. **Monitoring the Crab Pulsar.** Access the video "The Crab Pulsar" in Chapter 21 of the *Universe* Web site or eBook. View the video and use it to answer the following questions. For each part, explain how you determined your answer. (**a**) How many rotations does the neutron star complete during the duration of the video? (**b**) Is the neutron star visible at the beginning of the video? If not, explain why not. (**c**) How does the peak brightness of the Crab pulsar compare to the steady brightness of the nearby star? (**d**) What total amount of time is depicted in the video?

Activities

Observing Projects

54. If you did not take the opportunity to observe the Crab Nebula as part of the exercises in Chapter 20, do so now. The Crab Nebula is visible in the night sky from October through March. Its epoch 2000 coordinates are R.A. = $5^{h}\ 34.5^{m}$ and Decl. = 22° 00′, which is near the star marking the eastern horn of Taurus (the Bull). Be sure to schedule your observations for a moonless night. The larger the telescope you use, the better, because the Crab Nebula is quite dim.
55. Consult the World Wide Web to see if any novae have been sighted recently. If by good fortune one has been sighted, what is its apparent magnitude? Is it within reach of a telescope at your disposal? If so, arrange to observe it. Draw what you see through the eyepiece, noting the object's brightness in comparison with other stars in the field of view. If possible, observe the same object a few weeks or months later to see how its brightness has changed.
56. Use the *Starry Night Enthusiast*™ program to observe the sky in July 1054, when the supernova that spawned the Crab Nebula was visible from the American Southwest. Select **Viewing Location . . .** in the **Options** menu, click on the **Latitude/Longitude** tab, enter 36° N for latitude and 109° W for longitude, and click on the **Set Location** button. You are now near the location of the pictograph shown in Figure 21-1. In the toolbar, set the **Date** to **July 5, 1054 A.D.** and the **Time** to **5:00 A.M.** Remove the artificial satellites from the view by clicking on **View > Solar System > Satellites,** since these would not have been in the sky at this time in history. Use the **Find** pane to find and center on the **Crab Nebula.** Zoom in or out until you can see both the position of the nebula and the Moon. You may find it helpful to turn daylight on or off (Select **Show Daylight** or **Hide Daylight** in the **View** menu). (**a**) What is the phase of the Moon? (**b**) Investigate how the relative positions of the Moon and the Crab Nebula change when you set the date to July 4, 1054, or July 6, 1054. On which date do the relative positions of the Moon and the Crab Nebula give the best match to the pictograph in Figure 21-1? (**c**) Zoom in on the Crab Nebula to see this supernova remnant. *Starry Night Enthusiast*™ superimposes an X-ray image of this active region from the Chandra space telescope on to the visible light image. To compare the images from these two different wavelengths, you can remove the X-ray image by opening the **Options** pane, expanding the **Deep Space** layer and clicking in the box next to **Chandra.**
57. Use the *Starry Night Enthusiast*™ program to locate the Small Magellanic Cloud, the site of a large concentration of X-ray pulsars. Open the **Favourites** menu and click on **Deep Space > Local Universe** to display the Milky Way and its nearby galaxies, suitably labeled, as seen from 0.282 Mly away. If the Milky Way does not appear immediately, click once on one of the Zoom buttons. Remove the image of the astronaut's feet by clicking on **View > Feet.** You can zoom in or out using the buttons at the upper right of the toolbar. You can rotate the view by holding down the **Shift** key while holding down the mouse button and moving the mouse. (On a two-button mouse, hold down the left mouse button.) (**a**) Open the **Find** pane and use the menu button to the left of the label for the **Sun** to **Centre** the field of view on the Sun. Describe the position of the Small Magellanic Cloud relative to the Milky Way Galaxy and to our solar system. (**b**) Use the **Find** pane to center on the Small Magellanic Cloud. Rotate this irregular

galaxy to view it face-on and zoom in to see its component stars. Pulsars are produced by supernovae, and only certain types of stars become supernovae. What evidence do you see that these types of stars are present in the Small Magellanic Cloud? (Note that ***Starry Night Enthusiast™*** depicts the Small Magellanic Cloud as being rather flat but this galaxy is thought to be an irregular blob of stars with some thickness.)

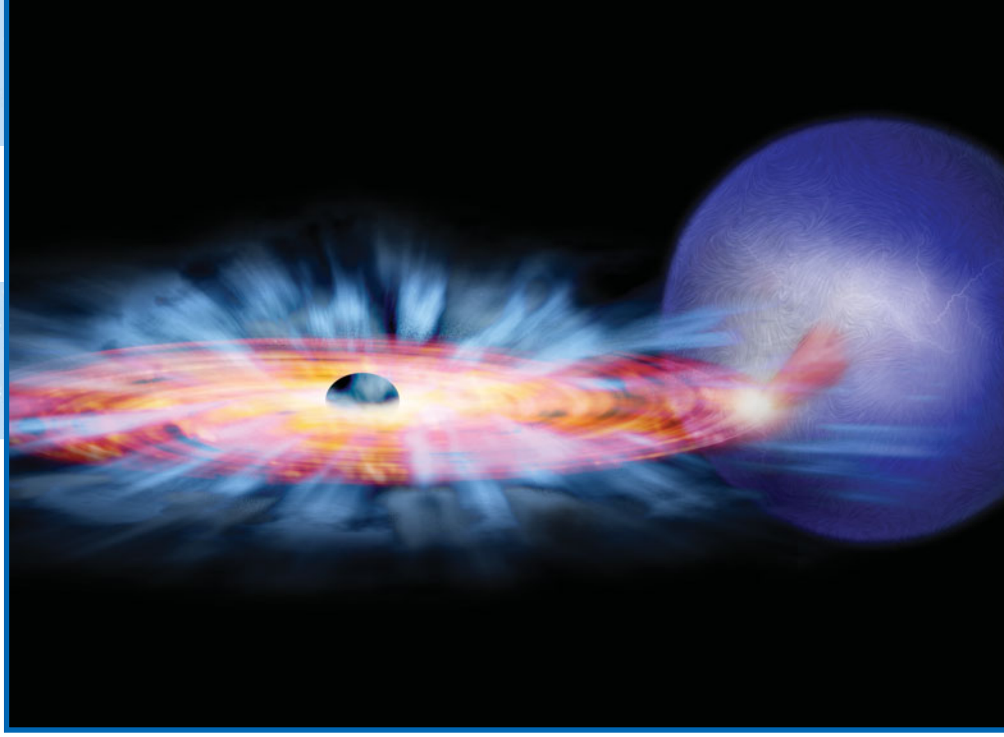
Collaborative Exercises

58. Consider the graph showing a recording of a pulsar in Figure 21-2. Sketch and label similar graphs that your group estimates for: (1) a rapidly spinning, professional ice skater holding a flashlight; (2) an emergency signal on an ambulance; and (3) a rotating beacon at an airport.

59. As stars go, pulsars are tiny, only about 20 km across. Name three specific things or places that have a size or a separation of about 20 km.

60. If the Crab Nebula is slowing such that its period is increasing at a rate of 3×10^{-8} seconds per day, how much slower is it going today than on the day the youngest member of your group was born?

22 Black Holes



An artist's impression of a close binary star system that includes a black hole. The accretion disk around the black hole is made of material drawn from the blue companion star. (NASA/CXC/M. Weiss)

WEB LINK 22.1

Imagine a swirling disk of gas and dust, orbiting around an object that has more mass than the Sun but is so dark that it cannot be seen. Imagine the material in this disk being compressed and heated as it spirals into the unseen object, reaching temperatures so high that the material emits X rays. And imagine that the unseen object has such powerful gravity that any material that falls into it simply disappears, never to be seen again.

Such hellish maelstroms, like the one shown in the illustration, really do exist. The disks are called *accretion disks,* and the unseen objects with immensely strong gravity are called *black holes.*

The matter that makes up a black hole has been so greatly compressed that it violently warps space and time. If you get too close to a black hole, the speed you would need to escape exceeds the speed of light. Because nothing can travel faster than light, nothing—not even light—can escape from a black hole.

Black holes, whose existence was predicted by Einstein's general theory of relativity (our best description of what gravity is and how it behaves), are both strange and simple. The structure of a black hole is completely specified by only three quantities—its mass, electric charge, and angular momentum. Perhaps strangest of all, there is compelling evidence that black holes really exist. In recent years, astronomers have found that certain binary star systems contain black holes. They have also found evidence that extremely massive stars form black holes at their centers, producing immense bursts of gamma rays in the process. Even more remarkable is the discovery that black holes of more than a million solar masses lie at the centers of many galaxies, including our own Milky Way.

22-1 The special theory of relativity changes our conceptions of space and time

The idea of a black hole is a truly strange one. To appreciate it, we must discard some "commonsense" notions about the nature of space and time.

The Special Theory of Relativity

To understand black holes, we must first grasp the nature of space and time as described by Einstein's special theory of relativity

According to the classical physics of Newton, space is perfectly uniform and fills the universe like a rigid framework. Similarly, time passes at a monotonous, unchanging rate. It is always possible to know exactly how fast you are moving through this rigid fabric of space and time.

WEB LINK 22.2

Those ideas were upset forever in 1905 when Albert Einstein proposed his **special theory of relativity.** This theory describes how motion affects our measurements

Learning Goals

By reading the sections of this chapter, you will learn

22-1 The main ideas of Einstein's special theory of relativity

22-2 How Einstein's general theory of relativity describes the nature of gravitation

22-3 The evidence for black holes in binary star systems

22-4 How the sudden formation of black holes can explain the mysterious gamma-ray bursters

22-5 How astronomers have detected supermassive black holes at the centers of galaxies

22-6 The simple structure of a nonrotating black hole

22-7 How just three numbers completely describe the properties of a black hole

22-8 What it might be like to approach a black hole

22-9 How black holes evaporate over time of distance and time. It is "special" in the sense of being specialized. In particular, it does not include the effects of gravity. The word "relativity" is used because one of the key ideas of the theory is that all measurements are made relative to an observer. Contrary to Newton, there is *no* absolute framework of space and time. In particular, the distance between two points is not an absolute, nor is the time interval between two events. Instead, the values that you measure for these qualities depend on how you are moving, and are relative to you. Someone moving in a different way would measure different values for these quantities.

Remarkably, Einstein's theory is based on just two basic principles. The first is quite simple:

Your description of physical reality is the same regardless of the constant velocity at which you move.

In other words, if you are moving in a straight line at a constant speed, you experience the same laws of physics as you would if you were moving at any other constant speed and in any other direction. As an example, suppose you were inside a closed railroad car moving due north in a straight line at 100 km/h. Any measurements you make inside the car—for example, how long your thumb is or how much time elapses between ticks of your watch—will give exactly the same results as if the car were moving in any other direction or at any other speed, or were not moving at all.

Einstein's second principle is much more bizarre:

Regardless of your speed or direction of motion, you always measure the speed of light to be the same.

To see what this implies, imagine that you are in a spaceship moving toward a flashlight. Even if you are moving at 99% of the speed of light, you will measure the photons from the flashlight to be moving at the same speed ($c = 3 \times 10^8$ m/s $=$ 3×10^5 km/s) as if your spaceship were motionless. This statement is in direct conflict with the Newtonian view that a stationary person and a moving person should measure different speeds (Figure 22-1).

Speed involves both distance and time. Since speed has a very different behavior in the special theory of relativity than in Newton's physics, it follows that both space and time behave differently as well. Indeed, in relativity time proves to be so intimately intertwined with the three dimensions of space that we regard them as a single *four*-dimensional entity called **spacetime.**

Length Contraction

Einstein expressed his ideas about spacetime in a mathematical form and used this to make a number of predictions about nature. All these predictions have been verified in innumerable experiments. One prediction is that the length you measure an object to have depends on how that object is moving; the faster it moves, the shorter its length along its direction of motion (Figure 22-2). This is called **length contraction.** In other words, if a railroad car moves past you at high speed, from your perspective on the ground you will actually measure it to be shorter than if it were at rest (Figure 22-2*a*). However, if you are on board the railroad car and moving with it, you will measure its length to be the same as measured on the ground when it was at rest. The word "relativity" emphasizes the importance of the relative speed between the observer and the object being measured.

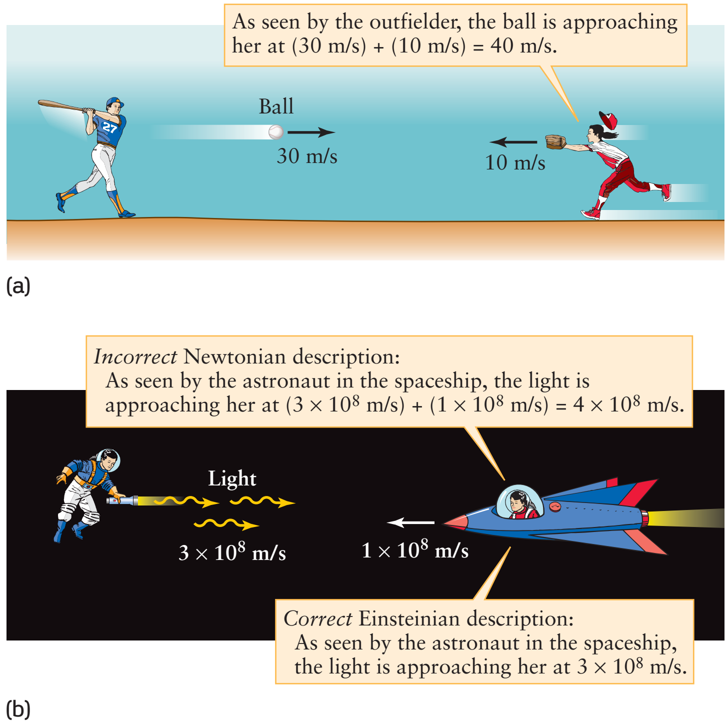


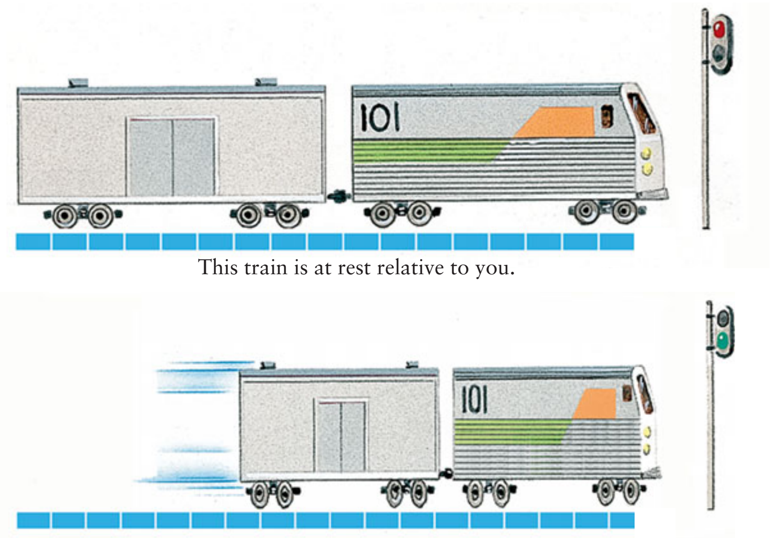
Figure 22-1

The Speed of Light Is the Same to All Observers (a) The speed you measure for ordinary objects depends on how you are moving. Thus, the batter sees the ball moving at 30 m/s, but the outfielder sees it moving at 40 m/s relative to her. **(b)** Einstein showed that this commonsense principle does not apply to light. No matter how fast or in what direction the astronaut in the spaceship is moving, she and the astronaut holding the flashlight will always measure light to have the same speed.

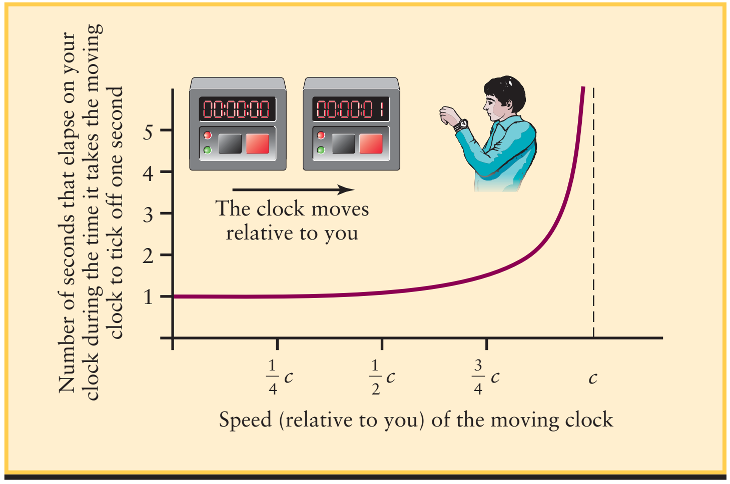
If the idea of length contraction seems outrageous, it is because this effect is noticeable only at very high speeds, comparable to that of light. But even the fastest spacecraft ever built by humans travels at a mere 1/25,000 of the speed of light. At this speed, a spacecraft 10 m long would be contracted in length by only 8 nm—a distance equal to the width of a single protein molecule! For moving cars, trains, and airplanes, length contraction is far too small to measure. As Box 22-1 describes, however, we can easily see the effects of length contraction for subatomic particles that do indeed travel at nearly the speed of light.

Time Dilation

A second result of relativity is no less strange: A clock moving past you runs more slowly than a clock that is at rest. Like length contraction, this **time dilation** is a very small effect unless the clock is moving at extremely high speeds (Figure 22-2*b*). Nonetheless,



(a) Length contraction



(b) Time dilation

Figure 22-2

Length Contraction and Time Dilation **(a)** The faster an object moves, the shorter it becomes along its direction of motion. Speed has no effect on the object's dimensions perpendicular to the direction of motion. **(b)** The faster a clock moves, the slower it runs. This graph shows how many seconds (as measured on your clock) it takes a clock that moves relative to you to tick off one second. The effect is pronounced only for speeds near the speed of light c.

physicists have confirmed the existence of time dilation by using an extremely accurate atomic clock carried on a jet airliner. When the airliner landed, they found that the on-board clock had actually ticked off slightly less time than an identical, stationary clock on the ground. For the passengers on board the airliner during this experiment, however, time flowed at a normal rate: The on-board clock ticked off the seconds as usual, their hearts beat at a normal rate, and so on. You only measure a clock (or a beating heart) to be running slow if it is *moving* relative to you. Box 22-1 discusses time dilation in more detail.

The special theory of relativity predicts that the astronaut with the flashlight in Figure 22-1*b* sees the astronaut flying in the spaceship as shortened along the direction of motion and as having slowly ticking clocks. Remarkably, it also says that the *flying* astronaut sees the astronaut with the flashlight (who is moving relative to her) as being shortened and as having slowly ticking clocks! This may seem contradictory, but it is not: Each spaceship is moving relative to the other, and in the special theory of relativity only relative motion matters. Each astronaut's measurements are correct relative to his or her frame of reference.

CAUTION! You may have the idea that the special theory of relativity implies that there are *no* absolutes in nature and that everything is relative. But, in fact, the special theory is based on the principles that the laws of physics and the speed of light in a vacuum *are* absolutes. Only certain quantities, such as distance and time, depend on your state of motion. You may also have the idea that length contraction and time dilation are just optical illusions caused by high speeds. But these effects are *real*. A moving spaceship doesn't just look shorter or seem to be shorter—it really *is* shorter. Likewise, a moving clock doesn't merely look like it's ticking slowly, or seem to be ticking slowly—it really *does* tick more slowly. Einstein's theory is supported by every experiment designed to test it. Relativity is very strange, but it is also very real!

Mass and Energy

Einstein's special theory of relativity predicts another famous relationship: Any object with mass (m) has energy (E) embodied in its mass. This is expressed by the well-known formula $E = mc^2$, where c is the speed of light. In thermonuclear reactions, part of the energy embodied in the mass of atomic nuclei is converted into other forms of energy. This released energy is what makes the Sun and the stars shine (see Section 18-1). The sunlight we see by day and the starlight that graces our nights are a resounding verification of the special theory of relativity.

The special theory of relativity also explains why it is impossible for a spaceship to move at the speed of light. If it could, then a light beam traveling in the same direction as the spaceship would appear to the ship's crew to be stationary. But this would contradict the second of Einstein's principles, which says that all observers, including those on a spaceship, must see light traveling at the speed of light. Therefore, it cannot be possible for a spaceship, or indeed any material object, to travel at the speed of light. This, too, has been verified by experiments. Subatomic particles can be made to travel at 99.99999999% of the speed of light c, but they can never make it all the way to c.

BOX 22-1 Tools of the Astronomer's Trade

Time Dilation and Length Contraction

The special theory of relativity describes how motion affects measurements of time and distance. By using the two basic principles of his theory—that no matter how fast you move, the laws of physics and the speed of light are the same—Einstein concluded that these measurements must depend on how the person making the measurements is moving.

To describe how measurements depend on motion, Einstein derived a series of equations named the **Lorentz transformations,** after the famous Dutch physicist Hendrik Antoon Lorentz. (Lorentz was a contemporary of Einstein who developed these equations independently but did not grasp their true meaning.) These equations tell us exactly how moving clocks slow down and how moving objects shrink.

To appreciate the Lorentz transformations, imagine two observers named Sergio and Majeeda. Sergio is at rest on Earth, while Majeeda is flying past in her spaceship at a speed v. Sergio and Majeeda both observe the same phenomenon on Earth—say, the beating of Sergio's heart or the ticking of Sergio's watch—which appears to occur over an interval of time. According to Sergio's clock (which is not moving relative to the phenomenon), the phenomenon lasts for T_0 seconds. This is called the **proper time** of the phenomenon. But according to Majeeda's clock (which *is* moving relative to the phenomenon), the same phenomenon lasts for a different length of time, T seconds. These two time intervals are related as follows:

Lorentz transformation for time

$$T = \frac{T_0}{\sqrt{1 - (v/c)^2}}$$

T = time interval measured by an observer moving relative to the phenomenon

T_0 = time interval measured by a observer *not* moving relative to the phenomenon (proper time)

v = speed of the moving observer relative to the phenomenon

c = speed of light

EXAMPLE: Sergio heats a cup of water in a microwave oven for 1 minute. According to Majeeda, who is flying past Sergio at 98% of the speed of light, how long does it take to heat the water?

Situation: The phenomenon in question is heating the water in the microwave oven. Sergio is not moving relative to this phenomenon (so he measures the *proper* time interval T_0). Majeeda is moving at speed $v = 0.98c$ relative to this phenomenon (so she measures a different time interval T). Our goal is to determine the value of T.

Tools: We use the Lorentz transformation for time to calculate T.

Answer: We have $v/c = 0.98$, so

$$T = \frac{T_0}{\sqrt{1 - (v/c)^2}} = \frac{1 \text{ minute}}{\sqrt{1 - (0.98)^2}} = 5 \text{ minutes}$$

Review: A phenomenon that lasts for $T_0 = 1$ minute on Sergio's clock is stretched out to $T = 5T_0 = 5$ minutes as measured on Majeeda's clock moving at 98% of the speed of light. Other phenomena are affected in the same way: As measured by *Majeeda,* it takes 5 seconds for Sergio's wristwatch to tick off one second, and the minute hand on Sergio's wristwatch takes 5 hours to make a complete sweep. The converse is also true. As measured by *Sergio,* the minute hand on Majeeda's wristwatch will take 5 hours to make a complete sweep. This phenomenon, in which events moving relative to an observer happen at a slower pace, is called *time dilation.*

The Lorentz transformation for time is plotted in Figure 22-2*b,* which shows how 1 second measured on a stationary clock (say, Sergio's) is stretched out when measured using a clock carried by a moving observer (such as Majeeda). For speeds less than about half the speed of light, the mathematical factor $\sqrt{1 - (v/c)^2}$ is not too different from 1, and there is little difference between the recordings of the stationary and moving clocks. At the fastest speeds that humans have ever traveled (en route from the Earth to the Moon), the difference between 1 and the factor $\sqrt{1 - (v/c)^2}$ is less than 10^{-9}. So, for any speed associated with human activities, stationary and slowly moving clocks tick at essentially the same rate. As the next example shows, however, time dilation is important for subatomic particles that travel at speeds comparable to c.

EXAMPLE: When unstable particles called *muons* (pronounced "mew-ons") are produced in experiments on Earth, they decay into other particles in an average time of 2.2×10^{-6} s. Muons are also produced by fast-moving protons from interstellar space when they collide with atoms in the Earth's upper atmosphere. These muons typically move at 99.9% of the speed of light and are formed at an altitude of 10 km. How long do such muons last before they decay?

Situation: The phenomenon in question is the life of a muon, which lasts a time $T_0 = 2.2 \times 10^{-6}$ s as measured by an observer not moving with respect to the muon. Our goal is to find the time interval T measured by an observer on Earth, which is moving at $v = 0.999c$ relative to the muon (the same speed at which the muon is moving relative to the Earth).

Tools: As in the previous example, we use the Lorentz transformation for time.

Answer: Using $v/c = 0.999$,

$$T = \frac{2.2 \times 10^{-6}\ \text{s}}{\sqrt{1 - (0.999)^2}} = 4.9 \times 10^{-5}\ \text{s}$$

Review: At this speed, the muon's lifetime is slowed down by time dilation by a factor of more than 22. Note that as measured by an observer on Earth, the time that it takes a muon produced at an altitude of 10 km = 10,000 m to reach the Earth's surface is

$$\frac{\text{distance}}{\text{speed}} = \frac{10{,}000\ \text{m}}{0.999 \times 3.00 \times 10^8\ \text{m/s}} = 3.3 \times 10^{-5}\ \text{s}$$

Were there no time dilation, such a muon would decay in just 2.2×10^{-6} s and would never reach the Earth's surface. But in fact, these muons *are* detected by experiments on the Earth's surface, because a muon moving at $0.999c$ lasts more than 3.3×10^{-5} s. The detection at the Earth's surface of muons from the upper atmosphere is compelling evidence for the reality of time dilation.

In the same terminology as "proper time," we say that a ruler at rest measures **proper length** or **proper distance** (L_0). According to the Lorentz transformations, distances perpendicular to the direction of motion are unaffected. However, a ruler of proper length L_0 held parallel to the direction of motion shrinks to a length L, given by

Lorentz transformation for length

$$L = L_0\sqrt{1 - (v/c)^2}$$

L = length of a moving object along the direction of motion

L_0 = length of the same object at rest (proper length)

v = speed of the moving object

c = speed of light

EXAMPLE: Again, imagine that Majeeda is traveling at 98% of the speed of light relative to Sergio. If Majeeda holds a 1-meter ruler parallel to the direction of motion, how long is this ruler as measured by Sergio?

Situation: The ruler is at rest relative to Majeeda, so she measures its *proper* length $L_0 = 1$ m. Our goal is to determine its length L as measured by Sergio, who is moving at $v = 0.98c$ relative to Majeeda and her ruler.

Tools: We use the Lorentz transformation for length.

Answer: With $v/c = 0.98$, we find

$$L = L_0\sqrt{1 - (v/c)^2} = (1\ \text{m})\sqrt{1 - (0.98)^2}$$
$$= (1\ \text{m}) \times (0.20) = 0.20\ \text{m} = 20\ \text{cm}$$

Review: We saw in the first example that according to *Sergio,* Majeeda's clocks are ticking only one-fifth as fast as his. This example shows that he also measures Majeeda's 1-meter ruler to be only one-fifth as long (20 cm) when held parallel to the direction of motion. Note that the converse is also true: If Sergio holds a 1-m ruler parallel to the direction of relative motion, *Majeeda* measures it to be only 20 cm long. This shrinkage of length is called *length contraction.*

EXAMPLE: Consider again the above example about muons created 10 km above the Earth's surface. If a muon is traveling straight down, what is the distance to the surface as measured by an observer riding along with the muon?

Situation: Imagine a ruler that extends straight up from the Earth's surface to where the muon is formed. This ruler is at rest relative to the Earth, so its length of 10 km is the proper length L_0. Our goal is to find the length L of this ruler as measured by the observer riding with the muon.

Tools: As in the previous example, we use the Lorentz transformation for length.

Answer: With $v/c = 0.999$, we calculate

$$L = (1\ \text{km})\sqrt{1 - (0.999)^2} = 0.45\ \text{km} = 450\ \text{m}$$

Review: The distance is contracted tremendously as measured by an observer riding with the muon. This result gives us another way to explain how such muons are able to reach the Earth's surface. As measured by the muon, the time required to travel the contracted distance is

$$\frac{\text{distance}}{\text{speed}} = \frac{450\ \text{m}}{0.999 \times 3.00 \times 10^8\ \text{m/s}} = 1.5 \times 10^{-6}\ \text{s}$$

This is less time than the 2.2×10^{-6} s that an average muon takes to decay. Hence, muons can successfully reach the Earth's surface.

22-2 The general theory of relativity predicts black holes

Einstein's special theory of relativity is a comprehensive description of the behavior of light and, by extension, of electricity and magnetism. (Recall from Section 5-2 that light is both electric and magnetic in nature.) Einstein's next goal was to develop an even more comprehensive theory that also explained gravity. This was the **general theory of relativity**, which he published in 1915.

The Equivalence Principle

According to Newton's theory of gravity, an apple falls to the floor because the force of gravity pulls the apple down. But Einstein pointed out that the apple would appear to behave in exactly the same way in space, far from the gravitational influence of any planet or star, if the floor were to accelerate upward (in other words, if the floor came up to meet the apple).

In Figure 22-3, two famous gentlemen are watching an apple fall toward the floor of their closed compartments. They have no way of telling who is at rest on the Earth and who is in the hypothetical elevator moving upward through empty space at a constantly increasing speed. This is an example of Einstein's **equivalence principle,** which states that in a small volume of space, the downward pull of gravity can be accurately and completely duplicated by an upward acceleration of the observer.

The equivalence principle is the key to the general theory of relativity. It allowed Einstein to focus entirely on motion, rather than force, in discussing gravity. A hallmark of gravity is that it causes the same acceleration no matter the mass of the object. For example, a baseball and a cannon ball have very different masses, but if you drop them side by side in a vacuum, they accelerate downward at exactly the same rate. To explain this, Einstein envisioned gravity as being caused by a *curvature of space*. In fact, his general theory of relativity describes gravity entirely in terms of the geometry of both space *and* time, that is, of spacetime. Far from a source of gravity, like a planet or a star, spacetime is "flat" and clocks tick at their normal rate. Closer to a source of gravity, however, space is curved and clocks slow down.

ANALOGY A useful analogy is to picture the spacetime near a massive object such as the Sun as being curved like the surface in Figure 22-4. Imagine a ball rolling along this surface. Far from the "well" that represents the Sun, the surface is fairly flat and the ball moves in a straight line. If the ball passes near the well, however, it curves in toward it. If the ball is moving at an appropriate speed, it might move in an orbit around the sides of the well. The curvature of the well has the same effect on a ball of any size, which explains why gravity produces the same acceleration on objects of different mass.

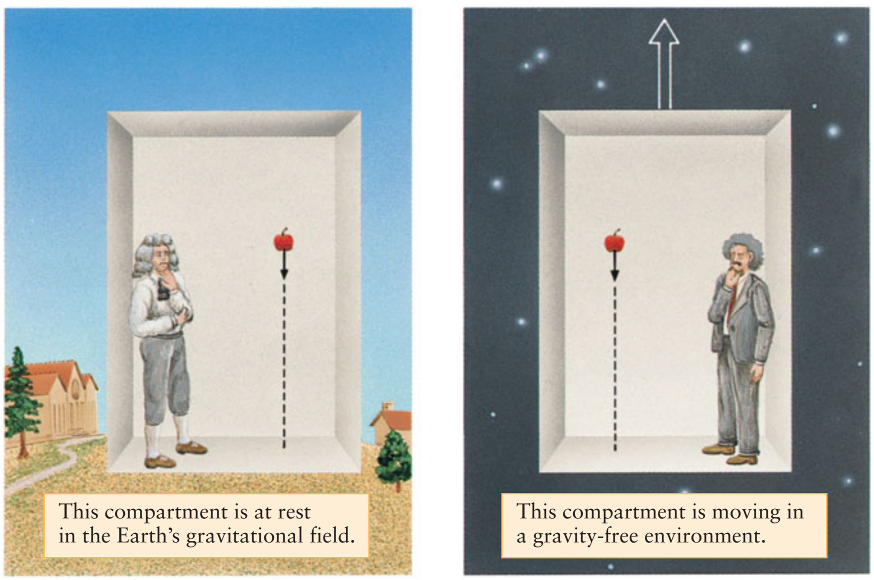
Testing the General Theory

WEB LINK 22.3

Einstein's general theory of relativity and its picture of curved spacetime have been tested in a variety of different ways. In what follows we discuss some key experimental tests of the theory.

Experimental Test 1: The gravitational bending of light. Light rays naturally travel in straight lines. But if the space through which the rays travel is curved, as happens when light passes near the surface of a massive object like the Sun, the paths of the rays will likewise be curved (Figure 22-5). In other words, gravity should bend light rays, an effect not predicted by Newtonian mechanics because light has no mass. This prediction was first tested in 1919 during a total solar eclipse. During totality, when the Moon blocked out

Careful experiments have verified the key ideas of Einstein's theory of gravity



(a) The apple hits the floor of the compartment because the Earth's gravity accelerates the apple downward.

(b) The apple hits the floor of the compartment because the compartment accelerates upward.

ANIMATION 22.1

Figure 22-3

The Equivalence Principle The equivalence principle asserts that you cannot distinguish between **(a)** being at rest in a gravitational field and **(b)** being accelerated upward in a gravity-free environment. This idea was an important step in Einstein's quest to develop the general theory of relativity.

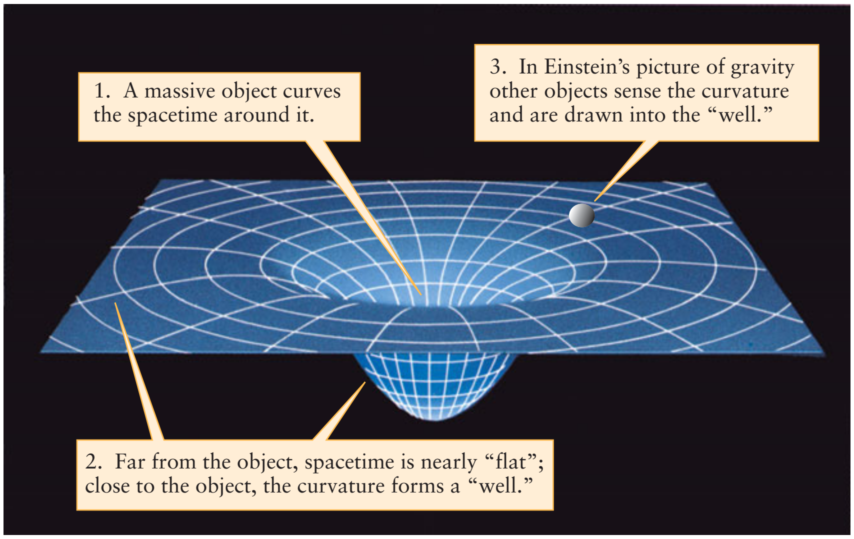


Figure 22-4

The Gravitational Curvature of Spacetime According to Einstein's general theory of relativity, spacetime becomes curved near a massive object. To help you visualize the curvature of four-dimensional spacetime, this figure shows the curvature of a *two*-dimensional space around a massive object.

the Sun's disk, astronomers photographed the stars around the Sun. Careful measurements afterward revealed that the stars were shifted from their usual positions by an amount consistent with Einstein's theory.

Experimental Test 2: The precession of Mercury's orbit. As the planet Mercury moves along its elliptical orbit, the orbit itself slowly changes orientation or *precesses* (Figure 22-6). Most of Mercury's precession is caused by the gravitational pull of the other planets, as explained by Newtonian mechanics. But once the effects of all the other planets had been accounted for, there remained an unexplained excess rotation of Mercury's major axis of 43 arcsec per century. Although this discrepancy may seem very small, it frustrated astronomers for half a century. Some astronomers searched for a missing planet even closer to the Sun that might be tugging on Mercury; none has ever been found. Einstein showed that at Mercury's position close to the Sun, the general theory of relativity predicts a small correction to Newton's description of gravity. This correction is just enough to account for the excess precession.

Experimental Test 3: The gravitational slowing of time and the gravitational red shift (Figure 22-7). In the general theory of relativity, a massive object such as the Earth warps time as well as space. Einstein predicted that clocks on the ground floor of a building should tick slightly more slowly than clocks on the top floor, which are farther from the Earth (Figure 22-7*a*).

A light wave can be thought of as a clock; just as a clock makes a steady number of ticks per minute, an observer sees a steady number of complete cycles of a light wave passing by each

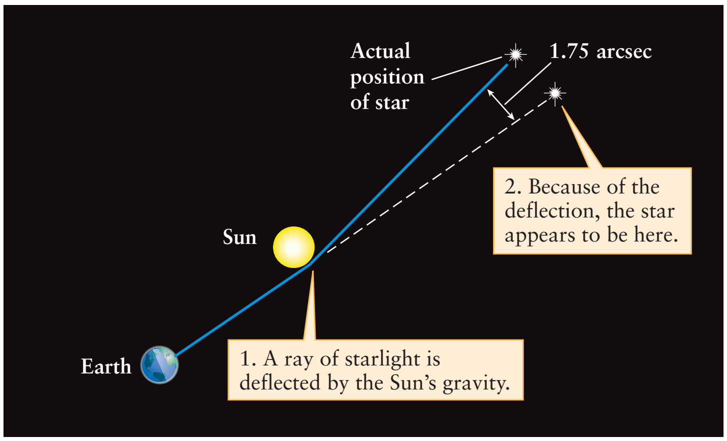


Figure 22-5

The Gravitational Deflection of Light Light rays are deflected by the curved spacetime around a massive object like the Sun. The maximum deflection is very small, only 1.75 arcsec for a light ray grazing the Sun's surface. By contrast, Newton's theory of gravity predicts *no* deflection at all. The deflection of starlight by the Sun was confirmed during a solar eclipse in 1919.

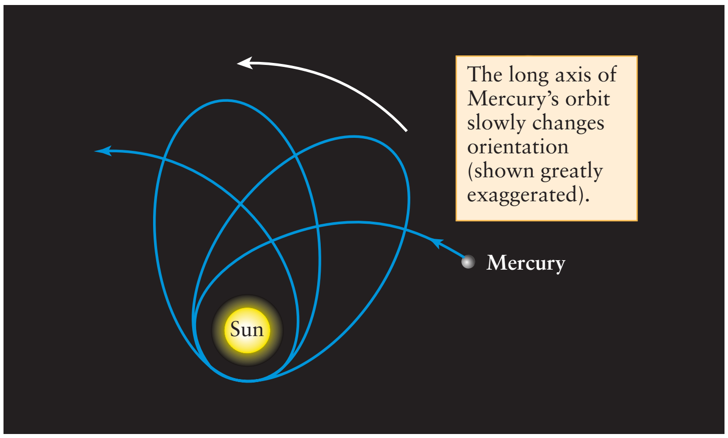


Figure 22-6

The Precession of Mercury's Orbit Mercury's orbit changes orientation at a rate of 574 arcsec (about one-sixth of a degree) per century. Newtonian mechanics predicts that the gravitational influences of the other planets should make the orientation change by only 531 arcsec per century. Einstein was able to explain the excess 43 arcsec per century using his general theory of relativity.

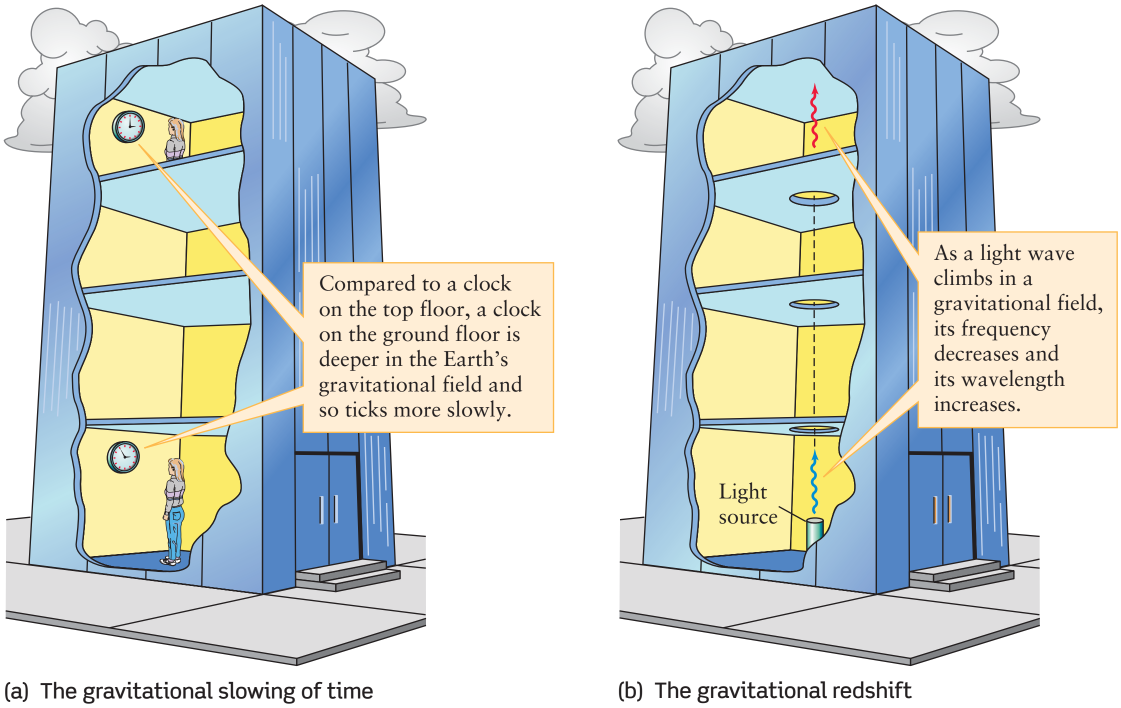


Figure 22-7

The Gravitational Slowing of Time and the Gravitational Redshift
(a) Clocks at different heights in a gravitational field tick at different rates. **(b)** The oscillations of a light wave constitute a type of clock. As a light wave climbs from the ground floor toward the top floor, its oscillation frequency becomes lower and its wavelength becomes longer.

second. If a light beam is aimed straight up from the ground floor of a building, an observer at the top floor will measure a "slow-ticking" light wave with a lower frequency, and thus a longer wavelength, than will an observer on the ground floor (Figure 22-7*b*). The increase in wavelength means that a photon reaching the top floor has less energy than when it left the ground floor. (You may want to review the discussion of frequency and wavelength in Section 5-2 as well as the description of photons in Section 5-5.) These effects, which have no counterpart in Newton's theory of gravity, are called the **gravitational redshift.**

CAUTION! Be careful not to confuse the gravitational redshift with a Doppler shift. In the Doppler effect, redshifts are caused by a light source moving away from an observer. Gravitational redshifts, by contrast, are caused by time flowing at different rates at different locations. No motion is involved.

The American physicists Robert Pound and Glen Rebka first measured the gravitational redshift in 1960 using gamma rays fired between the top and bottom of a shaft 20 meters tall. Because the Earth's gravity is relatively weak, the redshift that they measured was very small ($\Delta\lambda/\lambda = 2.5 \times 10^{-15}$) but was in complete agreement with Einstein's prediction.

Much larger shifts are seen in the spectra of white dwarfs, whose spectral lines are redshifted as light climbs out of the white dwarf's intense surface gravity. As an example, the gravitational redshift of the spectral lines of the white dwarf Sirius B (see Figure 20-8) is $\Delta\lambda/\lambda = 3.0 \times 10^{-4}$, which also agrees with the general theory of relativity.

LOOKING DEEPER 22.1

Experimental Test 4: Gravitational waves. Electric charges oscillating up and down in a radio transmitter's antenna produce electromagnetic radiation. In a similar way, the general theory of relativity predicts that oscillating massive objects should produce **gravitational radiation,** or **gravitational waves.** (Newton's theory of gravity makes no such prediction.)

WEB LINK 22.4

Gravitational radiation is exceedingly difficult to detect, because it is by nature much weaker than electromagnetic radiation. Although physicists have built a number of sensitive "antennas" for gravitational radiation, no confirmed detections have been made as of this writing (2007). But compelling *indirect* evidence for the existence of gravitational radiation has come from a binary system of two neutron stars. Russell Hulse and Joseph Taylor at the University of Massachusetts discovered that these stars are slowly spiraling toward each other and losing energy in the process (see Section 21-7). The rate at which they lose energy is just what would be expected if the two neutron stars are emitting gravitational radiation as predicted by Einstein. Hulse and Taylor shared the 1993 Nobel Prize in physics for their discovery.

The general theory of relativity has never made an incorrect prediction. It now stands as our most accurate and complete

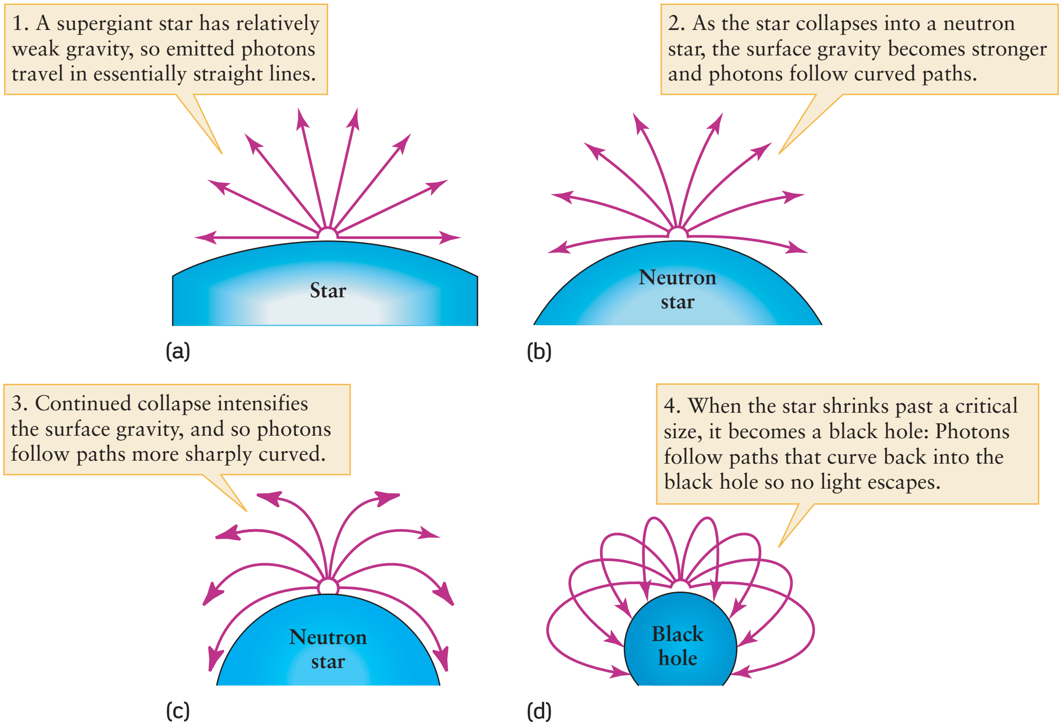


Figure 22-8

The Formation of a Black Hole
(a)–(c) These illustrations show four steps leading up to the formation of a black hole from a dying star. **(d)** When the star becomes a black hole, not even photons emitted directly upward from the surface can escape; they undergo an infinite gravitational redshift and disappear.

description of gravity. Einstein demonstrated that Newtonian mechanics is accurate only when applied to low speeds and weak gravity. If extremely high speeds or powerful gravity are involved, only a calculation using relativity will give correct answers.

Relativity and Black Holes

Perhaps the most dramatic prediction of the general theory of relativity concerns what happens when a large amount of matter is concentrated in a small volume. We have seen that if a dying star is not too massive, it ends up as a white dwarf star. If the dying star is more massive than the Chandrasekhar limit of about 1.4 $M_\odot$, it cannot exist as a stable white dwarf star and, instead, shrinks down to form a neutron star. But if the dying star has more mass than the maximum permissible for a neutron star, about 2 to 3 $M_\odot$, not even the internal pressure of neutrons can hold the star up against its own gravity, and the star contracts rapidly.

As the star's matter becomes compressed to enormous densities, the strength of gravity at the surface of this rapidly shrinking sphere also increases dramatically. According to the general theory of relativity, the space immediately surrounding the star becomes so highly curved that it closes on itself (Figure 22-8). Photons flying outward at an angle from the star's surface arc back inward, while photons that fly straight outward undergo such a strong gravitational redshift that they lose all their energy and cease to exist.

Ordinary matter can never travel as fast as light. Hence, if light cannot escape from the collapsing star, neither can anything else. An object from which neither matter nor electromagnetic radiation can escape is called a **black hole.** In a sense, a hole is punched in the fabric of the universe, and the dying star disappears into this cavity (Figure 22-9).

None of the star's mass is lost when it collapses to form a black hole, however. This mass gives the spacetime around the black hole its strong curvature. Thanks to this curvature, the black hole's gravitational influence can still be felt by other objects.

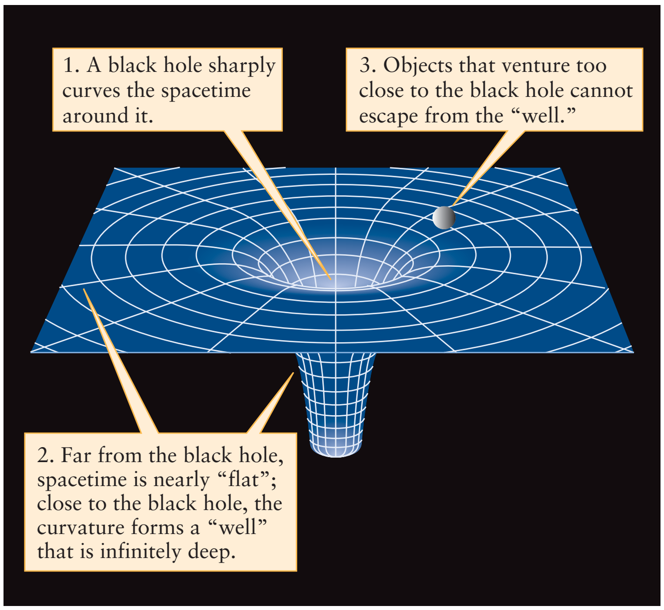


Figure 22-9

Curved Spacetime around a Black Hole This diagram suggests how spacetime is distorted by a black hole's mass. As in Figure 22-4, spacetime is represented as a two-dimensional surface. Unlike the situation in Figure 22-4, a black hole's gravitational "well" is infinitely deep.

CAUTION! Some low-quality science-fiction movies and books suggest that black holes are evil things that go around gobbling up everything in the universe. Not so! The bizarre effects created by highly warped spacetime are limited to a region quite near the hole. For example, the effects of the general theory of relativity predominate only within 1000 km of a 10-$M_\odot$ black hole. Beyond 1000 km, gravity is weak enough that Newtonian physics can adequately describe everything. If our own Sun somehow turned into a black hole (an event that, happily, seems to be quite impossible), the orbits of the planets would hardly be affected at all.

Perhaps the most remarkable aspect of black holes is that they really exist! As we will see, astronomers have located a number of black holes with masses a few times that of the Sun. What is truly amazing is that they have also discovered many truly immense black holes containing millions or billions of solar masses. These discoveries are a resounding confirmation of the ideas of the general theory of relativity. In the next three sections we will see how astronomers hunt down black holes in space.

22-3 Certain binary star systems probably contain black holes

Finding black holes is a difficult business. Because light cannot escape from inside the black hole, you cannot observe one directly in the same way that you can observe a star or a planet. The best you can hope for is to detect the effects of a black hole's powerful gravity.

Close binary star systems offer one way to find **stellar-mass black holes** (that is, ones with masses comparable to those of ordinary stars). Imagine that one of the stars in a binary system evolves into a black hole. If the black hole orbits close enough to the other, ordinary star in the system, tidal forces can draw matter from the ordinary star onto the black hole. If we can detect radiation coming from this "stolen" matter, we can infer the presence of the black hole.

One key to detecting black holes is to search for the radiation emitted by material as it falls into a hole

The Strange Case of Cygnus X-1

The first sign of such emissions from a binary system with a black hole came shortly after the launch of the *Uhuru* X-ray–detecting satellite in 1971. Astronomers became intrigued with an X-ray source designated Cygnus X-1. This source is quite unlike pulsating X-ray sources, which emit regular bursts of X rays every few seconds (see Section 21-8). Instead, the X-ray emissions from Cygnus X-1 are highly variable and irregular; they flicker on time scales as short as one-hundredth of a second. One of the fundamental concepts in physics is that nothing can travel faster than the speed of light (recall Section 22-1). Because of this limitation, an object cannot flicker faster than the time required for light to travel across the object. Because light travels 3000 km in a hundredth of a second, Cygnus X-1 can be no more than 3000 km across, or about a quarter the size of the Earth.

Cygnus X-1 occasionally emits bursts of radio radiation, and in 1971 radio astronomers used these bursts to show that the source was at the same location in the sky as the star HDE 226868 (Figure 22-10). Spectroscopic observations revealed that HDE 226868 is a B0 supergiant with a surface temperature of about 31,000 K. Such stars do not emit significant amounts of X rays, so HDE 226868 alone cannot be the X-ray source Cygnus X-1. Because binary stars are very common, astronomers began to suspect that the visible star and the X-ray source are in orbit about each other.

Further spectroscopic observations soon showed that the lines in the spectrum of HDE 226868 shift back and forth with a period of 5.6 days. This behavior is characteristic of a single-line spectroscopic binary (see Section 17-10); the companion of HDE 226868 is too dim to produce its own set of spectral lines. The clear implication is that HDE 226868 and Cygnus X-1 are the two components of a binary star system.

From what we know about the masses of other supergiant stars, HDE 226868 is estimated to have a mass of roughly 30 $M_\odot$. As a result, the unseen member of the binary system must have a mass of about 7 $M_\odot$ or more. Otherwise, it would not exert enough gravitational pull to make the B0 star wobble by the amount deduced from the periodic Doppler shifting of its spectral lines. Because the unseen companion does not emit visible light, it cannot be an ordinary star. Furthermore, because 7 solar masses is too large for either a white dwarf or a neutron star, Cygnus X-1 is likely to be a black hole.

The case for Cygnus X-1 being a black hole is not airtight. HDE 226868 might have a low mass for its spectral type, which would imply a somewhat lower mass for Cygnus X-1. In addition, uncertainties in the distance to the binary system could

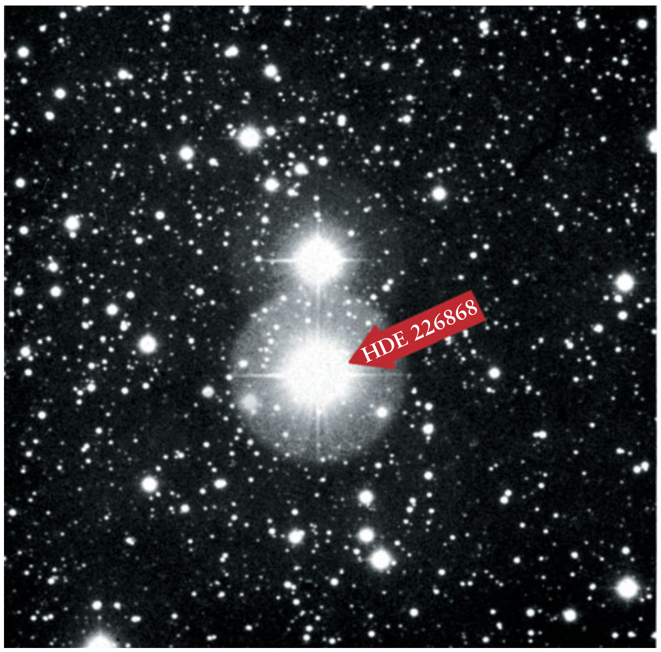
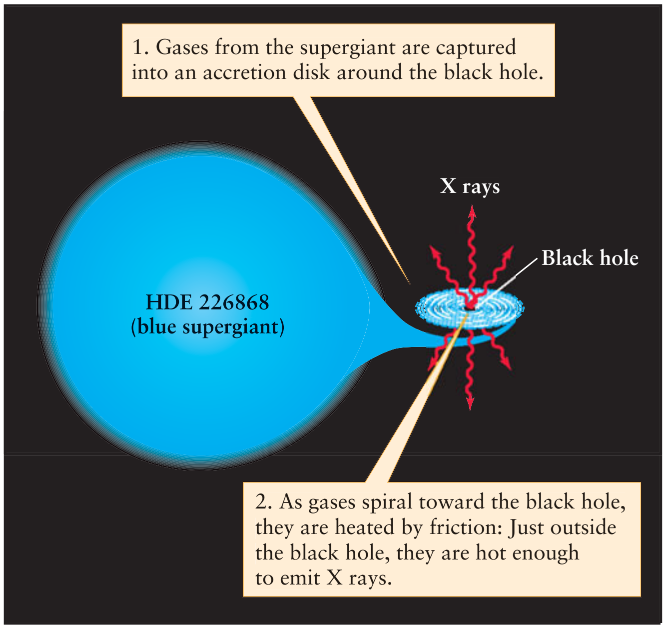


Figure 22-10 R I V U X G

HDE 226868 This photograph from the 5-m telescope on Palomar Mountain shows HDE 226868, the B0 blue supergiant star at the location of the X-ray source Cygnus X-1. This star is located about 2500 pc (87,000 ly) from Earth. The bright star directly above HDE 226868 is not part of the binary system. (Courtesy of J. Kristian, Carnegie Observatories)



(a) A schematic diagram of Cygnus X-1

(b) An artist's impression of Cygnus X-1

WEB LINK 22.5

Figure 22-11

The Cygnus X-1 System **(a)** The larger member of the Cygnus X-1 system is a B0 supergiant of about 30 $M_\odot$. The other, unseen member of the system has a mass of at least 7 $M_\odot$ and is probably a black hole. **(b)** This illustration shows how the Cygnus X-1 system might look at close range. (The illustration that opens this chapter depicts a similar system.) At even closer range, the black hole and its immediate surroundings might appear as shown in Figure 22-12. (Courtesy of D. Norton, Science Graphics)

further reduce estimates of the mass of Cygnus X-1. If all these uncertainties combined in just the right way, the estimated mass of Cygnus X-1 could be pushed down to about 3 $M_\odot$. Thus, there is a slim chance that Cygnus X-1 might contain the most massive possible neutron star rather than a black hole.

If Cygnus X-1 does contain a black hole, the X rays do not come from the black hole itself. Gas captured from HDE 226868 goes into orbit about the hole, forming an accretion disk about 4 million kilometers in diameter (Figure 22-11). As material in the disk gradually spirals in toward the hole, friction heats the gas to temperatures approaching 2×10^6 K. In the final 200 kilometers above the hole, these extremely hot gases emit the X rays that our satellites detect. Presumably the X-ray flickering is caused by small hot spots on the rapidly rotating inner edge of the accretion disk. In this way, the black hole's existence is announced by doomed gases just before they plunge to oblivion.

Other Black Hole Candidates

Astronomers have found more than 20 other black hole candidates like Cygnus X-1. All are compact X-ray sources orbiting ordinary stars in a spectroscopic binary system. One of the best candidates is V404 Cygni in the constellation Cygnus. Doppler shift measurements reveal that as the visible star moves around its 6.47-day orbit, its radial velocity (see Section 17-10) varies by more than 400 km/s. These data give a firm lower limit of 6.26 $M_\odot$ for the mass of the unseen companion. It is probably impossible for a neutron star to be more massive than 3 $M_\odot$, so V404 Cygni must almost certainly be a black hole.

Another particularly convincing case is the flickering X-ray source A0620-00 in the constellation Monoceros. (The A refers to the British satellite *Ariel 5,* which discovered the source; the numbers denote its position in the sky.) The visible companion to A0620-00 is an orange K5 main-sequence star called V616 Monocerotis, which orbits the X-ray source every 7.75 hours. Because this star is relatively faint, it is possible to observe the shifting spectral lines from *both* the visible companion and the X-ray source as they orbit each other. With this more complete information about the orbits, astronomers have determined the mass of A0620-00 to be more than 3.2 $M_\odot$, and more probably about 9 $M_\odot$.

WEB LINK 22.6

Astronomers have seen jets of hot, glowing material extending several light-years from some black hole candidates. The ejected material emerges from the vicinity of the black hole at speeds approaching the speed of light. It is thought that these jets are formed by strong electric and magnetic fields in the material around a *rotating* black hole, much as occurs for rotating protostars (see Figure 18-16) and for rotating neutron stars (see Figure 21-7). Figure 22-12 is an artist's impression of the immediate vicinity of such a rotating black hole. (We will discuss the curious features of rotating black holes in Section 22-7.)

If there really are black holes in close binary systems, how did they get there? One possibility is that one member of the binary explodes as a Type II supernova, leaving a burned-out core whose mass exceeds 3 $M_\odot$. This core undergoes gravitational collapse to form a black hole.

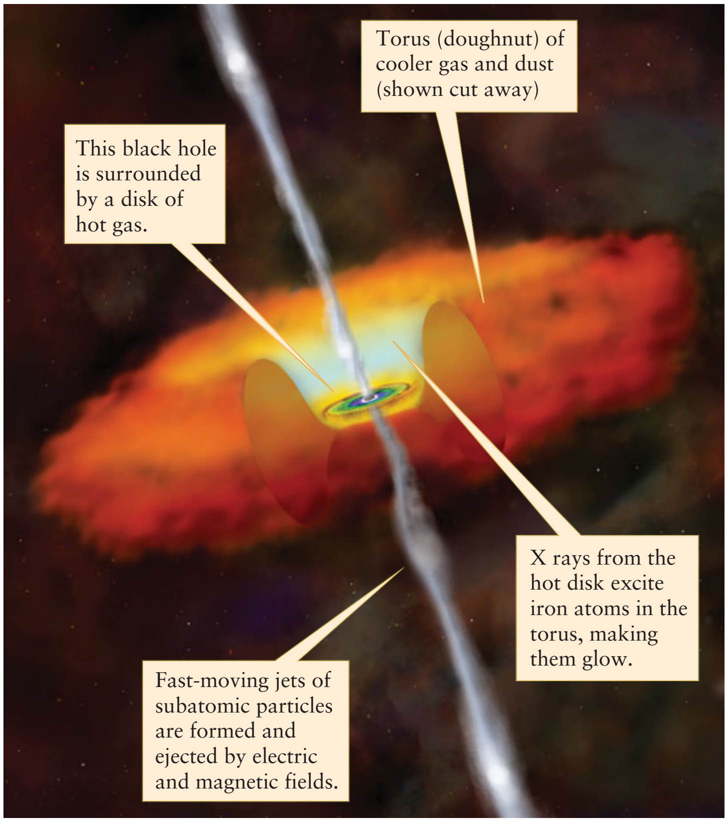


Figure 22-12

The Environment of an Accreting Black Hole If a black hole is rotating, it can generate strong electric and magnetic fields in its immediate vicinity. These fields draw material from the accretion disk around the black hole and accelerate it into oppositely directed jets along the black hole's rotation axis. This illustration also shows other features of the material surrounding such a black hole. (CXC/M. Weiss)

A white dwarf or a neutron star in a binary system could also become a black hole if it accretes enough matter from its companion star. This transformation can occur when the companion star becomes a red giant and dumps a significant part of its mass over its Roche lobe (see Figure 19-21*b*).

Finally, another possibility is two dead stars coalescing to form a black hole. For example, imagine two neutron stars orbiting each other, like the binary systems discussed in Section 21-7 and Section 22-2. Because the two stars emit gravitational radiation, they gradually spiral in toward each other and eventually merge. If their total mass exceeds 2 to 3 $M_\odot$, the entire system may become a black hole.

22-4 The most intense radiation bursts in the universe may be caused by the formation of black holes

WEB LINK 22.7

As we have seen, bursts of X rays from binary systems such as Cygnus X-1 lead us to the remarkable conclusion that black holes are real, and that certain stars can evolve into black holes. Even more remarkable is that astronomers may actually be able to observe the moment that a black hole forms from a dying star. The key to this comes from **gamma-ray bursters,** mysterious objects that emit the most powerful bursts of high-energy radiation ever measured. To understand the connection between these objects and black holes, we begin by looking at how gamma-ray bursters were discovered and how astronomers struggled to determine their true nature.

Gamma-ray bursters are incredibly bright flashes of radiation from remarkably distant galaxies

A Gamma-Ray Mystery

Gamma-ray bursters were discovered in the late 1960s by the orbiting Vela satellites, whose detectors noticed flashes of gamma rays coming from random parts of the sky at random intervals. This discovery was an unexpected consequence of the Cold War. The Vela satellites were originally placed in orbit by the United States to look for high-energy photons coming from above-ground tests of nuclear weapons by the Soviet Union, tests that had been banned by treaty since 1963. (No such tests were ever detected.) More than 3000 of these cosmic flashes have been detected since 1967. With a new generation of gamma-ray telescopes currently in orbit (see Section 6-7), new gamma-ray bursters are being found at a rate of about one per day.

Gamma-ray bursters fall into two types. *Long-duration* gamma-ray bursters, which are more common, last from about 2 to about 1000 seconds before fading to invisibility. The less common *short-duration* gamma-ray bursters last from a few hundredths of a second to about 2 seconds, and tend to emit photons of shorter wavelength and hence higher energy. Unlike X-ray bursters (see Section 21-9), gamma-ray bursters of both types appear to emit only one burst in their entire history.

What are gamma-ray bursters, and how far away are they? These questions plagued astronomers for almost 30 years. One important clue is that gamma-ray bursters are seen with roughly equal probability in all parts of the sky, as Figure 22-13*a* shows. This suggests that they are not in the disk of our Galaxy, because then most gamma-ray bursters would be found in the plane of the Milky Way (Figure 22-13*b*). One idea was that gamma-ray bursters are relatively close to us and lie in a spherical halo surrounding the Milky Way. Alternatively, gamma-ray bursters could be strewn across space like galaxies, with some of them billions of light-years away.

Long-Duration Gamma-Ray Bursters and Supernovae

For many years there was no convincing way to decide between these competing models. This state of affairs changed dramatically in 1997, thanks to the Italian-Dutch BeppoSAX satellite. Unlike previous gamma-ray satellites, which could determine the position of a burst only to within a few degrees, BeppoSAX could pin down its position to within an arcminute. This made it far easier for astronomers using optical telescopes to search for a counterpart to the burster. Just 21 hours after BeppoSAX observed a long-duration gamma-ray burster on February 28, 1997, astronomers led by Jan van Paradis of the University of Amsterdam located a visible-light afterglow of the burst. Since then,

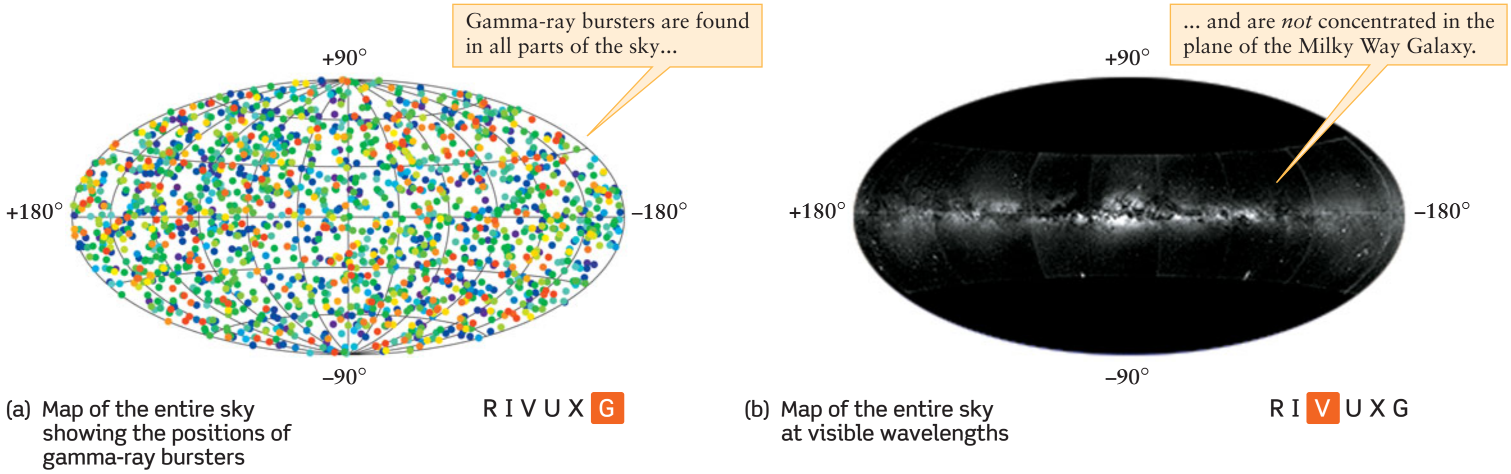


Figure 22-13

Gamma-Ray Bursters (a) This map shows the locations of 1776 gamma-ray bursters detected by the Compton Gamma Ray Observatory (see Section 6-7). The entire celestial sphere is mapped onto an oval. The colors in the order of the rainbow indicate the total amount of energy detected from each burst; bright bursts appear in red and weak bursts appear in violet. **(b)** This map shows the same sky as in part (a) but at visible wavelengths. Comparing part (b) with part (a) shows that unlike X-ray bursters, which originate in the disk of the Milky Way Galaxy, gamma-ray bursters are seen in all parts of the sky.

astronomers using large visible-light and infrared telescopes have detected the afterglow of many gamma-ray bursters, and have found that they occur within galaxies (Figure 22-14). As of this writing (2007), the most distant gamma-ray burster yet seen lies within a galaxy some 13 billion (1.3×10^{10}) light-years away. (In Chapter 24 we will see how astronomers determine the distances to remote galaxies.) To be seen at such immense distances a gamma-ray burster must release a truly stupendous amount of energy in the form of radiation.

WEB LINK 22.8

The relationship between gamma-ray bursters and galaxies strongly suggests that a burster is a cataclysmic event involving one of a galaxy's stars. Observations of a relatively bright (and hence relatively close) long-duration gamma-ray burster have confirmed this idea. On March 29, 2003, the orbiting gamma-ray telescope HETE-2 (short for High Energy Transient Explorer) detected this burster—denoted GRB 030329 for the date it was observed—in the constellation Leo. (GRB 030329 was so intense that its high-energy photons temporarily ionized part of the Earth's atmosphere—the only event of its kind ever caused by an object beyond our own Galaxy.) Within 90 minutes a visible-light afterglow of GRB 030329 was found by astronomers in Australia and Japan. The afterglow was bright enough that astronomers in the United States and Chile were able to measure its detailed spectrum, which turned out to be that of a Type Ic supernova at a distance of some 2.65 billion (2.65×10^9) light-years. We saw in Section 20-9 that a Type Ic supernova results when a massive star undergoes core collapse after having first shed the hydrogen and helium from its outer layers (see Figure 20-20*c*). Subsequent observations with the Hubble Space Telescope revealed an expanding shell of gas at the position of the gamma-ray burster. This shell is just what we would expect from a supernova explosion. Another long-duration gamma-ray burster with an associated Type Ic supernova was seen in 2006.

Beamed Radiation

A long-duration gamma-ray burster cannot simply be an ordinary Type Ic core-collapse supernova, however. Such supernovae release about 10^{46} joules of energy, of which only about 0.03% is

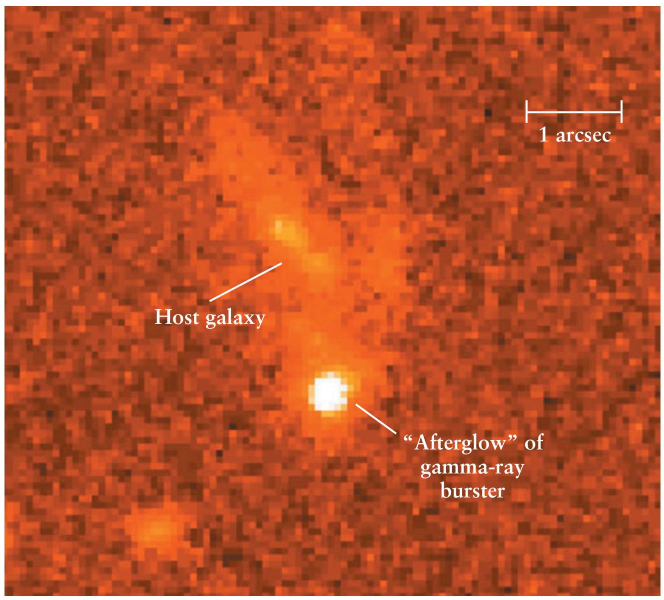


Figure 22-14 R I V U X G

The Host Galaxy of a Gamma-Ray Burster This false-color image was made on February 8, 1999, 16 days after a gamma-ray burster was observed at this location. The host galaxy of the gamma-ray burster has a very blue color (not shown in this image), indicating the presence of many recently formed stars. The gamma-ray burst may have been produced when one of the most massive of these stars became a supernova. (Andrew Fruchter, STScI; and NASA) released as electromagnetic radiation. (Most of the remaining energy goes into neutrinos, while some goes into accelerating the debris that expands away from the supernova.) This radiation streams outward from the supernova in all directions equally, like light from the Sun. If we assume that a gamma-ray burster also emits light equally in all directions, the inverse-square law that relates brightness, luminosity, and distance (see Section 17-2) tells us that the most energetic bursters would have to emit about 3×10^{47} joules of radiation in less than a minute. It would take 100,000 supernovae going off simultaneously to release this much radiation!

This dilemma can be resolved if we imagine a type of supernova that emits most of its radiation in narrow beams. If such a beam happened to be aimed toward Earth, we would detect a far more intense burst of radiation than we would from an ordinary supernova.

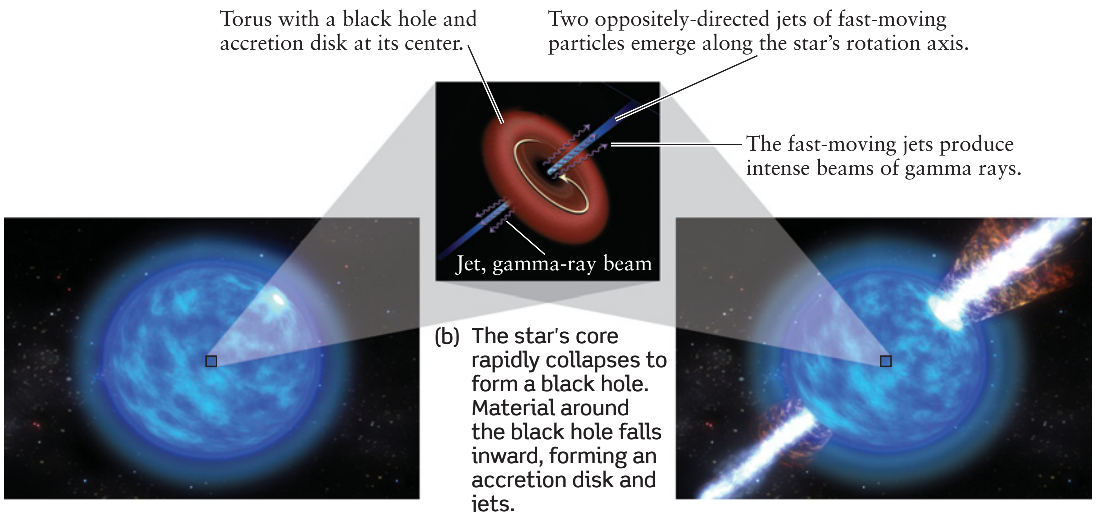
ANALOGY An ordinary flashlight is an example of beamed radiation. The lightbulb in a flashlight is very small and produces only a weak light if you remove it from the flashlight housing. But the flashlight's curved mirror channels the bulb's light into a narrow beam. Thus, a flashlight produces an intense beam of light with only a small input of energy from the batteries.

Collapsars and the Birth of a Black Hole

One theoretical model that would produce beamed radiation invokes a special type of supernova called a **collapsar** (also called a *hypernova*). These objects are thought to result from progenitor stars that are very massive (more than about 30 $M_\odot$), have lost their outer layers of hydrogen and helium, and are rotating rapidly. The core of such a star is too massive to be a white dwarf or a neutron star, so when thermonuclear reactions cease the core will become a black hole. Figure 22-15 depicts what happens when thermonuclear reactions cease in the core of such a star. The black hole forms before the material outside the core has a chance to contract very much (Figure 22-15*a*). As a result, the black hole quickly forms an accretion disk from the surrounding stellar material as shown in Figure 22-15*b*. This process is aided because the progenitor star was rotating rapidly, which makes it easier for the infalling material to form a rotating disk around the black hole. Some of the infalling material does not fall into the black hole, but is ejected in powerful, back-to-back jets along the rotation axis of the accretion disk. (Much the same mechanism acts in close binary star systems in which one member is a black hole, as we described in Section 22-3.)

The jets are so energetic that they reach and break through the surface of the star within 5 to 10 seconds after being formed, as Figure 22-15*c* shows. (It helps that the star has already lost much of its outer layers. This makes it easier for the jets to break through what remains of the star.) As they travel outward, the jets produce powerful shock waves that blow the star to pieces.

If one of the jets from a collapsar happens to be aimed toward Earth, we see an intense burst of gamma rays. These are produced as the relativistic particles in the jet slow down and convert their kinetic energy (energy of motion) into radiation. The burst is short-lived because the jets have only a brief existence: It takes the black hole only about 20 seconds to accrete the entire inner core of the star, after which there is no longer an accretion disk to produce the jets. We do not see a burster if the jets are aimed away from Earth. Hence, for each gamma-ray burster that we observe, there may be hundreds or thousands of collapsars that go undetected.



(a) After shedding its outer layers of hydrogen and helium, a rapidly rotating supergiant star of more than 30 $M_\odot$ reaches the end of its lifetime.

(b) The star's core rapidly collapses to form a black hole. Material around the black hole falls inward, forming an accretion disk and jets.

(c) The jets blast through what remains of the supergiant star. If one of the jets and its beam is directed towards Earth, we see a gamma-ray burster.

ANIMATION 22.2

Figure 22-15

The Collapsar Model of a Long-Duration Gamma-Ray Burster These illustrations show the final few seconds in the life of a massive, rapidly rotating supergiant star that has lost its outer layers of hydrogen and helium. After the jets have subsided, what remains is a Type Ic supernova with a black hole at its center. (a, c: NASA; b: NASA and A. Field, STScI)

The collapsar model cannot explain short-duration gamma-ray bursters, whose lifetimes of less than 2 seconds are far shorter than those of a collapsar's jets. A number of models for short-duration bursters have been proposed, such as the energy released by the merger of two neutron stars or a particularly intense magnetar burst (see Section 21-6).

The newest tool in the search for answers about gamma-ray bursters is a NASA satellite called Swift, which was launched in 2004. In addition to sensitive gamma-ray detectors to identify bursters, Swift carries telescopes that are sensitive to X rays, ultraviolet light, and visible light. Thus, instead of having to wait for ground-based telescopes to make follow-up observations of gamma-ray bursters, Swift can make these observations itself as swiftly as possible (hence the spacecraft's name). Observations from Swift may help us better understand the nature of long-duration bursters, and provide important clues about the still enigmatic short-duration bursters.

22-5 Supermassive black holes exist at the centers of most galaxies

Stellar evolution makes it possible to form black holes with masses several times that of the Sun. But calculations suggest other ways to form black holes that are either extremely large or extraordinarily small.

Supermassive Black Holes

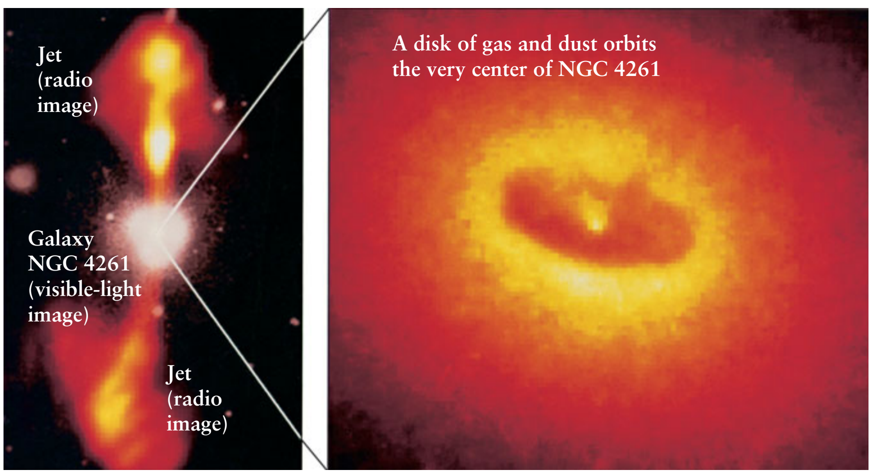
As we will discuss in Chapter 26, galaxies formed in the early universe by the coalescence of gas clouds. During this formation process, some of a galaxy's gas could have plunged straight toward the center of the galaxy, where it collected and compressed together under its own gravity. If this central clump became sufficiently dense, it could have formed a black hole. Typical galaxies have masses of 10^{11} $M_\odot$, so even a tiny percentage of the galaxy's gas collected at its center would give rise to a **supermassive black hole** with a truly stupendous mass.

The largest black holes have billions of times the mass of the Sun

Since the 1970s, astronomers have found evidence suggesting the existence of supermassive black holes in other galaxies. With the advent of a new generation of optical telescopes in the 1990s, it became possible to see the environments of these suspected black holes with unprecedented detail (Figure 22-16). Figure 22-16*a* shows radio emissions from two jets that extend some 15,000 pc (50,000 ly) from the center of the galaxy NGC 4261. These resemble an immensely magnified version of the jets seen emerging from some stellar-mass black holes (see Section 22-3). The highly magnified Hubble Space Telescope image in Figure 22-16*b* shows a disk some 250 pc (800 ly) in diameter at the very center of this galaxy. The jets are perpendicular to the plane of the disk, just as in Figure 22-12.

By measuring the Doppler shifts of light coming from the two sides of the disk, astronomers found that the disk material was orbiting around the bright object at the center of the disk at speeds of hundreds of kilometers per second. (By comparison, the Earth orbits the Sun at 30 km/s.) Given the size of the material's orbit, and using Newton's form of Kepler's third law (see Section 4-7 and Box 4-4), they were able to calculate the mass of the central bright object. The answer is an amazing 1.2 *billion* (1.2×10^9) solar masses! What is more, the observations show that this object can be no larger than our solar system. The only possible explanation is that the object at the center of NGC 4261 is a black hole.

Dozens of other black holes in the centers of galaxies have been identified by their gravitational effect on surrounding gas and dust. Surveys of galaxies have shown that supermassive black holes are not at all unusual; most large galaxies appear to have them at their centers. As we will see in the next chapter, a black hole with several million solar masses lies at the center of our own Milky Way Galaxy, some 8000 pc (26,000 ly) from Earth.



(a) Galaxy NGC 4261 (b) Evidence for a supermassive black hole in NGC 4261

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

Figure 22-16

A Supermassive Black Hole The galaxy NGC 4261 lies some 30 million pc (100 million ly) from Earth in the constellation Virgo. **(a)** This composite view superimposes a visible-light image of the galaxy (white) with a radio image of the galaxy's immense jets (orange). **(b)** This Hubble Space Telescope image shows a disk of gas and dust about 250 pc (800 ly) across at the center of NGC 4261. Observations indicate that a supermassive black hole is at the center of the disk. (NASA, ESA)

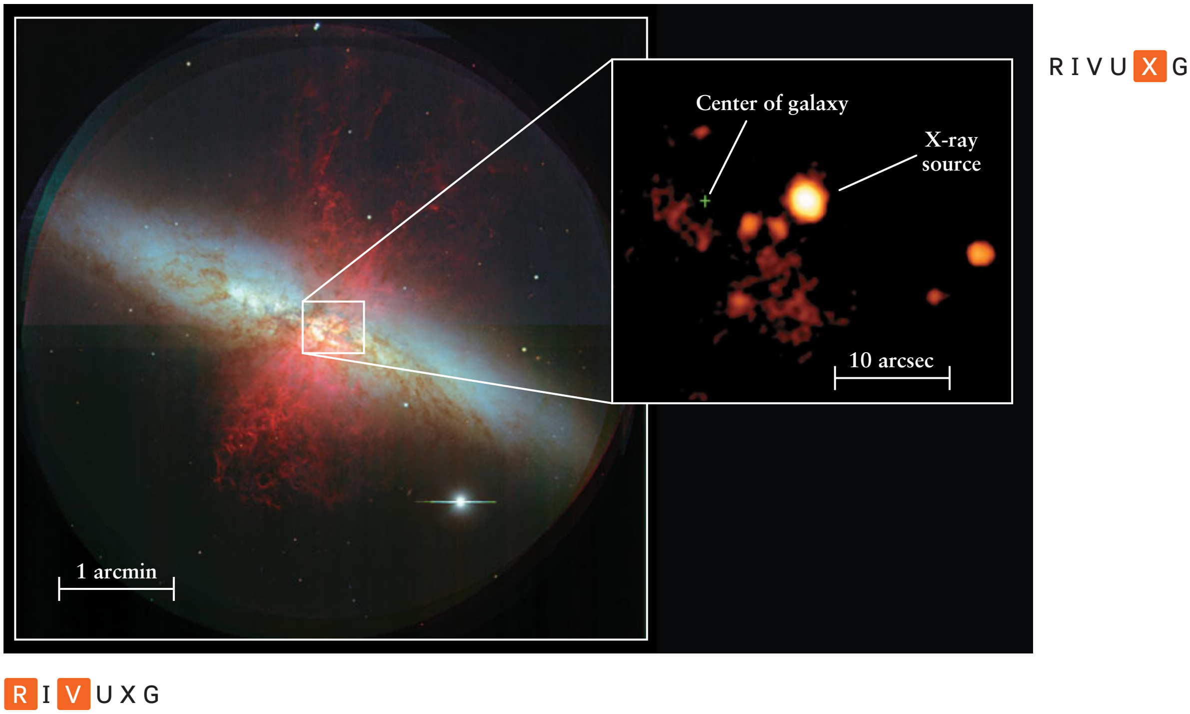


Figure 22-17

An Intermediate-mass Black Hole? M82 is an unusual galaxy in the constellation Ursa Major. (The cover of this book shows another image of M82.) The inset is an image of the central region of this galaxy from the Chandra X-ray Observatory. The bright, compact X-ray source varies in its light output over a period of months. The properties of this source suggest that it may be a black hole of roughly 500 solar masses or more. (Subaru Telescope, National Astronomical Observatory of Japan; inset: NASA/SAO/CXC)

VIDEO 22.1

Other Black Holes

Black holes seem to fall into two very different groups depending on their size. Supermassive black holes have masses from 10^6 to 10^9 $M_\odot$, while stellar-mass black holes like those discussed in Section 22-3 have masses around 10 $M_\odot$. In 2000 a team of astronomers from Japan, the United States, and Great Britain reported finding a "missing link" between these two groups.

WEB LINK 22.11

The new discovery was a fluctuating X-ray source located some 200 pc (600 ly) from the center of the galaxy M82 (Figure 22-17). Images do not reveal material orbiting around this source, so we cannot determine the source's mass using Newton's form of Kepler's third law. But by relating the source's X-ray luminosity to the rate at which matter would have to fall onto the source to produce that luminosity, it can be shown that the source must have a mass of at least 500 $M_\odot$ or so. The rapid fluctuations of the source show that it has a very small diameter, which suggests that it is a black hole (see Section 22-3). Hence, the source in M82 may be an **intermediate-mass black hole** (also called a **mid-mass black hole**) Additional intermediate-mass black hole candidates have been found in other galaxies. Such black holes might have formed by the coalescence of many normal stars or by the direct merger of stellar-mass black holes.

During the early 1970s, Stephen Hawking of Cambridge University proposed the existence of yet another type of black hole. He suggested that during the Big Bang some 13.7 billion years ago, local pockets of the universe could have been dense enough and under sufficient pressure to form small black holes. According to his theory, these **primordial black holes** could have had masses as large as the Earth or as small as 5×10^{-8} kg, about 1/40 the mass of a single raindrop. To date, however, no evidence has been found for the existence of primordial black holes in nature.

22-6 A nonrotating black hole has only a "center" and a "surface"

Understanding in detail how black holes form is a challenging task. But once a black hole forms, it has a remarkably simple structure.

Surrounding a black hole, where the escape speed from the hole just equals the speed of light, is the **event horizon.** You can think of this sphere as the "surface" of the black hole, although the black hole's mass all lies well inside this "surface." Once a massive dying star collapses to within its event horizon, it disappears permanently from the universe. The term "event horizon" is quite appropriate, because this surface is like a horizon beyond which we cannot see any events.

BOX 22-2 Tools of the Astronomer's Trade

The Schwarzschild Radius

The Schwarzschild radius (R_{Sch}) is the distance from the center of a nonrotating black hole to its event horizon. You can think of it as the "size" of the black hole. The Schwarzschild radius is related to the mass M of the black hole by

Schwarzschild radius

$$R_{Sch} = \frac{2GM}{c^2}$$

R_{Sch} = Schwarzschild radius of a black hole

G = universal constant of gravitation

M = mass of black hole

c = speed of light

When using this formula, be careful to express the mass in kilograms, not solar masses. That is because of the units in which G and c are commonly expressed: $G = 6.67 \times 10^{-11}\ \mathrm{N \cdot m^2/kg^2}$ and $c = 3.00 \times 10^8$ m/s. If you use M in kilograms, the answer that you get for R_{Sch} will be in meters.

EXAMPLE: Find the Schwarzschild radius (in kilometers) of a black hole with 10 times the mass of the Sun.

Situation: We are given the mass of the black hole (10 $M_\odot$) and wish to determine its Schwarzschild radius R_{Sch}.

Tools: We use the above formula for R_{Sch}, being careful to first convert the mass from solar masses to kilograms.

Answer: One solar mass is 1.99×10^{30} kg, so in this case $M = 10 \times 1.99 \times 10^{30}\ \mathrm{kg} = 1.99 \times 10^{31}$ kg. The Schwarzschild radius of this black hole is then

$$R_{Sch} = \frac{2GM}{c^2} = \frac{2(6.67 \times 10^{-11})(1.99 \times 10^{31})}{(3.00 \times 10^8)^2}$$
$$= 3.0 \times 10^4\ \mathrm{m}$$

One kilometer is 10^3 meters, so the Schwarzschild radius of this 10-$M_\odot$ black hole is

$$R_{Sch} = (3.0 \times 10^4\ \mathrm{m}) \times \frac{1\ \mathrm{km}}{10^3\ \mathrm{m}} = 30\ \mathrm{km, or}\ 19\ \mathrm{mi}$$

Review: The Schwarzschild radius is directly proportional to the mass of the black hole. Thus, a black hole with 10 times the mass (100 $M_\odot$) would have a Schwarzschild radius 10 times larger ($R_{Sch} = 10 \times 30\ \mathrm{km} = 300$ km); a black hole with half the mass (5 $M_\odot$) would have half as large a Schwarzschild radius ($R_{Sch} = 1/2 \times 30\ \mathrm{km} = 15$ km), and so on.

If the dying star is not rotating before it collapses, the black hole will likewise not be rotating. The distance from the center of a nonrotating black hole to its event horizon is called the **Schwarzschild radius** (denoted R_{Sch}), after the German physicist Karl Schwarzschild who first determined its properties. Box 22-2 describes how to calculate the Schwarzschild radius, which depends only on the black hole's mass. The more massive the black hole, the larger its event horizon.

Inside a Black Hole

Once a nonrotating star has contracted inside its event horizon, nothing can prevent its complete collapse. The star's entire mass is crushed to zero volume—and hence infinite density—at a single point, known as the **singularity**, at the center of the black hole. We now can see that the structure of a nonrotating black hole is quite simple. As drawn in Figure 22-18, it has only two parts: a singularity at the center surrounded by a spherical event horizon.

All of the mass of a nonrotating black hole is concentrated in a point of zero size and infinitely high density

To understand why the complete collapse of such a doomed star is inevitable, think about your own life on the Earth, far from any black holes. You have the freedom to move as you wish through the three dimensions of space: up and down, left and right, or forward and back. But you do not have the freedom to move at will through the dimension of time. Whether we like it or not, we are all carried inexorably from the cradle to the grave.

Inside a black hole, gravity distorts the structure of spacetime so severely that the directions of space and time become interchanged. In a sense, inside a black hole you acquire a limited ability to affect the passage of time. This seeming gain does you no good, however, because you lose a corresponding amount of freedom to move through space. Whether you like it or not, you will be dragged inexorably from the event horizon toward the singularity. Just as no force in the universe can prevent the forward march of time from past to future outside a black hole, no force in the universe can prevent the inward march of space from event horizon to singularity inside a black hole. Once an object dropped into a black hole crosses the event horizon, it is gone forever, like an object dropped into a bottomless pit.

The same is true for a light beam aimed at the black hole. Because all this light will be absorbed and none reflected back, a

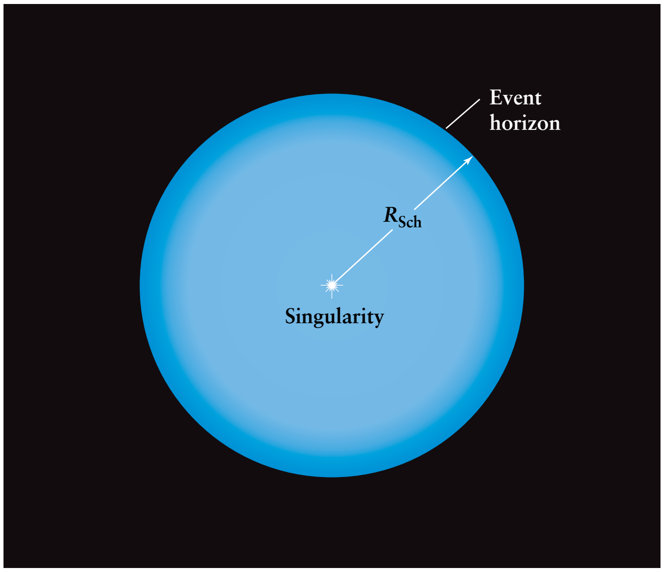


Figure 22-18

The Structure of a Nonrotating (Schwarzschild) Black Hole

A nonrotating black hole has only two parts: a singularity, where all of the mass is located, and a surrounding event horizon. The distance from the singularity to the event horizon is called the Schwarzschild radius (R_{Sch}). The event horizon is a one-way surface: Things can fall in, but nothing can get out.

black hole is totally and utterly black. (By contrast, even the blackest object on Earth reflects *some* light when you shine a flashlight on it. A black hole would reflect none.)

At a black hole's singularity, the strength of gravity is infinite, so the curvature of spacetime there is infinite. Space and time at the singularity are thus all jumbled up. They do not exist as separate, distinctive entities.

This confusion of space and time has profound implications for what goes on inside a black hole. All the laws of physics require a clear, distinct background of space and time. Without this identifiable background, we cannot speak rationally about the arrangement of objects in space or the ordering of events in time. Because space and time are all jumbled up at the center of a black hole, the singularity there does not obey the laws of physics. The singularity behaves in a random, capricious fashion, totally devoid of rhyme or reason.

Fortunately, we are shielded from the singularity by the event horizon. In other words, although random things do happen at the singularity, none of their effects manages to escape beyond the event horizon. Consequently, the outside universe remains understandable and predictable.

The chaotic, random behavior of the singularity has been so disturbing to physicists that in 1969 the British mathematician Roger Penrose and his colleagues proposed the **law of cosmic censorship:** "Thou shalt not have naked singularities." In other words, every singularity must be completely surrounded by an event horizon, because an exposed singularity could affect the universe in an unpredictable and random way.

Although black holes are really very simple objects, there are some common misconceptions about their nature. The *Cosmic Connections* figure exposes two of these misconceptions.

22-7 Just three numbers completely describe the structure of a black hole

In addition to shielding us from singularities, the event horizon prevents us from ever knowing much about anything that falls into a black hole. A black hole is, in fact, an "information sink." Many properties of matter falling into a black hole, such as its chemical composition, texture, color, shape, and size, simply vanish as soon as the matter crosses the event horizon.

When an object falls into a black hole, all information about that object disappears except its mass, electric charge, and angular momentum

Because this information has completely vanished, it cannot affect the structure or properties of the hole. For example, consider two hypothetical black holes, one made from the gravitational collapse of 10 $M_\odot$ of iron and the other made from the gravitational collapse of 10 $M_\odot$ of peanut butter. Obviously, quite different substances went into the creation of the two holes. Once the event horizons of these two black holes have formed, however, both the iron and the peanut butter will have permanently disappeared from the universe. As seen from the outside, the two holes are absolutely identical, making it impossible for us to tell which ate the peanut butter and which ate the iron. A black hole is thus unaffected by the information it destroys.

The Three Properties of a Black Hole

Because a black hole is indeed an information sink, it is reasonable to wonder whether we can determine anything at all about a black hole. What properties *do* characterize a black hole?

First, we can measure the *mass* of a black hole. One way to do this would be by placing a satellite into orbit around the hole. After measuring the size and period of the satellite's orbit, we could use Newton's form of Kepler's third law (recall Section 4-7 and Box 4-4) to determine the mass of the black hole. This is equal to the total mass of all the material that has gone into the hole.

Second, we can also measure the total *electric charge* possessed by a black hole. Like gravity, the electric force acts over long distances—it is a long-range interaction that is felt in the space around the hole. Appropriate equipment on a space probe passing near the hole could measure the strength of the electric force, and the electric charge could thus be determined.

In actuality, we would not expect a black hole to possess much, if any, electric charge. For example, if a hole did happen to start off with a sizable positive charge, it would vigorously attract vast numbers of negatively charged electrons from the interstellar medium, which would soon neutralize the hole's charge. For this reason, astronomers usually neglect electric charge when discussing real black holes.

Although a black hole might theoretically have a tiny electric charge, it can have no magnetic field of its own whatsoever. (We discussed in Section 22-3 how magnetic fields are involved in producing jets from black holes. However, these fields are associated with the accretion disk around the black hole, not the black hole itself.)

When a black hole is created, however, the collapsing star from which it forms may possess an appreciable magnetic field. The star must therefore radiate this magnetic field away before it

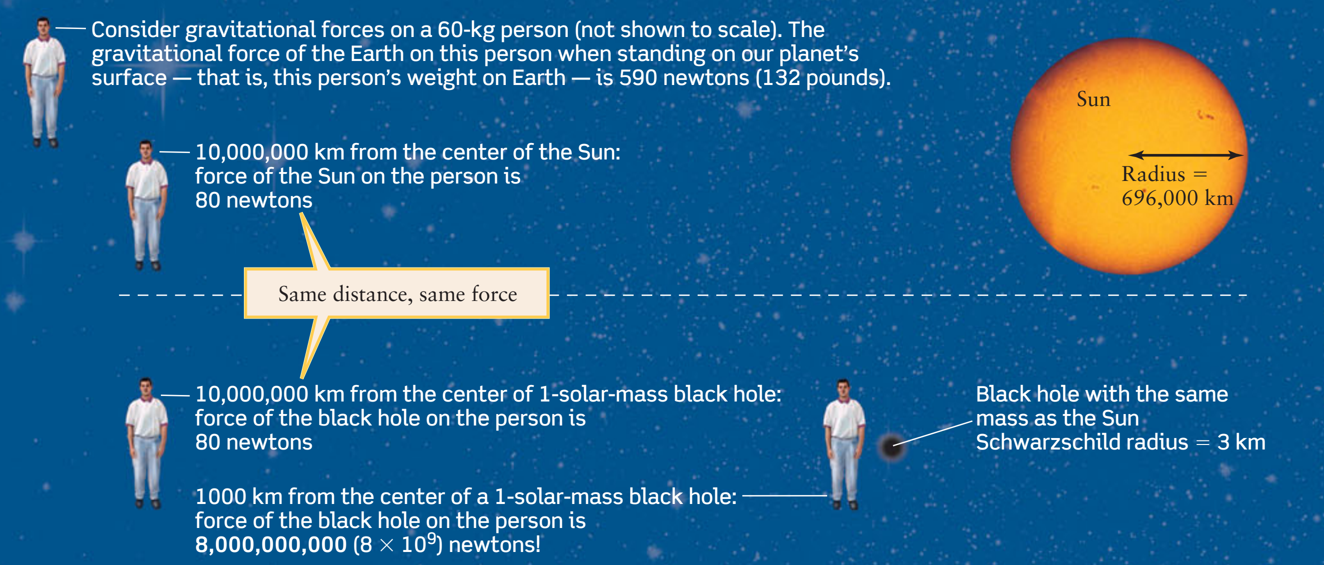
COSMIC CONNECTIONS

Black Hole "Urban Legends"

Black holes are such intriguing objects that a folklore has developed about them. As often is the case with folklore, myth and reality can become confused. These illustrations depict two misconceptions or "urban legends" about black holes.

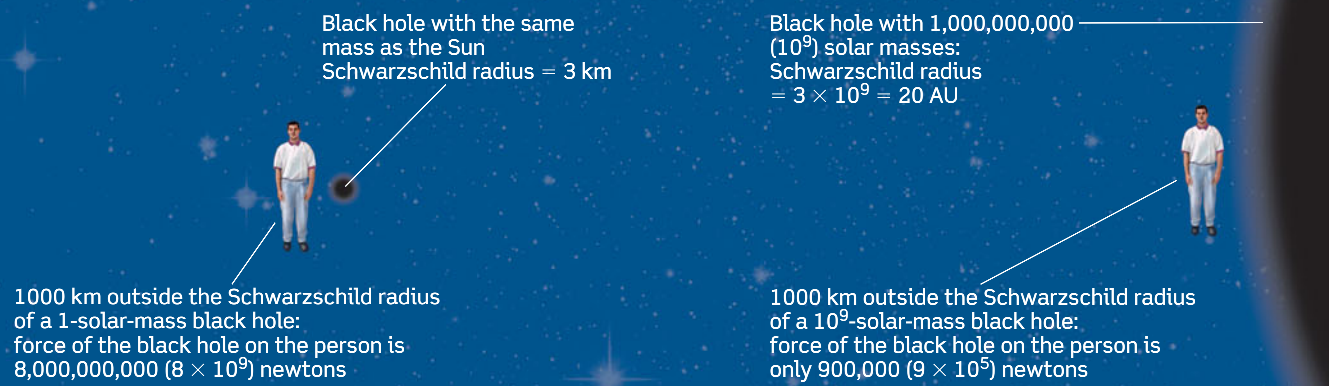
Urban Legend #1:
When a star becomes a black hole, its gravitational pull becomes stronger.

Reality: At great distances from a black hole, the gravitational force that it exerts is exactly the *same* as that exerted by an ordinary star of the same mass. The truly stupendous gravitational effects of a black hole appear only if you venture close to the black hole's event horizon.



Urban Legend #2:
The larger a black hole, the more powerful the gravity near its Schwarzschild radius.

Reality: If you increase the mass of a black hole, its Schwarzschild radius increases by the same factor. Newton's law of universal gravitation, F = Gm_1m_2Ir2, tells us that the force that an object of mass m_1 exerts on a second object of mass m_2 is directly proportional to mass m_1 but *inversely* proportional to the *square* of the distance r between the two objects. Hence the force that a black hole exerts on an object a given distance outside its Schwarzschild actually *decreases* as the black hole's mass and Schwarzschild radius increase.



can settle down inside its event horizon. Theory predicts that the star does this by emitting electromagnetic and gravitational waves. As we saw in Section 22-2, gravitational waves are ripples in the overall geometry of space. Various experiments soon to be in operation may detect bursts of gravitational radiation emitted by massive stars as they collapse to form black holes.

Third, we can detect the effects of a black hole's rotation, that is, measure its *angular momentum*. An object's angular momentum is related to how fast it rotates and how the object's mass is distributed over its volume. As a dead star collapses into a black hole, its rotation naturally speeds up as its mass moves toward the center, just as a figure skater rotates faster when she pulls her arms and legs in. This same effect explains the rotation of the solar nebula (see Section 8-4), as well as why neutron stars spin so fast (see Section 21-3). We expect a black hole that forms from a rotating star to be spinning very rapidly.

Rotating Black Holes

When the matter that collapses to form a black hole is rotating, that matter does not compress to a point. Instead, it collapses into a ring-shaped singularity located between the center of the hole and the event horizon (Figure 22-19). The structure of such rotating black holes was first worked out in 1963 by the New Zealand mathematician Roy Kerr. Most rotating black holes should be spinning thousands of times per second, even faster than the most rapid pulsars (see Section 21-7).

WEB LINK 22.12

If a rotating black hole is surrounded by an accretion disk with a magnetic field, it may be possible for the field to steal some of the rotational energy and angular momentum from the black hole and transfer it to the disk. The magnetic field acts as a "brake" that retards the black hole's rotation and makes the disk's rotation speed up. (On Earth, magnetic braking is used to slow locomotives, amusement park rides, and hybrid automobiles to a smooth stop without generating excess heat.) Figure 22-20 shows an arching magnetic field that connects an accretion disk with a rotating black hole in just this way. This process may be taking place with the supermassive black hole at the center of the galaxy MCG–6-15-30. Observations with the XMM-Newton telescope (see Section 6-7) indicate that unusually intense X-rays are coming from the accretion disk around this black hole. The suspicion is that the energy to power this radiation may be extracted from the black hole's rotation. (This transfer of energy and angular momentum can also go the other way. The image that opens this chapter depicts a system in which magnetic fields are thought to transfer angular momentum *to* a black hole *from* its accretion disk. Robbing the accretion disk of its rotation makes it easier for its material to fall into the black hole.)

WEB LINK 22.13

Even a rotating black hole without an accretion disk can transfer energy to other objects. This is possible according to Einstein's general theory of relativity, which makes the startling prediction that a rotating body drags spacetime around it. (Very precise spacecraft measurements indicate that the rotating Earth drags spacetime in just this way.) Surrounding the event horizon of every rotating black hole is a region where this dragging of space and time is so severe that it is impossible to stay in the same place. No matter what you do, you get pulled around the hole, along with the rotating geometry of space and time. This region, where it is impossible to be at rest, is called the **ergoregion** (see Figure 22-19).

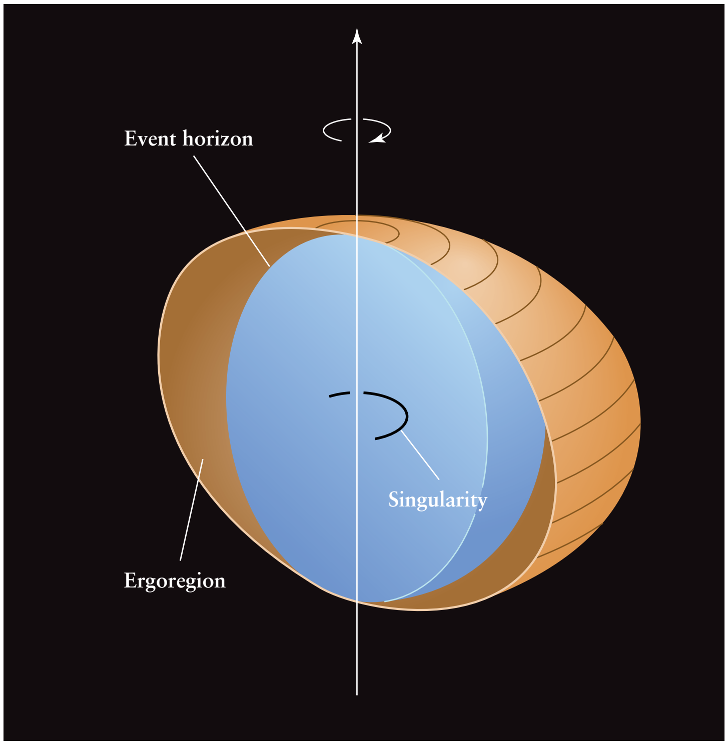


Figure 22-19

The Structure of a Rotating (Kerr) Black Hole The singularity of a rotating black hole is an infinitely thin ring centered on the geometrical center of the hole. (It appears as an arc in this cutaway diagram.) Outside the spherical event horizon is the doughnut-shaped ergoregion, where the dragging of spacetime around the hole is so severe that nothing can remain at rest. The tan-colored surface marks the outer boundary of the ergoregion.

To measure a black hole's angular momentum, we could hypothetically place two satellites in orbit about the hole. Suppose that one satellite circles the hole in the same direction the hole rotates and the other in the opposite direction. One satellite is thus carried along by the geometry of space and time, but the other is constantly fighting its way "upstream." The two satellites will thus have different orbital periods. By comparing these two periods, astronomers can deduce the total angular momentum of the hole.

Because the ergoregion is outside the event horizon, this bizarre region is accessible to us, and spacecraft could travel through it without disappearing into the black hole. According to detailed calculations, objects grazing the ergoregion could be catapulted back out into space at tremendous speeds. In other words, the ejected object could leave the ergoregion with more energy than it had initially, having extracted added energy from the hole's rotation. This is called the *Penrose process*, after Roger Penrose, the British mathematician who proposed it.

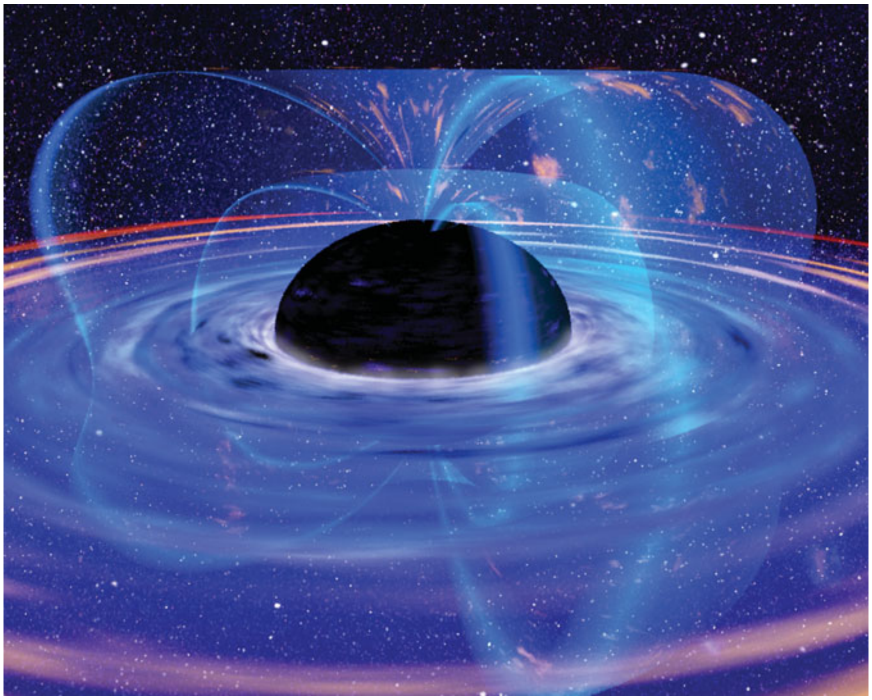


Figure 22-20

A Rotating Supermassive Black Hole This artist's impression shows the accretion disk around the supermassive black hole at the heart of the galaxy MCG-6-15-30. The arching magnetic field allows the accretion disk to extract energy and angular momentum from the black hole. (XMM-Newton/ESA/NASA)

Mass, charge, and angular momentum are the only three properties that a black hole possesses. This simplicity is the essence of the **no-hair theorem,** first formulated in the early 1970s: "Black holes have no hair." Once matter has fallen into the hole, any and all additional properties ("hair") carried by the matter have disappeared from the universe. Hence, these properties can have no effect on the structure of the hole.

22-8 Falling into a black hole is an infinite voyage

WEB LINK 22.14

Imagine that you are on board a spacecraft at a safe distance from a 5-$M_\odot$ black hole. A distance of 1000 Schwarzschild radii, or 15,000 km, would suffice. You now release a space probe and let it fall into the black hole. To make the probe easier to track, it is coated with a phosphorescent paint that emits a blue glow. You also equip the probe with a video camera that transmits images of the view of the approaching black hole. What will you see as the probe falls?

As Figure 22-21*a* shows, at 1000 Schwarzschild radii from the black hole the video camera would send back a rather normal view of space. But as the probe approaches the black hole, the bending of light by the black hole (see Section 22-2) becomes more pronounced. Light rays passing close to the back hole are deflected so severely that at close range, the camera will actually see multiple images of the same stars (Figure 22-21*b*).

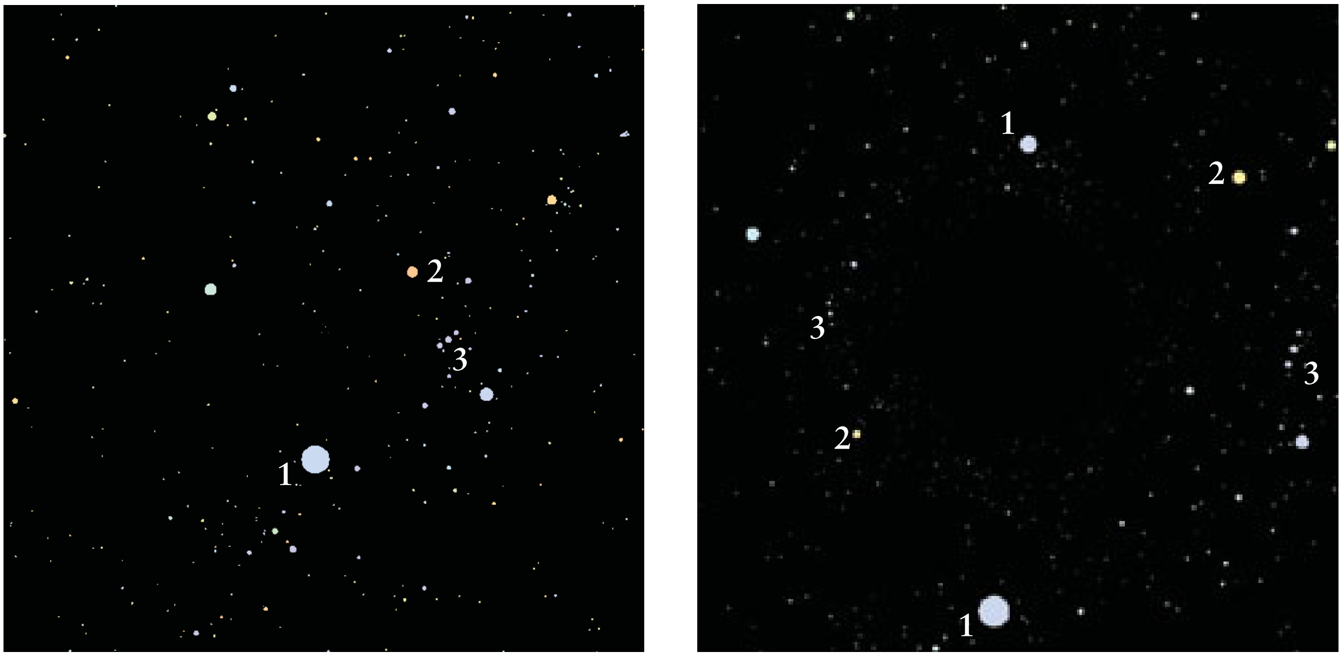
The gravitational field of a black hole distorts the image of other stars

How will the probe itself look from your vantage point at a safe distance? You might expect that as the probe falls, its speed should continue to increase. This is true up to a point. But as the probe approaches the event horizon, where the black hole's gravity is extremely strong, the gravitational slowing of time becomes so pronounced that the probe will appear to slow down! From your point of view, the spacecraft takes an infinite amount of time to reach the event horizon, where it will appear to remain suspended for all eternity.

To watch these effects, however, you will need special equipment. The reason is the gravitational redshift: As the probe falls, the light reaching you from the probe is shifted to longer and longer wavelengths. The probe's blue glow will turn green, then yellow, then red, and eventually fade into infrared wavelengths that your unaided eye cannot detect.

If you could view the falling probe with infrared goggles, you would see it come to an unpleasant end. Near the event horizon, the strength of the black hole's gravity increases dramatically as the probe moves just a short distance closer to the hole. In fact, the side of the probe nearest the black hole feels a much stronger gravitational pull than the side opposite the hole. These *tidal forces* are like those that the Moon exerts on the Earth (see Section 4-8), but tremendously stronger. Furthermore, the sides of the probe are pulled together, since the hole's gravity makes them fall in straight lines aimed at the center of the hole. The net effect is that the probe will be stretched out along the line pointing toward the hole, and squeezed together along the perpendicular directions (Figure 22-22). The stretching is so great that it will rip the probe into atoms, and even rip the atoms apart.

Someone foolhardy enough to ride along with the probe would have a very different view of the fall than yours. From this



(a) Looking directly toward the black hole from a distance of 1000 Schwarzschild radii: Note positions of stars 1, 2, and 3.

(b) Looking directly toward the black hole from a distance of 10 Schwarzschild radii: Light bending causes multiple images.

Figure 22-21

The View Approaching a Black Hole These images are taken from a computer simulation of the appearance of the sky behind a black hole. **(a)** The bending of light by the black hole is negligible at large distances, but **(b)** becomes evident as you approach the black hole. (Courtesy Robert J. Nemiroff, Michigan Technological University)

astronaut's point of view, there is no slowing of time, and the probe continues its fall through the event horizon into the singularity. The astronaut is unlikely to enjoy the ride, however, since he will be disintegrated by tidal forces as he falls inward toward the singularity.

Could a black hole somehow be connected to another part of spacetime, or even some other universe? General relativity predicts that such connections, called **wormholes,** can exist for rotating black holes. Box 22-3 discusses the possibility that such wormholes really exist.

22-9 Black holes evaporate

With all the mass of a black hole hidden behind its event horizon and collapsed into a singularity, it may seem that there is no way of getting mass from the black hole back out into the universe. But in fact there is, as Stephen Hawking pointed out in the 1970s. To understand how this is possible, we must look at the behavior of matter at the submicroscopic scale of nuclei, and electrons.

The **Heisenberg uncertainty principle** is a basic tenet of submicroscopic physics. This principle states that you cannot determine

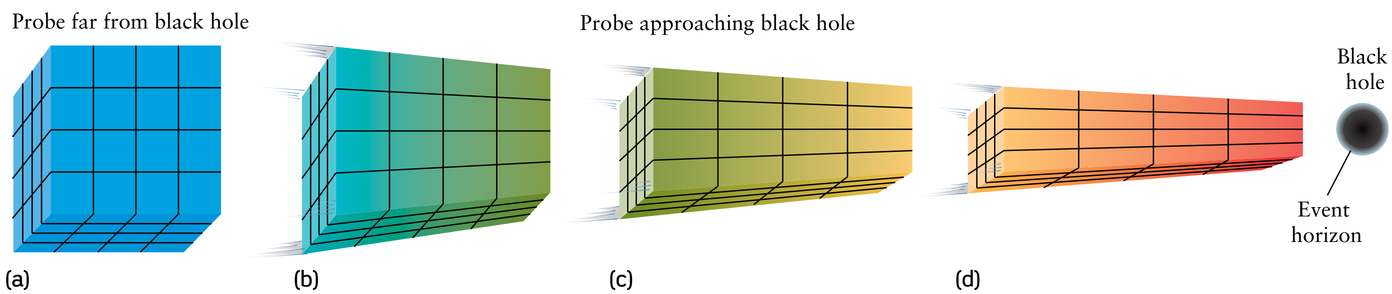


Figure 22-22

Falling into a Black Hole (a) A cube-shaped probe is dropped from a distance of 1000 Schwarzschild radii from a 5-$M_\odot$ black hole. **(b), (c), (d)** As the probe approaches the event horizon, it is distorted into a long, thin shape by the black hole's extreme gravity. A distant observer sees the probe change color as photons from the probe undergo a strong gravitational redshift.

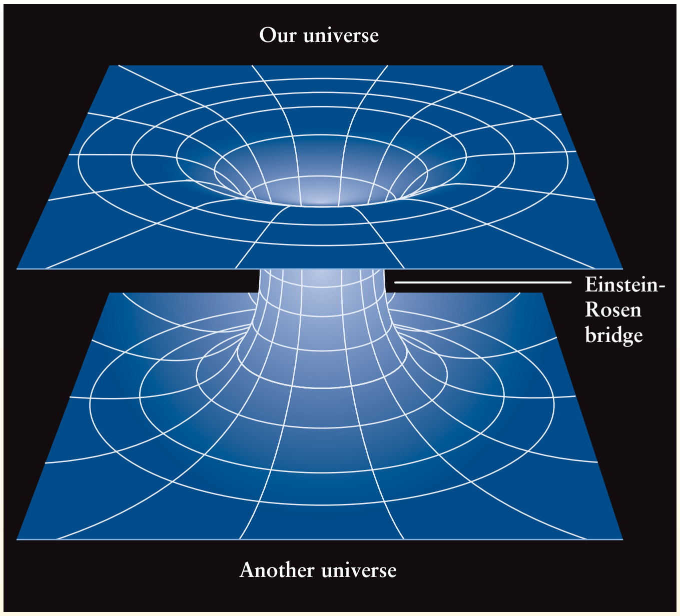
BOX 22-3 The Heavens on the Earth

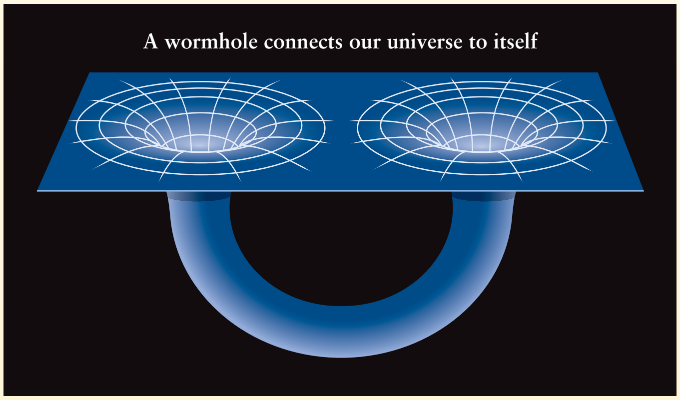
Wormholes and Time Machines

Wormholes have become a staple of science fiction. They are certainly a convenient plot device: The gallant crew flies their spaceship into one end of a wormhole, and a short time later they emerge many light-years away. Time machines, too, have been part of science fiction ever since H. G. Wells published his classic novel *The Time Machine* at the end of the nineteenth century. What would it be like to go back in time and watch famous historical events—or even change history?

Wormholes and time machines challenge our normal ideas about space and time. So, too, does the general theory of relativity. Could it be that this theory really makes it possible to travel through wormholes and to travel in time?

In the 1930s, Einstein and his colleague Nathan Rosen discovered that a black hole could possibly connect our universe with a second domain of space and time that is separate from ours. The first diagram shows this connection, called an *Einstein-Rosen bridge*. You could think of the upper surface as our universe and the lower surface as a "parallel universe." Alternatively, the upper and lower surfaces could be different regions of our own universe. In that case, an Einstein-Rosen bridge would connect our universe with itself, forming a wormhole, shown in the second diagram.





A wormhole may seem like a shortcut to distant places in our universe, but detailed calculations reveal a major obstacle. The powerful gravity of a black hole causes the wormhole to collapse almost as soon as it forms. As a result, to get from one side of a wormhole to the other, you would have to travel faster than the speed of light, which is not possible.

Caltech physicists Kip Thorne, Michael Morris, and Ulvi Yurtsever have proposed a scheme that might get around the difficulty of a collapsing wormhole. According to general relativity, they point out, pressure as well as mass can be a source of gravity. Normally we do not see the gravitational effects of pressure because they are so small. Thorne and his colleagues speculate that a technologically advanced civilization might someday be able to use pressure to produce *antigravity* strong enough to keep the wormhole open.

If a wormhole could be held open, it could also be a time machine. To see how, imagine you take one end of a wormhole and move it around for a while at speeds very near the speed of light. As we saw in Box 22-1, such motion causes clocks to slow down. Thus, when you stop moving that end of the wormhole, you find that it has not aged as much as the stationary end. In other words, one side of the wormhole has a different time from the other. As a result, you could go into one end of the wormhole at a late time and come out at an early time. For example, you might go in at 10 A.M. and come out at 9 A.M.

Time machines challenge ordinary logic. If you could get back from a trip an hour before you left, you could meet yourself and tell yourself what a nice journey you had. Then both of you could take the trip. If you and your twin return just before you both left, there would be four of you. And all four of you could take the trip again. And then all eight of you. Then all sixteen of you . . .

Making copies of yourself is an example of how time machines violate *causality,* the notion that effects must follow their causes. To date, we have never seen a violation of causality. Many scientists would therefore like to show that time machines cannot exist. The British astrophysicist Stephen Hawking points to one strong bit of observational evidence against time machines: We are not being visited by hordes of tourists from the future. If we could discover why nature seems to preclude time machines, we would have a much deeper understanding of the nature of space and time.

precisely both the position and the speed of a subatomic particle. Over extremely short distances or times, a certain amount of "fuzziness" is built into the nature of reality.

The Heisenberg uncertainty principle leads to the concept of **virtual pairs:** At every point in space, pairs of particles and antiparticles are constantly being created and destroyed. An antiparticle is quite like an ordinary particle except that it has an opposite electric charge and can annihilate an ordinary particle so that both disappear, usually leaving a pair of photons in their place. In the case of virtual pairs, the process of creation and annihilation occurs over such incredibly brief time intervals that these virtual particles and antiparticles cannot be observed directly.

Think about a tiny black hole. Furthermore, think about the momentary creation of a virtual pair of an electron and a positron (the antiparticle of an electron) just outside the hole's event horizon. It may happen that one of the particles falls into the black hole. Its partner is then deprived of a counterpart with which to annihilate and must therefore become a real particle. To accomplish this conversion, some of the energy of the black hole's gravity must be converted into matter, according to $E = mc^2$. This decreases the mass of the black hole by a corresponding amount, and the particle is free to escape from the hole. In this way, particles can quantum-mechanically "leak" out of a black hole, carrying some of the hole's mass with them (Figure 22-23).

The less mass a black hole has, the more easily particles can leak out through its event horizon. Jacob Bekenstein and Stephen Hawking proved mathematically that you can speak of the *temperature* of a black hole as a way of describing the amounts of energy carried away by particles leaking out of it. For example, a 1-trillion-ton (10^{15} kg) black hole emits particles and energy as if it were a blackbody with a temperature of nearly 10^9 K.

A fundamental uncertainty in physical knowledge makes it possible for black holes to emit particles and radiation into space

For stellar-mass black holes, such as Cygnus X-1, this effect is negligible over time spans of billions of years. The reason is that the temperature of a black hole is inversely proportional to the black hole's mass. Compared to a trillion-ton black hole, a 5-$M_\odot$ (10^{31} kg) hole has a mass that is 10^{16} times greater and a temperature T that is only 10^{-16} as much:

$$T = 10^{-16} \times 10^9 \text{ K} = 10^{-7} \text{ K}$$

This temperature is barely above absolute zero. In other words, ordinary black holes have such low temperatures that particles hardly ever manage to escape from them.

The story is different for low-mass black holes, such as the primordial black holes we discussed in Section 22-5. As particles escape from a small black hole, the mass of the black hole decreases, making its temperature go up. As its temperature rises, still more particles escape, further decreasing the hole's mass and forcing the temperature still higher. This runaway process of **black hole evaporation** finally causes the hole to vanish altogether. During its final seconds of evaporation, the hole gives up the last of its mass in a violent burst of energy equal to the detonation of a billion megaton hydrogen bombs.

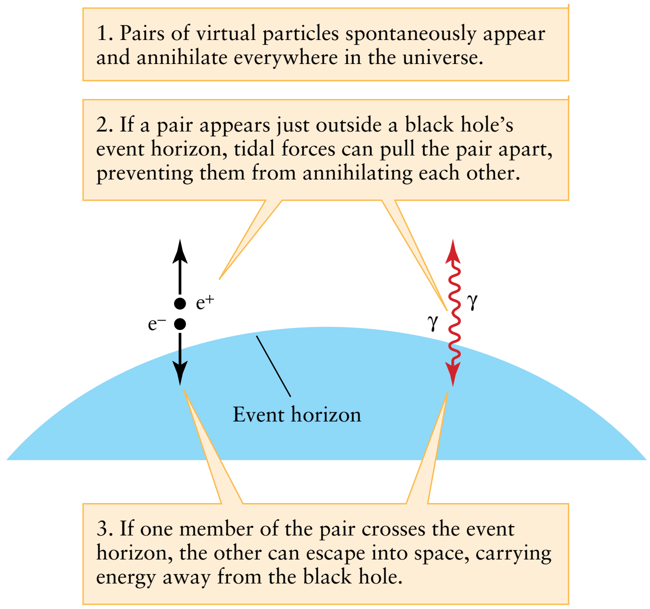


Figure 22-23

Evaporation of a Black Hole This illustration shows two pairs of virtual particles—an electron (e^-) and a positron (e^+), and a pair of photons (γ)—appearing just outside the event horizon of a black hole. If one member of the pair escapes and carries energy away from the black hole, the black hole decreases in mass and the event horizon shrinks.

A 10^{10}-kg primordial black hole (comparable to the mass of Mount Everest) would take about 15 billion years to evaporate. This is close to the present age of the universe. If primordial black holes were formed in the Big Bang, we would expect to see some of them going through the explosive final stages of evaporation. To date, however, there is no compelling evidence that we have seen such evaporations taking place, and many astronomers suspect that there are no primordial black holes.

By contrast, a 5-$M_\odot$ black hole would take more than 10^{62} years to evaporate, and a supermassive black hole of 5 million solar masses would take more than 10^{80} years. These time spans are far longer than the age of the universe. We can safely predict that the black holes we have observed to date will remain as black holes for the foreseeable future.

Key Words

Terms preceded by an asterisk (*) are discussed in the Boxes.

black hole, p. 585
black hole evaporation, p. 600
collapsar, p. 590
equivalence principle, p. 582
ergoregion, p. 596
event horizon, p. 592
gamma-ray burster, p. 588
general theory of relativity, p. 582
gravitational radiation (gravitational waves), p. 584
gravitational redshift, p. 584
Heisenberg uncertainty principle, p. 598
intermediate-mass (mid-mass) black hole, p. 592
law of cosmic censorship, p. 594
length contraction, p. 578

*Lorentz transformations, p. 580
mid-mass (intermediate-mass) black hole, p. 592
no-hair theorem, p. 597
primordial black hole, p. 592
*proper length (proper distance), p. 581
*proper time, p. 580
Schwarzschild radius (R_{Sch}), p. 593
singularity, p. 593
spacetime, p. 578
special theory of relativity, p. 577
stellar-mass black hole, p. 586
supermassive black hole, p. 591
time dilation, p. 578
virtual pairs, p. 600
wormhole, p. 598

Key Ideas

The Special Theory of Relativity: This theory, published by Einstein in 1905, is based on the notion that there is no such thing as absolute space or time.

• The speed of light is the same to all observers, no matter how fast they are moving.

• An observer will note a slowing of clocks and a shortening of rulers that are moving with respect to the observer. This effect becomes significant only if the clock or ruler is moving at a substantial fraction of the speed of light.

• Space and time are not wholly independent of each other, but are aspects of a single entity called spacetime.

The General Theory of Relativity: Published by Einstein in 1915, this is a theory of gravity. Any massive object causes space to curve and time to slow down, and these effects manifest themselves as a gravitational force. These distortions of space and time are most noticeable in the vicinity of large masses or compact objects.

• The general theory of relativity is our most accurate description of gravitation. It predicts a number of phenomena, including the bending of light by gravity and the gravitational redshift, whose existence has been confirmed by observation and experiment.

• The general theory of relativity also predicts the existence of gravitational waves, which are ripples in the overall geometry of space and time produced by moving masses. Gravitational waves have been detected indirectly, and specialized antennas are under construction to make direct measurement of the gravitational waves from cosmic cataclysms.

Black Holes: If a stellar corpse has a mass greater than about 2 to 3 $M_\odot$, gravitational compression will overwhelm any and all forms of internal pressure. The stellar corpse will collapse to such a high density that its escape speed exceeds the speed of light.

Observing Black Holes: Black holes have been detected using indirect methods.

• Some binary star systems contain a black hole. In such a system, gases captured from the companion star by the black hole emit detectable X rays.

• Many galaxies have supermassive black holes at their centers. These are detected by observing the motions of material around the black hole.

Gamma-Ray Bursters: Short, intense bursts of gamma rays are observed at random times coming from random parts of the sky.

• By observing the afterglow of long-duration gamma-ray bursters, astronomers find that these objects have very large redshifts and appear to be located within distant galaxies. The bursts are correlated with supernovae, and may be due to an exotic type of supernova called a collapsar.

• The origin of short-duration gamma-ray bursters is unknown.

INTERACTIVE EXERCISE 22.1

Properties of Black Holes: The entire mass of a black hole is concentrated in an infinitely dense singularity.

• The singularity is surrounded by a surface called the event horizon, where the escape speed equals the speed of light. Nothing—not even light—can escape from inside the event horizon.

• A black hole has only three physical properties: mass, electric charge, and angular momentum.

• A rotating black hole (one with angular momentum) has an ergoregion around the outside of the event horizon. In the ergoregion, space and time themselves are dragged along with the rotation of the black hole.

• Black holes can evaporate, but in most cases at an extremely slow rate.

Questions

Review Questions

1. You drop a ball inside a car traveling at a steady 50 km/h in a straight line on a smooth road. Does it fall in the same way as it does inside a stationary car? How does this question relate to Einstein's special theory of relativity?
2. In Einstein's special theory of relativity, two different observers moving at different speeds will measure the same value of the speed of light. Will these same observers measure the same value of, say, the speed of an airplane? Explain.
3. A friend summarizes the special theory of relativity by saying "Everything is relative." Explain why this statement is inaccurate.
4. Serena flies past Michael in her spaceship at nearly the speed of light. According to Michael, Serena's clock runs slow. According to Serena, does Michael's clock run slow, fast, or at the normal rate? Explain.
5. Ole flies past Lena in a spherical spaceship at nearly the speed of light. According to Lena, how does the distance from the front to the back of Ole's spaceship (that is, measured along the direction of motion) compare to the distance from the top to the bottom (that is, measured perpendicular to the direction of motion)? Explain.
6. Why does the speed of light represent an ultimate speed limit?
7. Why is Einstein's general theory of relativity a better description of gravity than Newton's universal law of gravitation? Under what circumstances is Newton's description of gravity adequate?
8. Describe two different predictions of the general theory of relativity and how these predictions were tested experimentally. Do the results of the experiments agree with the theory?

9. How does a gravitational redshift differ from a Doppler shift?
10. In what circumstances are degenerate electron pressure and degenerate neutron pressure incapable of preventing the complete gravitational collapse of a dead star?
11. Should we worry about the Earth being pulled into a black hole? Why or why not?
12. How does rapid flickering of an X-ray source provide evidence that the source is small?
13. All the stellar-mass black hole candidates mentioned in the text are members of very short-period binary systems. Explain how this makes it possible to detect the presence of the black hole.
14. Astronomers cannot actually see the black hole candidates in close binary systems. How, then, do they know that these candidates are not white dwarfs or neutron stars?
15. Describe two ways in which a member of a binary star system could become a black hole.
16. What is a gamma-ray burster? What is the evidence that gamma-ray bursters are not located in the disk of our Galaxy or in a halo surrounding our Galaxy?
17. Summarize the evidence that gamma-ray bursters result from a process involving a star in a distant galaxy.
18. What is a collapsar? How does the collapsar model account for the existence of long-duration gamma-ray bursters?
19. How do astronomers locate supermassive black holes in galaxies?
20. What is an intermediate-mass black hole? How are such objects thought to form?
21. When we say that the Moon has a radius of 1738 km, we mean that this is the smallest radius that encloses all of the Moon's material. In this sense, is it correct to think of the Schwarzschild radius as the radius of a black hole? Why or why not?
22. A twenty-third-century instructor at Starfleet Academy tells her students, "If someday your starship falls into a black hole, it'll be your own fault." Explain why it would require careful piloting to direct a spacecraft into a black hole.
23. In what way is a black hole blacker than black ink or a black piece of paper?
24. If the Sun suddenly became a black hole, how would the Earth's orbit be affected? Explain.
25. According to the general theory of relativity, why can't some sort of yet-undiscovered degenerate pressure prevent the matter inside a black hole from collapsing all the way down to a singularity?
26. What is the law of cosmic censorship?
27. Is it possible to tell which chemical elements went into a black hole? Why or why not?
28. Why is it unlikely that a black hole has an electric charge?
29. What kind of black hole is surrounded by an ergoregion? What happens inside the ergoregion?
30. What is the no-hair theorem?
31. As seen by a distant observer, how long does it take an object dropped from a great distance to fall through the event horizon of a black hole? Explain.
32. If even light cannot escape from a black hole, how is it possible for black holes to evaporate?

Advanced Questions

Questions preceded by an asterisk (*) involve topics discussed in the Boxes.

> **Problem-solving tips and tools**
>
> Remember that the time to travel a certain distance is equal to the distance divided by the speed, and that the density of an object is its mass divided by its volume. The volume of a sphere of radius r is $4\pi r^3/3$. Section 4-7 describes Newton's law of universal gravitation. Box 4-4 shows how to use Newton's formulation of Kepler's third law, which explicitly includes masses; when using this formula, note that the period P must be expressed in seconds, the semimajor axis a in meters, and the masses in kilograms. For another version of this formula, in which period is in years, semimajor axis in AU, and masses in solar masses, see Section 17-9.

*33. A spaceship flies from Earth to a distant star at a constant speed. Upon arrival, a clock on board the spaceship shows a total elapsed time of 8 years for the trip. An identical clock on the Earth shows that the total elapsed time for the trip was 10 years. What was the speed of the spaceship relative to the Earth?

*34. An unstable particle called a positive pion (pronounced "pie-on") decays in an average time of 2.6×10^{-8} s. On average, how long will a positive pion last if it is traveling at 95% of the speed of light?

*35. How fast should a meter stick be moving in order to appear to be only 60 cm long?

*36. An astronaut flies from the Earth to a distant star at 80% of the speed of light. As measured by the astronaut, the one-way trip takes 15 years. (**a**) How long does the trip take as measured by an observer on the Earth? (**b**) What is the distance from the Earth to the star (in light-years) as measured by an Earth observer? As measured by the astronaut?

37. In the binary system of two neutron stars discovered by Hulse and Taylor (Section 22-2), one of the neutron stars is a pulsar. The distance between the two stars varies between 1.1 and 4.8 times the radius of the Sun. The time interval between pulses from the pulsar is not constant: It is greatest when the two stars are closest to each other and least when the two stars are farthest apart. Explain why this is consistent with the gravitational slowing of time (Figure 22-7*a*).

38. Find the total mass of the neutron star binary system discovered by Hulse and Taylor (Section 22-2), for which the orbital period is 7.75 hours and the average distance between the neutron stars is 2.8 solar radii. Is your result reasonable for a pair of neutron stars? Explain.

39. Estimate how long it will be until the two neutron stars that make up the binary system discovered by Hulse and Taylor collide with each other. Assume that the distance between the two stars will continue to decrease at its present rate of 3 mm every 7.75 hours, and use the data given in Question 38. (You can assume that the two stars are very small, so they will collide when the distance between them is equal to zero.)

40. The orbital period of the binary system containing A0620-00 is 0.32 day, and Doppler shift measurements reveal that the radial velocity of the X-ray source peaks at 457 km/s (about 1 million miles per hour). (**a**) Assuming that the orbit of the X-ray source is a circle, find the radius of its orbit in kilometers. (This is actually an estimate of the semimajor axis of the orbit.) (**b**) By using Newton's form of Kepler's third law, prove that the mass of the X-ray source must be at least 3.1 times the mass of the Sun. (*Hint:* Assume that the mass of the K5V visible star—about 0.5 $M_\odot$ from the mass-luminosity relationship—is negligible compared to that of the invisible companion.)
41. Contrast gamma-ray bursters with X-ray bursters (discussed in Section 21-9). From our models of what causes these energetic phenomena, explain why X-ray bursters emit repeated pulses but gamma-ray bursters apparently emit just once.
42. Long-duration gamma-ray bursters are only observed in galaxies where there is ongoing star formation. Explain how this is consistent with the collapsar model of how these bursters occur.
43. The spectrum of a Type Ic supernova lacks absorption lines of hydrogen and helium. This means that when a black hole formed at the center of the progenitor star, the resulting jets were more easily able to escape into space. Explain why this is so.
44. Find the orbital period of a star moving in a circular orbit of radius 500 AU around the supermassive black hole in the galaxy NGC 4261 (Section 22-5).

*45. Find the Schwarzschild radius for an object having a mass equal to that of the planet Saturn.

*46. What is the Schwarzschild radius of a black hole whose mass is that of (**a**) the Earth, (**b**) the Sun, (**c**) the supermassive black hole in NGC 4261 (Section 22-5)? In each case, also calculate what the density would be if the matter were spread uniformly throughout the volume of the event horizon.

*47. What is the mass in kilograms of a black hole whose Schwarzschild radius is 11 km?

*48. To what density must the matter of a dead 8-$M_\odot$ star be compressed in order for the star to disappear inside its event horizon? How does this compare with the density at the center of a neutron star, about 3×10^{18} kg/m^3?

*49. Prove that the density of matter needed to produce a black hole is inversely proportional to the square of the mass of the hole. If you wanted to make a black hole from matter compressed to the density of water (1000 kg/m^3), how much mass would you need?

Discussion Questions

50. The speed of light is the same for all observers, regardless of their motion. Discuss why this requires us to abandon the Newtonian view of space and time.
51. Describe the kinds of observations you might make in order to locate and identify black holes.
52. Speculate on the effects you might encounter on a trip to the center of a black hole (assuming that you could survive the journey).

Web/eBook Questions

53. Search the World Wide Web for information about a stellar-mass black hole candidate named V4641 Sgr. In what ways does it resemble other black hole candidates such as Cygnus X-1 and V404 Cygni? In what ways is it different and more dramatic? How do astronomers explain why V4641 Sgr is different?
54. Search the World Wide Web for information about supernova SN 2006aj, which was associated with gamma-ray burster GRB 060218. In what ways were this supernova and gamma-ray burster unusual? Are the observations of these objects consistent with the collapsar model?
55. Search the World Wide Web for information about the intermediate-mass black hole candidate in M82. Is this still thought to be an intermediate-mass black hole? What new evidence has been used to either support or oppose the idea that this object is an intermediate-mass black hole?
56. **The Equivalence Principle.** Access the animation "The Equivalence Principle" in Chapter 22 of the *Universe* Web site or eBook. View the animation and answer the following questions. (**a**) Describe what is happening as viewed from the frame of reference of the elevator. What causes the apple to fall to the floor of each elevator? (**b**) Describe what is happening as viewed from the frame of reference of the stars. What causes the apple to fall to the floor of each elevator? (**c**) Think of another experiment you could perform with the apple. Describe what would happen during this experiment as seen by Newton (in the left-hand box) and by Einstein (in the right-hand box).

ANIMATION 22.1

Activities

Observing Projects

57. You cannot see a black hole with a telescope. Nevertheless, you might want to observe the visible companion of Cygnus X-1. The epoch 2000 coordinates of this ninth-magnitude star are R.A. = $19^h\ 58.4^m$ and Decl. = $+35°\ 12'$, which is quite near the bright star η (eta) Cygni. Compare what you see with the photograph in Figure 22-10.
58. Use the *Starry Night Enthusiast™* program to plan observations of Cygnus X-1. Use the **Find** pane to **Centre** the view on Cygnus X-1. Select **View > Show Daylight** from the menu. (**a**) Using the time controls in the toolbar, determine when Cygnus X-1 rises and sets on today's date from your home location. (**b**) At approximately what time on today's date is Cygnus X-1 highest in the sky? Is tonight a good night for observing this star with a visible-light telescope? Would it be better placed in the sky for observation six months from now? Explain how you determined this.

STARRY NIGHT

59. Use the *Starry Night Enthusiast™* program to examine X-ray images of galaxies with supermassive black holes at their centers. Open the **Options** pane and expand the **Deep Space** layer. Select **Chandra Images** and deselect all of the other options in this

STARRY NIGHT

layer. Use the **Find** pane and Zoom controls to examine each of the following galaxies: (i) NGC 4261; (ii) Virgo A (M87); (iii) M31. Open the Options pane again and select Messier Objects and deselect Chandra Images and compare the visual images of Virgo A (M87) and M31. Suggest why supermassive black holes were discovered in these galaxies only after relatively recent advances were made in telescope and detector technology.

Collaborative Exercise

60. Using Einstein's theory of relativity, estimate (1) the length of your pencil or pen at constant speed at the speeds of a bicycle rider, a car traveling on the highway, and a commercial jet liner at cruising altitude; and (2) the speed of a light beam emitted by a spaceshift traveling at 200,000 kilometers per second toward another spaceship traveling at the same speed.

23
Our Galaxy

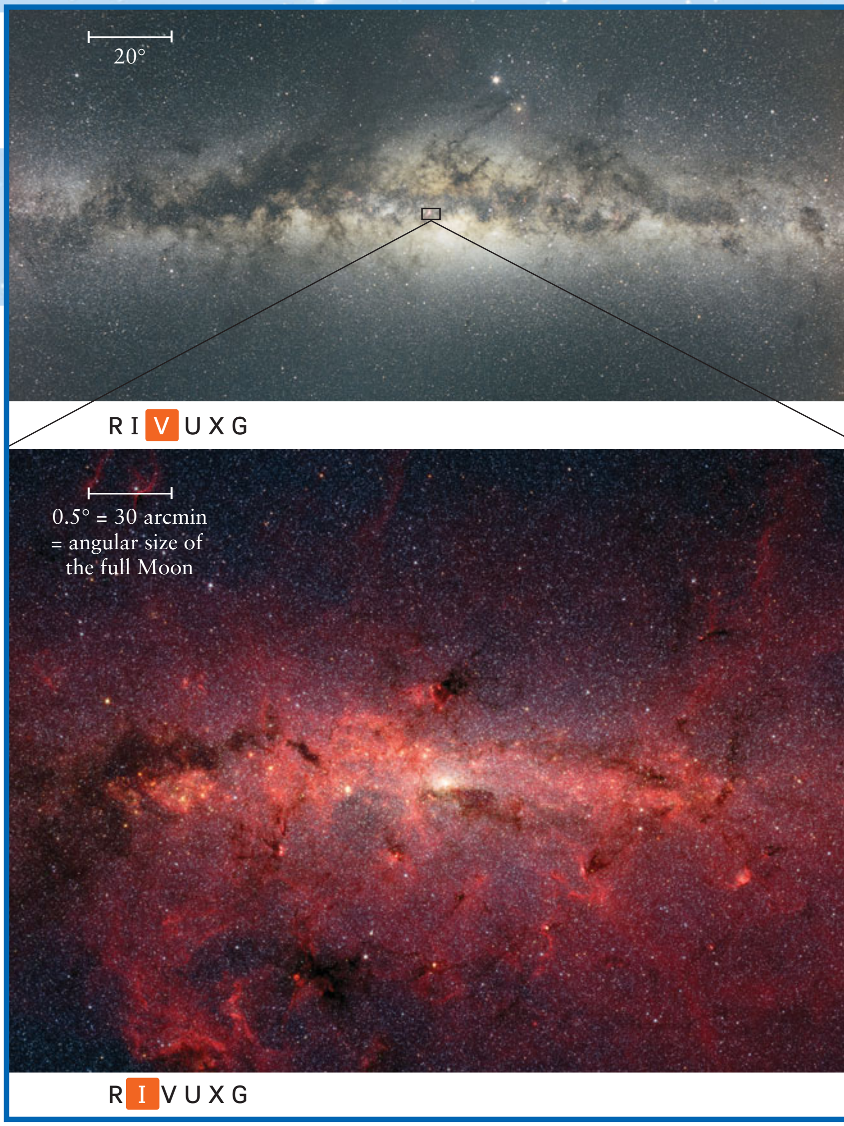
On a clear, moonless night, away from the glare of city lights, you can often see a hazy, luminous band stretching across the sky. This band, called the Milky Way, extends all the way around the celestial sphere. The upper of the two accompanying photographs is centered on the brightest part of the Milky Way, in the constellation Sagittarius.

Galileo, the first person to view the Milky Way with a telescope, discovered that it is composed of countless dim stars. Today, we realize that the Milky Way is actually a disk tens of thousands of parsecs across containing hundreds of billions of stars—one of which is our own Sun—as well as vast quantities of gas and dust. This vast assemblage of matter is collectively called the Milky Way Galaxy.

Just as Galileo's telescope revealed aspects of the Milky Way that the naked eye could not, modern astronomers use telescopes at nonvisible wavelengths to peer through our Galaxy's obscuring dust and observe what visible-light telescopes never could. For example, the lower of the two accompanying photographs is an infrared image that shows hundreds of thousands of stars near the center of the Galaxy. As we will see, radio, infrared, and X-ray observations reveal that at the very center of the Galaxy lies a black hole with a mass of 3.7 million Suns.

Modern astronomers have also discovered that most of the Milky Way's mass is not in its stars, gas, or dust, but in a halo of *dark matter* that emits no measurable radiation. What the character of this dark matter could be remains one of the greatest open questions in astronomy and physics.

The Milky Way is just one of myriad *galaxies,* or systems of stars and interstellar matter, that are spread across the observable universe. By studying our home Milky Way Galaxy, we begin to explore the universe on a grand scale. Instead of focusing on individual stars, we look at the overall arrangement and history of a huge stellar community of which the Sun is a member. In this way, we gain insights into galaxies in general and prepare ourselves to ask fundamental questions about the cosmos.



WEB LINK 23.1

Two views of the Milky Way: a wide-angle visible-light mosaic (upper) and a close-up infrared image from the Spitzer Space Telescope (lower). (upper: Dirk Hoppe; lower: NASA/JPL-Caltech/S. Stolovy, Spitzer Science Center/Caltech)

Learning Goals

By reading the sections of this chapter, you will learn

23-1 How astronomers discovered the solar system's location within the Milky Way Galaxy

23-2 The shape and size of our Galaxy

23-3 How the Milky Way's spiral structure was discovered

23-4 The evidence for the existence of dark matter in our Galaxy

23-5 What causes the Milky Way's spiral arms to form and persist

23-6 How astronomers discovered a supermassive black hole at the galactic center

23-1 The Sun is located in the disk of our Galaxy, about 8000 parsecs from the galactic center

Eighteenth-century astronomers were the first to suspect that because the Milky Way completely encircles us, all the stars in the sky are part of an enormous disk of stars—the **Milky Way Galaxy.** As we learned in Section 1-4, a **galaxy** is an immense collection of stars and interstellar matter, far larger than a star cluster. We have an edge-on view from inside the disk of our own Milky Way Galaxy, which is why the Milky Way appears as a band around the sky (Figure 23-1).

Locating the Sun Within the Galaxy: Early Attempts

But where within this disk is our own Sun? Until the twentieth century, the prevailing opinion was that the Sun and planets lie at the Galaxy's center. One of the first to come to this conclusion was the eighteenth-century English astronomer William Herschel, who discovered the planet Uranus and was a pioneering cataloger of binary star systems (see Section 17-9). Herschel's approach to determining the Sun's position within the Galaxy was to count the number of stars in each of 683 regions of the sky. He reasoned that he should see the greatest number of stars toward the Galaxy's center and a lesser number toward the Galaxy's edge.

Herschel found approximately the same density of stars all along the Milky Way. Therefore, he concluded that we are at the center of our Galaxy (Figure 23-2). In the early 1900s, the Dutch astronomer Jacobus Kapteyn came to essentially the same conclusion by analyzing the brightness and proper motions of a large number of stars. According to Kapteyn, the Milky Way is about 17 kpc (17 kiloparsecs = 17,000 parsecs, or 55,000 light-years) in diameter, with the Sun near its center.

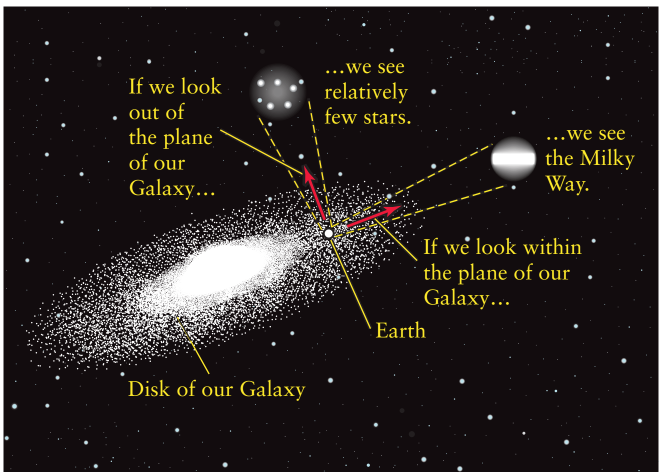
The Problem: Interstellar Extinction

Both Herschel and Kapteyn were wrong about the Sun being at the center of our Galaxy. The reason for their mistake was discerned in 1930 by Robert J. Trumpler of Lick Observatory. While studying star clusters, Trumpler discovered that the more remote clusters appear unusually dim—more so than would be expected from their distances alone. As a result, Trumpler concluded that interstellar space must not be a perfect vacuum: It must contain dust that absorbs or scatters light from distant stars.

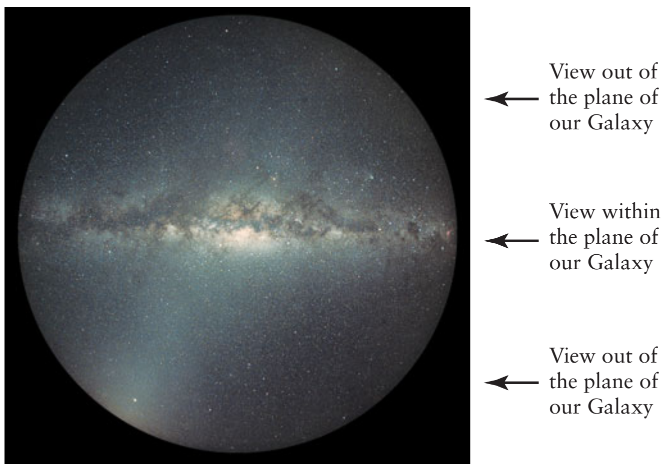
Like the stars themselves, interstellar dust is concentrated in the plane of the Galaxy (see Section 18-2). As a result, it obscures our view within the plane and makes distant objects appear dim, an effect called **interstellar extinction.** Great patches of interstellar dust are clearly visible in wide-angle photographs such as the one that opens this chapter. Thanks to interstellar extinction, Herschel and Kapteyn were actually seeing only the nearest stars in the Galaxy. Hence, they had no idea of either the enormous size of the Galaxy or of the vast number of stars concentrated around the galactic center.

ANALOGY Herschel and Kapteyn faced much the same dilemma as a lost motorist on a foggy night. Unable to see more than a city block in any direction, he would have a hard time deciding what part of town he was in. If the fog layer were relatively shallow, however, our motorist would be able to see the lights from tall buildings that extend above the fog, and in that way he could determine his location (Figure 23-3*a*).

The same principle applies to our Galaxy. While interstellar dust in the plane of our Galaxy hides the sky covered by the Milky Way, we have an almost unobscured view out of the plane (that is, to either side of the Milky Way). To find our location in



(a)



(b) R I V U X G

Figure 23-1

Our View of the Milky Way (a) The Milky Way Galaxy is a disk-shaped collection of stars. When we look out at the night sky in the plane of the disk, the stars appear as a band of light that stretches all the way around the sky. When we look perpendicular to the plane of the Galaxy, we see only those relatively few stars that lie between us and the "top" or "bottom" of the disk. **(b)** This wide-angle photograph shows a 180° view of the Milky Way centered on the constellation Sagittarius (compare with the photograph that opens this chapter). The dark streaks across the Milky Way are due to interstellar dust in the plane of our Galaxy. (Courtesy of Dennis di Cicco)

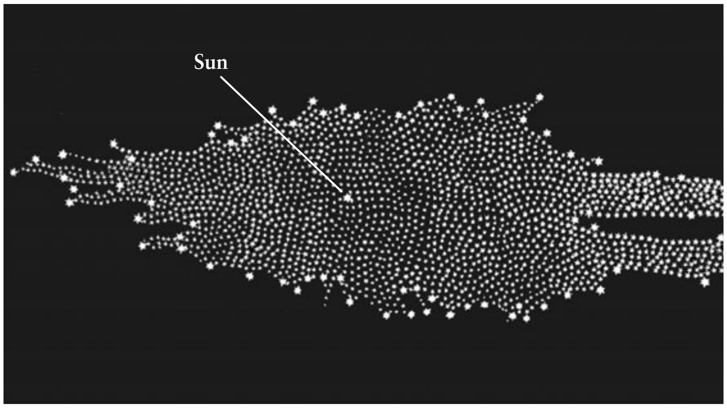


Figure 23-2

Herschel's Map of Our Galaxy In a paper published in 1785, the English astronomer William Herschel presented this map of the Milky Way Galaxy. He determined the Galaxy's shape by counting the numbers of stars in various parts of the sky. Herschel's conclusions were flawed because interstellar dust blocked his view of distant stars, leading him to the erroneous idea that the Sun is at the center of the Galaxy. (Yerkes Observatory)

the Galaxy, we need to locate bright objects that are part of the Galaxy but lie outside its plane in unobscured regions of the sky.

The Breakthrough: Globular Clusters and Variable Stars

WEB LINK 23.2

Fortunately, bright objects of the sort we need do in fact exist. They are the **globular clusters,** a class of star clusters associated with the Galaxy but which lie outside its plane (Figure 23-3*b*). As we saw in Section 19-4, a typical globular cluster is a spherical distribution of roughly 10^6 stars packed in a volume only a few hundred light-years across (see Figure 19-12).

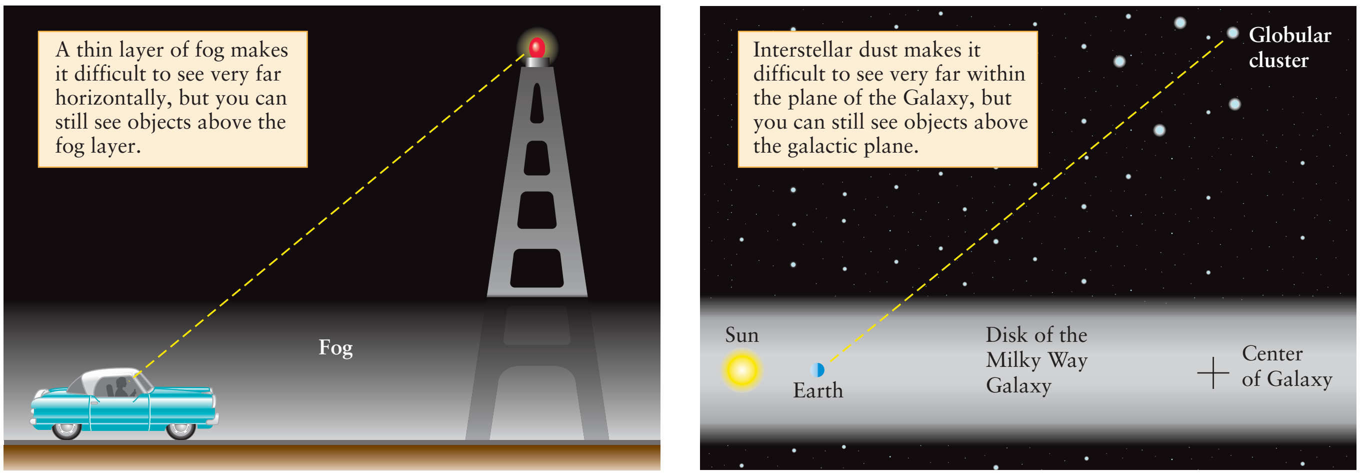
However, to use globular clusters to determine our location in the Galaxy, we must first determine the distances from the Earth to these clusters. (Think again of our lost motorist—glimpsing the lights of a skyscraper through the fog may be useful to the motorist, but only if he can tell how far away the skyscraper is.) Pulsating variable stars in globular clusters provide the distances, giving astronomers the key to the dimensions of our Galaxy.

Observations of pulsating variable stars revealed the immense size of the Milky Way

In 1912, the American astronomer Henrietta Leavitt reported her important discovery of the period-luminosity relation for Cepheid variables. As we saw in Section 19-6, Cepheid variables are pulsating stars that vary periodically in brightness (see Figure 19-18). Leavitt studied numerous Cepheids in the Small Magellanic Cloud (a small galaxy near the Milky Way) and found their periods to be directly related to their average luminosities. Figure 23-4 shows that the longer a Cepheid's period, the greater its luminosity.

The period-luminosity law is an important tool in astronomy because it can be used to determine distances. For example, suppose you find a Cepheid variable in the sky. By measuring its period and using a graph like Figure 23-4, you can determine the star's average luminosity. Knowing the star's average luminosity, you can find out how far away the star must be in order to give the observed brightness. (Box 17-2 explains how this is done.)

Shortly after Leavitt's discovery of the period-luminosity law, Harlow Shapley, a young astronomer at the Mount Wilson Observatory in California, began studying a family of pulsating stars closely related to Cepheid variables called **RR Lyrae variables.** The light curve of an RR Lyrae variable is similar to that of a



(a) Determining your position in the fog

(b) Determining your position in the Galaxy

Figure 23-3

Finding the Center of the Galaxy **(a)** A motorist lost on a foggy night can determine his location by looking for tall buildings that extend above the fog. **(b)** In the same way, astronomers determine our location in the Galaxy by observing globular clusters that are part of the Galaxy but lie outside the obscuring material in the galactic disk. The globular clusters form a spherical halo centered on the center of the Galaxy.

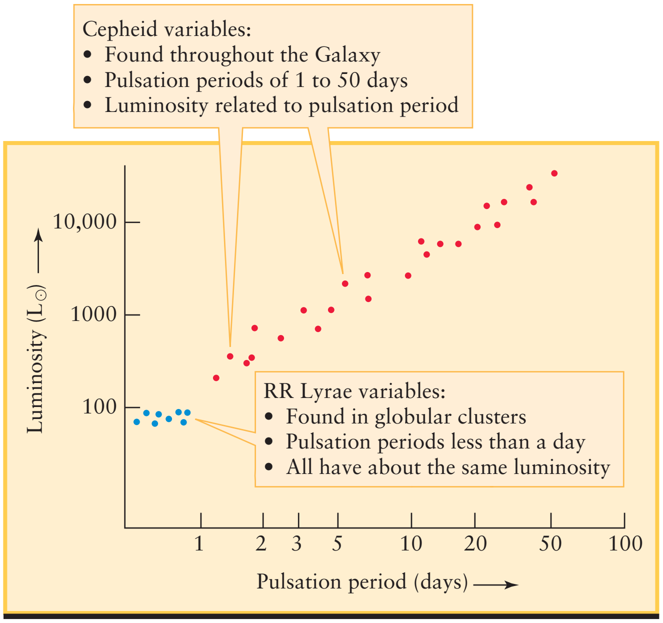
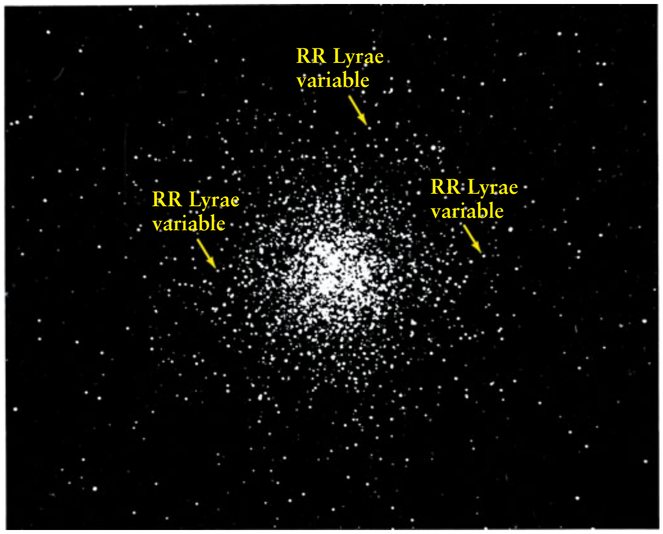


Figure 23-4

Period and Luminosity for Cepheid and RR Lyrae Variables This graph shows the relationship between period and luminosity for Cepheid variables and RR Lyrae variables. Cepheids come in a broad range of luminosities: The more luminous the Cepheid, the longer its pulsation period. By contrast, RR Lyrae variables are horizontal branch stars that all have roughly the same average luminosity of about 100 $L_\odot$.



WEB LINK 23.3 VIDEO 23.1 **Figure 23-5** R I V U X G

RR Lyrae Variables in a Globular Cluster The arrows point to three RR Lyrae variables in the globular cluster M55, located in the constellation Sagittarius. From the average apparent brightness (as seen in this photograph) and average luminosity (known to be roughly 100 $L_\odot$) of these variable stars, astronomers have deduced that the distance to M55 is 6500 pc (20,000 ly). (Harvard Observatory)

Cepheid, but RR Lyrae variables have shorter pulsation periods and lower peak luminosities (see Figure 23-4).

The tremendous importance of RR Lyrae variables is that they are commonly found in globular clusters (Figure 23-5). By using the period-luminosity relationship for these stars, Shapley was able to determine the distances to the 93 globular clusters then known. He found that some of them were more than 100,000 light-years from Earth. The large values of these distances immediately suggested that the Galaxy was much larger than Herschel or Kapteyn had thought.

Another striking property of globular clusters is how they are distributed across the sky. Ordinary stars and open clusters of stars (see Section 18-6, especially Figures 18-18 and 18-19) are rather uniformly spread along the Milky Way. However, the majority of the 93 globular clusters that Shapley studied are located in one-half of the sky, widely scattered around the portion of the Milky Way that is in the constellation Sagittarius.

From the directions to the globular clusters and their distances from us, Shapley mapped out the three-dimensional distribution of these clusters in space. In 1920 he concluded that the globular clusters form a huge spherical distribution centered not on the Earth but rather about a point in the Milky Way several kiloparsecs away in the direction of Sagittarius (see Figure 23-3*b*). This point, reasoned Shapley, must coincide with the center of our Galaxy, because of gravitational forces between the disk of the Galaxy and the "halo" of globular clusters. Therefore, by locating the center of the distribution of globular clusters, Shapley was in effect measuring the location of the galactic center.

Modern-Day Measurements

Since Shapley's pioneering observations, many astronomers have measured the distance from the Sun to the **galactic nucleus**, the center of our Galaxy. Shapley's estimate of this distance was too large by about a factor of 2, because he did not take into account the effects of interstellar extinction (which were not well understood at the time). Today, the generally accepted distance to the galactic nucleus is about 8 kpc (26,000 ly); the actual distance could be greater or less than that value by about 1 kpc (3300 ly).

Just as Copernicus and Galileo showed that the Earth was not at the center of the solar system, Shapley and his successors showed that the solar system lies nowhere near the center of the Galaxy. We see that the Earth indeed occupies no special position in the universe.

23-2 Observations at nonvisible wavelengths reveal the shape of the Galaxy

At visible wavelengths, light suffers so much interstellar extinction that the galactic nucleus is totally obscured from view. But the amount of interstellar extinction is roughly inversely proportional to wavelength. In other words, the longer the wavelength, the farther radiation can travel through interstellar dust without being scattered or absorbed. As a result, we can see farther into the plane of the Milky Way at infrared wavelengths than at visible wavelengths, and radio waves can traverse the Galaxy freely. For this reason, telescopes sensitive to these nonvisible wavelengths are important tools for studying the structure of our Galaxy.

WEB LINK 23.4

Exploring the Milky Way in the Infrared

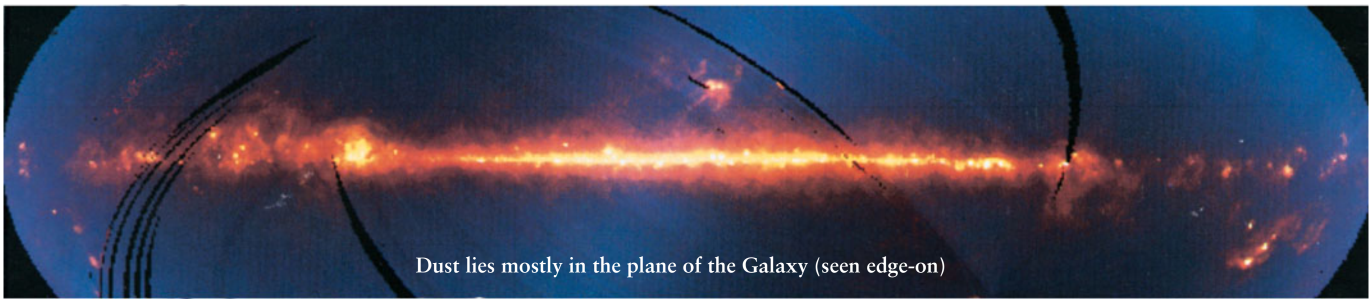
Infrared light is particularly useful for tracing the location of interstellar dust in the Galaxy. Starlight warms the dust grains to temperatures in the range of 10 to 90 K; thus, in accordance with Wien's law (see Section 5-4), the dust emits radiation predominately at wavelengths from about 30 to 300 μm. These are called **far-infrared** wavelengths, because they lie in the part of the infrared spectrum most different in wavelength from visible light (see Figure 5-7). At these wavelengths, interstellar dust radiates more strongly than stars, so a far-infrared view of the sky is principally a view of where the dust is. In 1983 the Infrared Astronomical Satellite (IRAS) scanned the sky with a 60-cm reflecting telescope at far-infrared wavelengths, giving the panoramic view of the Milky Way's dust shown in Figure 23-6*a*.

Our Galaxy's dust and stars—including the Sun—lie mostly in a relatively thin disk

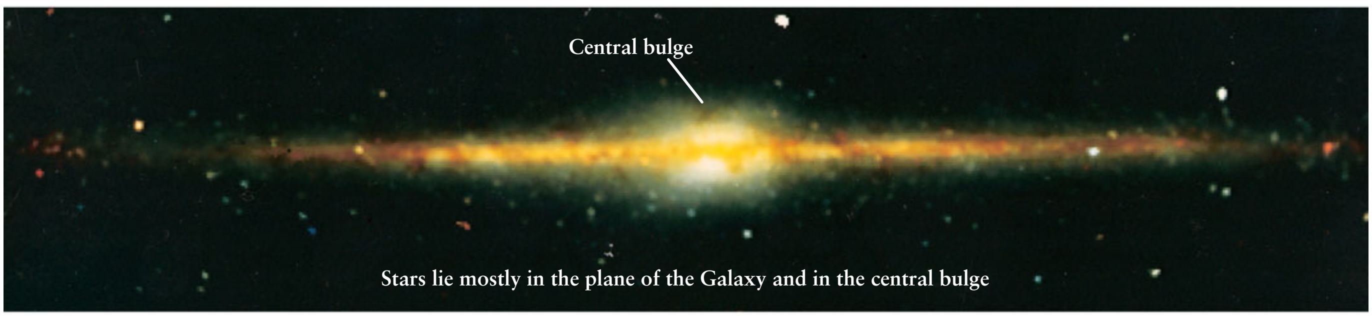
In 1990 an instrument on the Cosmic Background Explorer (COBE) satellite scanned the sky at **near-infrared** wavelengths, that is, relatively short wavelengths closer to the visible spectrum. Figure 23-6*b* shows the resulting near-infrared view of the plane of the Milky Way. At near-infrared wavelengths, interstellar dust does not emit very much light. Hence, the light sources in Figure 23-6*b* are stars, which do emit strongly in the near-infrared (especially the cool stars, such as red giants and supergiants). Because interstellar dust causes little interstellar extinction in the near-infrared, many of the stars whose light is recorded in Figure 23-6*b* lie deep within the Milky Way.

Observations such as those shown in Figure 23-6, along with the known distance to the center of the Galaxy, have helped astronomers to establish the dimensions of the Galaxy. The **disk** of our Galaxy is about 50 kpc (160,000 ly) in diameter and about 0.6 kpc (2000 ly) thick, as shown in Figure 23-7. The center of the Galaxy is surrounded by a distribution of stars, called the **central bulge,** which is about 2 kpc (6500 ly) in diameter. This central bulge is clearly visible in Figure 23-6*b*. The spherical distribution of globular clusters traces the **halo** of the Galaxy.

This structure is not unique to our Milky Way Galaxy. Figure 23-8 shows another galaxy whose dust and stars lie in a disk and that has a central bulge of stars, just like the Milky Way.



(a) Infrared emission from dust at wavelengths of 25, 60, and 100 μm



(b) Infrared emission from dust at wavelengths of 1.2, 2.2, and 3.4 μm

WEB LINK 23.5

Figure 23-6 R **I** V U X G

The Infrared Milky Way **(a)** This view was constructed from observations made at far-infrared wavelengths by the IRAS spacecraft. Interstellar dust, which is mostly confined to the plane of the Galaxy, is the principal source of radiation in this wavelength range. **(b)** Observing at near-infrared wavelengths, as in this composite of COBE data, allows us to see much farther through interstellar dust than we can at visible wavelengths. Light in this wavelength range comes mostly from stars in the plane of the Galaxy and in the bulge at the Galaxy's center. (NASA)

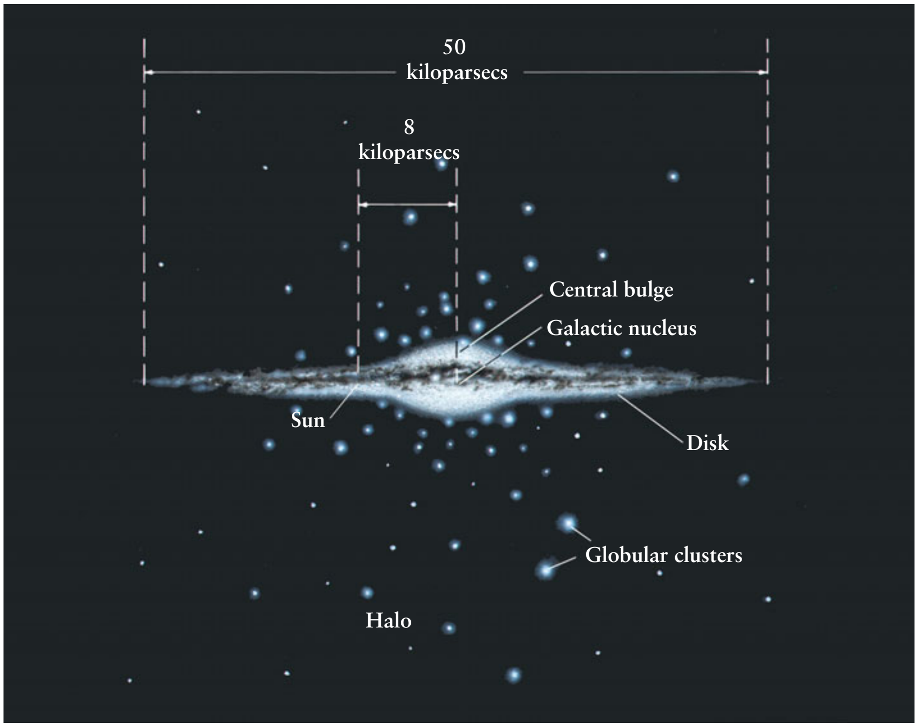


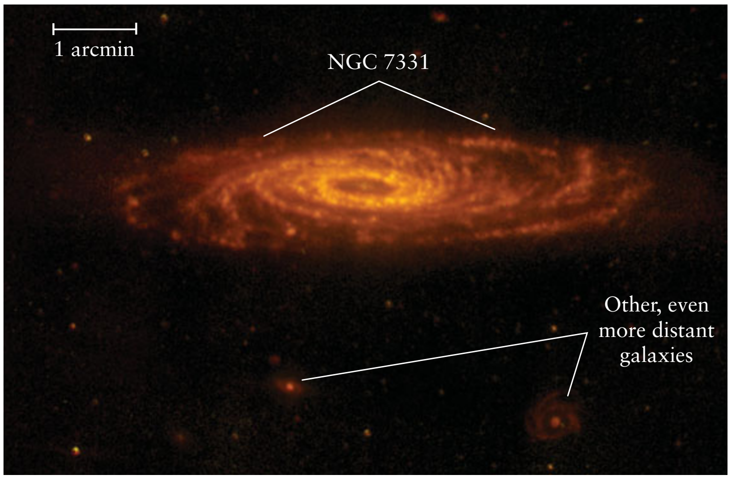
Figure 23-7

Our Galaxy (Schematic Edge-on View) There are three major components of our Galaxy: a disk, a central bulge, and a halo. The disk contains gas and dust along with metal-rich (Population I) stars. The halo is composed almost exclusively of old, metal-poor (Population II) stars. The central bulge is a mixture of Population I and Population II stars.

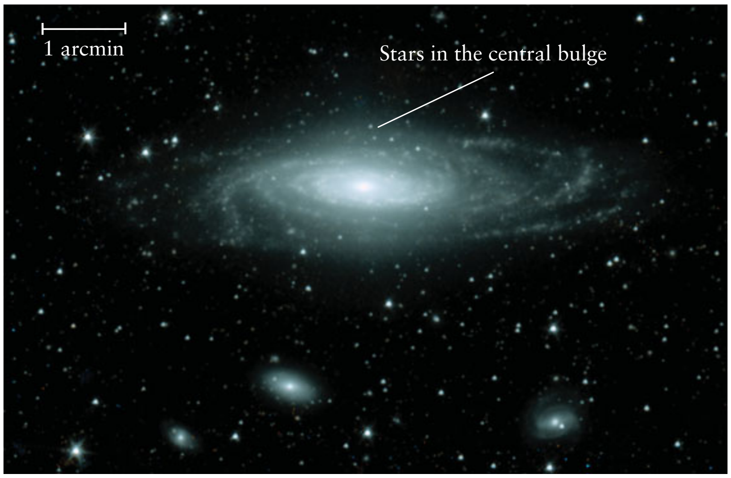
In the same way that our Sun is a rather ordinary member of the stellar community that makes up the Milky Way, the Milky Way turns out to be a rather common variety of galaxy.

The Milky Way's Distinct Stellar Populations

It is estimated that our Galaxy contains about 200 billion (2×10^{11}) stars. Remarkably, different kinds of stars are found in the various components of the Galaxy. The globular clusters in the halo are composed of old, metal-poor, Population II stars (see Section 19-5). Although these clusters are conspicuous, they contain only about 1% of the total number of stars in the halo. Most halo stars are single Population II stars in isolation, called **high-velocity stars** because of their high speed relative to the Sun. These ancient stars orbit the Galaxy along paths tilted at random



(a) Infrared emission from dust in NGC 7331 at 5.8 and 8.0 μm



(b) Infrared emission from stars in NGC 7331 at 3.6 and 4.5 μm

WEB LINK 23.6

Figure 23-8 R I V U X G

NGC 7331: A Near-Twin of the Milky Way If we could view our Galaxy from a great distance, it would probably look like this galaxy in the constellation Pegasus. As in Figure 23-6, the far-infrared image **(a)** reveals the presence of dust in the galaxy's plane, while the near-infrared image **(b)** shows the distribution of stars. These images of NGC 7331, which is about 15 million pc (50 million ly) from Earth, were made with the Spitzer Space Telescope (from the Spitzer Space Telescope (see Section 6-7, especially Figure 6-26). (NASA; JPL-Caltech; M. Regan (STScI); and the SINGS Team)

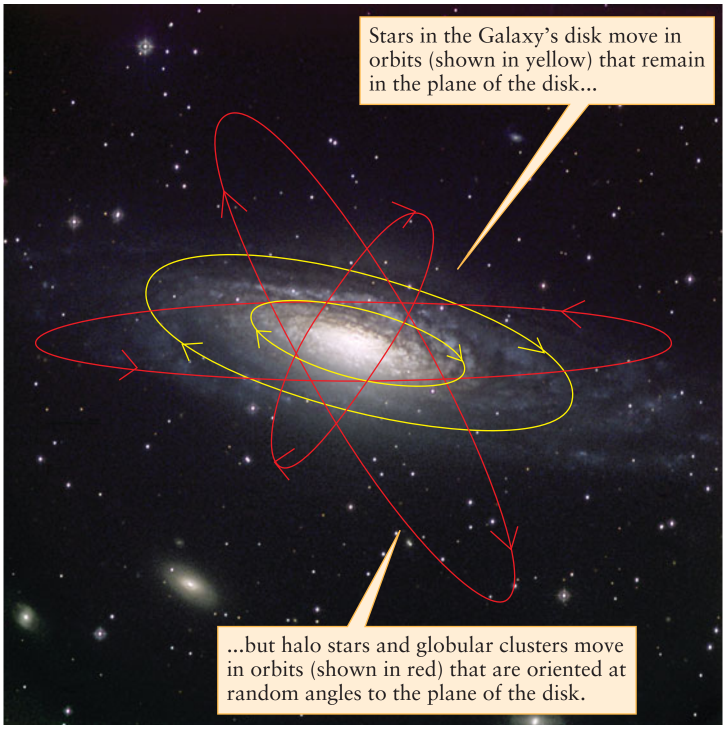


Figure 23-9 R I V U X G

Star Orbits in the Milky Way The different populations of stars in our Galaxy travel along different sorts of orbits. The galaxy in this visible-light image is the Milky Way's near-twin NGC 7331, the same galaxy shown at infrared wavelengths in Figure 23-8. (Daniel Bramich (ING) and Nik Szymanek)

angles to the disk of the Milky Way, as do the globular clusters. By contrast, stars in the disk travel along orbits that remain in the disk (Figure 23-9).

The stars in the disk are mostly young, metal-rich, Population I stars like the Sun. The disk of a galaxy like the Milky Way appears bluish because its light is dominated by radiation from hot O and B main-sequence stars. Such stars have very short main-sequence lifetimes (see Section 19-1, especially Table 19-1), so they must be quite young by astronomical standards. Hence, their presence shows that there must be active star formation in the galactic disk. By contrast, no O or B stars are present in the halo, which implies that star formation ceased there long ago.

The central bulge contains both Population I stars and metal-poor Population II stars. Since Population II stars are thought to have formed in the history of the universe, this suggests that some of the stars in the bulge are quite ancient while others were created more recently. The central bulge looks yellowish or reddish because it contains many red giants and red supergiants (see Figure 1-7), but does *not* contain luminous, short-lived, blue O or B stars. Hence, there cannot be ongoing star formation in the central bulge. The same is true for other galaxies whose structure is similar to that of the Milky Way (Figure 23-10).

Why are there such different populations of stars in the halo, disk, and central bulge? Why has star formation stopped in some regions of the Galaxy but continues in other regions? The answers to these questions are related to the way that stars, as well as the gas and dust from which stars form, move within the Galaxy.

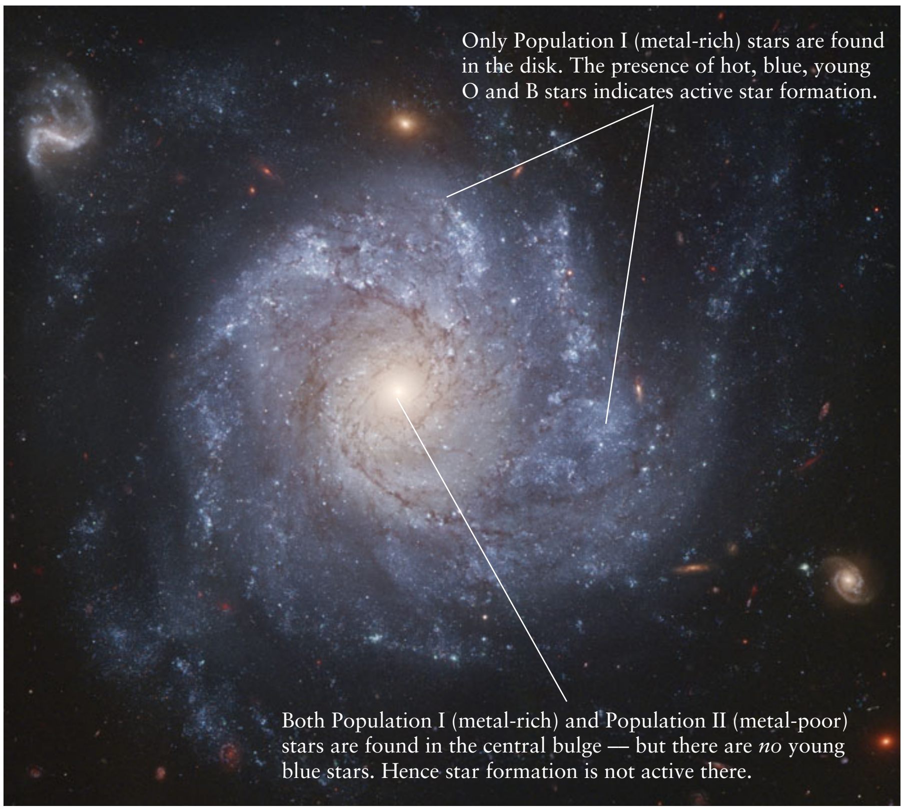


Figure 23-10 R I V U X G

Stellar Populations: Disk Versus Central Bulge The disk and central bulge of the Milky Way contain rather different populations of stars. The same is true for the galaxy NGC 1309, which has a similar structure to the Milky Way Galaxy and happens to be oriented face-on to us. NGC 1309 lies about 30 million pc (100 million ly) from us in the constellation Eridanus. (NASA; The Hubble Heritage Team; and A. Riess, STScI)

23-3 Observations of cold hydrogen clouds and star-forming regions reveal that our Galaxy has spiral arms

The galaxies shown in Figure 23-8 and Figure 23-10 both have **spiral arms,** spiral-shaped concentrations of gas and dust that extend outward from the center in a shape reminiscent of a pinwheel. This would lead us to suspect that our own Milky Way Galaxy also has spiral arms. However, because interstellar dust obscures our visible-light view in the plane of our Galaxy, a detailed understanding of the structure of the galactic disk had to wait until the development of radio astronomy. Thanks to their long wavelengths, radio waves can penetrate the interstellar medium even more easily than infrared light and can travel without being scattered or absorbed. As we shall see in this section, both radio and optical observations reveal that our Galaxy does indeed have spiral arms.

Mapping Hydrogen in the Milky Way

LOOKING DEEPER 23.1

Hydrogen is by far the most abundant element in the universe (see Figure 8-4 in Section 8-2). Hence, by looking for concentrations of hydrogen gas, we should be able to detect important clues about the distribution of matter in our Galaxy. Unfortunately, ordinary visible-light telescopes are of little use in this quest. This is so because hydrogen atoms can only emit visible light if they are first excited to high energy levels (see Section 5-8, and especially Figure 5-24). This is quite unlikely to occur in the cold depths of interstellar space. Furthermore, even if there are some hydrogen atoms that glow strongly at visible wavelengths, interstellar extinction due to dust (see Section 23-1) would make it impossible to see this glow from distant parts of the Galaxy.

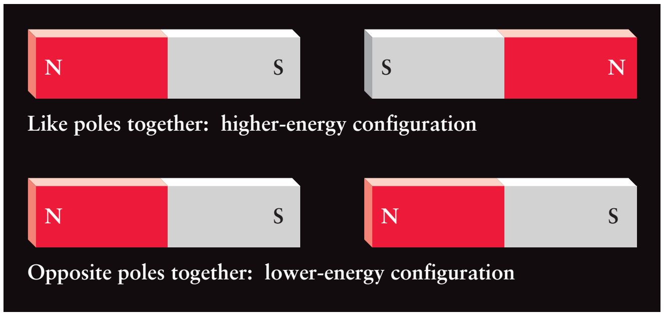
The magnetic nature of subatomic particles makes it possible to detect cold clouds of gas strewn throughout the galactic disk

What makes it possible to map out the distribution of hydrogen in our Galaxy is that even cold hydrogen clouds emit *radio* waves. As we saw in Section 23-2, radio waves can easily penetrate the interstellar medium, so we can detect the radio emission from such cold clouds no matter where in the Galaxy they lie. The hydrogen in these clouds is neutral—that is, not ionized—and is called **H I.** (This distinguishes it from ionized hydrogen, which is designated H II.) To understand how H I clouds can emit radio waves, we must probe a bit more deeply into the structure of protons and electrons, the particles of which hydrogen atoms are made.

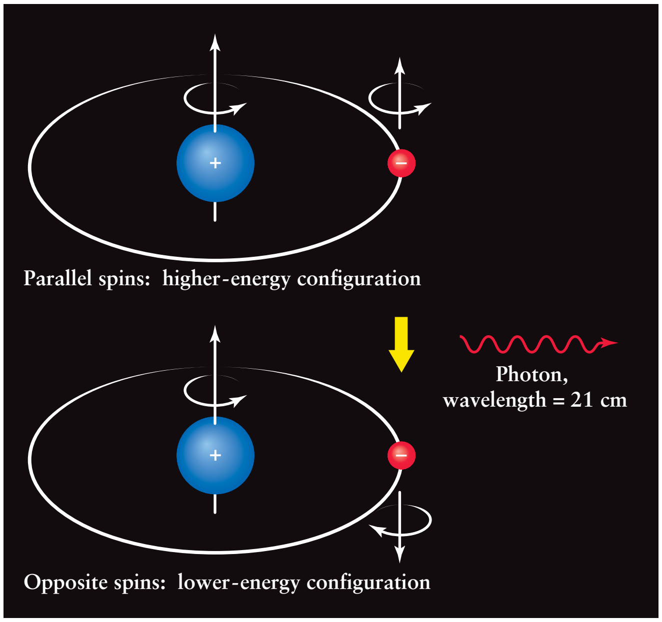
In addition to having mass and charge, particles such as protons and electrons possess a tiny amount of angular momentum (that is, rotational motion) commonly called **spin.** Very roughly, you can visualize a proton or electron as a tiny, electrically charged sphere that spins on its axis. Because electric charges in motion generate magnetic fields, a proton or electron behaves like a tiny magnet with a north pole and a south pole (Figure 23-11).

If you have ever played with magnets, you know that two magnets attract when the north pole of one magnet is next to the south pole of the other and repel when two like poles (both north or both south) are next to each other (Figure 23-11*a*). In other words, the energy of the two magnets is least when opposite poles are together and highest when like poles are together. Hence, as shown in Figure 23-11*b*, the energy of a hydrogen atom is slightly different depending on whether the spins of the proton and electron are in the same direction or opposite directions. (According to the laws of quantum mechanics, these are the only two possibilities; the spins cannot be at random angles.)

If the spin of the electron changes its orientation from the higher-energy configuration to the lower-energy one—called a



(a) The magnetic energy of two bar magnets depends on their relative orientation



(b) The magnetic energy of a proton and electron depends on their relative spin orientation

Figure 23-11

Magnetic Interactions in the Hydrogen Atom (a) The energy of a pair of magnets is high when their north poles or their south poles are near each other, and low when they have opposite poles near each other. **(b)** Thanks to their spin, electrons and protons are both tiny magnets. When the electron flips from the higher-energy configuration (with its spin in the same direction as the proton's spin) to the lower-energy configuration (with its spin opposite to the proton's spin), the atom loses a tiny amount of energy and emits a radio photon with a wavelength of 21 cm.

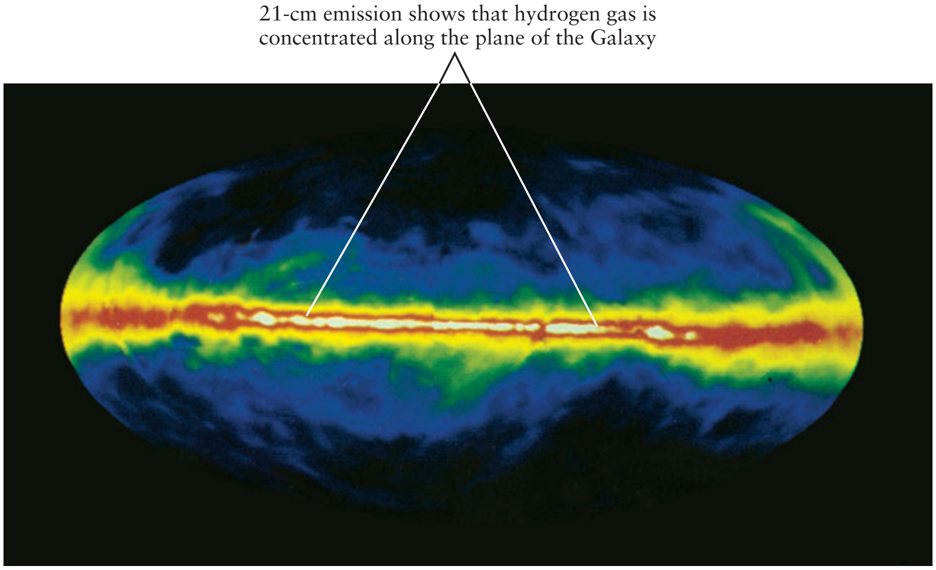


Figure 23-12 R I V U X G

The Sky at 21 Centimeters This image was made by mapping the sky with radio telescopes tuned to the 21-cm wavelength emitted by neutral interstellar hydrogen (H I). The entire sky has been mapped onto an oval, and the plane of the Galaxy extends horizontally across the image as in Figure 23-6. Black and blue represent the weakest emission, and red and white the strongest. (Courtesy of C. Jones and W. Forman, Harvard-Smithsonian Center for Astrophysics)

spin-flip transition—a photon is emitted. The energy difference between the two spin configurations is very small, only about 10^{-6} as great as those between different electron orbits (see Figure 5-24). Therefore, the photon emitted in a spin-flip transition between these configurations has only a small energy, and thus its wavelength is a relatively long 21 cm—a radio wavelength.

The spin-flip transition in neutral hydrogen was first predicted in 1944 by the Dutch astronomer Hendrik van de Hulst. His calculations suggested that it should be possible to detect the **21-cm radio emission** from interstellar hydrogen, although a very sensitive radio telescope would be required. In 1951, Harold Ewen and Edward Purcell at Harvard University first succeeded in detecting this faint emission from hydrogen between the stars.

Figure 23-12 shows the results of a more recent 21-cm survey of the entire sky. Neutral hydrogen gas (H I) in the plane of the Milky Way stands out prominently as a bright band across the middle of this image.

LOOKING DEEPER 23.2 The distribution of gas in the Milky Way is not uniform but is actually quite frothy. In fact, our Sun lies near the edge of an irregularly shaped region within which the interstellar medium is very thin but at very high temperatures (about 10^6 K). This region, called the **Local Bubble,** is several hundred parsecs across. The Local Bubble may have been carved out by a supernova that exploded nearby some 300,000 years ago.

Remarkably, spin-flip transitions are used not only to map our Galaxy but also to map the internal structure of the human body. Box 23-1 on page 614 discusses this application, called magnetic resonance imaging.

Detecting Our Galaxy's Spiral Arms

The detection of 21-cm radio radiation was a major breakthrough that permitted astronomers to reveal the presence of spiral arms in the galactic disk. Figure 23-13 shows how this was done. Suppose that you aim a radio telescope along a particular line of sight across the Galaxy. Your radio receiver, located at *S* (the position of the solar system), picks up 21-cm emission from H I clouds at points 1, 2, 3, and 4. However, the radio waves from these various clouds are Doppler shifted by slightly different amounts, because the clouds are moving at different speeds as they travel with the rotating Galaxy.

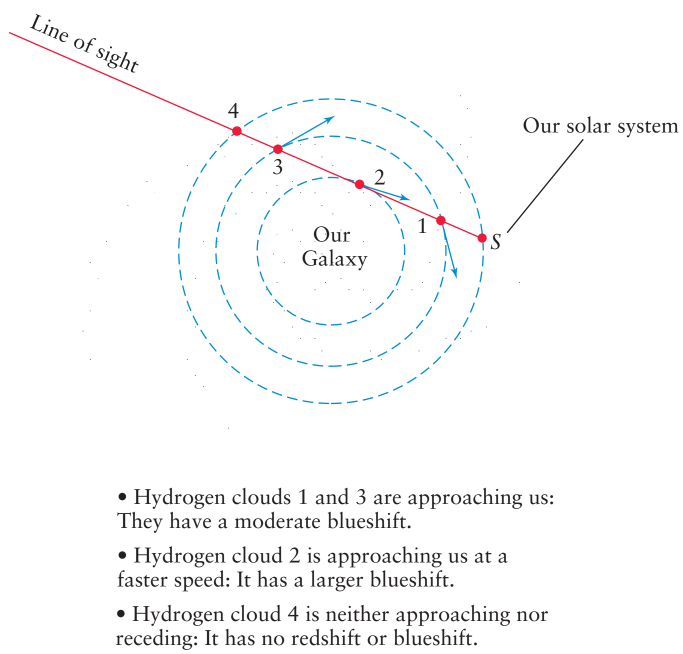


Figure 23-13

A Technique for Mapping Our Galaxy If we look within the plane of our Galaxy from our position at *S*, hydrogen clouds at different locations (shown as 1, 2, 3, and 4) along our line of sight are moving at slightly different speeds relative to us. As a result, radio waves from these various gas clouds are subjected to slightly different Doppler shifts. This permits radio astronomers to sort out the gas clouds and thus map the Galaxy.

BOX 23-1 The Heavens on the Earth

Spin-Flip Transitions in Medicine

WEB LINK 23.7

Thanks to their spin, protons and electrons act like microscopic bar magnets. In a hydrogen atom, the interaction between the magnetism of the electron and that of the proton gives rise to the 21-cm radio emission. But these particles can also interact with outside magnetic fields, such as that produced by a large electromagnet. This is the physical principle behind **magnetic resonance imaging (MRI),** an important diagnostic tool of modern medicine.

Much of the human body is made of water. Every water molecule has two hydrogen atoms, each of which has a nucleus made of a single proton. If a person's body is placed in a strong magnetic field, the spins of the protons in the hydrogen atoms of their body can either be in the same direction as the field ("aligned") or in the direction opposite to the field ("opposed"). The aligned orientation has lower energy, and therefore the majority of protons end up with their spins in this orientation. But if a radio wave of just the right wavelength is now sent through the person's body, an aligned proton can absorb a radio photon and flip its spin into the higher-energy, opposed orientation. How much of the radio wave is absorbed depends on the number of protons in the body, which in turns depends on how much water (and, thus, how much water-containing tissue) is in the body.

In magnetic resonance imaging, a magnetic field is used whose strength varies from place to place. The difference in energy between the opposed and aligned orientations of a proton depends on the strength of the magnetic field, so radio waves will only be absorbed at places where this energy difference is equal to the energy of a radio photon. (This equality is called *resonance,* which is how magnetic resonance imaging gets its name.) By varying the magnetic field strength over the body and the wavelength of the radio waves, and by measuring how much of the radio wave is absorbed by different parts of the body, it is possible to map out the body's tissues. The accompanying false-color image shows such a map of a patient's head.

R I V U X G

(Scott Camazine, Photo Researchers)

Unlike X-ray images, which show only the densest parts of the body, such as bones and teeth, magnetic resonance imaging can be used to view less dense (but water-containing) soft tissue. Just as the 21-cm radio emission has given astronomers a clear view of what were hidden regions of our Galaxy, magnetic resonance imaging allows modern medicine to see otherwise invisible parts of the human body.

It is important to remember that the Doppler shift reveals only motion along the line of sight (review Figure 5-26). In Figure 23-13, cloud 2 has the highest speed along our line of sight, because it is moving directly toward us. Consequently, the radio waves from cloud 2 exhibit a larger Doppler shift than those from the other three clouds along our line of sight. Because clouds 1 and 3 are at the same distance from the galactic center, they have the same orbital speed. The fraction of their velocity parallel to our line of sight is also the same, so their radio waves exhibit the same Doppler shift, which is less than the Doppler shift of cloud 2. Finally, cloud 4 is the same distance from the galactic center as the Sun. This cloud is thus orbiting the Galaxy at the same speed as the Sun, resulting in no net motion along the line of sight. Radio waves from cloud 4, as well as from hydrogen gas near the Sun, are not Doppler shifted at all.

These various Doppler shifts cause radio waves from gases in different parts of the Galaxy to arrive at our radio telescopes with wavelengths slightly different from 21 cm. It is therefore possible to sort out the various gas clouds and thus produce a map of the Galaxy like that shown in Figure 23-14.

Figure 23-14 shows that neutral hydrogen gas is not spread uniformly around the disk of the Galaxy but is concentrated into numerous arched lanes. Similar features are seen in other galaxies beyond the Milky Way. As an example, the galaxy in Figure 23-15*a* has prominent spiral arms outlined by hot, luminous, blue main-sequence stars and the red emission nebulae (H II regions) found near many such stars. Stars of this sort are very short-lived, so these features indicate that spiral arms are sites of active, ongoing star formation. The 21-cm radio image of this same galaxy, shown in Figure 23-15*b*, shows that spiral arms are also regions where neutral hydrogen gas is concentrated, similar to the structures in our own Galaxy visible in Figure 23-14. This similarity is a strong indication that our Galaxy also has spiral arms.

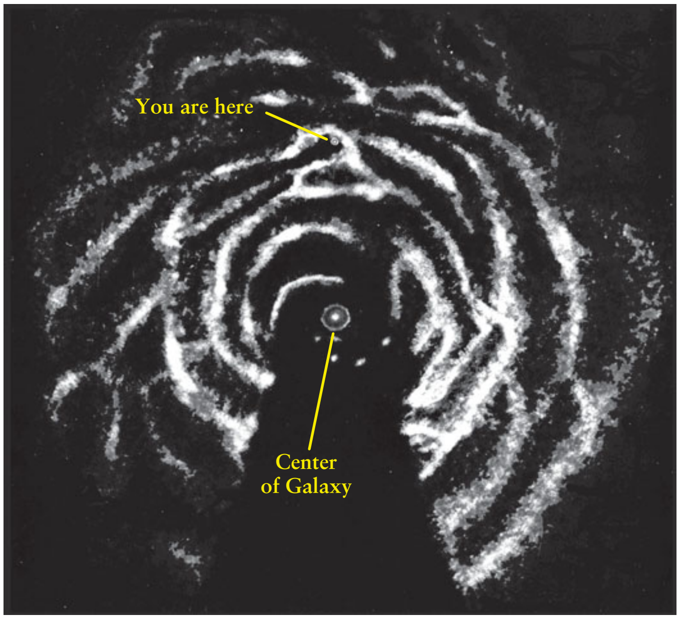


Figure 23-14 R I V U X G

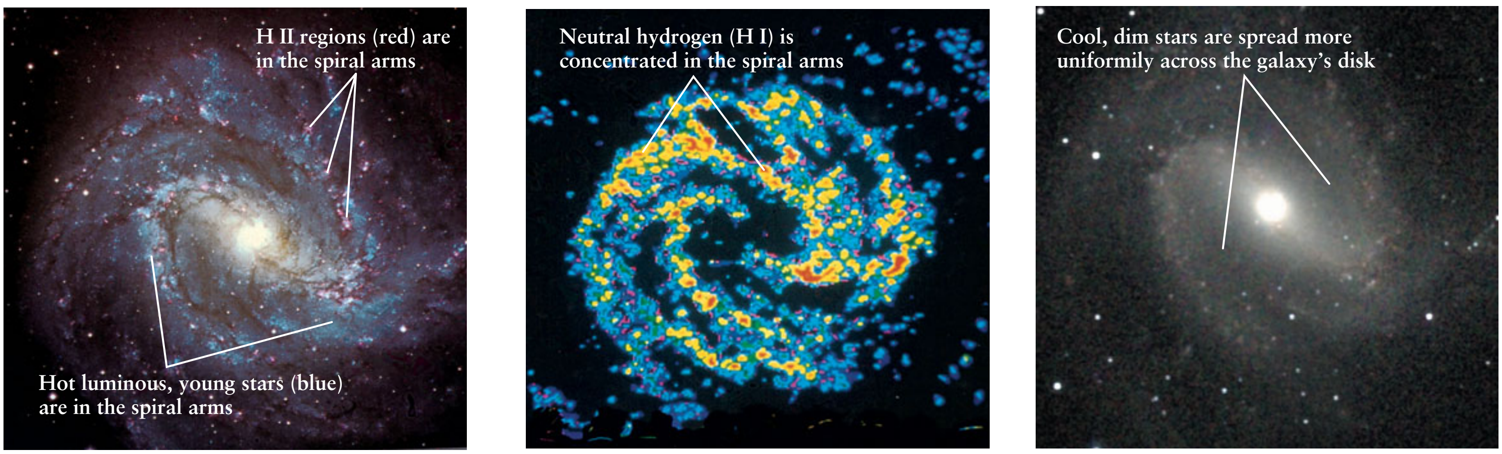
A Map of Neutral Hydrogen in Our Galaxy This map, constructed from radio-telescope surveys of 21-cm radiation, shows the distribution of hydrogen gas in a face-on view of our Galaxy. The map suggests a spiral structure. Details in the blank, wedge-shaped region at the bottom of the map are unknown. Gas in this part of the Galaxy is moving perpendicular to our line of sight and thus does not exhibit a detectable Doppler shift. (Courtesy of G. Westerhout)

CAUTION Photographs such as Figure 23-15*a* can lead to the impression that there are very few stars between the spiral arms of a galaxy. Nothing could be further from the truth! In fact, stars are distributed rather uniformly throughout the disk of a galaxy like the one in Figure 23-15*a;* the density of stars in the spiral arms is only about 5% higher than in the rest of the disk. The spiral arms stand out nonetheless because they are where hot, blue O and B stars are found. One such star is about 10^4 times more luminous than an average star in the disk, so the light from O and B stars completely dominates the visible appearance of a spiral galaxy. An infrared image such as Figure 23-15*c* gives a better impression of how stars of all kinds are distributed through a spiral galaxy's disk.

Mapping the Spiral Arms and the Central Bulge

Figure 23-15*a* suggests that we can confirm the presence of spiral structure in our own Galaxy by mapping the locations of star-forming regions. Such regions are marked by OB associations, H II regions, and molecular clouds (see Section 18-7). Unfortunately, the first two of these are best observed using visible light, and interstellar extinction limits the range of visual observations in the plane of the Galaxy to less than 3 kpc (10,000 ly) from the Earth. But there are enough OB associations and H II regions within this range to plot the spiral arms in the vicinity of the Sun.

Molecular clouds are easier to observe at great distances, because molecules of carbon monoxide (CO) in these clouds emit radio waves that are relatively unaffected by interstellar extinction. Hence, the positions of molecular clouds have been plotted even in remote regions of the Galaxy, as Figure 18-21 shows. (We saw in Section 18-7 that CO molecules in molecular clouds emit more strongly than the hydrogen atoms do, even though hydrogen is the principal constituent of these clouds.)

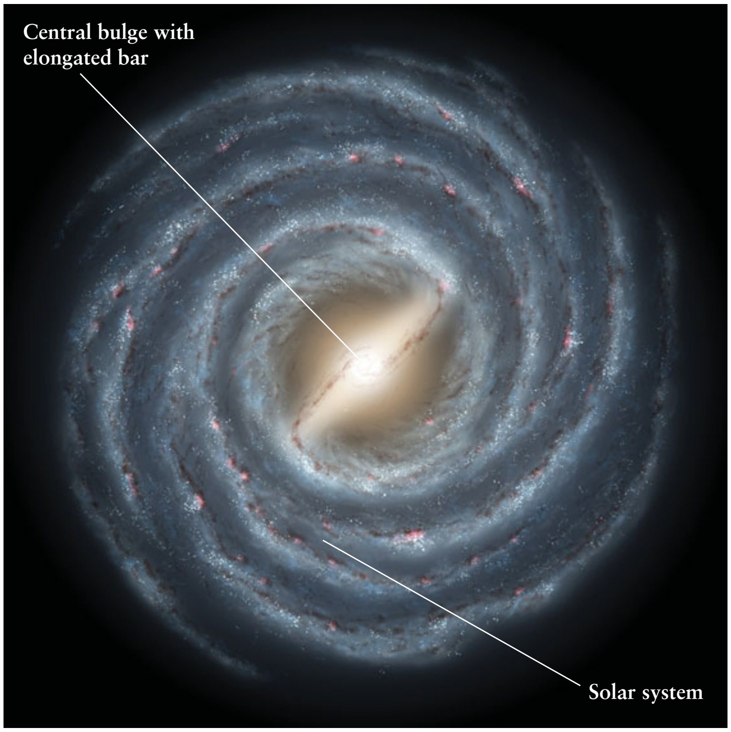


(a) Visible-light view of M83 R I V U X G
(b) 21-cm radio view of M83 R I V U X G
(c) Near-infrared view of M83 R I V U X G

WEB LINK 23.8

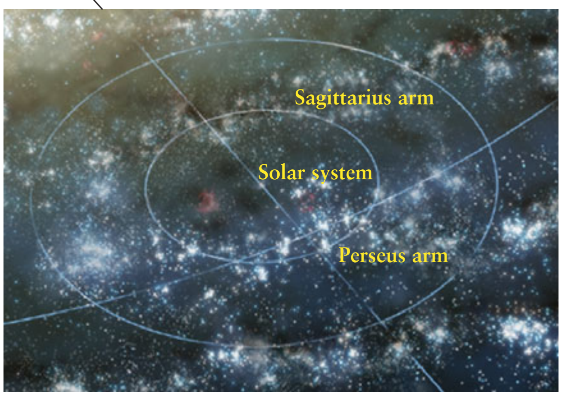
Figure 23-15

A Spiral Galaxy The galaxy M83 lies in the southern constellation Hydra about 5 million pc (15 million ly) from Earth. **(a)** This visible-light image clearly shows the spiral arms. The presence of young stars and H II regions indicates that star formation takes place in spiral arms. **(b)** This radio view at a wavelength of 21 cm shows the emission from neutral interstellar hydrogen gas (H I). Note that essentially the same pattern of spiral arms is traced out in this image as in the visible-light photograph. **(c)** M83 has a much smoother appearance in this near-infrared view. This shows that cooler stars, which emit strongly in the infrared, are spread more uniformly across the galaxy's disk. Note the elongated bar shape of the central bulge. (a: Anglo-Australian Observatory; b: VLA, NRAO; c: S. Van Dyk/IPAC)



(a) The structure of the Milky Way's disk

To the center of the Galaxy



(b) Closeup of the Sun's galactic neighborhood

Figure 23-16

Our Galaxy Seen Face-on: Artist's Impressions **(a)** The Galaxy's diameter is about 50,000 pc (160,000 ly), and our solar system is about 8000 pc (26,000 ly) from the galactic center. The elongated central bulge is about 8300 pc (27,000 ly) long and is oriented at approximately 45° to a line running from the solar system to the galactic center. **(b)** Our solar system is located between the Sagittarius and Perseus arms, two of the major spiral arms in the Milky Way. (a: NASA/JPL-Caltech/R. Hurt, SSC; b: National Geographic)

Taken together, all these observations demonstrate that our Galaxy has at least four major spiral arms as well as several short arm segments (Figure 23-16). The Sun is located just outside a relatively short arm segment called the Orion arm, which includes the Orion Nebula and neighboring sites of vigorous star formation in that constellation.

Two major spiral arms border either side of the Sun's position. The Sagittarius arm is on the side toward the galactic center. This is the arm you see on June and July nights when you look at the portion of the Milky Way stretching across Scorpius and Sagittarius, near the center of the upper photograph on the first page of this chapter. In December and January, when our nighttime view is directed away from the galactic center, we see the Perseus arm. The other major spiral arms cannot be seen at visible wavelengths thanks to the obscuring effects of dust.

Figure 23-16*a* also shows that the central bulge of the Milky Way is not spherical, but is elongated like a bar. This is unlike the galaxy NGC 7331 shown in Figure 23-8, but similar to the galaxy M83 depicted in Figure 23-15. The elongated shape of the central bulge had been suspected since the 1980s; this was confirmed in 2005 using the Spitzer Space Telescope, which was used to survey the infrared emissions from some 3 million stars in the central bulge. Thus the artist's impression shown in Figure 23-16*a* is based on observations using both radio wavelengths (for the spiral arms) and infrared wavelengths (for the central bulge). We will see in Section 23-5 that this elongated shape may play a crucial role in sustaining the Galaxy's spiral structure.

Why are the young stars, star-forming regions, and clouds of neutral hydrogen in our Galaxy all found predominantly in the spiral arms? To answer this question, we must understand why spiral arms exist at all. Spiral arms are essentially cosmic "traffic jams," places where matter piles up as it orbits around the center of the Galaxy. This orbital motion, which is essential to grasping the significance of spiral arms, is our next topic as we continue our exploration of the Galaxy.

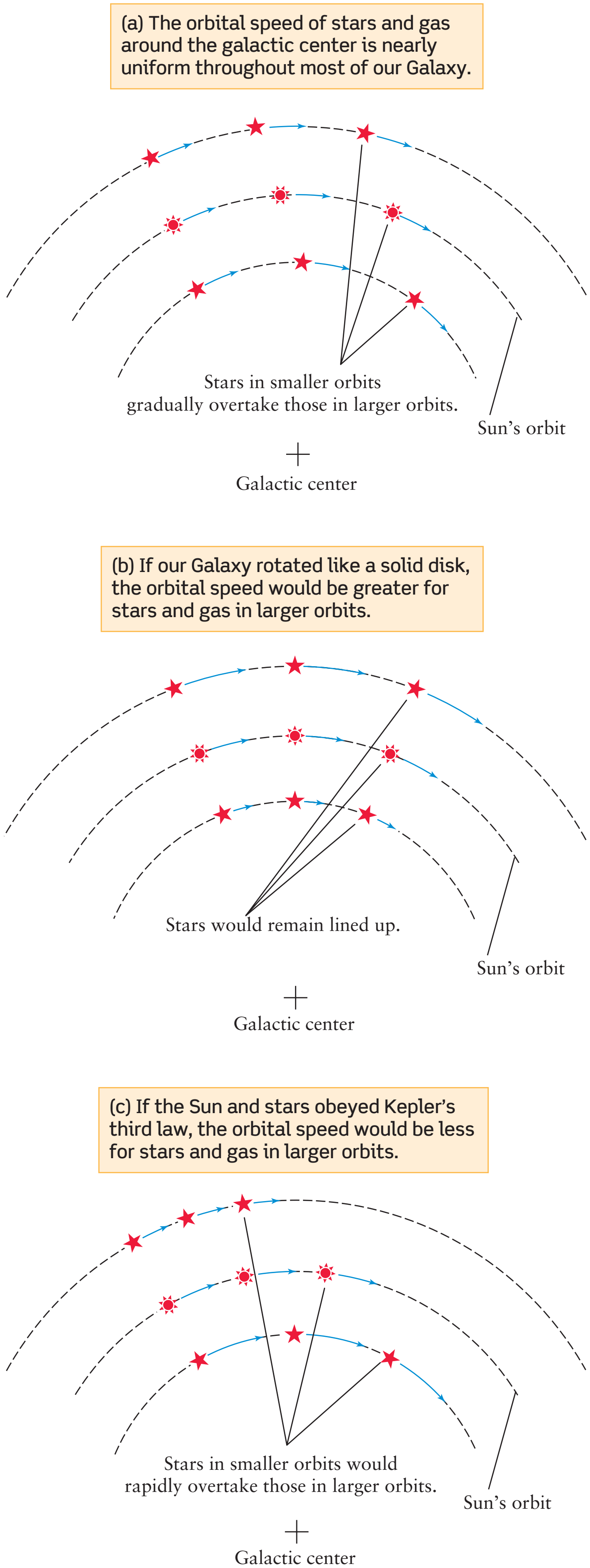
23-4 The rotation of our Galaxy reveals the presence of dark matter

The spiral arms in the disk of our Galaxy suggest that the disk rotates. This means that the stars, gas, and dust in our Galaxy are all orbiting the galactic center. Indeed, if this were not the case, mutual gravitational attraction would cause the entire Galaxy to collapse into the galactic center. In the same way, the Moon is kept from crashing into the Earth and the planets from crashing into the Sun because of their motion around their orbits (see Section 4-7).

Measuring the rotation of our Galaxy accurately is a difficult business. But such challenging measurements have been made, as we shall see, and the results lead to a remarkable conclusion: Most of the mass of the Galaxy is in the form of *dark matter,* a mysterious sort of material that emits no light at all.

Measuring How the Milky Way Rotates

Radio observations of 21-cm radiation from hydrogen gas provide important clues about our Galaxy's rotation. Doppler shift measurements of this radiation indicate that stars and gas all orbit in the same direction around the galactic center, just as the planets all orbit in the same direction around the Sun. Measurements also show that the orbital speed of stars and gas about the



INTERACTIVE EXERCISE 23.1

Figure 23-17

The Rotation of Our Galaxy **(a)** This schematic diagram shows three stars (the Sun and two others) orbiting the center of the Galaxy. Although they start off lined up, the stars become increasingly separated as they move along their orbits. Stars inside the Sun's orbit overtake and move ahead of the Sun, while stars far from the galactic center lag behind the Sun. **(b)** The stars would remain lined up if the Galaxy rotated like a solid disk. This is not what is observed. **(c)** If stars orbited the galactic center in the same way that planets orbit the Sun, stars inside the Sun's orbit would overtake us faster than they are observed to do.

galactic center is fairly uniform throughout much of the Galaxy's disk (Figure 23-17). As a result, stars inside the Sun's orbit complete a trip around the galactic center more quickly than the Sun, because the stars have a shorter distance to travel. Conversely, stars outside the Sun's orbit take longer to go once around the galactic center because they have farther to travel. As seen by Earth-based astronomers moving along with the Sun, stars inside the Sun's orbit overtake and pass us, while we overtake and pass stars outside the Sun's orbit (Figure 23-17*a*).

CAUTION Note that when we say that objects in different parts of the Galaxy orbit at the same speed, we do *not* mean that the Galaxy rotates like a solid disk. All parts of a rotating solid disk—a CD or DVD, for example—take the same time to complete one rotation. Because the outer part of the disk has to travel around a larger circle than the inner part, the speed (distance per time) is greater in the outer part (Figure 23-17*b*). By contrast, the orbital speed of material in our Galaxy is roughly the *same* at all distances from the galactic center.

The most familiar examples of orbital motion are the motions of the planets around the Sun. As we saw in Section 4-7, the farther a planet is from the Sun, the less gravitational force it experiences and the slower the speed it needs to have to remain in orbit. The same would be true for the orbits of stars and gas in the Galaxy *if* they were held in orbit by a single, massive object at the galactic center (Figure 23-17*c*). Hence, the 21-cm observations of our Galaxy, which show that the speed does *not* decrease with increasing distance from the galactic center, demonstrate that there is no such single, massive object holding objects in their galactic orbits.

Instead, what keeps a star in its orbit around the center of the Galaxy is the combined gravitational force exerted on it by *all* of the mass (including stars, gas, and dust) that lies within the star's orbit. (It turns out that the gravitational force from matter *outside* a star's orbit has little or no net effect on the star's motion around the galactic center.) This gives us a tool for determining the Galaxy's mass and how that mass is distributed.

The Sun's Orbital Motion and the Mass of the Galaxy

An important example is the orbital motion of the Sun (and the solar system) around the center of the Galaxy. If we know the semimajor axis and period of the Sun's orbit, we can use Newton's form of Kepler's third law (described in Section 4-7) to determine the mass of that portion of the Galaxy that lies within the orbit. We saw in Section 23-2 that the Sun is about 8000 pc (26,000 ly) from the galactic center; this is the semimajor axis of the Sun's orbit. The orbit is in fact nearly circular, so we can regard 8000 pc as the radius r of the orbit. In one complete trip around the Galaxy, the Sun travels a distance equal to the circumference of its orbit, which is $2\pi r$. The time required for one orbit, or orbital period P, is equal to the distance traveled divided by the Sun's orbital speed v:

Period of the Sun's orbit around the galactic center

$$P = \frac{2\pi r}{v}$$

P = orbital period of the Sun

r = distance from the Sun to the galactic center

v = orbital speed of the Sun

Unfortunately, we cannot tell the Sun's orbital speed from 21-cm observations, since these reveal only how fast things are moving relative to the Sun. Instead, we need to measure how the Sun is moving relative to a background that is not rotating along with the rest of the Galaxy. Such a background is provided by distant galaxies beyond the Milky Way and by the globular clusters. (Since globular clusters lie outside the plane of the Galaxy, they do not take part in the rotation of the disk.) By measuring the Doppler shifts of these objects and averaging their velocities, astronomers deduce that the Sun is moving along its orbit around the galactic center at about 220 km/s—about 790,000 kilometers per hour or 490,000 miles per hour!

Using this information, we find that the Sun's orbital period is

$$P = \frac{2\pi \times 8000 \text{ pc}}{220 \text{ km/s}} \times \frac{3.09 \times 10^{13} \text{ km}}{1 \text{ pc}} = 7.1 \times 10^{15} \text{ s}$$
$$= 2.2 \times 10^{8} \text{ years}$$

Traveling at 790,000 kilometers per hour, it takes the Sun about 220 million years to complete one trip around the Galaxy. (In the 65 million years since the demise of the dinosaurs, our solar system has traveled less than a third of the way around its orbit.) The Galaxy is a very large place!

Box 23-2 shows how to combine the radius and period of the Sun's orbit to calculate the total mass of all the matter that lies inside the Sun's orbit. Such calculations give an answer of 9.0×10^{10} $M_\odot$ (90 billion solar masses). As Figure 23-7 shows, the Galaxy extends well beyond the Sun's orbit, so the mass of the entire Galaxy must be larger than this.

BOX 23-2 Tools of the Astronomer's Trade

Estimating the Mass Inside the Sun's Orbit

The force that keeps the Sun in orbit around the center of the Galaxy is the gravitational pull of all the matter interior to the Sun's orbit. We can estimate the total mass of all of this matter using Newton's form of Kepler's third law (see Section 4-7 and Box 4-4):

$$P^2 = \frac{4\pi^2 a^3}{G(M + M_\odot)}$$

In this equation P is the orbital period of the Sun, a is the semimajor axis of the Sun's orbit around the galactic center, G is the universal constant of gravitation, M is the amount of mass inside the Sun's orbit, and $M_\odot$ is the mass of the Sun.

Because the Sun is only one of more than 10^{11} stars in the Galaxy, the Sun's mass is minuscule compared to M. Hence, we can safely replace the sum $M + M_\odot$ in the above equation by simply M. If we now assume that the Sun's orbit is a circle, the semimajor axis a of the orbit is just the radius of this circle, which we call r. As we saw in Section 25-4, the period P of the orbit is equal to $2\pi r/v$, where v is the Sun's orbital speed. You can then show that

$$M = \frac{rv^2}{G}$$

(We leave the derivation of this equation as an exercise at the end of this chapter.)

Now we can insert known values to obtain the mass inside the Sun's orbit. Being careful to express distance in meters and speed in meters per second, we have $v = 220$ km/s $= 2.2 \times 10^5$ m/s, $G = 6.67 \times 10^{-11}$ newton $\cdot$ m^2/kg^2, and

$$r = 26{,}000 \text{ light-years} \times \frac{9.46 \times 10^{12} \text{ km}}{1 \text{ light-year}} \times \frac{10^3 \text{ m}}{1 \text{ km}}$$
$$= 2.5 \times 10^{20} \text{ m}$$

Hence, we find that

$$M = \frac{2.5 \times 10^{20} \times (2.2 \times 10^5)^2}{6.67 \times 10^{-11}} = 1.8 \times 10^{41} \text{ kg}$$

or, in terms of the mass of the Sun,

$$M = 1.8 \times 10^{41} \text{ kg} \times \frac{1 M_\odot}{1.99 \times 10^{30} \text{ kg}}$$
$$= 9.0 \times 10^{10} M_\odot$$

This estimate involves only mass that is interior to the Sun's orbit. Matter outside the Sun's orbit has no net gravitational effect on the Sun's motion and thus does not enter into Kepler's third law. (This is strictly true only if the matter outside our orbit is distributed over a sphere rather than a disk. In fact, the dark matter that dominates our Galaxy seems to have a spherical distribution.)

Rotation Curves and the Mystery of Dark Matter

In recent years, astronomers have been astonished to discover how much matter may lie outside the Sun's orbit. The clues come from 21-cm radiation emitted by hydrogen in spiral arms that extend to the outer reaches of the Galaxy. Because we know the true speed of the Sun, we can convert the Doppler shifts of this radiation into actual speeds for the spiral arms. This calculation gives us a **rotation curve,** a graph of the speed of galactic rotation measured outward from the galactic center (Figure 23-18). We would expect that for gas clouds beyond the confines of most of the Galaxy's mass, the orbital speed should decrease with increasing distance from the Galaxy's center, just as the orbital speeds of the planets decrease with increasing distance from the Sun (see Figure 23-17*c*). But as Figure 23-18 shows, the Galaxy's rotation curve is quite flat, indicating roughly uniform orbital speeds well beyond the visible edge of the galactic disk.

To explain these nearly uniform orbital speeds in the outer parts of the Galaxy, astronomers conclude that a large amount of matter must lie outside the Sun's orbit. When this matter is included, the total mass of our Galaxy could exceed 10^{12} $M_\odot$ or more, of which about 10% is in the form of stars. This implies that our Galaxy contains roughly 200 billion stars.

Unlike the Galaxy's stars and dust, its dark matter forms a roughly spherical halo

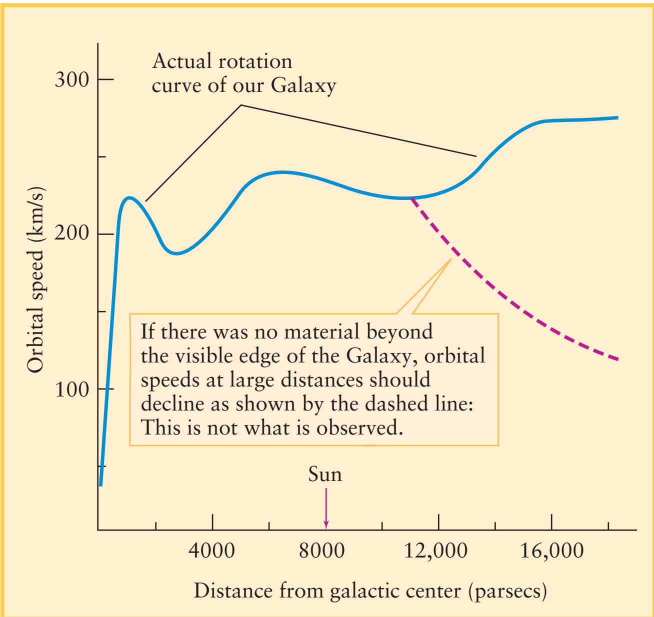


Figure 23-18

The Galaxy's Rotation Curve The blue curve shows the orbital speeds of stars and gas in the disk of the Galaxy out to a distance of 18,000 parsecs from the galactic center. (Very few stars are found beyond this distance.) The dashed red curve indicates how this orbital speed should decline beyond the confines of most of the Galaxy's visible mass. Because there is no such decline, there must be an abundance of invisible dark matter that extends to great distances from the galactic center.

WEB LINK 23.9

These observations lead to a profound mystery. Stars, gas, and dust account for only about 10% of the Galaxy's total mass. What, then, makes up the remaining 90% of the matter in our Galaxy? Whatever it is, it is dark. It does not show up on photographs, nor indeed in images made in any part of the electromagnetic spectrum. This unseen material, which is by far the predominant constituent of our Galaxy, is called **dark matter.** We sense its presence only through its gravitational influence on the orbits of stars and gas clouds.

CAUTION Be careful not to confuse dark *matter* with dark *nebulae.* A dark nebula like the one in Figure 18-4 emits no visible light, but does radiate at longer wavelengths. By contrast, no electromagnetic radiation of any kind has yet been discovered coming from dark matter.

Observations of star groupings outside the Milky Way suggest that our Galaxy's dark matter forms a spherical halo centered on the galactic nucleus, like the halo of globular clusters and high-velocity stars shown in Figure 23-7. However, the dark matter halo is much larger; it may extend to a distance of 100–200 kpc from the center of our Galaxy, some 2 to 4 times the extent of the visible halo (Figure 23-19). Analysis of the rotation curve in Figure 23-18 shows that the density of the dark matter halo decreases with increasing distance from the center of the Galaxy.

Dark Matter Speculations

WEB LINK 23.10

What is the nature of this mysterious dark matter? One proposal is that the dark matter halo is composed, at least in part, of dim objects with masses less than 1 $M_\odot$. These objects, which could include brown dwarfs, white dwarfs, or black holes, are called **massive compact halo objects,** or **MACHOs.** Astronomers have searched for MACHOs by monitoring the light from distant stars. If a MACHO passes between us and the star, its gravity will bend the light coming from the star. (In Section 24-2 we described how gravity can bend starlight.) As Figure 23-20 shows, the MACHO's gravity acts like a lens that focuses the light from the star. This effect, called **microlensing,** makes the star appear to brighten substantially for a few days.

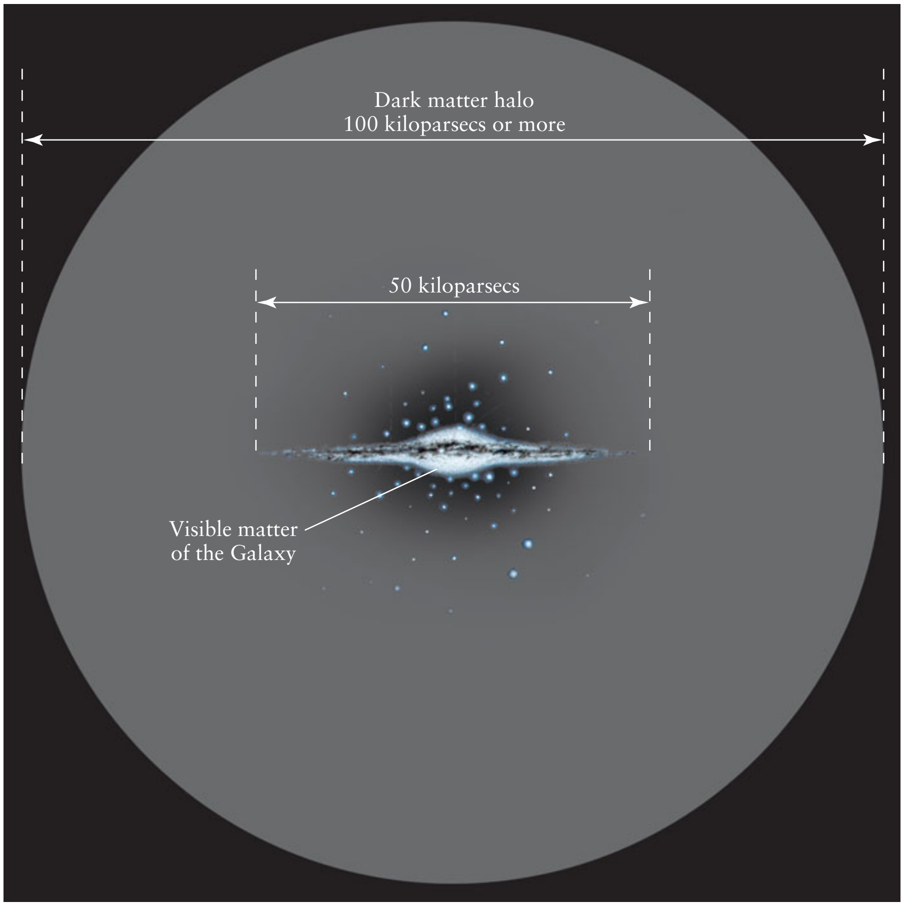
Astronomers have indeed detected MACHOs in this way, but not enough to completely solve the dark matter mystery. MACHOs with very low mass (10^{-6} to 0.1 $M_\odot$ each) do not appear to be a significant part of the dark matter halo. MACHOs of roughly 0.5 $M_\odot$ are more prevalent, but account for only about half of the dark matter halo.

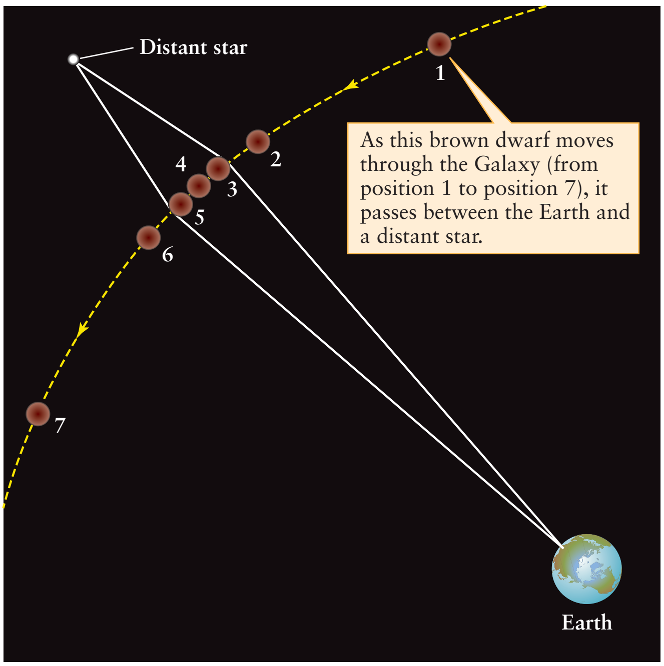
The remainder of the dark matter is thought to be much more exotic. One candidate is a neutrino with a small amount of mass. If these neutrinos are sufficiently massive, and if enough of them are present in the halo of the Galaxy, they might constitute a reasonable fraction of the dark matter. As we saw in Section 16-4, one type of neutrino can transform into another. These transformations can take place only if neutrinos have a nonzero amount of mass. Thus, neutrinos must comprise at least part of the dark matter, though it is not known how much.

Another speculative possibility is a new class of subatomic particle called **weakly interacting massive particles,** or **WIMPs.** These particles, whose existence is suggested by certain theories but has not yet been confirmed experimentally, would not emit

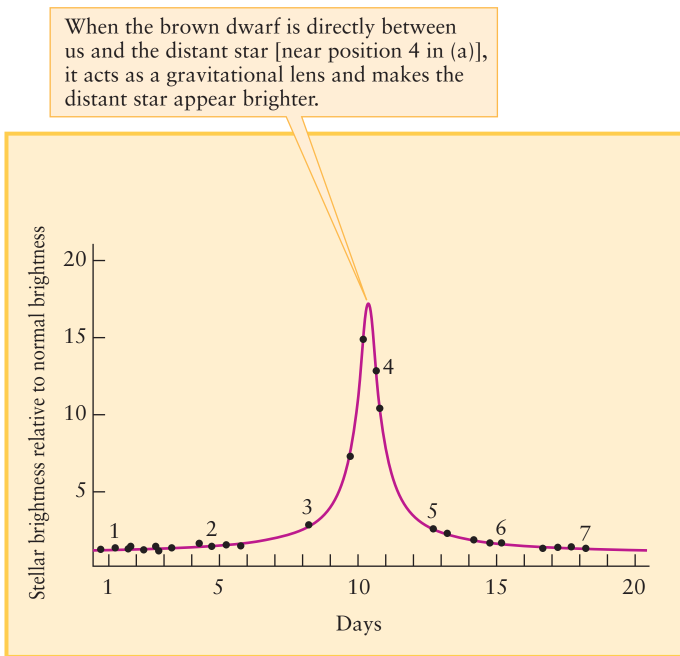
Figure 23-19

The Galaxy and Its Dark Matter Halo The dark matter in our Galaxy forms a spherical halo whose center is at the center of the visible Galaxy. The extent of the dark matter halo is unknown, but its diameter is at least 100 kiloparsecs. The total mass of the dark matter halo is at least 10 times the combined mass of all of the stars, dust, gas, and planets in the Milky Way.





(a)



(b)

Figure 23-20

Microlensing by Dark Matter in the Galactic Halo **(a)** If a dense object such as a brown dwarf or black hole passes between the Earth and a distant star, the gravitational curvature of space around the dense object deflects the starlight and focuses it in our direction. This effect is called microlensing. **(b)** This light curve shows the gravitational microlensing of light from a star in the Galaxy's central bulge. Astronomers do not know the nature of the object that passed between the Earth and this star to cause the microlensing. (Courtesy of the MACHO and GMAN Collaborations) or absorb electromagnetic radiation. Physicists are attempting to detect these curious particles, which would have masses 10 to 10,000 times greater than a proton or neutron, by using a large crystal cooled almost to absolute zero. If a WIMP should enter this crystal and collide with one of its atoms, the collision will deposit a tiny but measurable amount of heat in the crystal.

As yet, the true nature of dark matter remains a mystery. Furthermore, this mystery is not confined to our own Galaxy. In Chapter 24 we will find that other galaxies have the same sort of rotation curve as in Figure 23-18, indicating that they also contain vast amounts of dark matter. Indeed, dark matter appears to make up most of the mass in our universe. Hence, the quest to understand dark matter is one of the most important in modern astronomy.

23-5 Spiral arms are caused by density waves that sweep around the Galaxy

The disk shape of our Galaxy is not difficult to understand. In Section 8-4 we described what happens when a large number of objects are put into orbit around a common center: Over time the objects tend naturally to orbit in the same plane. This is what happened when our solar system formed from the solar nebula. There a giant cloud of material eventually organized itself into planets, all of which orbit in nearly the same plane. In like fashion, the disk of our Galaxy, which is also made up of a large number of individual objects orbiting a common center, is very flat (see Figure 23-7). Understanding why our Galaxy has spiral arms presents more of a challenge.

Our Sun and its solar system were spawned when a cloud of gas and dust passed through a spiral arm

The Winding Dilemma

One early explanation for the Galaxy's spiral structure was that the material in the Galaxy somehow condensed into a spiral pattern from the very start. In this view, once stars, gas, and dust had become concentrated within the spiral arms, the pattern would remain fixed. This would be possible only if the Galaxy rotated like a solid disk (see Figure 23-17*b*); the fixed pattern would be like the spokes on a rotating bicycle wheel. But the reality is that the Galaxy is not a solid disk. As we have seen, stars, gas, and dust all orbit the galactic center with approximately the same speed, as shown in Figure 23-17*a*. Let us see why this makes it impossible for a rigid spiral pattern to persist.

Imagine four stars, *A*, *B*, *C*, and *D*, that originally lie on a line extending outward from the galactic center (Figure 23-21*a*). In a given amount of time, each of the stars travels the same distance around its orbit. But because the innermost star has a smaller orbit than the others, it takes less time to complete one orbit. As a result, a line connecting the four stars is soon bent into a spiral (Figure 23-21*b*). Moreover, the spiral becomes tighter and tighter with the passage of time (Figures 23-21*c* and 23-21*d*). This "winding up" of the spiral arms causes the spiral structure to disappear completely after a few hundred million years—a very brief time compared to the age of our Galaxy, thought to be about 13.5 billion (1.35×10^{10}) years.

Figure 23-21 suggests that the Milky Way's spiral arms ought to have disappeared by now. The fact that they have not is called the **winding dilemma.** It shows that the spiral arms cannot simply be assemblages of stars and interstellar matter that travel around the Galaxy together, like a troop of soldiers marching in formation around a flagpole. In other words, the spiral arms cannot be made of anything *material.* What, then, can the spiral arms be?

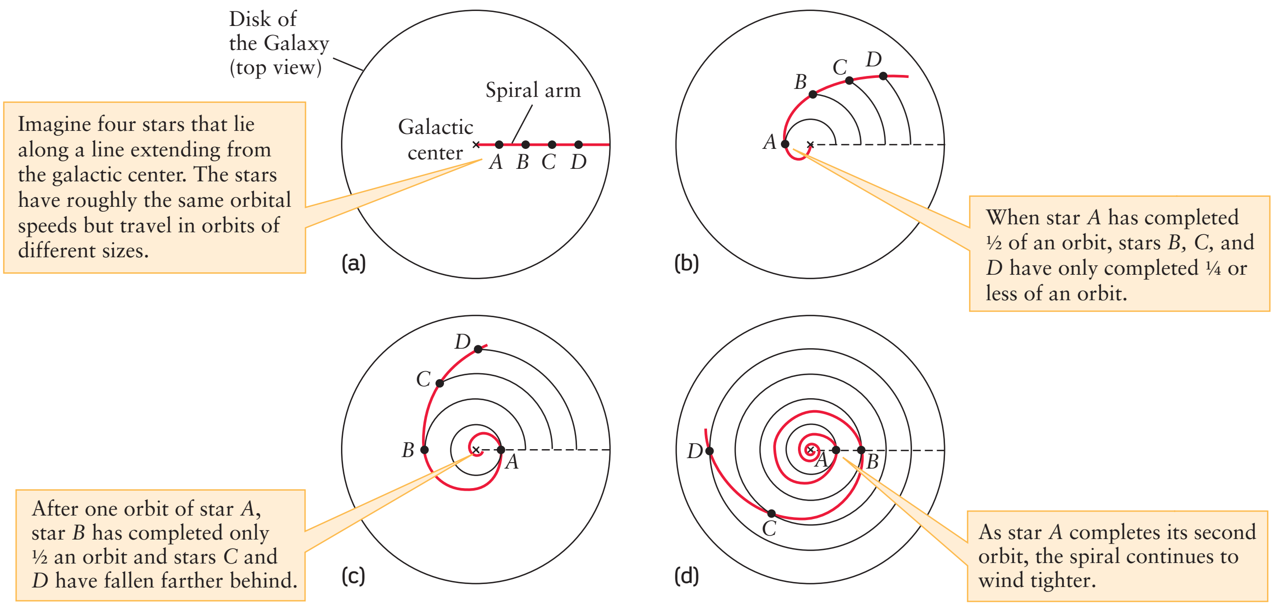


Figure 23-21

The Winding Dilemma This series of drawings shows that spiral arms in galaxies like the Milky Way cannot simply be assemblages of stars. If they were, the spiral arms would "wind up" and disappear in just a few hundred million years.

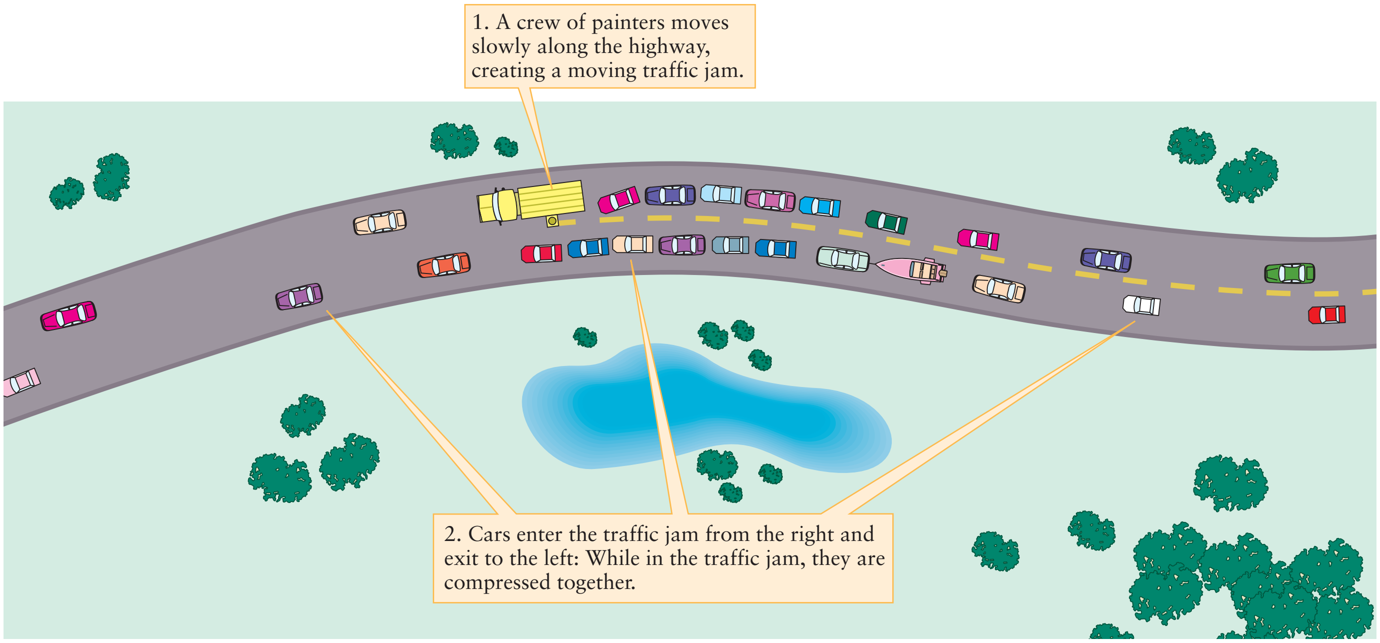


Figure 23-22

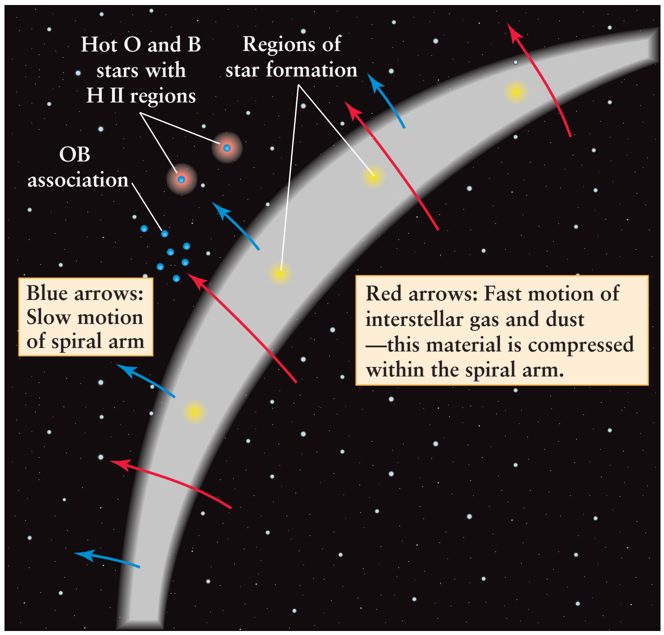
A Density Wave on the Highway A density wave in a spiral galaxy is analogous to a crew of painters moving slowly along the highway, creating a moving traffic jam. Like such a traffic jam, a density wave in a spiral galaxy is a slow-moving region where stars, gas, and dust are more densely packed than in the rest of the galaxy. As the material of the galaxy passes through the density wave, it is compressed. This triggers star formation, as Figure 23-23 shows.

The Density-Wave Model

In the 1940s, the Swedish astronomer Bertil Lindblad proposed that the spiral arms of a galaxy are actually a pattern that moves through the Galaxy like ripples on water. This idea was greatly enhanced and embellished in the 1960s by the American astronomers Chia Chiao Lin and Frank Shu. In this picture, spiral arms are a kind of wave, like the waves that move across the surface of a pond when you toss a stone into the water. Water molecules pile up at a crest of the wave but spread out again when the crest passes. By analogy, Lindblad, Lin, and Shu pictured a pattern of **density waves** sweeping around the Galaxy. These waves make matter pile up in the spiral arms, which are the crests of the waves. Individual parts of the Galaxy's material are compressed only temporarily when they pass through a spiral arm. The pattern of spiral arms persists, however, just as the waves made by a stone dropped in the water can persist for quite awhile after the stone has sunk.

To understand better how a density wave operates in a galaxy, think again about a water wave in a pond. If one part of the pond is disturbed by dropping a stone into it, the molecules in that part will be displaced a bit. They will nudge the molecules next to them, causing those molecules to be displaced and to nudge the molecules beyond them. In this way the wave disturbance spreads throughout the pond.

In a galaxy, stars play the role of water molecules. Although stars and interstellar clouds of gas and dust are separated by vast distances, they can nonetheless exert forces on each other because they are affected by each other's gravity. If a region of above-average density should form, its gravitational attraction will draw



INTERACTIVE EXERCISE 23.2

Figure 23-23

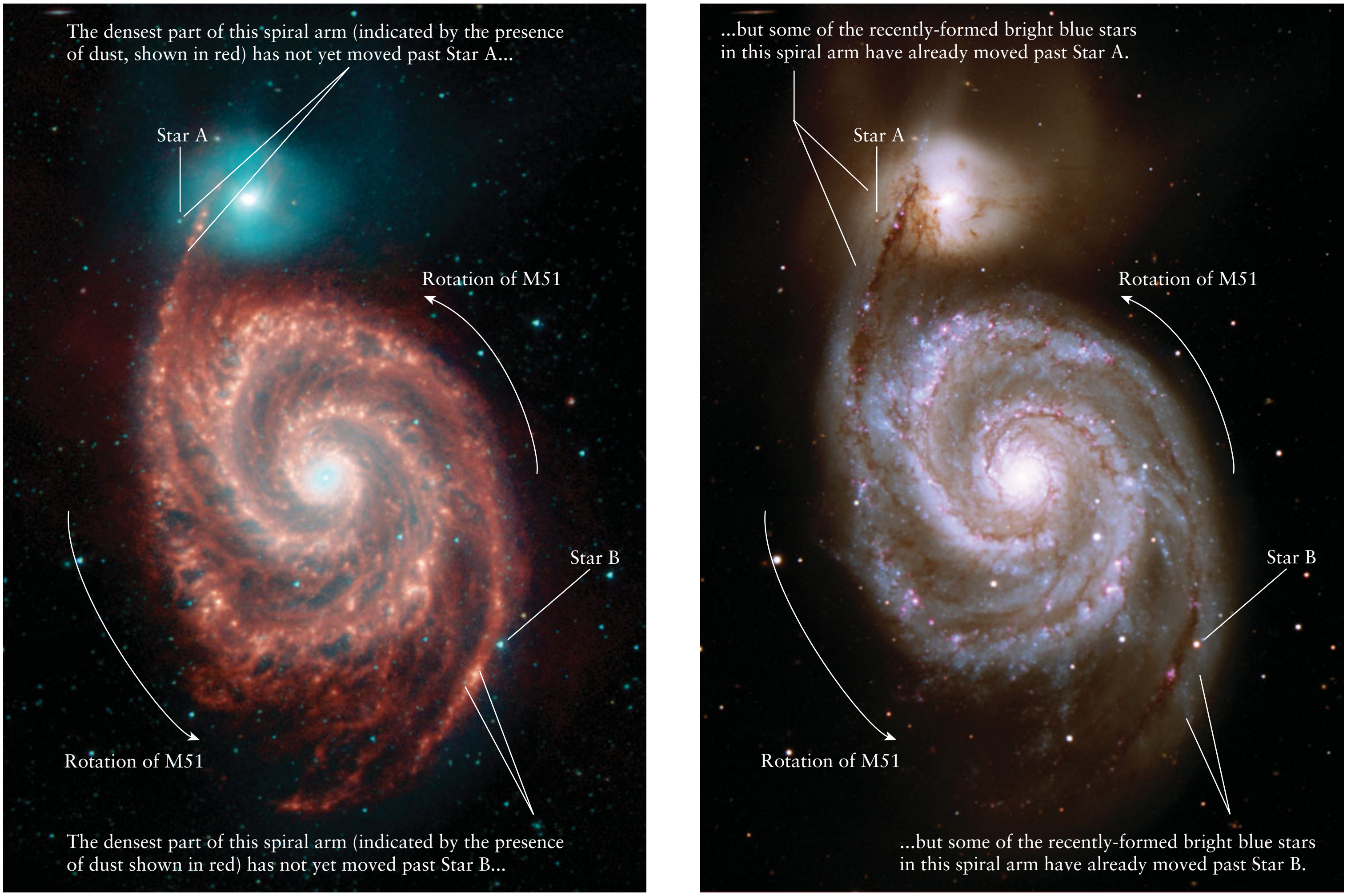
Star Formation in the Density-Wave Model A spiral arm is a region where the density of material is higher than in the surrounding parts of a galaxy. Interstellar matter moves around the galactic center rapidly (shown by the red arrows) and is compressed as it passes through the slow-moving spiral arms (whose motion is shown by the blue arrows). This compression triggers star formation in the interstellar matter, so that new stars appear on the "downstream" side of the densest part of the spiral arms.

nearby material into it. The displacement of this material will change the gravitational force that it exerts on other parts of the galaxy, causing additional displacements. In this way a spiral-shaped density wave can travel around the disk of a galaxy.

ANALOGY A key feature of density waves is that they move more slowly around a galaxy than do stars or interstellar matter. To visualize this, imagine workers painting a line down a busy freeway (Figure 23-22). The cars normally cruise along the freeway at high speed, but the crew of painters is moving much more slowly. When the cars come up on the painters, they must slow down temporarily to avoid hitting anyone. As seen from the air, cars are jammed together around the painters. An individual car spends only a few moments in the traffic jam before resuming its usual speed, but the traffic jam itself lasts all day, inching its way along the road as the painters advance.

A similar crowding takes place when interstellar matter enters a spiral arm. This crowding plays a key role in the formation of stars and the recycling of the interstellar medium. As interstellar gas and dust moves through a spiral arm, it is compressed into new nebulae (Figure 23-23). This compression begins the process by which new stars form, which we described in Section 18-3.

These freshly formed stars continue to orbit around the center of their galaxy, just like the matter from which they formed. The most luminous among these are the hot, massive, blue O and B stars, which may have emission nebulae (H II regions) associated with them. These stars have main-sequence lifetimes of only



(a) An infrared view of M51 shows the locations of dust R I V U X G

(b) A visible-light view of M51 shows the locations of young stars R I V U X G

WEB LINK 23.11

Figure 23-24

Star Formation in the Whirlpool Galaxy The spiral galaxy M51 (called the Whirlpool) is a real-life example of the density-wave model illustrated in Figure 23-23. **(a)** This infrared image shows where dust has piled up as the material within M51 passes through its spiral arms. Radio images of M51 show that hydrogen gas also piles up in the same locations, thus beginning the formation of new stars. **(b)** By the time stars complete their formation process, their motion around the galaxy has swept them "downstream" of the positions of greatest dust density, just as depicted in Figure 23-23. (a: NASA, JPL-Caltech, and R. Kennicutt (Univ. of Arizona); b: DSS)

3 to 15 million years (see Table 19-1), which is very short compared to the 220 million years required for the Sun to make a complete orbit around the Galaxy.

As a result, these luminous O and B stars can travel only a relatively short distance before dying off. Therefore, these stars, and their associated H II regions, are only seen in or slightly "downstream" of the spiral arm in which they formed. Figure 23-24 illustrates this for the spiral galaxy M51. Less massive stars have much longer main-sequence lifetimes, and thus their orbits are able to carry them all around the galactic disk. These less-luminous stars are found throughout the disk, including between the spiral arms (see Figure 23-15*c*).

The density-wave model of spiral arms explains why the disk of our Galaxy is dominated by metal-rich Population I stars. Because the material left over from the death of ancient stars is enriched in heavy elements, new generations of stars formed in spiral arms are likely to be more metal-rich than their ancestors. The *Cosmic Connections* figure illustrates this cycle of star birth and death in the disk of our Galaxy.

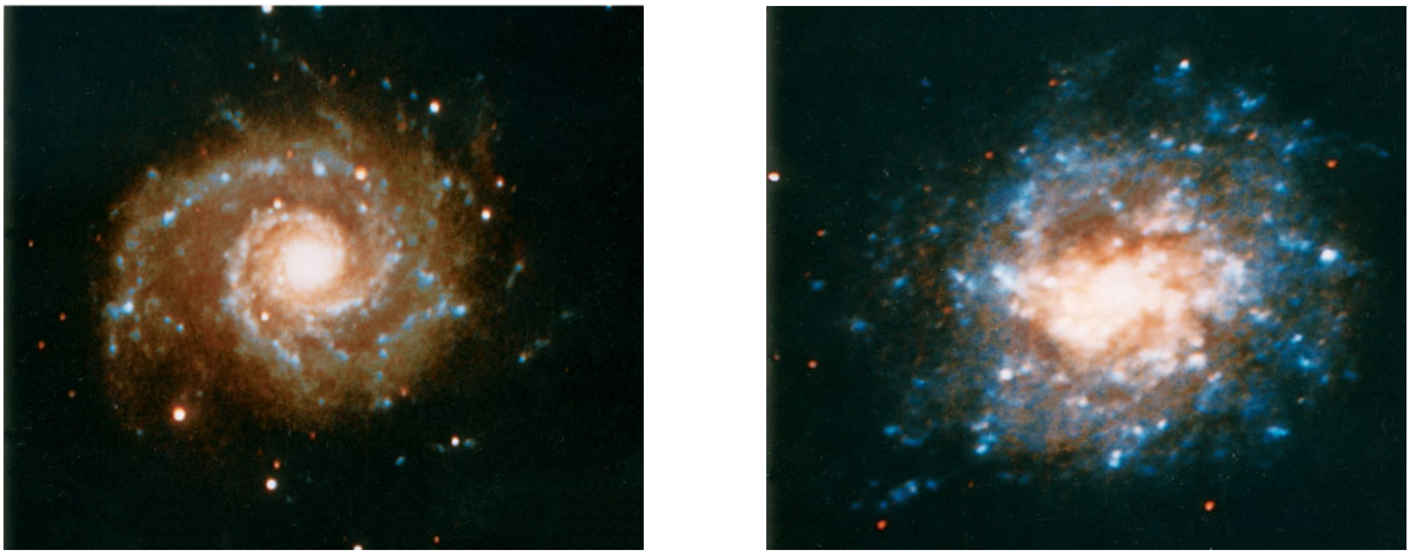
The density-wave model is still under development. One problem is finding a driving mechanism that keeps density waves going in spiral galaxies. After all, density waves expend an enormous amount of energy to compress the interstellar gas and dust. Hence, we would expect that density waves should eventually die away, just as do ripples on a pond. The American astronomers Debra and Bruce Elmegreen have suggested that gravity can supply that needed energy. As mentioned in Section 23-3, the central bulge of our Galaxy is elongated into a bar shape, much like the central bulge of the galaxy M83 (see Figure 23-15*a* and Figure 23-16*a*). The asymmetric gravitational field of such a bar pulls on the stars and interstellar matter of a galaxy to generate density waves. Another factor that may help to generate and sustain spiral arms is the gravitational interactions *between* galaxies. We will discuss this in Chapter 24.

The Self-Propagating Star-Formation Model

Spiral density waves may not be the whole story behind spiral arms in our Galaxy and other galaxies. The reason is that spiral density waves should produce very well defined spiral arms. We do indeed see many so-called **grand-design spiral galaxies** (Figure 23-25*a*), with thin, graceful, and well-defined spiral arms. But in some galaxies, called **flocculent spiral galaxies** (Figure 23-25*b*), the spiral arms are broad, fuzzy, chaotic, and poorly defined. ("Flocculent" means "resembling wool.")

To explain such flocculent spirals, M. W. Mueller and W. David Arnett in 1976 proposed a theory of **self-propagating star formation.** Imagine that star formation begins in a dense interstellar cloud within the disk of a galaxy that does not yet have spiral arms. As soon as hot, massive stars form, their radiation and stellar winds compress nearby matter, triggering the formation of additional stars in that gas. When massive stars become supernovae, they produce shock waves that further compress the surrounding interstellar medium, thus encouraging still more star formation.

Although all parts of this broad, star-forming region have approximately the same orbital speed about the galaxy's center, the inner regions have a shorter distance to travel to complete one orbit than the outer regions. As a result, the inner edges of the star-forming region move ahead of the outer edges as the Galaxy rotates. The bright O and B stars and their nearby glowing nebulae soon become stretched out in the form of a spiral arm. These spiral arms come and go essentially at random across a galaxy. Bits and pieces of spiral arms appear where star formation has recently begun but fade and disappear at other locations where all



(a) Grand-design spiral galaxy

(b) Flocculent spiral galaxy

Figure 23-25 R I V U X G

Variety in Spiral Arms The differences from one spiral galaxy to another suggest that more than one process can create spiral arms. **(a)** NGC 628 is a grand-design spiral galaxy with thin, well-defined spiral arms. **(b)** NGC 7793 is a flocculent spiral galaxy with fuzzy, poorly defined spiral arms. (Courtesy of P. Seiden, D. Elmegreen, B. Elmegreen, and A. Mobarak; IBM)

COSMIC CONNECTIONS

Stars in the Milky Way

Different populations of stars are found in different neighborhoods of our home galaxy. (The galaxy shown here is another spiral galaxy similar to our own.) The variations from one galactic region to another are due to the presence or absence of ongoing star formation.

7. Also orbiting the Galaxy are the globular clusters. These star clusters have no mechanism to trigger star formation, so they contain only very old, metal-poor Population II stars. Studies of these clusters suggest that the first stars formed in them about 13.6 billion years ago — a mere 100 million years after the Big Bang.

8. Not visible to any telescope is the Galaxy's dark matter, which emits no electromagnetic radiation of any kind. The spherical halo of dark matter that envelops our Galaxy has at least 10 times the combined mass of all of the Galaxy's stars, planets, gas, and dust. Its fundamental nature remains a mystery.

(Image of spiral galaxy M101: NASA and ESA; image of globular cluster M3: S. Kafka and K. Honeycutt, Indiana University/WIYN/NOAO/NSF) the massive stars have died off. Self-propagating star formation therefore tends to produce flocculent spiral galaxies that have a chaotic appearance with poorly defined spiral arms, like the galaxy in Figure 23-25*b*.

The two theories presented here are very different in character. In the density-wave model, star formation is caused by the spiral arms; in the self-propagating star formation model, by contrast, the spiral arms are caused by star formation. A complete and correct description of spiral arms in our Galaxy remains a topic of active research.

23-6 Infrared, radio, and X-ray observations are used to probe the galactic center

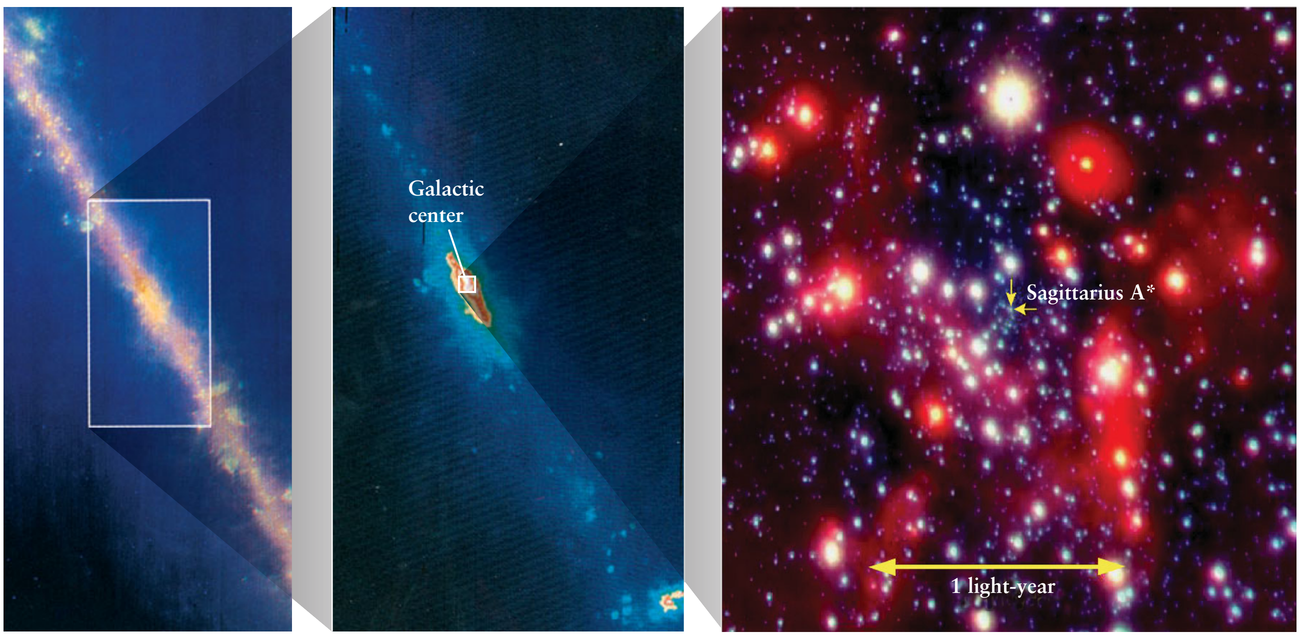
The innermost region of our Galaxy is an active, crowded place. If you lived on a planet near the galactic center, you could see a million stars as bright as Sirius, the brightest single star in our own night sky. The total intensity of starlight from all those nearby stars would be equivalent to 200 of our full moons. In effect, night would never really fall on a planet near the center of the Milky Way. At the center of this empire of light, however, lies the darkest of all objects in the universe—a black hole millions of times more massive than the Sun.

At the very center of the Milky Way Galaxy is a maelstrom of activity centered on a supermassive black hole

Sagittarius A*: Heart of Darkness

Because of the severe interstellar absorption at visual wavelengths, most of our information about the galactic center comes from infrared and radio observations. Figure 23-26 shows three infrared views of the center of our Galaxy. Figure 23-26*a* is a wide-angle view covering a 50° segment of the Milky Way from Sagittarius through Scorpius. (The photograph that opens this chapter shows this same region at visible wavelengths, viewed at a different angle.) The prominent reddish band through the center of this false-color infrared image is a layer of dust in the plane of the Galaxy. Figure 23-26*b* is an IRAS view of the galactic center. It is surrounded by numerous streamers of dust (shown in blue). The strongest infrared emission (shown in white) comes from a grouping of several powerful sources of radio waves. One of these sources, **Sagittarius A*** (say "A star"), lies at the very center of the Galaxy. (Its position, pinpointed with simultaneous observations by radio telescopes scattered around the world, seems to be very near the gravitational center of the Galaxy.) The high-resolution infrared view in Figure 23-26*c*, made using adaptive optics (see Section 6-3), shows hundreds of stars crowded within 1 ly (0.3 pc) of Sagittarius A*. Compare this to our region of the Galaxy, where the average distance between stars is more than a light-year.

Sagittarius A* itself does not appear in infrared images. Nonetheless, astronomers have used infrared observations to make truly startling discoveries about this object. Since the 1990s, two research groups—one headed by Reinhard Genzel of the Max



(a) A wide-angle (50°) infrared view

(b) A close-up view shows a more luminous region at the galactic center

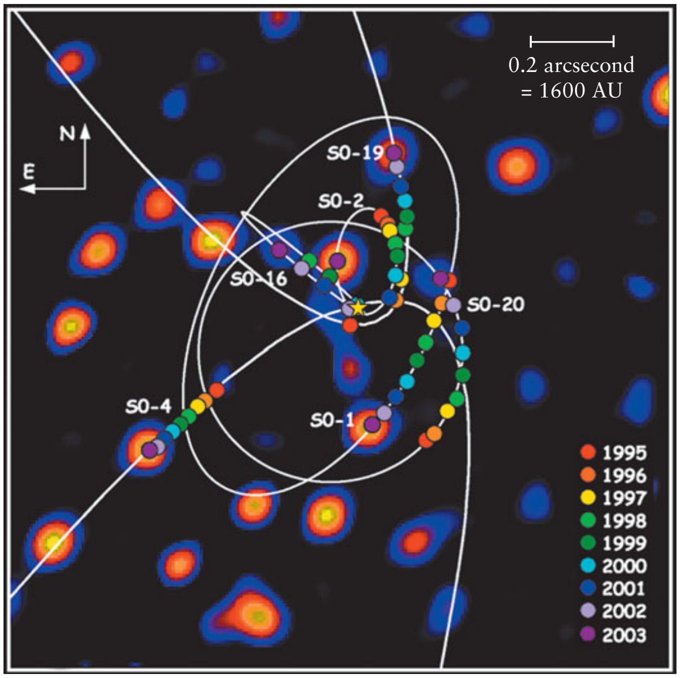
(c) An extreme close-up view centered on Sagittarius A*, a radio source at the very center of the Milky Way Galaxy, shows hundreds of stars within 1 ly (0.3 pc)

Figure 23-26 R I V U X G

The Galactic Center (a) In this false-color infrared image, the reddish band is dust in the plane of the Galaxy and the fainter bluish blobs are interstellar clouds heated by young O and B stars. **(b)** This close-up infrared view covers the area outlined by the white rectangle in (a). **(c)** Adaptive optics reveals stars densely packed around the galactic center. (a, b: NASA; c: R. Schödel et al., MPE/ESO)

Planck Institute for Extraterrestrial Physics in Garching, Germany, and another led by Andrea Ghez at the University of California, Los Angeles—have been using infrared detectors to monitor the motions of stars in the immediate vicinity of Sagittarius A*. They have found a number of stars orbiting around Sagittarius A* at speeds in excess of 1500 km/s (Figure 23-27). (By comparison, the Earth orbits the Sun at a lackadaisical 30 km/s.) In 2000 the UCLA group observed one such star, called SO-16, as its elliptical orbit brought it within a mere 45 AU from Sagittarius A* ($1\frac{1}{2}$ times the distance from the Sun to Neptune). At its closest approach, SO-16 was traveling at a breathtaking speed of 12,000 km/s, or 4% of the speed of light!

In order to keep stars like SO-16 in such small, rapid orbits, Sagittarius A* must exert a powerful gravitational force and hence must be very massive. By applying Newton's form of Kepler's third law to the motions of these stars around Sagittarius A*, the UCLA group calculates the mass of Sagittarius A* to be a remarkable 3.7 *million* solar masses ($3.7 \times 10^6\ M_\odot$). Furthermore, the small separation between SO-16 and Sagittarius A* at closest approach shows that Sagittarius A* can be no more than 45 AU in radius. An object this massive and this compact can only be one thing: a supermassive black hole (see Section 22-4).



WEB LINK 23.12 VIDEO 23.2

Figure 23-27 R I V U X G

Stars Orbiting Sagittarius A* The colored dots superimposed on this infrared image show the motion of six stars in the vicinity of the unseen massive object (denoted by the yellow five-pointed star) at the position of the radio source Sagittarius A*. The orbits were measured over an 8-year period. This plot indicates that the stars are held in orbit by a black hole of 3.7×10^6 solar masses. (UCLA Galactic Center Group)

X-rays from Around a Supermassive Black Hole

ANIMATION 23.1

Evidence in favor of this picture comes from the Chandra X-ray Observatory, which has observed X-ray flares coming from Sagittarius A*. The flares brighten dramatically over the space of just 10 minutes, which shows that the size of the flare's source can be no larger than the distance that light travels in 10 minutes. (We used a similar argument in Section 22-3 to show that the flickering X-ray source Cygnus X-1 must be very small.) In 10 minutes light travels a distance of 1.8×10^8 km or 1.2 AU, and only a black hole could pack a mass of $3.7 \times 10^6\ M_\odot$ into a volume that size or smaller. The X-ray flares were presumably emitted by blobs of material that were compressed and heated as they fell into the black hole (see Section 22-3).

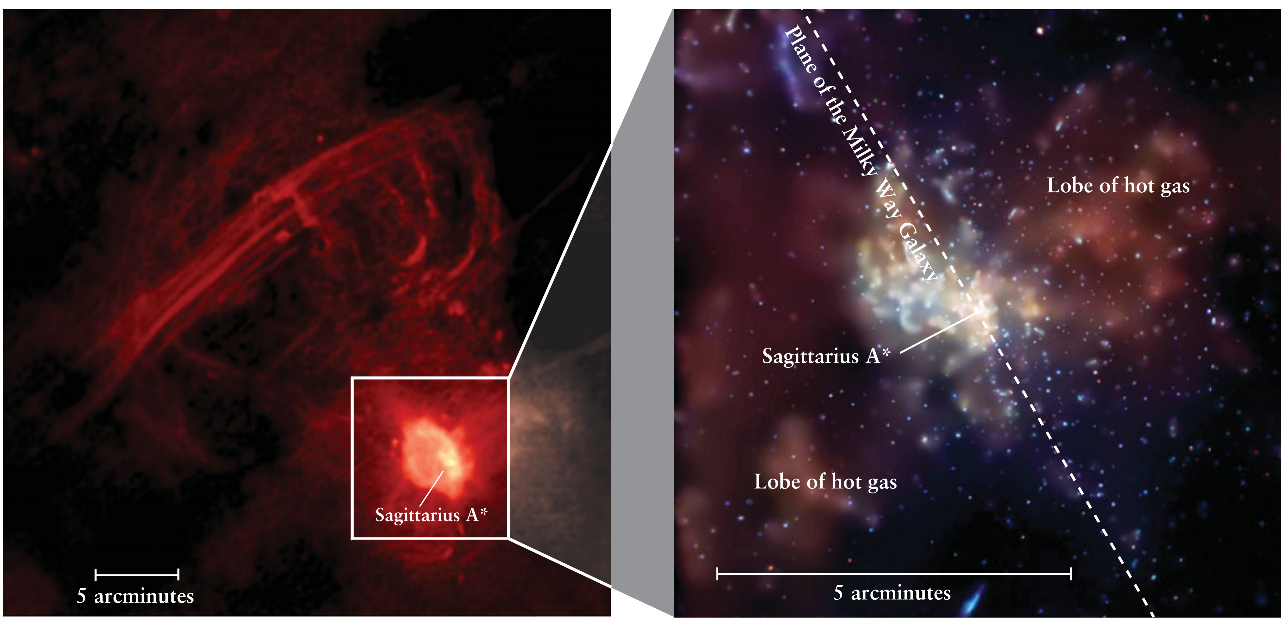
The X-ray flares from Sagittarius A* are relatively feeble, which suggests that the supermassive black hole is swallowing only relatively small amounts of material. But the region around Sagittarius A* is nonetheless an active and dynamic place. Figure 23-28*a* is a wide-angle radio image of the galactic center covering an area more than 60 pc (200 ly) across. Huge filaments of gas stretch for 20 pc (65 ly) northward of the galactic center (to the right and upward in Figure 23-28*a*), then abruptly arch southward (down and to the left in the figure). The orderly arrangement of these filaments is reminiscent of prominences on the Sun (see Section 18-10, especially Figure 18-28). This suggests that, as on the Sun, there is ionized gas at the galactic center that is being controlled by a powerful magnetic field. Indeed, much of the radio emission from the galactic center is synchrotron radiation: As we saw in Section 23-4, such radiation is produced by high-energy electrons spiraling in a magnetic field.

The false-color X-ray image in Figure 23-28*b* shows the immediate vicinity of Sagittarius A*. The black hole is flanked by lobes of hot, ionized, X-ray–emitting gas that extend for dozens of light-years. These are thought to be the relics of immense explosions that may have taken place over the past several thousand years. Perhaps these past explosions cleared away much of the material around Sagittarius A*, leaving only small amounts to fall into the black hole. This could explain why the X-ray flares from Sagittarius A* are so weak.

WEB LINK 23.13

There is nonetheless evidence that Sagittarius A* has been more active in the recent past. Between 2002 and 2005, Michael Muno and his colleagues at Caltech observed that certain nebulae within about 50 ly of Sagittarius A* suddenly became very bright at X-ray wavelengths. The character of the X-ray light was such that it could not have originated from the nebulae themselves. Muno and colleagues concluded that X-rays emitted from Sagittarius A* about 50 years earlier had struck and excited the nebulae, causing the nebulae to glow intensely at X-ray wavelengths. To produce such an intense X-ray glow, an object the size of the planet Mercury must have fallen into the supermassive black hole.

The supermassive black hole at the center of our Galaxy is not unique. Observations show that such titanic black holes are a feature of most large galaxies. In Chapter 25 we will see how black holes of this kind power *quasars*, the most luminous sustained light sources in the cosmos.



(a) A radio view of the galactic center R I V U X G

(b) An X-ray view of the galactic center R I V U X G

Figure 23-28

The Energetic Center of the Galaxy **(a)** The area shown in this radio image has the same angular size as the full moon. Sagittarius A*, at the very center of the Galaxy, is one of the brightest radio sources in the sky. Magnetic fields shape nearby interstellar gas into immense, graceful arches. **(b)** This composite of images at X-ray wavelengths from 0.16 to 0.62 nm shows lobes of gas on either side of Sagittarius A*. The character of the X-ray emission shows that the gas temperature is as high as 2×10^7 K. (a: NRAO/VLA/F. Zadeh et al.; b: NASA/CXC/MIT/F. K. Baganoff et al.)

Astronomers are still groping for a better understanding of the galactic center. With future developments in very-long-baseline interferometry (described in Section 6-6), it may be possible to actually obtain a picture of the supermassive black hole lurking there. During the coming years, observations from Earth-orbiting satellites as well as from radio and infrared telescopes on the ground will certainly add to our knowledge of the core of the Milky Way.

Key Words

Word preceded by an asterisk (*) is discussed in Box 23-1.

central bulge (of a galaxy), p. 609
dark matter, p. 619
density wave, p. 622
disk (of a galaxy), p. 609
far-infrared, p. 609
flocculent spiral galaxy, p. 624
galactic nucleus, p. 608
galaxy, p. 606
globular cluster, p. 607
grand-design spiral galaxy, p. 624
H I, p. 612
halo (of a galaxy), p. 609
high-velocity star, p. 610
interstellar extinction, p. 606
Local Bubble, p. 613
*magnetic resonance imaging (MRI), p. 614
massive compact halo object (MACHO), p. 619
microlensing, p. 619
Milky Way Galaxy, p. 606
near-infrared, p. 609
rotation curve, p. 619
RR Lyrae variable, p. 607
Sagittarius A*, p. 626
self-propagating star formation, p. 624
spin (of a particle), p. 612
spin-flip transition, p. 613
spiral arm, p. 612
21-cm radio emission, p. 613
weakly interacting massive particle (WIMP), p. 619
winding dilemma, p. 621

Key Ideas

The Shape and Size of the Galaxy: Our Galaxy has a disk about 50 kpc (160,000 ly) in diameter and about 600 pc (2000 ly) thick, with a high concentration of interstellar dust and gas in the disk.

• The galactic center is surrounded by a large distribution of stars called the central bulge. This bulge is not perfectly symmetrical, but may have a bar or peanut shape.

• The disk of the Galaxy is surrounded by a spherical distribution of globular clusters and old stars, called the galactic halo.

• There are about 200 billion (2×10^{11}) stars in the Galaxy's disk, central bulge, and halo.

The Sun's Location in the Galaxy: Our Sun lies within the galactic disk, some 8000 pc (26,000 ly) from the center of the Galaxy.

• Interstellar dust obscures our view at visible wavelengths along lines of sight that lie in the plane of the galactic disk. As a result, the Sun's location in the Galaxy was unknown for many years. This dilemma was resolved by observing parts of the Galaxy outside the disk.

• The Sun orbits around the center of the Galaxy at a speed of about 790,000 km/h. It takes about 220 million years to complete one orbit.

The Rotation of the Galaxy and Dark Matter: From studies of the rotation of the Galaxy, astronomers estimate that the total mass of the Galaxy is about 10^{12} $M_{\odot}$. Only about 10% of this mass is in the form of visible stars, gas, and dust. The remaining 90% is in some nonvisible form, called dark matter, that extends beyond the edge of the luminous material in the Galaxy.

• Our Galaxy's dark matter may be a combination of MACHOs (dim, star-sized objects), massive neutrinos, and WIMPs (relatively massive subatomic particles).

The Galaxy's Spiral Structure: OB associations, H II regions, and molecular clouds in the galactic disk outline huge spiral arms.

• Spiral arms can be traced from the positions of clouds of atomic hydrogen. These can be detected throughout the galactic disk by the 21-cm radio waves emitted by the spin-flip transition in hydrogen. These emissions easily penetrate the intervening interstellar dust.

Theories of Spiral Structure: There are two leading theories of spiral structure in galaxies.

• According to the density-wave theory, spiral arms are created by density waves that sweep around the Galaxy. The gravitational field of this spiral pattern compresses the interstellar clouds through which it passes, thereby triggering the formation of the OB associations and H II regions that illuminate the spiral arms.

• According to the theory of self-propagating star formation, spiral arms are caused by the birth of stars over an extended region in a galaxy. Differential rotation of the galaxy stretches the star-forming region into an elongated arch of stars and nebulae.

The Galactic Nucleus: The innermost part of the Galaxy, or galactic nucleus, has been studied through its radio, infrared, and X-ray emissions (which are able to pass through interstellar dust).

• A strong radio source called Sagittarius A* is located at the galactic center. This marks the position of a supermassive black hole with a mass of about 3.7×10^6 $M_{\odot}$.

Questions

Review Questions

1. Why do the stars of the Galaxy appear to form a bright band that extends around the sky?
2. How did interstellar extinction mislead astronomers into believing that we are at the center of our Galaxy?
3. How did observations of globular clusters help astronomers determine our location in the Galaxy?
4. What are RR Lyrae stars? Why are they useful for determining the distance from our solar system to the center of the Galaxy?
5. Why are infrared telescopes useful for exploring the structure of the Galaxy? Why is it important to make observations at both near-infrared and far-infrared wavelengths?
6. The galactic halo is dominated by Population II stars, whereas the galactic disk contains predominantly Population I stars. In which of these parts of the Galaxy has star formation taken place recently? Explain.
7. O or B main-sequence stars are found in the galactic disk but not in globular clusters. Why is this so?
8. What must happen within a hydrogen atom for it to emit a photon of wavelength 21 cm?
9. Most interstellar hydrogen atoms emit only radio waves at a wavelength of 21 cm, but some hydrogen clouds emit profuse amounts of visible light (see, for example, Figure 18-1*b* and Figure 18-2). What causes this difference?
10. How do astronomers determine the distances to H I (neutral hydrogen) clouds?
11. The radio map in Figure 23-14 has a large gap on the side of the Galaxy opposite to ours. Why is this?
12. In a spiral galaxy, are stars in general concentrated in the spiral arms? Why are spiral arms so prominent in visible-light images of spiral galaxies?
13. Many classic black-and-white photographs of spiral galaxies were taken using film that was most sensitive to blue light. Explain why the spiral arms were particularly prominent in such photographs.
14. What kinds of objects (other than H I clouds) do astronomers observe to map out the Galaxy's spiral structure? What is special about these objects? Which of these can be observed at great distances?
15. Why don't astronomers detect 21-cm radiation from the hydrogen in giant molecular clouds?
16. In what way are the orbits of stars in the galactic disk different from the orbits of planets in our solar system? What does this difference imply about the way that matter is distributed in the Galaxy?
17. How do astronomers determine how fast the Sun moves in its orbit around the Galaxy? How does this speed tell us about the amount of mass inside the Sun's orbit? Does this speed tell us about the amount of mass outside the Sun's orbit?
18. How do astronomers conclude that vast quantities of dark matter surround our Galaxy? How is this dark matter distributed in space?
19. Another student tells you that the Milky Way Galaxy is made up "mostly of stars." Is this statement accurate? Why or why not?
20. What is the difference between dark matter and dark nebulae?
21. What proposals have been made to explain the nature of dark matter? What experiments or observations have been made to investigate these proposals? What are the results of this research?
22. What is the winding dilemma? What does it tell us about the nature of spiral arms?
23. Do density waves form a stationary pattern in a galaxy? If not, do they move more rapidly, less rapidly, or at the same speed as stars in the disk?
24. In our Galaxy, why are stars of spectral classes O and B only found in or near the spiral arms? Is the same true for stars of other spectral classes? Explain why or why not.
25. Compare the kinds of spiral arms produced by density waves with those produced by self-propagating star formation. By examining Figure 23-16, cite evidence that both processes may occur in our Galaxy.
26. What is the evidence that there is a supermassive black hole at the center of our Galaxy? How is it possible to determine the mass of this black hole?
27. What is the evidence that material has been falling into the supermassive black hole at the galactic center?

Advanced Questions

Questions preceded by an asterisk (*) involve topics discussed in the Boxes.

Problem-solving tips and tools

Several of the following questions make extensive use of Newton's form of Kepler's third law, and you might find it helpful to review Box 4-4. Another useful version of Kepler's third law is given in Section 17-9. Box 19-2 discusses the relationship between luminosity, apparent brightness, and distance. We discussed the relationship between the energy and wavelength of a photon in Section 5-5. According to the Pythagorean theorem, an isosceles right triangle has a hypotenuse that is longer than its sides by a factor of $\sqrt{2} = 1.414$. The formula for the Schwarzschild radius of a black hole is given in Box 22-2. You will find the small-angle formula in Box 1-1 useful. It is also helpful to remember that an object 1 AU across viewed at a distance of 1 parsec has an angular size of 1 arcsecond. Remember, too, that the volume of a cylinder is equal to its height multiplied by the area of its base, the area of a circle of radius r is πr^2, and the volume of a sphere of radius r is $4\pi r^3/3$. You can find other geometrical formulae in Appendix 8.

28. Discuss how the Milky Way would appear to us if the Sun were relocated to (**a**) the edge of the Galaxy; (**b**) the galactic halo; (**c**) the galactic bulge.

29. Explain why globular clusters spend most of their time in the galactic halo, even though their eccentric orbits take them close to the galactic center.

30. The disk of the Galaxy is about 50 kpc in diameter and 600 pc thick. (**a**) Find the volume of the disk in cubic parsecs. (**b**) Find the volume (in cubic parsecs) of a sphere 300 pc in radius centered on the Sun. (**c**) If supernovae occur randomly throughout the volume of the Galaxy, what is the probability that a given supernova will occur within 300 pc of the Sun? If there are about three supernovae each century in our Galaxy, how often, on average, should we expect to see one within 300 pc of the Sun?

***31.** An RR Lyrae star whose peak luminosity is 100 $L_\odot$ is in a globular cluster. At its peak luminosity, this star appears from Earth to be only 1.47×10^{-18} as bright as the Sun. Determine the distance to this globular cluster (**a**) in AU and (**b**) in parsecs.

32. A typical hydrogen atom in interstellar space undergoes a spin-flip transition only once every 10^7 years. How, then, is it at all possible to detect the 21-cm radio emission from interstellar hydrogen?

33. Calculate the energy of the photon emitted when a hydrogen atom undergoes a spin-flip transition. How many such photons would it take to equal the energy of a single H_α photon of wavelength 656.3 nm?

34. Suppose you were to use a radio telescope to measure the Doppler shift of 21-cm radiation in the plane of the Galaxy. (**a**) If you observe 21-cm radiation from clouds of atomic hydrogen at an angle of 45° from the galactic center, you will see the highest Doppler shift from a cloud that is as far from the galactic center as it is from the Sun. Explain this statement using a diagram. (**b**) Find the distance from the Sun to the particular cloud mentioned in (a).

35. Calculate approximately how many times our solar system has orbited the center of our Galaxy since the Sun and planets were formed 4.56×10^9 years ago.

36. Sketch the rotation curve you would obtain if the Galaxy were rotating like a rigid body.

37. The mass of our Galaxy interior to the Sun's orbit is calculated from the radius of the Sun's orbit and its orbital speed. By how much would this estimate be in error if the calculated distance to the galactic center were off by 10%? By how much would this estimate be in error if the calculated orbital velocity were off by 10%? Explain your reasoning.

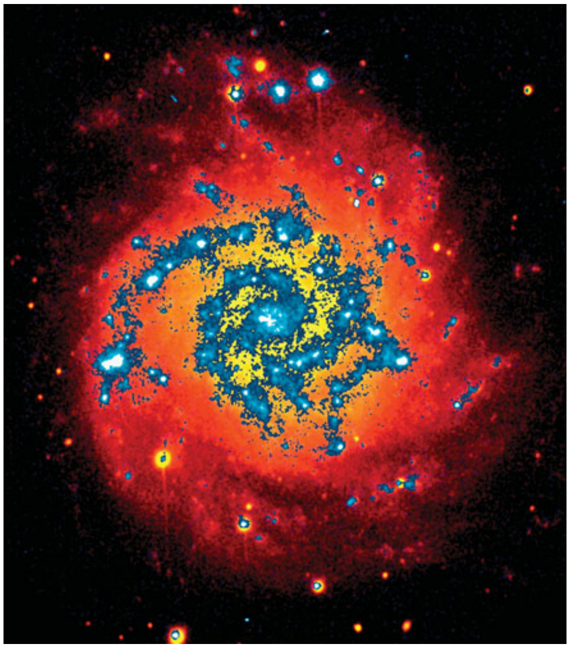
38. A gas cloud located in the spiral arm of a distant galaxy is observed to have an orbital velocity of 400 km/s. If the cloud is 20,000 pc from the center of the galaxy and is moving in a circular orbit, find (**a**) the orbital period of the cloud and (**b**) the mass of the galaxy contained within the cloud's orbit.

39. According to the Galaxy's rotation curve in Figure 23-18, a star 16 kpc from the galactic center has an orbital speed of about 270 km/s. Calculate the mass within that star's orbit.

40. Speculate on the reasons for the rapid rise in the Galaxy's rotation curve (see Figure 23-18) at distances close to the galactic center.

***41.** Show that the form of Kepler's third law stated in Box 23-2, $P^2 = 4\pi^2 a^3/G(M + M_\odot)$, is equivalent to $M = rv^2/G$, provided the orbit is a circle. (*Hint:* The mass of the Sun ($M_\odot$) is much less than the mass of the Galaxy inside the Sun's orbit (M).)

42. The image below shows the spiral galaxy M74, located about 55 million light-years from the Earth in the constellation



R I V U X G

(Astronomical Society of the Pacific)

Pisces (the Fish). It is actually a superposition of two false-color images: The red portion is an optical image taken at visible wavelengths, while the blue portion is an ultraviolet image made by NASA's Ultraviolet Imaging Telescope, which was carried into orbit by the space shuttle *Columbia* during the *Astro-1* mission in 1990. Compare the visible and ultraviolet images and, from what you know about stellar evolution and spiral structure, explain the differences you see.

43. The figure at the bottom of this page shows infrared images of two spiral galaxies. Explain which of these is a grand-design spiral galaxy and which is a flocculent spiral galaxy. Explain your reasoning.

***44.** (**a**) Calculate the Schwarzschild radius of a supermassive black hole of mass $3.7 \times 10^6\ M_\odot$, the estimated mass of the black hole at the galactic center. Give your answer in both kilometers and astronomical units. (**b**) What is the angular diameter of such a black hole as seen at a distance of 8 kpc, the distance from the Earth to the galactic center? Give your answer in arcseconds. Observing an object with such a small angular size will be a challenge indeed! (**c**) What is the angular diameter of such a black hole as seen from a distance of 45 AU, the closest that the star SO-16 comes to Sagittarius A*? Again, give your answer in arcseconds. Would it be discernible to the naked eye at that distance? (A normal human eye can see details as small as about 60 arcseconds.)

***45.** (**a**) The scale bar in Figure 23-27 shows that at the distance of Sagittarius A*, a length of 1600 AU has an angular size of 0.2 arcsecond. Use this information to calculate the distance to Sagittarius A*. (**b**) The star SO-16 moves around Sagittarius A* in an elliptical orbit with semimajor axis 1680 AU. Use this and the information given in the caption to Figure 23-27 to find the orbital period of SO-16. (**c**) Given the period and the semimajor axis of the star's orbit, is it possible to calculate the mass of SO-16 itself? If it is, explain how this could be done; if not, explain why not.

***46.** The stars S0-2 and S0-19 orbit Sagittarius A* with orbital periods of 14.5 and 37.3 years, respectively. (**a**) Assuming that the supermassive black hole in Sagittarius A* has a mass of $3.7 \times 10^6\ M_\odot$, determine the semimajor axes of the orbits of these two stars. Give your answers in AU. (**b**) Calculate the angular size of each orbit's semimajor axis as seen from Earth. (See Section 23-1 for the distance from the Earth to the center of the Galaxy.) Explain why extremely high-resolution infrared images are required to observe the motions of these stars.

47. Consider a star that orbits around Sagittarius A* in a circular orbit of radius 530 AU. (**a**) If the star's orbital speed is 2500 km/s, what is its orbital period? Give your answer in years. (**b**) Determine the sum of the masses of Sagittarius A* and the star. Give your answer in solar masses. (Your answer is an estimate of the mass of Sagittarius A*, because the mass of a single star is negligibly small by comparison.)

Discussion Questions

48. From what you know about stellar evolution, the interstellar medium, and the density-wave theory, explain the

(a)

(b)

R I V U X G

(NASA; JPL-Caltech; R. Kennicutt (University of Arizona); and the SINGS Team) appearance and structure of the spiral arms of grand-design spiral galaxies.

49. What observations would you make to determine the nature of the dark matter in our Galaxy's halo?

50. Describe how the appearance of the night sky might change if dark matter were visible to our eyes.

51. Discuss how a supermassive black hole could have formed at the center of our Galaxy.

Web/eBook Questions

52. Some scientists have suggested that the rotation curve of the Galaxy can be explained by modifying the laws of physics rather than by positing the existence of dark matter. Search the World Wide Web for information about these proposed modifications, called MOND (for *MO*dified *N*ewtonian *D*ynamics). Under what circumstances do MOND and conventional physics predict different behavior? What evidence is there that MOND might be correct?

53. Fast-Moving Stars at the Galactic Center. Access and view the video "Stars Orbiting Sagittarius A*" in Chapter 23 of the *Universe* Web site or eBook. Explain how you can tell which of the stars in the video are actually close to Sagittarius A* and which just happen to lie along our line of sight to the center of the Galaxy.

Activities

Observing Projects

54. Use the *Starry Night Enthusiast*™ program to observe the Milky Way. (**a**) Display the entire celestial sphere by selecting **Favourites > Guides > Atlas.** Select **View > Stars > Milky Way** to display this galaxy. Select **Options > Stars > Milky Way,** move the **Brightness** slide-bar to the far right to brighten the Milky Way and click **OK.** In the **View** menu, ensure that the **Scrollbars** are activated and use them to look at different parts of the Milky Way. Can you identify the direction toward the galactic nucleus? In this direction the Milky Way appears broadest. Open the **Find** pane, enter **Sagittarius** in the **Query** box and press **Enter** to center on this constellation to check your identification. (**b**) Use this full-sky view to determine the orientation of the plane of the Galaxy with respect to the celestial sphere. Move the vertical scrollbar to its central position to display the Celestial Equator as a horizontal line across the lower part of the view. Move the horizontal scrollbar until the Milky Way is centered upon the view. Estimate the angle between the Milky Way and the celestial equator on the screen. How well aligned is the plane of the Milky Way with the plane of the Earth's equator? (**c**) A third plane of interest is that of the ecliptic, which is shown as a green line. Use the scrollbars to adjust the view so that the ecliptic appears as a straight line rather than as a curve, thereby ensuring that you are viewing in a direction that lies in the ecliptic plane. Use the horizontal scrollbar to move the view to where you can see where the ecliptic crosses the Milky Way. Estimate the angle between the Milky Way and the ecliptic on the screen. How well aligned is the plane of the Milky Way to the ecliptic, the plane of the Earth's orbit around the Sun? (**d**) Click on **Home** in the toolbar to return to your home view, stop **Time Flow** and set the local time to midnight (12:00:00 A.M.). Select **Options > Stars > Milky Way,** move the **Brightness** slide-bar to the far right to brighten the Milky Way and click **OK.** Adjust the date to January 1, then February 1, and so on. In which month is the galactic nucleus highest in the sky at midnight, so that it is most easily seen from your location?

55. Use the *Starry Night Enthusiast*™ program to measure the dimensions of the Milky Way Galaxy. Select **Favourites > Stars > Sun in Milky Way** to move to a position 0.15 Mly above the galactic plane, directly above the position of the Sun. (If the Milky Way does not appear immediately, click once on the + Zoom button.) Remove the image of the astronaut's feet by clicking on **View > Feet.** You can rotate the Milky Way Galaxy by putting the mouse cursor over the image, holding down the **Shift** key and the mouse button, and moving the mouse. (**a**) Rotate the Galaxy so that you are seeing it face-on. Use the Hand Tool to measure the distance on the screen from the position of the Sun to the center of the Galaxy, noting this value in both linear distance in light-years and in the subtended angle when viewed from the observer's position. This measurement will calibrate angular measurements in terms of linear distance when seen from this viewpoint. (**b**) Use the Hand Tool to measure several distances from galactic center to the fringes of the galaxy. (Note: These distances will often be displayed in both angular and linear dimensions.) Estimate the diameter of the galaxy in light-years. (**c**) Rotate the Galaxy so that you are seeing it edge-on. Use the Hand Tool to measure the angle subtended from the center of the galaxy to the upper edge of the central bulge of the Galaxy. Use the above measurements of the Sun-Galactic Center in both angle and linear dimension and simple proportions to calculate the half-width of the galactic bulge and multiply this number in light years by 2 to determine the bulge diameter. You can use the relationship between parsecs and light-years (1 pc = 3.26 light-years) to calculate this diameter in parsecs. (**d**) Calculate the thickness-to-diameter ratio of the Milky Way Galaxy.

Collaborative Exercises

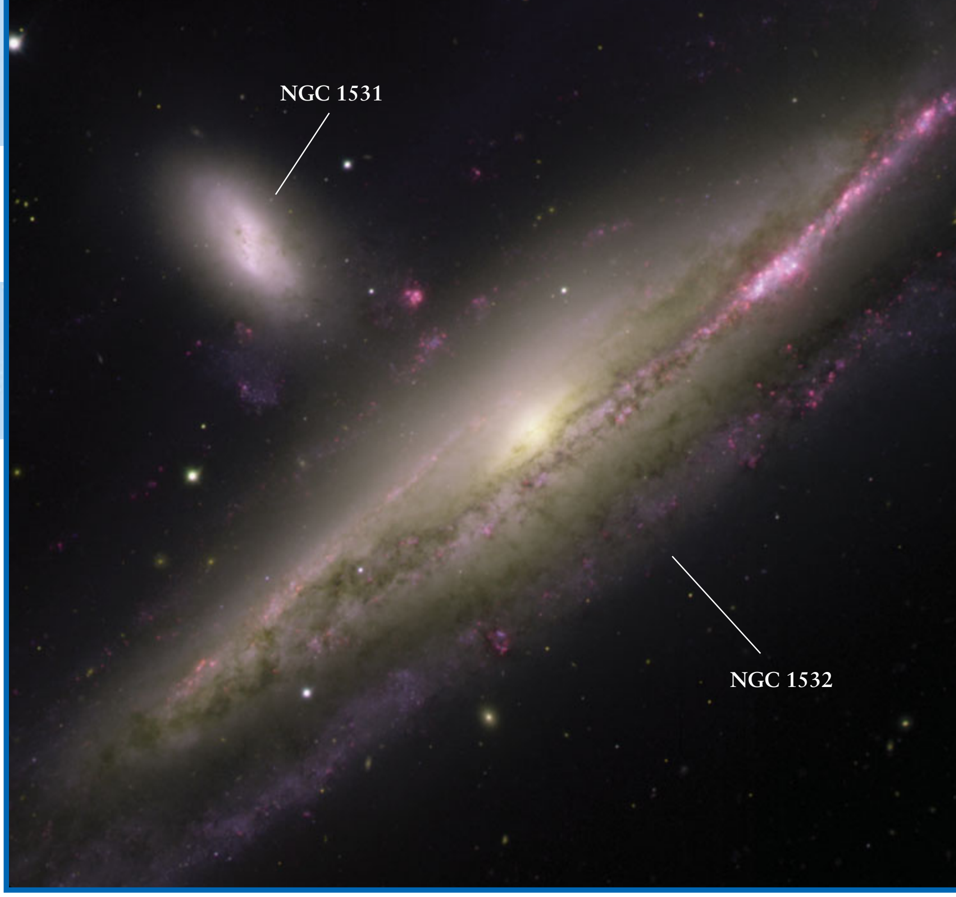
56. Student book bags often contain a wide collection of odd-shaped objects. Each person in your group should rummage through their own book bags and find one object that is most similar to the Milky Way Galaxy in shape. List the items from each group member's belongings and describe what about the items is similar to the shape of our Galaxy and what about the items is not similar, then indicate which of the items is the closest match.

57. One strategy for identifying a central location is called *triangulation.* In triangulation, a central position can be pinpointed by knowing the distance from each of three different places. First, on a piece of paper, create a rough map showing where each person in your group lives. Second, create a circle around each person's home that has a radius equal to the distance that each home is from your classroom. Label the place where the circles intersect as your classroom. Why can you not identify the position of the classroom with only two people's circles?

58. Figure 23-13 shows how emission spectra from hydrogen clouds would be shifted due to their motion around the Galaxy. Create a similar sketch showing an oval automobile racetrack with four cars moving on the track and a stationary observer outside the track at one end. Position and label the four moving cars all sounding their horns: (1) one that would have its horn sound shifted to longer wavelengths; (2) one that would have its horn sound shifted to shorter wavelengths; and (3) two cars moving in opposite directions so that their horn sounds would have no Doppler shift at all.

This page intentionally left blank

24
Galaxies



WEB LINK 24.1

RI V UXG

The two galaxies NGC 1531 and NGC 1532 are so close together that they exert strong gravitational forces on each other. Both galaxies are about 17 million pc (55 million ly) from us in the constellation Eridanus. (Gemini Observatory/Travis Rector, University of Alaska, Anchorage)

A century ago, most astronomers thought that the entire universe was only a few thousand light-years across and that nothing lay beyond our Milky Way Galaxy. One of the most important discoveries of the twentieth century was that this picture was utterly wrong. We now understand that the Milky Way is just one of billions of galaxies strewn across billions of light-years. The accompanying image shows two of them, denoted by rather mundane catalog numbers (NGC 1531 and NGC 1532) that give no hint to these galaxies' magnificence.

Some galaxies are spirals like NGC 1532 or the Milky Way, with arching spiral arms that are active sites of star formation. (The bright pink bands in NGC 1532 are H II regions, clouds of excited hydrogen that are set aglow by ultraviolet radiation from freshly formed massive stars.) Others, like NGC 1531, are featureless, ellipse-shaped agglomerations of stars, virtually devoid of interstellar gas and dust. Some galaxies are only one-hundredth the size and one ten-thousandth the mass of the Milky Way. Others are giants, with 5 times the size and 50 times the mass of the Milky Way. Only about 10% of a typical galaxy's mass emits radiation of any kind; the remainder is made up of the mysterious dark matter.

Just as most stars are found within galaxies, most galaxies are located in groups and clusters. These clusters of galaxies stretch across the universe, forming huge, lacy patterns. Remarkably, remote clusters of galaxies are receding from us; the greater their distance, the more rapidly they are moving away. This relationship between distance and recessional velocity, called the Hubble law, reveals that our immense universe is expanding. In Chapters 26 and 27 we will learn what this implies about the distant past and remote future of the universe.

Learning Goals

By reading the sections of this chapter, you will learn

24-1 How astronomers first observed other galaxies

24-2 How astronomers determined the distances to other galaxies

24-3 The basic types of galaxies

24-4 What techniques astronomers use to determine distances to remote galaxies

24-5 How the spectra of remote galaxies tell us that the universe is expanding

24-6 How galaxies are grouped into clusters and larger structures

24-7 What happens when galaxies collide

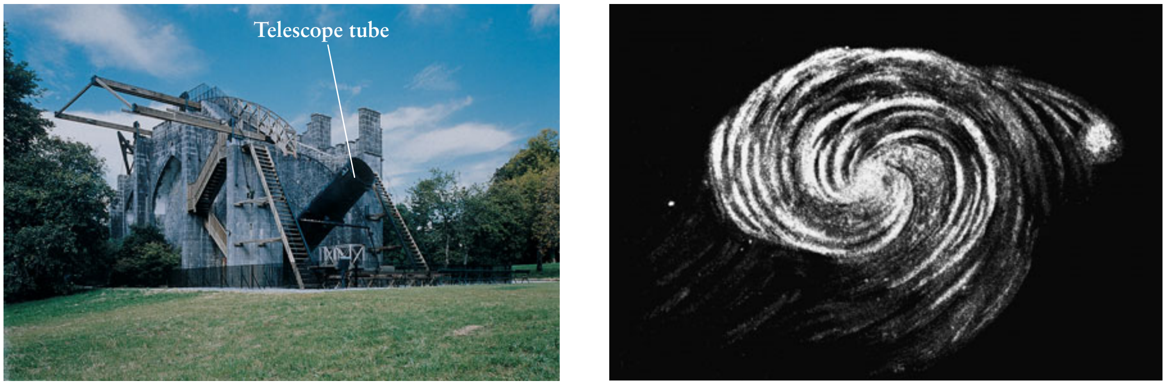
24-8 What observations indicate the presence of dark matter in other galaxies

24-9 How galaxies formed and evolved

24-1 When galaxies were first discovered, it was not clear that they lie far beyond the Milky Way

As early as 1755, the German philosopher Immanuel Kant suggested that vast collections of stars lie outside the confines of the Milky Way. Less than a century later, an Irish astronomer observed the structure of some of the "island universes" that Kant proposed.

William Parsons, the third Earl of Rosse, was a wealthy amateur astronomer who used his fortune to build immense telescopes. His pièce de résistance, completed in February 1845, had an objective mirror 1.8 meters (6 feet) in diameter (Figure 24-1). The mirror was mounted at one end of a 60-foot tube controlled by cables, straps, pulleys, and cranes (Figure 24-1*a*). For many



(a) Rosse's "Leviathan of Parsonstown" R I V U X G (b) M51 as viewed through the "Leviathan"

Figure 24-1

A Pioneering View of Another Galaxy (a) Built in 1845, this structure housed the largest telescope of its day. The telescope itself (the black cylinder pointing at a 45° angle above the horizontal) was restored to its original state during 1996–1998. **(b)** Using his telescope, Lord Rosse made this sketch of spiral structure in M51. This galaxy, whose angular size is 8 × 11 arcminutes (about a third the angular size of the full moon), is today called the Whirlpool Galaxy because of its distinctive appearance. (a: Birr Castle Demesne; b: Courtesy of Lund Humphries)

years, this triumph of nineteenth-century engineering was the largest telescope in the world.

With this new telescope, Rosse examined many of the nebulae that had been discovered and cataloged by William Herschel. He observed that some of these nebulae have a distinct spiral structure. One of the best examples is M51, also called NGC 5194. (The "M" designations of galaxies and nebulae come from a catalog compiled by the French astronomer Charles Messier between 1758 and 1782. The "NGC" designations come from the much more extensive "New General Catalogue" of galaxies and nebulae published in 1888 by J. L. E. Dreyer, a Danish astronomer who lived and worked in Ireland.)

Lacking photographic equipment, Lord Rosse had to make drawings of what he saw. Figure 24-1*b* shows a drawing he made of M51, which compares favorably with modern photographs (Figure 24-2). Views such as this inspired Lord Rosse to echo Kant's proposal that such nebulae are actually "island universes."

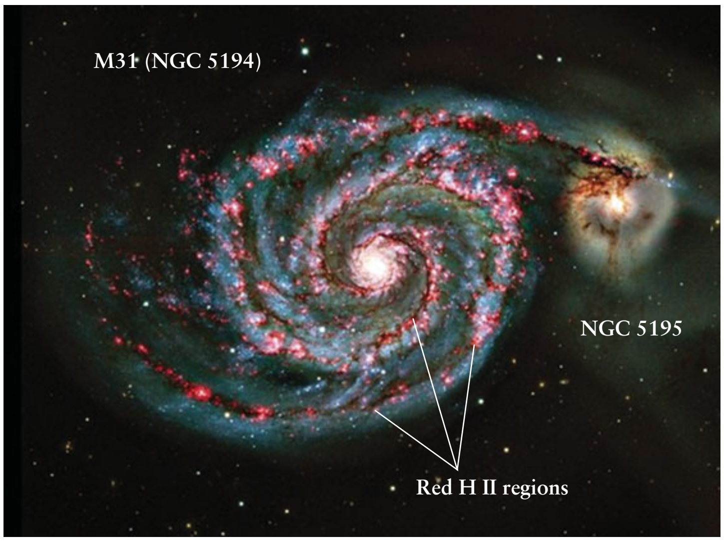
Many astronomers of the nineteenth century disagreed with this notion of island universes. A considerable number of nebulae are in fact scattered throughout the Milky Way. (Figures 18-2 and 18-4 show some examples.) It therefore seemed reasonable that "spiral nebulae," even though they are very different in shape from other sorts of nebulae, could also be components of our Galaxy.

As late as the 1920s it was unclear whether spiral nebulae were very remote "island universes" or simply nearby parts of our Galaxy

WEB LINK 24.3

The astronomical community became increasingly divided over the nature of the spiral nebulae. In April 1920, two opposing ideas were presented before the National Academy of Sciences in Washington, D.C. On one side was Harlow Shapley from the Mount Wilson Observatory, renowned for his recent determination of the size of the Milky Way Galaxy (see Section 23-1). Shapley thought the spiral nebulae were relatively small, nearby objects scattered around our Galaxy like the globular clusters he had studied. Opposing Shapley was Heber D. Curtis of the University of California's Lick Observatory. Curtis championed the island universe theory, arguing that each of these spiral nebulae is a rotating system of stars much like our own Galaxy.

The Shapley-Curtis "debate" generated much heat but little light. Nothing was decided, because no one could present conclusive evidence to demonstrate exactly how far away the spiral nebulae are. Astronomy desperately needed a definitive determination of the distance to a spiral nebula. Such a measurement became the first great achievement of a young man who studied astronomy at the Yerkes Observatory, near Chicago. His name was Edwin Hubble.



WEB LINK 24.2

Figure 24-2 R I V U X G

A Modern View of the Spiral Galaxy M51 This galaxy, also called NGC 5194, has spiral arms that are outlined by glowing H II regions. These reveal the sites of star formation (see Section 18-2). One spiral arm extends toward the companion galaxy NGC 5195. Figure 24-23*a* shows these same galaxies at infrared wavelengths. (CFHT)

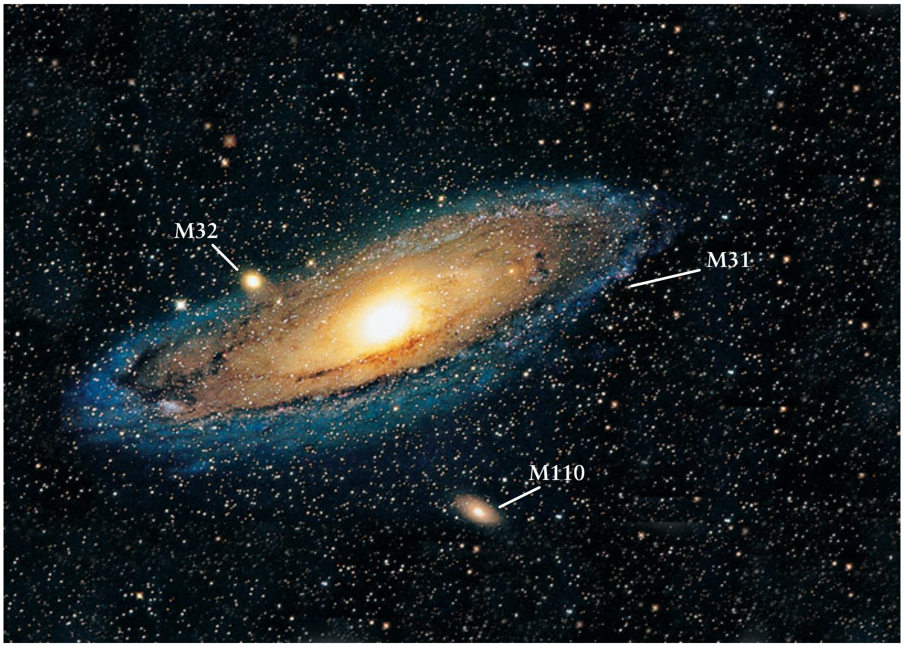


Figure 24-3 R I V U X G

The Andromeda "Nebula" The Great Nebula in Andromeda, also known as M31, can be seen with even a small telescope. Edwin Hubble was the first to demonstrate that M31 is actually a galaxy that lies far beyond the Milky Way. M32 and M110 are two small satellite galaxies that orbit M31. (Bill & Sally Fletcher/Tom Stack & Associates)

24-2 Hubble proved that the spiral nebulae are far beyond the Milky Way

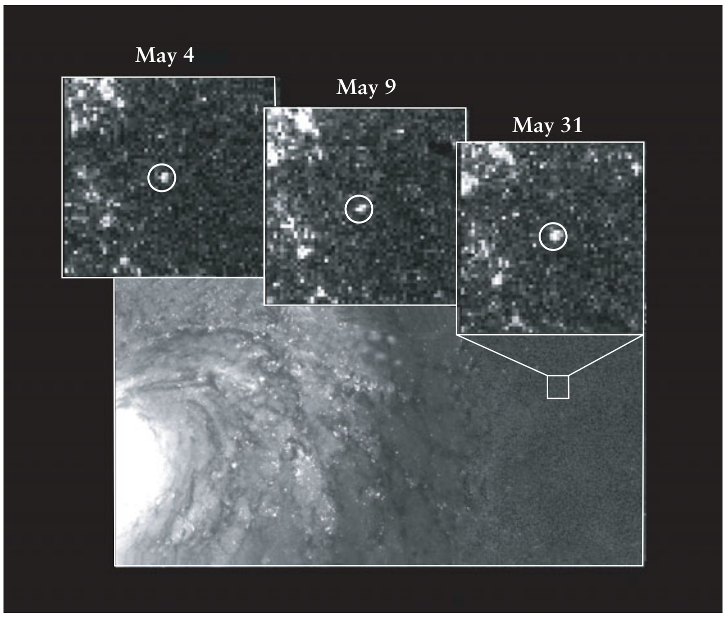
After completing his studies, Edwin Hubble joined the staff of the Mount Wilson Observatory in Pasadena, California. On October 6, 1923, he took a historic photograph of the Andromeda "Nebula," one of the spiral nebulae around which controversy raged. (Figure 24-3 is a modern photograph of this object.)

Hubble carefully examined his photographic plate and discovered what he at first thought to be a nova. Referring to previous plates of that region, he soon realized that the object was actually a Cepheid variable star. Further scrutiny of additional plates over the next several months revealed several more Cepheids. Figure 24-4 shows modern observations of a Cepheid in another "spiral nebula."

Just as RR Lyrae variable stars demonstrated the extent of the Milky Way, Cepheid variables revealed the immense distances to other galaxies

As we saw in Section 19-6, Cepheid variables help astronomers determine distances. An astronomer begins by carefully measuring the variations in apparent brightness of a Cepheid variable, then recording the results in the form of a plot of brightness versus time, or light curve (see Figure 19-18*a*). This graph gives the variable star's period and average brightness. Given the star's period, the astronomer then uses the period-luminosity relation shown in Figure 19-19 to find the Cepheid's average luminosity. (The astronomer must also examine the spectrum of the star to determine whether it is a metal-rich Type I Cepheid or a metal-poor Type II Cepheid. As Figure 19-19 shows, these have somewhat different period-luminosity relations.) Knowing both the apparent brightness and luminosity of the Cepheid, the astronomer can then use the inverse-square law to calculate the distance to the star (see Box 17-2). Box 24-1 presents an example of how this is done. This procedure is very similar to that used by Harlow Shapley to measure the distances to the Milky Way's globular clusters using RR Lyrae variable stars (see Section 23-1).

Cepheid variables are intrinsically quite luminous, with average luminosities that can exceed 10^4 $L_{\odot}$. Hubble realized that for these luminous stars to appear as dim as they were on his



VIDEO 24.1

Figure 24-4 R I V U X G

Measuring Galaxy Distances with Cepheid Variables By observing Cepheid variable stars in M100, the galaxy shown here, astronomers have found that it is about 17 Mpc (56 million ly) from Earth. The insets show one of the Cepheids in M100 at different stages in its brightness cycle, which lasts several weeks. (Wendy L. Freedman, Carnegie Institution of Washington, and NASA)

BOX 24-1 Tools of the Astronomer's Trade

Cepheids and Supernovae as Indicators of Distance

Because their periods are directly linked to their luminosities, Cepheid variables are one of the most reliable tools astronomers have for determining the distances to galaxies. To this day, astronomers use this link—much as Hubble did back in the 1920s—to measure intergalactic distances. More recently, they have begun to use Type Ia supernovae, which are far more luminous and thus can be seen much farther away, to determine the distances to very remote galaxies.

EXAMPLE: In 1992 a team of astronomers observed Cepheid variables in a galaxy called IC 4182 to deduce this galaxy's distance from the Earth. One such Cepheid has a period of 42.0 days and an average apparent magnitude (m) of +22.0. (See Box 19-3 for an explanation of the apparent magnitude scale.) By comparison, the dimmest star you can see with the naked eye has $m = +6$; this Cepheid in IC 4182 appears less than one one-millionth as bright. The star's spectrum shows that it is a metal-rich Type I Cepheid variable.

According to the period-luminosity relation shown in Figure 21-17, such a Type I Cepheid with a period of 42.0 days has an average luminosity of 33,000 $L_\odot$. An equivalent statement is that this Cepheid has an average absolute magnitude (M) of −6.5. (This compares to $M = +4.8$ for the Sun.) Use this information to determine the distance to IC 4182.

Situation: We are given the apparent magnitude $m = +22.0$ and the absolute magnitude $M = -6.5$ of the Cepheid variable star in IC 4182. Our goal is to calculate the distance to this star, and hence the distance to the galaxy of which it is part.

Tools: We use the relationship between apparent magnitude, absolute magnitude, and distance given in Box 19-3.

Answer: In Box 19-3, we saw that the apparent magnitude of a star is related to its absolute magnitude and distance in parsecs (d) by

$$m - M = 5 \log d - 5$$

This can be rewritten as

$$d = 10^{(m-M+5)/5} \text{ parsecs}$$

We have $m - M = (+22.0) - (-6.5) = 22.0 + 6.5 = 28.5$. (Recall from Box 19-3 that $m - M$ is called the *distance modulus.*) Hence, our equation becomes

$$d = 10^{(28.5+5)/5} \text{ parsecs} = 10^{6.7} \text{ parsecs} = 5 \times 10^6 \text{ parsecs}$$

(A calculator is needed to calculate the quantity $10^{6.7}$.)

Review: Our result tells us that the galaxy is 5 million parcsecs, or 5 Mpc, from Earth (1 Mpc = 10^6 pc). This distance can also be expressed as 16 million light-years.

EXAMPLE: Astronomers are interested in IC 4182 because a Type Ia supernova was observed there in 1937. All Type Ia supernovae are exploding white dwarfs that reach nearly the same maximum brightness at the peak of their outburst (see Section 22-9). Once astronomers know the peak absolute magnitude of Type Ia supernovae, they can use these supernovae as distance indicators. Because the distance to IC 4182 is known from its Cepheids, the 1937 observations of the supernova in that galaxy help us calibrate Type Ia supernovae as distance indicators.

At maximum brightness, the 1937 supernova reached an apparent magnitude of $m = +8.6$. What was its absolute magnitude at maximum brightness?

Situation: We are given the supernova's apparent magnitude m, and we know its distance from the previous example. Our goal is to calculate its absolute magnitude M.

Tools: We again use the relationship $m - M = 5 \log d - 5$.

Answer: We could plug in the value of d found in the previous example. But it is simpler to note that the distance modulus $m - M$ has the same value no matter whether it refers to a Cepheid, a supernova, or any other object, just so it is at the same distance d. From the Cepheid example we have $m - M = 28.5$ for IC 4182, so

$$M = m - (m - M) = 8.6 - (28.5) = -19.9$$

This absolute magnitude corresponds to a remarkable peak luminosity of 10^{10} $L_\odot$.

Review: Whenever astronomers find a Type Ia supernova in a remote galaxy, they can combine this absolute magnitude with the observed maximum apparent magnitude to get the galaxy's distance modulus, from which the galaxy's distance can be easily calculated (just as we did above for the Cepheids in IC 4182). This technique has been used to determine the distances to galaxies hundreds of millions of parsecs away (see Section 26-4).

photographs of the Andromeda "Nebula," they must be extremely far away. Straightforward calculations using modern data reveal that M31 is some 750 kiloparsecs (2.5 million light-years) from Earth. Based on its angular size, M31 has a diameter of 70 kiloparsecs—larger than the diameter of our own Milky Way Galaxy!

These results prove that the Andromeda "Nebula" is actually an enormous stellar system, far beyond the confines of the Milky Way. Today, this system is properly called the Andromeda Galaxy. (Under good observing conditions, you can actually see this galaxy's central bulge with the naked eye. If you could see the entire

Andromeda Galaxy, it would cover an area of the sky roughly 5 times as large as the full moon.) Galaxies are so far away that their distances from us are usually given in millions of parsecs, or *megaparsecs* (Mpc): 1 Mpc = 10^6 pc. For example, the distance to the galaxies in the image that opens this chapter is 17 Mpc.

Hubble's results, which were presented at a meeting of the American Astronomical Society on December 30, 1924, settled the Shapley-Curtis "debate" once and for all. The universe was recognized to be far larger and populated with far bigger objects than anyone had seriously imagined. Hubble had discovered the realm of the galaxies.

CAUTION! In everyday language, many people use the words "galaxy" and "universe" interchangeably. It is true that before Hubble's discoveries our Milky Way Galaxy was thought to constitute essentially the entire universe. But we now know that the universe contains literally billions of galaxies. A single galaxy, vast though it may be, is just a tiny part of the entire observable universe.

24-3 Galaxies are classified according to their appearance

Millions of galaxies are visible across every unobscured part of the sky. Although all galaxies are made up of large numbers of stars, they come in a variety of shapes and sizes.

WEB LINK 24.4

Hubble classified galaxies into four broad categories based on their appearance. These categories form the basis for the **Hubble classification,** a scheme that is still used today. The four classes of galaxies are the spirals, classified S; barred spirals, or SB; ellipticals, E; and irregulars, Irr. Table 24-1 summarizes some key properties of each class. These various types of galaxies differ not only in their shapes but also in the kinds of processes taking place within them.

Spiral Galaxies: Stellar Birthplaces

VIDEO 24.2

M51 (Figure 24-2), M31 (Figure 24-3), and M100 (Figure 24-4) are examples of **spiral galaxies.** Figure 24-5 shows that spiral galaxies are characterized by arched lanes of stars, just as is our own Milky Way Galaxy (see Section 23-3). The spiral arms contain young, hot, blue stars and their associated H II regions, indicating ongoing star formation.

Thermonuclear reactions within stars create *metals,* that is, elements heavier than hydrogen or helium (see Section 17-5). These metals are dispersed into space as the stars evolve and die. So, if new stars are being formed from the interstellar matter in spiral galaxies, they will incorporate these metals and be Population I stars (see Section 19-5 and Section 23-2). Indeed, the visible-light spectrum of the disk of a spiral galaxy has strong metal absorption lines. Such a spectrum is a composite of the spectra of many stars and shows that the stars in the disk are principally of Population I. By contrast, there is relatively little star formation in the central bulges of spiral galaxies, and these regions are dominated by old Population II stars that have a low metal content. This also explains why the central bulges of spiral galaxies have a yellowish or reddish color; as a population of stars ages, the massive, luminous blue stars die off first, leaving only the longer-lived, low-mass red stars.

Hubble further classified spiral galaxies according to the texture of the spiral arms and the relative size of the central bulge. Spirals with smooth, broad spiral arms and a fat central bulge are called Sa galaxies, for spiral type *a* (Figure 24-5*a*); those with moderately well-defined spiral arms and a moderate-sized central bulge, like M31 and M51, are Sb galaxies (Figure 24-5*b*); and galaxies with narrow, well-defined spiral arms and a tiny central bulge are Sc galaxies (Figure 24-5*c*).

The differences between Sa, Sb, and Sc galaxies may be related to the relative amounts of gas and dust they contain. Observations with infrared telescopes (which detect the emission from interstellar dust) and radio telescopes (which detect radiation from interstellar gases such as hydrogen and carbon monoxide) show that about 4% of the mass of a Sa galaxy is in the form of gas and dust. This percentage is 8% for Sb galaxies and 25% for Sc galaxies.

Interstellar gas and dust is the material from which new stars are formed, so an Sc galaxy has a greater proportion of its mass involved in star formation than an Sb or Sa galaxy. Hence, a Sc galaxy has a large disk (where star formation occurs) and a small central bulge (where there is little or no star formation). By comparison, a Sa galaxy, which has relatively little gas and dust, and

INTERACTIVE EXERCISE 24.1

Table 24-1 Some Properties of Galaxies

	Spiral (S) and barred spiral (SB) galaxies	Elliptical galaxies (E)	Irregular galaxies (Irr)
Mass ($M_\odot$)	10^9 to 4×10^{11}	10^5 to 10^{13}	10^8 to 3×10^{10}
Luminosity ($L_\odot$)	10^8 to 2×10^{10}	3×10^5 to 10^{11}	10^7 to 10^9
Diameter (kpc)	5 to 250	1 to 200	1 to 10
Stellar populations	Spiral arms: young Population I Nucleus and throughout disk: Population II and old Population I	Population II and old Population I	mostly Population I
Percentage of observed galaxies	77%	20%*	3%

This percentage does not include dwarf elliptical galaxies that are as yet too dim and distant to detect. Hence, the actual percentage of galaxies that are ellipticals may be higher than shown here.

(a) Sa (NGC 1357) (b) Sb (M81) (c) Sc (NGC 4321)

Figure 24-5 R I V U X G

Spiral Galaxies Edwin Hubble classified spiral galaxies according to the texture of their spiral arms and the relative size of their central bulges. Sa galaxies have smooth, broad spiral arms and the largest central bulges, while Sc galaxies have narrow, well-defined arms and the smallest central bulges. (a: STScI Digital Sky Survey; b: N. A. Sharp/AURA/NOAO/NSF; c: AngloAustralian Observatory)

thus less material from which to form stars, has a large central bulge and only a small star-forming disk.

Barred Spiral Galaxies: Spirals with an Extra Twist

In **barred spiral galaxies,** such as those shown in Figure 24-6, the spiral arms originate at the ends of a bar-shaped region running through the galaxy's nucleus rather than from the nucleus itself. As with ordinary spirals, Hubble subdivided barred spirals according to the relative size of their central bulge and the character of their spiral arms. A SBa galaxy has a large central bulge and thin, tightly wound spiral arms (Figure 24-6*a*). Likewise, a SBb galaxy is a barred spiral with a moderate central bulge and moderately wound spiral arms (Figure 24-6*b*), while a SBc galaxy has lumpy, loosely wound spiral arms and a tiny central bulge (Figure 24-6*c*). As for ordinary spiral galaxies, the difference between SBa, SBb, and SBc galaxies may be related to the amount of gas and dust in the galaxy.

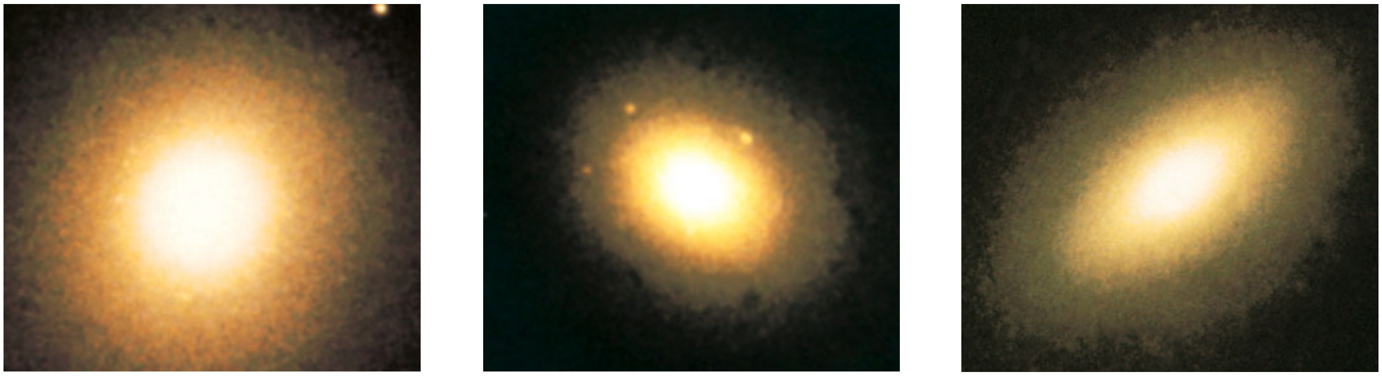
Bars appear to form naturally in many spiral galaxies. This conclusion comes from computer simulations of galaxies, which set hundreds of thousands of simulated "stars" into orbit around a common center. As the "stars" orbit and exert gravitational forces on one another, a bar structure forms in most cases. Indeed, barred spiral galaxies outnumber ordinary spirals by about two to one. (As we saw in Section 23-2, there is evidence that the Milky Way Galaxy is a barred spiral.)

Why don't all spiral galaxies have bars? According to calculations by Jeremiah Ostriker and P. J. E. Peebles of Princeton Uni-

(a) SBa (NGC 4650) (b) SBb (M83) (c) SBc (NGC 1365)

Figure 24-6 R I V U X G

Barred Spiral Galaxies As with spiral galaxies, Hubble classified barred spirals according to the texture of their spiral arms (which correlates to the sizes of their central bulges). SBa galaxies have the smoothest spiral arms and the largest central bulges, while SBc galaxies have narrow, well-defined arms and the smallest central bulges. (a: AURA/NOAO/NSF; b, c: Anglo-Australian Observatory)



(a) E0 (M105) (b) E3 (NGC 4365) (c) E6 (NGC 3377)

Figure 24-7 R I V U X G

Elliptical Galaxies Hubble classified elliptical galaxies according to how round or flattened they look. A galaxy that appears round is labeled E0, and the flattest-appearing elliptical galaxies are designated E7. (Courtesy of J. D. Wray; McDonald Observatory)

versity, a bar will not develop if a galaxy is surrounded by a sufficiently massive halo of nonluminous *dark matter*. (In Section 23-4 we saw evidence that our Milky Way Galaxy is surrounded by such a dark matter halo.) The difference between barred spirals and ordinary spirals may thus lie in the amount of dark matter the galaxy possesses. In Section 24-8 we will see compelling evidence for the existence of dark matter in spiral galaxies.

Elliptical Galaxies: From Giants to Dwarfs

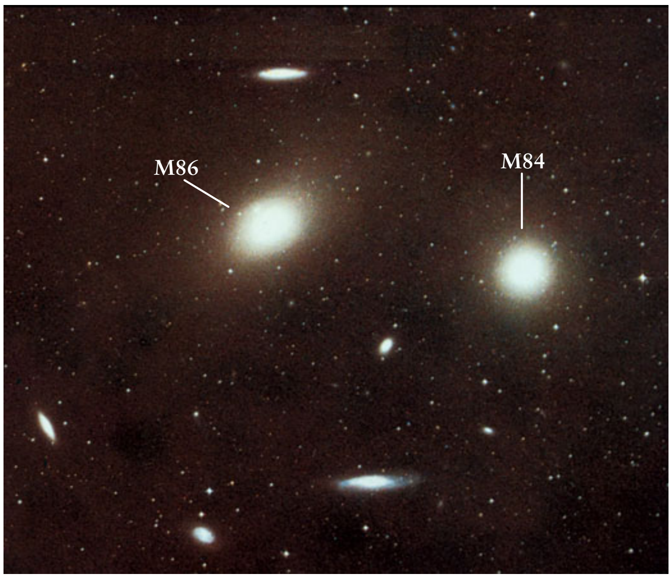
Elliptical galaxies, so named because of their distinctly elliptical shapes, have no spiral arms. Hubble subdivided these galaxies according to how round or flattened they look. The roundest elliptical galaxies are called E0 galaxies and the flattest, E7 galaxies. Elliptical galaxies with intermediate amounts of flattening are given designations between these extremes (Figure 24-7).

CAUTION! Unlike the designations for spirals and barred spirals, the classifications E0 through E7 may not reflect the true shape of elliptical galaxies. An E1 or E2 galaxy might actually be a very flattened disk of stars that we just happen to view face-on, and a cigar-shaped E7 galaxy might look spherical if seen end-on. The Hubble scheme classifies galaxies entirely by how they *appear* to us on the Earth.

Elliptical galaxies look far less dramatic than their spiral and barred spiral cousins. The reason, as shown by radio and infrared observations, is that ellipticals are virtually devoid of interstellar gas and dust. Consequently, there is little material from which stars could have recently formed, and indeed there is no evidence of young stars in most elliptical galaxies. For the most part, star formation in elliptical galaxies ended long ago. Hence, these galaxies are composed of old, red, Population II stars with only small amounts of metals.

Elliptical galaxies come in a wide range of sizes and masses. Both the largest and the smallest galaxies in the known universe are elliptical. Figure 24-8 shows two **giant elliptical galaxies** that are about 20 times larger than an average galaxy. These giant ellipticals are located near the middle of a large cluster of galaxies in the constellation Virgo. (We will discuss this and other clusters of galaxies in Section 24-6.)

Giant ellipticals are rather rare, but **dwarf elliptical galaxies** are quite common. Dwarf ellipticals are only a fraction the size of their normal counterparts and contain so few stars—only a few million, compared to more than 100 billion (10^{11}) stars in our Milky Way Galaxy—that these galaxies are completely transparent. You can actually see straight through the center of a dwarf galaxy and out the other side, as Figure 24-9 shows.



WEB LINK 24.5

Figure 24-8 R I V U X G

Giant Elliptical Galaxies The Virgo cluster is a rich, sprawling collection of more than 2000 galaxies about 17 Mpc (56 million ly) from Earth. Only the center of this huge cluster appears in this photograph. The two largest members of this cluster are the giant elliptical galaxies M84 and M86. These galaxies have angular sizes of 5 to 7 arcmin. (Royal Observatory, Edinburgh)

Figure 24-9 R I V U X G

A Dwarf Elliptical Galaxy This diffuse cloud of stars is a nearby E4 dwarf elliptical called Leo I. It actually orbits the Milky Way at a distance of about 180 kpc (600,000 ly). Leo I is about 1 kpc (3000 ly) in diameter but contains so few stars that you can see through the galaxy's center. (Anglo-Australian Observatory)

The visible light from a galaxy is emitted by its stars, so the visible-light spectrum of a galaxy has absorption lines. But because a galaxy's stars are in motion, with some approaching us and others moving away, the Doppler effect smears out and broadens the absorption lines. The average motions of stars in a galaxy can be deduced from the details of this broadening.

For elliptical galaxies, studies of this kind show that star motions are quite random. In a very round (E0) elliptical galaxy, this randomness is **isotropic,** meaning "equal in all directions." Because the stars are whizzing around equally in all directions, the galaxy is genuinely spherical. In a flattened (E7) elliptical galaxy, the randomness of the stellar motions is **anisotropic,** which means that the range of star speeds is different in different directions.

Hubble also identified galaxies that are midway in appearance between ellipticals and the two kinds of spirals. These are denoted as S0 and SB0 galaxies, also called **lenticular galaxies.** Although they look somewhat elliptical, lenticular ("lens-shaped") galaxies have both a central bulge and a disk like spiral galaxies, but no discernible spiral arms. They are therefore sometimes referred to as "armless spirals." Figure 24-10 shows an example of an SB0 lenticular galaxy.

Edwin Hubble summarized his classification scheme for spiral, barred spiral, and elliptical galaxies in a diagram, now called the **tuning fork diagram** for its shape (Figure 24-11).

CAUTION! When Hubble first drew his tuning fork diagram, he had the idea that it represented an evolutionary sequence. He thought that galaxies evolved over time from the left to the right of the diagram, beginning as ellipticals and eventually becoming either spiral or barred spiral galaxies. We now understand that this is not the case at all! For one thing, elliptical galaxies have little or no overall rotation, while spiral and barred spiral galaxies have a substantial amount of overall rotation. There is no way that an elliptical galaxy could suddenly start rotating, which means that it could not evolve into a spiral galaxy.

A more modern interpretation of the Hubble tuning fork diagram is that it is an arrangement of galaxies according to their overall rotation. A rapidly rotating collection of matter in space tends to form a disk, while a slowly rotating collection does not. Thus, the elliptical galaxies at the far left of the tuning fork diagram have little internal rotation and hence no disk. Sa and SBa galaxies have enough overall rotation to form a disk, though their central bulges are still dominant. The galaxies with the greatest amount of overall rotation are Sc and SBc galaxies, in which the central bulges are small and most of the gas, dust, and stars are in the disk.

Irregular Galaxies: Deformed and Dynamic

Galaxies that do not fit into the scheme of spirals, barred spirals, and ellipticals are usually referred to as **irregular galaxies.** They are generally rich in interstellar gas and dust, and have both young and old stars. For lack of any better scheme, the irregular galaxies are sometimes placed between the ends of the tines of the Hubble tuning fork diagram, as in Figure 24-11.

Figure 24-10 R I V U X G

A Lenticular Galaxy NGC 2787 is classified as a lenticular galaxy because it has a disk but no discernible spiral arms. Its nucleus displays a faint bar (not apparent in this image), so NGC 2787 is denoted as an SB0 galaxy. It lies about 7.4 Mpc (24 million ly) from Earth in the constellation Ursa Major. (NASA and The Hubble Heritage Team, STScI/AURA)

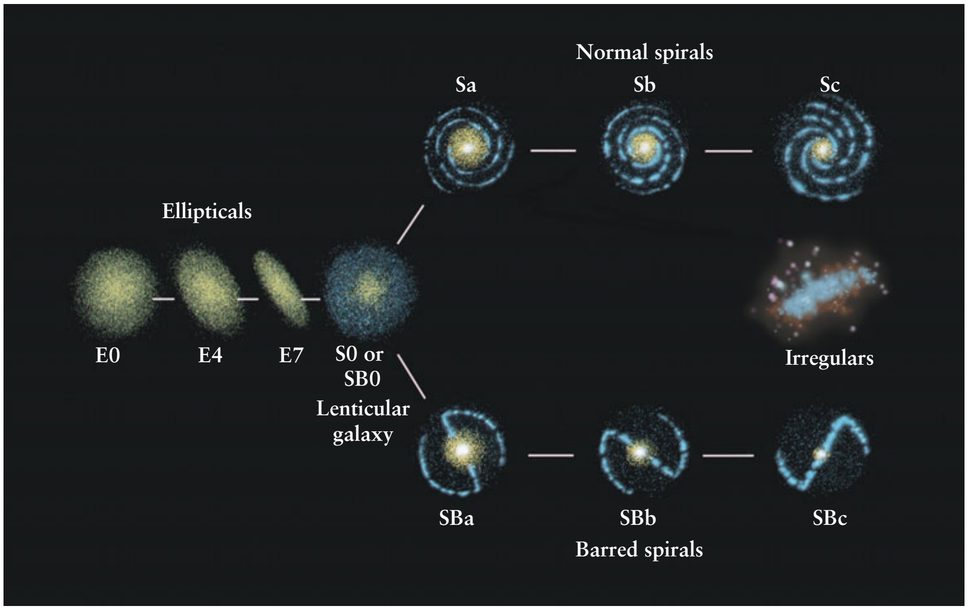
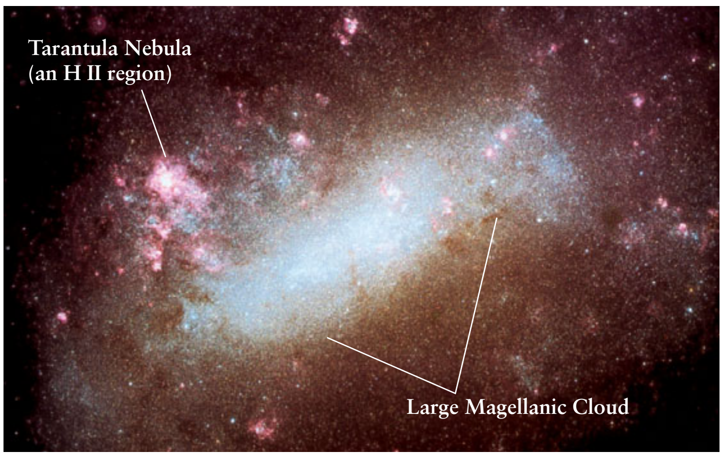


Figure 24-11

Hubble's Tuning Fork Diagram Edwin Hubble's classification of regular galaxies is shown in his tuning fork diagram. An elliptical galaxy is classified by how flattened it appears. A spiral or barred spiral galaxy is classified by the texture of its spiral arms and the size of its central bulge. A lenticular galaxy, is intermediate between ellipticals and spirals. Irregular galaxies do not fit into this simple classification scheme.

Hubble defined two types of irregulars. Irr I galaxies have only hints of organized structure, and have many OB associations and H II regions. The best-known examples of Irr I galaxies are the Large Magellanic Cloud (Figure 24-12) and the Small Magellanic Cloud. Both are nearby companions of our Milky Way and can be seen with the naked eye from southern latitudes. Both these galaxies contain substantial amounts of interstellar gas. Tidal forces exerted on these irregular galaxies by the Milky Way help to compress the gas, which is why both of the Magellanic Clouds are sites of active star formation.



WEB LINK 24.6

Figure 24-12 R I V U X G

The Large Magellanic Cloud (LMC) At a distance of only 55 kpc (179,000 ly), this Irr I galaxy is the third closest known companion of our Milky Way Galaxy. About 19 kpc (62,000 ly) across, the LMC spans 22° across the sky, or about 50 times the size of the full moon. One sign that star formation is ongoing in the LMC is the Tarantula Nebula, whose diameter of 250 pc (800 ly) and mass of $5 \times 10^6\ M_\odot$ make it the largest known H II region. (Anglo-Australian Observatory)

The other type of irregulars, called Irr II galaxies, have asymmetrical, distorted shapes that seem to have been caused by collisions with other galaxies or by violent activity in their nuclei. M82, shown in Figure 22-17 and in the image that opens Chapter 6, is an example of an Irr II galaxy.

24-4 Astronomers use various techniques to determine the distances to remote galaxies

WEB LINK 24.7

A key question that astronomers ask about galaxies is "How far away are they?" Unfortunately, many of the techniques that are used to measure distances within our Milky Way Galaxy cannot be used for the far greater distances to other galaxies. The extremely accurate parallax method that we described in Section 17-1 can be used only for stars within about 500 pc. Beyond that distance, parallax angles become too small to measure. Spectroscopic parallax, in which the distance to a star or star cluster is found with the help of the H-R diagram (see Section 17-8), is accurate only out to roughly 10 kpc from Earth; more distant stars or clusters are too dim to give reliable results.

The various methods of distance determination are interrelated because one is used to calibrate another

Standard Candles: Variable Stars and Type Ia Supernovae

To determine the distance to a remote galaxy, astronomers look instead for a **standard candle**—an object, such as a star, that lies within that galaxy and for which we know the luminosity (or, equivalently, the absolute magnitude, described in Section 17-3). By measuring how bright the standard candle appears, astronomers can calculate its distance—and hence the distance to the galaxy of which it is part—using the inverse-square law.

The challenge is to find standard candles that are luminous enough to be seen across the tremendous distances to galaxies. To be useful, standard candles should have four properties:

1. They should be luminous, so we can see them out to great distances.

2. We should be fairly certain about their luminosities, so we can be equally certain of any distance calculated from a standard candle's apparent brightness and luminosity.

3. They should be easily identifiable—for example, by the shape of the light curve of a variable star.

4. They should be relatively common, so that astronomers can use them to determine the distances to many different galaxies.

For nearby galaxies, Cepheid variable stars make reliable standard candles. These variables can be seen out to about 30 Mpc (100 million ly) using the Hubble Space Telescope, and their luminosity can be determined from their period through the period-luminosity relation depicted in Figure 19-19. Box 24-1 gives an example of using Cepheid variables to determine distances. RR Lyrae stars, which are Population II variable stars often found in globular clusters, can be used as standard candles in a similar way. (We saw in Section 25-1 how RR Lyrae variables helped determine the size of our Galaxy.) Because they are less luminous than Cepheids, RR Lyrae variables can be seen only out to 100 kpc (300,000 ly).

Beyond about 30 Mpc even the brightest Cepheid variables, which have luminosities of about $2 \times 10^4\ L_\odot$, fade from view. Astronomers have tried to use even more luminous stars such as blue supergiants to serve as standard candles. However, this idea is based on the assumption that there is a fixed upper limit on the luminosities of stars, which may not be the case. Hence, these standard candles are not very "standard," and distances measured in this way are somewhat uncertain.

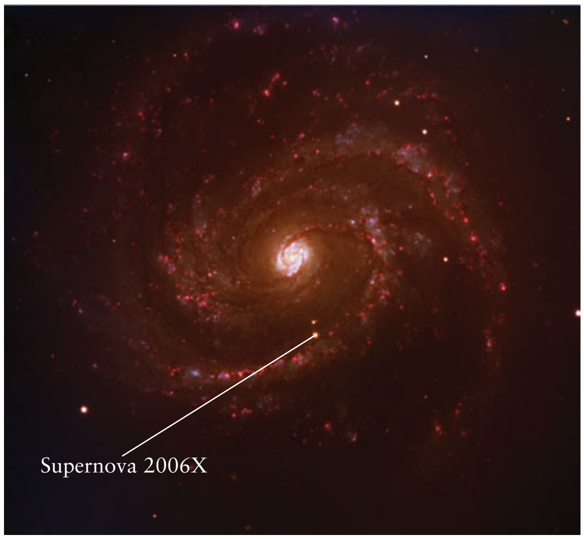
WEB LINK 24.8

One class of standard candles that astronomers have used beyond 30 Mpc is Type Ia supernovae. As we described in Section 20-9, these occur when a white dwarf in a close binary system accretes enough matter from its companion to blow itself apart in a thermonuclear conflagration. A Type Ia supernova can reach a maximum luminosity of about $3 \times 10^9\ L_\odot$ (Figure 24-13). If a Type Ia supernova is seen in a distant galaxy and its maximum apparent brightness measured, the inverse-square law can be used to find the galaxy's distance (see Box 24-1).

One complication is that not all Type Ia supernovae are equally luminous. Fortunately, there is a simple relationship between the peak luminosity of a Type Ia supernova and the rate at which the luminosity decreases after the peak: The more slowly the brightness decreases, the more luminous the supernova. Using this relationship, astronomers have measured distances to supernovae more than 1000 Mpc (3 billion ly) from Earth.

Unfortunately, this technique can be used only for galaxies in which we happen to observe a Type Ia supernova. But telescopic surveys now identify many dozens of these supernovae every year, so the number of galaxies whose distances can be measured in this way is continually increasing.

(a) M100 in March 2002 R I V U X G



(b) M100 in February 2006, showing Supernova 2006X R I V U X G

Figure 24-13

A Supernova in a Spiral Galaxy These images from the Very Large Telescope show the spiral galaxy M100 **(a)** before and **(b)** after a Type Ia supernova exploded within the galaxy in 2006. (The two images were made with different color filters, which gives them different appearances.) Such luminous supernovae, which can be seen at extreme distances, are important standard candles used to determine the distances to faraway galaxies. The distance to M100 is also known from observations of Cepheid variables (see Figure 24-4), so this particular supernova can help calibrate Type Ia supernovae as distance indicators. (European Southern Observatory)

Distance Determination without Standard Candles

Other methods for determining the distances to galaxies do not make use of standard candles. One was discovered in the 1970s by the astronomers Brent Tully and Richard Fisher. They found that the width of the hydrogen 21-cm emission line of a spiral galaxy (see Section 23-3) is related to the galaxy's luminosity. This correlation is the **Tully-Fisher relation**—the broader the line, the more luminous the galaxy.

Such a relationship exists because radiation from the approaching side of a rotating galaxy is blueshifted while that from the galaxy's receding side is redshifted. Thus, the 21-cm line is Doppler broadened by an amount directly related to how fast a galaxy is rotating. Rotation speed is related to the galaxy's mass by Newton's form of Kepler's third law. The more massive a galaxy, the more stars it contains and thus the more luminous it is. Consequently, the width of a galaxy's 21-cm line is directly related to its luminosity.

Because line widths can be measured quite accurately, astronomers can use the Tully-Fisher relation to determine the luminosity of a distant spiral galaxy. By combining this information with measurements of apparent brightness, they can calculate the distance to the galaxy. This technique can be used to measure distances of 100 Mpc or more.

Elliptical galaxies do not rotate, so the Tully-Fisher relation cannot be used to determine their distances. But in 1987 the American astronomers Marc Davis and George Djorgovski pointed out a correlation between the size of an elliptical galaxy, the average motions of its stars, and how the galaxy's brightness appears distributed over its surface.

In geometry, three points define a plane, so the relationship among size, motion, and brightness is called the **fundamental plane.** By measuring the last two quantities, an astronomer can use the fundamental plane relationship to determine a galaxy's actual size. And by comparing this to the galaxy's apparent size, the astronomer can calculate the distance to the galaxy using the small-angle formula (Box 1-1). Ellipticals can be substantially larger and more luminous than spirals, so the fundamental plane can be used at somewhat greater distances than the Tully-Fisher relation.

The Distance Ladder

Figure 24-14 shows the ranges of applicability of several important means of determining astronomical distances. Because these ranges overlap, one technique can be used to calibrate another. As an example, astronomers have studied Cepheids in nearby galaxies that have been host to Type Ia supernovae. The Cepheids provide the distances to these nearby galaxies, making it possible to determine the peak luminosity of each supernova using its maximum apparent brightness and the inverse-square law. Once the peak luminosity is known, it can be used to determine the distance to Type Ia supernovae in more distant galaxies. Because one measuring technique leads us to the next one like rungs on a ladder, the techniques shown in Figure 24-14 (along with others) are referred to collectively as the **distance ladder.**

ANALOGY If you give a slight shake to the bottom of a tall ladder, the top can wobble back and forth alarmingly. A change in distance-measuring techniques used for nearby objects can also have substantial effects on the distances to remote galaxies. For instance, if astronomers discovered that the distances to nearby Cepheids were in error, distance measurements using any technique that is calibrated by Cepheids would be affected as well. (As an example, the distance to the galaxy M100 shown in Figure 24-4 is determined using Cepheids. A Type Ia supernova

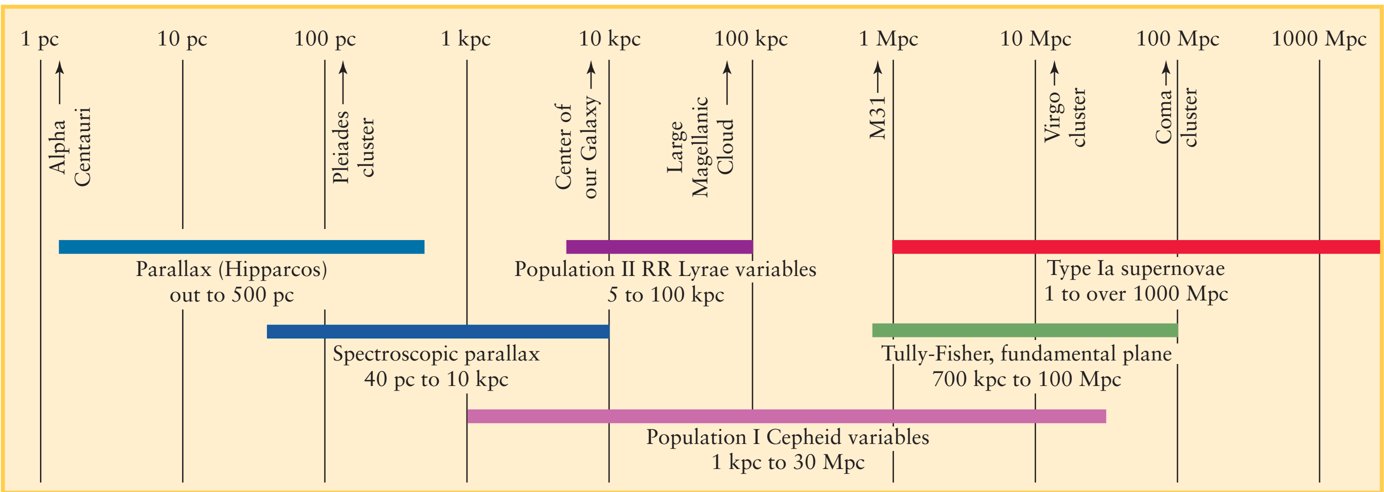


Figure 24-14

The Distance Ladder Astronomers employ a variety of techniques for determining the distances to objects beyond the solar system. Because their ranges of applicability overlap, one technique can be used to calibrate another. The arrows indicate distances to several important objects. Note that each division on the scale indicates a tenfold increase in distance, such as from 1 to 10 Mpc.

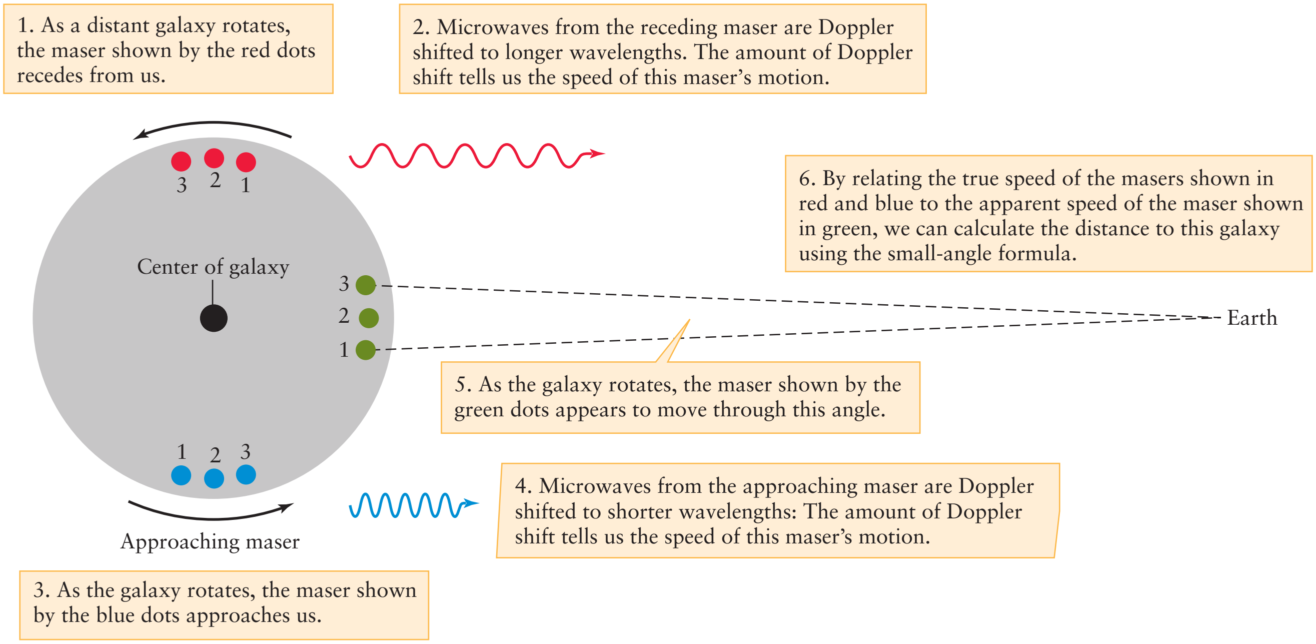


Figure 24-15

Measuring the Distance to a Galaxy Using Masers This drawing shows interstellar clouds called masers (the colored dots) moving from position 1 to 2 to 3 as they orbit the center of a galaxy. The redshift and blueshift of microwaves from the masers shown in red and blue tell us their orbital speed. By relating this to the angle through which the masers shown in green appear to move in a certain amount of time, we can calculate the distance to the galaxy.

has been seen in M100, as Figure 24-13 shows, and its luminosity is determined using the Cepheid-derived distance to M100. Any change in the calculated distance to M100 would change the calculated luminosity of the Type Ia supernova, and so would have an effect on all distances derived from observations of how bright these supernovae appear in other galaxies.) For this reason, astronomers go to great lengths to check the accuracy and reliability of their standard candles.

WEB LINK 24.9

One distance-measuring technique that has broken free of the distance ladder uses observations of molecular clouds called **masers.** ("Maser" is an acronym for "microwave amplification by stimulated emission of radiation.") Just as an electric current stimulates a laser to emit an intense beam of visible light, nearby luminous stars can stimulate water molecules in a maser to emit intensely at microwave wavelengths. This radiation is so intense that masers can be detected millions of parsecs away.

During the 1990s, Jim Herrnstein and his collaborators used the Very Long Baseline Array (see Section 6-6) to observe a number of masers orbiting in a disk around the center of the spiral galaxy M106. They determined the orbital speed of the masers from the Doppler shift of masers near the edges of the disk, where they are moving most directly toward or away from Earth. They also measured the apparent change in position of masers moving across the face of M106. By relating this apparent speed to the true speed determined from the Doppler shift, they calculated that the masers and the galaxy of which they are part are 7.2 Mpc (23 million ly) from Earth (Figure 24-15).

The maser technique is still in its infancy. But because it is independent of all other distance measuring methods, it is likely to play an important role in calibrating the rungs of the distance ladder.

24-5 The Hubble law relates the redshifts of remote galaxies to their distances from the Earth

Whenever an astronomer finds an object in the sky that can be seen or photographed, the natural inclination is to attach a spectrograph to a telescope and record the spectrum. As long ago as 1914, Vesto M. Slipher, working at the Lowell Observatory in Arizona, began taking spectra of "spiral nebulae." He was surprised to discover that of the 15 spiral nebulae he studied, the spectral lines of 11 were shifted toward the red end of the spectrum, indicating that they were moving away from the Earth.

This marked dominance of redshifts was presented by Curtis in the 1920 Shapley-Curtis "debate" as evidence that these spiral nebulae could not be ordinary nebulae in our Milky Way Galaxy. It was only later that astronomers realized that the redshifts of spiral nebulae—that is, galaxies—reveal a basic law of our expanding universe.

Redshift, Distance, and the Hubble Law

WEB LINK 24.10

During the 1920s, Edwin Hubble and Milton Humason photographed the spectra of many galaxies with the 100-inch (2.5-meter) telescope on Mount Wilson in

California. By observing the apparent brightnesses and pulsation periods of Cepheid variables in these galaxies, they were also able to measure the distance to each galaxy (see Section 24-2). Hubble and Humason found that most galaxies show a redshift in their spectrum. They also found a direct correlation between the distance to a galaxy and its redshift:

The greater the redshift of a distant galaxy, the farther away it is

The more distant a galaxy, the greater its redshift and the more rapidly it is receding from us.

In other words, nearby galaxies are moving away from us slowly, and more distant galaxies are rushing away from us much more rapidly. Figure 24-16 shows this relationship for five representative elliptical galaxies. This universal recessional movement is referred to as the **Hubble flow.**

Hubble estimated the distances to a number of galaxies and the redshifts of those galaxies. The **redshift,** denoted by the symbol z, is found by taking the wavelength (λ) observed for a given spectral line, subtracting from it the ordinary, unshifted wavelength of that line (λ_0) to get the wavelength difference ($\Delta\lambda$), and then dividing that difference by λ_0:

Redshift of a receding object

$$z = \frac{\lambda - \lambda_0}{\lambda_0} = \frac{\Delta\lambda}{\lambda_0}$$

z = redshift of an object

λ_0 = ordinary, unshifted wavelength of a spectral line

λ = wavelength of that spectral line that is actually observed from the object

From the redshifts, Hubble used the Doppler formula to calculate the speed at which these galaxies are receding from us. Box 24-2 describes how this is done. Plotting the data on a graph of distance versus speed, Hubble found that the points lie near a straight line. Figure 24-17 is a modern version of Hubble's graph based on recent data.

This relationship between the distances to galaxies and their redshifts was one of the most important astronomical discoveries of the twentieth century. As we will see in Chapter 28, this relationship tells us that we are living in an expanding universe. In 1929, Hubble published this discovery, which is now known as the **Hubble law.** The Hubble law is most easily stated as a formula:

The Hubble law

$$v = H_0 d$$

v = recessional velocity of a galaxy

H_0 = Hubble constant

d = distance to the galaxy

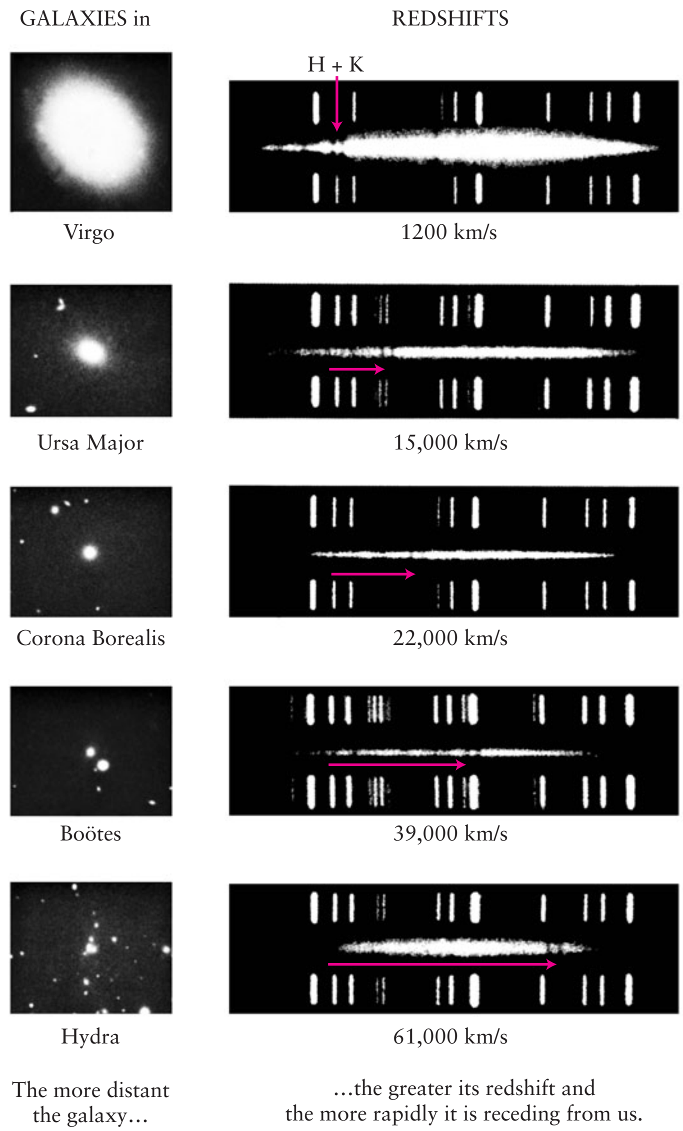


Figure 24-16 R I V U X G

Relating the Distances and Redshifts of Galaxies These five galaxies are arranged, from top to bottom, in order of increasing distance from us. All are shown at the same magnification. Each galaxy's spectrum is a bright band with dark absorption lines; the bright lines above and below it are a comparison spectrum of a light source at the observatory on Earth. The horizontal red arrows show how much the H and K lines of singly ionized calcium are redshifted in each galaxy's spectrum. Below each spectrum is the recessional velocity calculated from the redshift. The more distant a galaxy is, the greater its redshift. (Carnegie Observatories)

This formula is the equation for the straight line displayed in Figure 24-17, and the **Hubble constant** H_0 is the slope of this straight line. From the data plotted on this graph we find that H_0 = 73 km/s/Mpc (say "73 kilometers per second per megaparsec"). In other words, for each million parsecs to a galaxy, the galaxy's speed away from us increases by 73 km/s. For example, a galaxy located 100 million parsecs from the Earth should be

BOX 24-2 Tools of the Astronomer's Trade

The Hubble Law and the Relativistic Redshift

Suppose that you aim a telescope at an extremely distant galaxy. You take a spectrum of the galaxy and find that the spectral lines are shifted toward the red end of the spectrum. For example, a particular spectral line whose normal wavelength is λ_0 appears in the galaxy's spectrum at a longer wavelength λ. The spectral line has thus been shifted by an amount $\Delta\lambda = \lambda - \lambda_0$. The redshift of the galaxy, z, is given by

$$z = \frac{\lambda - \lambda_0}{\lambda_0} = \frac{\Delta\lambda}{\lambda_0}$$

The redshift means that the galaxy is receding from us. According to the Hubble law, the recessional velocity v of a galaxy is related to its distance d from the Earth by

$$v = H_0 d$$

where H_0 is the Hubble constant. We can rewrite this as

$$d = \frac{v}{H_0}$$

Given the value of H_0, we can find the distance d to the galaxy if we know how to determine the recessional velocity v from the redshift z.

If the redshift and recessional velocity are not too great, we can ignore the effects of the special theory of relativity and use the following equation:

$$z = \frac{v}{c} \text{ (valid for low speeds only)}$$

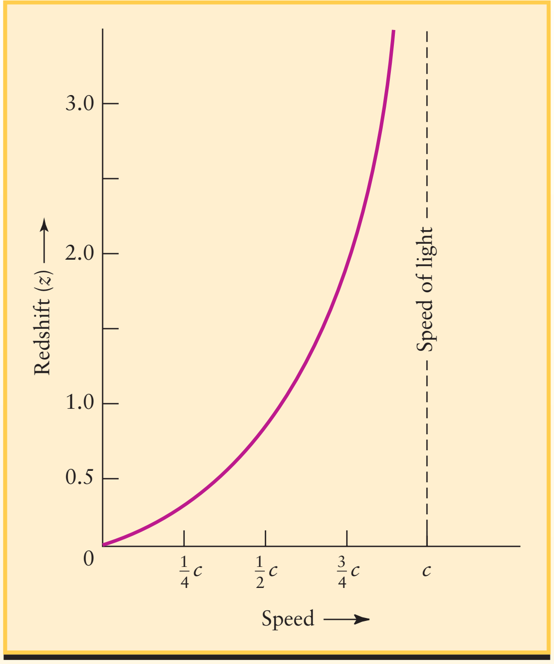
where c is the speed of light. For example, a 5% shift in wavelength ($z = 0.05$) corresponds to a recessional velocity of 5% of the speed of light ($v = 0.05c$). For redshifts greater than about 0.1, however, we must use the relativistic equation for the Doppler shift:

$$z = \sqrt{\frac{c + v}{c - v}} - 1$$

A useful way to rewrite this relationship is as follows:

$$\frac{v}{c} = \frac{(z + 1)^2 - 1}{(z + 1)^2 + 1} \text{ (valid for all speeds)}$$

The accompanying graph displays this relationship between redshift z and speed v. Note that z approaches infinity as v approaches the speed of light.



EXAMPLE: When measured in a laboratory on Earth, the so-called K line of singly ionized calcium has a wavelength $\lambda_0 = 393.3$ nm. But when you observe the spectrum of the giant elliptical galaxy NGC 4889, you find this spectral line at $\lambda = 401.8$ nm. Using $H_0 = 73$ km/s/Mpc, find the distance to this galaxy.

Situation: We are given the values of λ and λ_0 for a line in this galaxy's spectrum. Our goal is to determine the galaxy's distance d.

Tools: We use the relationship $z = (\lambda - \lambda_0)/\lambda_0$ to determine the redshift. We then use the appropriate formula to determine the galaxy's recessional velocity v, and finally use the Hubble law to determine the distance to the galaxy.

Answer: The redshift of this galaxy is

$$z = \frac{401.8 \text{ nm} - 393.3 \text{ nm}}{393.3 \text{ nm}} = 0.0216$$

This is substantially less than 0.1, so we can safely use the low-speed relationship between recessional speed and redshift: $z = v/c$. So NGC 4889 is moving away from us with a speed

$$v = zc = (0.0216)(3 \times 10^5 \text{ km/s}) = 6480 \text{ km/s}$$

Using H_0 = 73 km/s/Mpc in the Hubble law, the distance to this galaxy is

$$d = \frac{v}{H_0} = \frac{6480 \text{ km/s}}{73 \text{ km/s/Mpc}} = 89 \text{ Mpc}$$

Review: This galaxy is receding from us at 0.0216 (2.16%) of the speed of light, and is 89 megaparsecs (290 million light-years) away. Thus the light we see from NGC 4889 today left the galaxy 290 million years ago, even before the first dinosaurs appeared on Earth.

EXAMPLE: In late 1997 astronomers observed a Type Ia supernova (called SN 1997ff) with a redshift $z = 1.7$. Use the Hubble law to find the distance to this supernova.

Situation: We are to use the redshift z to determine the distance d.

Tools: Since the redshift is large, we use the relativistic formula relating recessional velocity v and redshift z to calculate the value of v. We then use the Hubble law to determine the galaxy's distance d.

Answer: From the above relativistic formula,

$$\frac{v}{c} = \frac{(1.7 + 1)^2 - 1}{(1.7 + 1)^2 + 1} = \frac{(2.7)^2 - 1}{(2.7)^2 + 1} = \frac{6.29}{8.29} = 0.76$$

The supernova's recessional velocity is 76% of the speed of light, or $v = 0.76c = (0.76)(3 \times 10^5 \text{ km/s}) = 2.3 \times 10^5$ km/s. From the Hubble law, the distance to the supernova is

$$d = \frac{v}{H_0} = \frac{2.3 \times 10^5 \text{ km/s}}{73 \text{ km/s/Mpc}} = 3200 \text{ Mpc} = 10^{10} \text{ ly}$$

This supernova is a remarkable 10 *billion* light-years away.

Review: For this supernova the value of z is greater than 1. Had we used the low-speed relationship $z = v/c$, we would have come to the erroneous conclusion that $v = 1.7c$—that is, that the supernova is receding from us faster than the speed of light. Using the correct relativistic formula tells us that the recessional velocity is only 76% of the speed of light.

In Chapter 26 we will see how observations of very distant supernovae like SN 1997ff reveal important deviations from the Hubble law. In particular, these observations show that the expansion of the universe is not proceeding at a constant rate, but is actually accelerating.

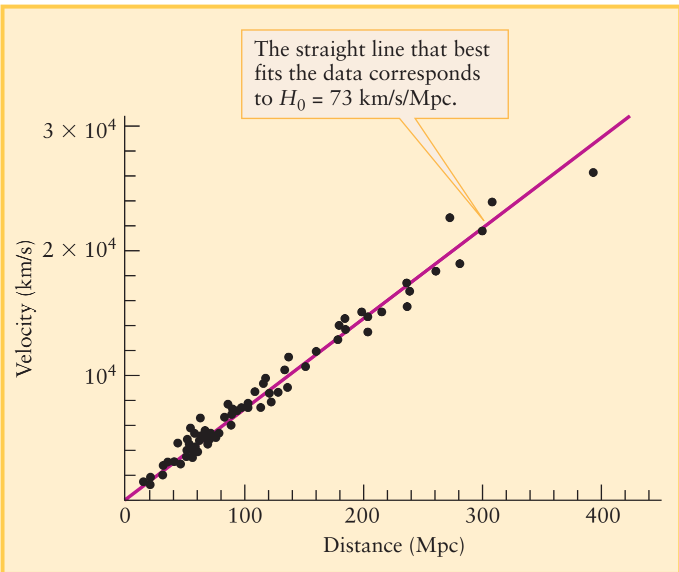


Figure 24-17

The Hubble Law This graph plots the distances and recessional velocities of a sample of galaxies. The straight line is the best fit to the data. This linear relationship between distance and recessional velocity is called the Hubble law.

rushing away from us with a speed of 7300 km/s. (In other books you may see the units of the Hubble constant written with exponents: 73 km s^{-1} Mpc^{-1}.)

CAUTION! A common misconception about the Hubble law is that *all* galaxies are moving away from the Milky Way. The reality is that galaxies have their own motions relative to one another, thanks to their mutual gravitational attraction. For nearby galaxies, the speed of the Hubble flow is small compared to these intrinsic velocities. Hence, some of the nearest galaxies, including M31 (shown in Figure 24-3), are actually approaching us and have blueshifts rather than redshifts. But for distant galaxies, the Hubble speed $v = H_0 d$ is much greater than any intrinsic motion that the galaxies might have. Even if the intrinsic velocity of such a distant galaxy is toward the Milky Way, the fast-moving Hubble flow sweeps that galaxy away from us.

Pinning Down the Hubble Constant

The exact value of the Hubble constant has been a topic of heated debate among astronomers for several decades. The problem is that while redshifts are relatively easy to measure in a reliable way, distances to galaxies (especially remote galaxies) are not, as we saw in Section 24-4. Hence, astronomers who use different methods of determining galactic distances have obtained different values of H_0. To see why this is so, it is helpful to rewrite the Hubble law as

$$H_0 = \frac{v}{d}$$

This shows that if galaxies of a given recessional velocity (v) are far away (so d is large), the Hubble constant H_0 must be relatively small. But if these galaxies are relatively close (so d is small), then H_0 must have a larger value.

In the past, astronomers who used Type Ia supernovae for determining galactic distances found galaxies to be farther away than their colleagues who employed the Tully-Fisher relation. Therefore, the supernova adherents found values of H_0 in the range from about 40 to 65 km/s/Mpc, while the values from the Tully-Fisher relation ranged from about 80 to 100 km/s/Mpc.

In the past few years, the Hubble Space Telescope has been used to observe Cepheid variables with unprecedented precision and in galaxies as far away as 30 Mpc (100 million ly). These observations and others suggest a value of the Hubble constant of about 73 km/s/Mpc, with an uncertainty of no more than 10%. At the same time, reanalysis of the supernova and Tully-Fisher results have brought the values of H_0 from these techniques closer to the Hubble Space Telescope Cepheid value. We will adopt the value H_0 = 73 km/s/Mpc in this book.

Determining the value of H_0 has been an important task of astronomers for a very simple reason: The Hubble constant is one of the most important numbers in all astronomy. It expresses the rate at which the universe is expanding and, as we will see in Chapter 26, even helps give the age of the universe. Furthermore, the Hubble law can be used to determine the distances to extremely remote galaxies. If the redshift of a galaxy is known, the Hubble law can be used to determine its distance from the Earth, as Box 24-2 shows. Thus, the value of the Hubble constant helps determine the distances of the most remote objects in the universe that astronomers can observe.

Because the value of H_0 remains somewhat uncertain, astronomers often express the distance to a remote galaxy simply in terms of its redshift z (which can be measured very accurately). Given the redshift, the distance to this galaxy can be calculated from the Hubble law, but the distance obtained in this way will depend on the particular value of H_0 adopted. Rather than going through these calculations, an astronomer might simply say that a certain galaxy is "at z = 0.128." From the Hubble law relating redshift and distance, this makes it clear that the galaxy in question is more distant than one at z = 0.120 but not as distant as one at z = 0.130. When astronomers use redshift to describe distance, they are making use of the following general rule:

The greater the redshift of a distant galaxy, the greater its distance.

24-6 Galaxies are grouped into clusters and superclusters

WEB LINK 24.11

Galaxies are not scattered randomly throughout the universe but are found in **clusters.** Figure 24-18 shows one such cluster. Like stars within a star cluster, the members of a cluster of galaxies are in continual motion around one another. They appear to be at rest only because they are so distant from us.

WEB LINK 24.12

Figure 24-18 R I V U X G

The Hercules Cluster This irregular cluster of galaxies is about 200 Mpc (650 million ly) from Earth. The Hercules cluster contains many spiral galaxies, often associated in pairs and small groups. (NOAO)

Clusters of Galaxies: Rich and Poor, Regular and Irregular

A cluster is said to be either **poor** or **rich,** depending on how many galaxies it contains. Poor clusters, which far outnumber rich ones, are often called **groups.** For example, the Milky Way Galaxy, the Andromeda Galaxy (M31), and the Large and Small Magellanic Clouds belong to a poor cluster familiarly known as the **Local Group.** The Local Group contains more than 40 galaxies, most of which are dwarf ellipticals (see Figure 24-9). Figure 24-19 is a map of a portion of the Local Group.

In recent years, astronomers have discovered several previously unknown dwarf elliptical galaxies in the Local Group. As of this writing (early 2007), the nearest and most recently discovered is the Canis Major Dwarf, so named for the constellation in whose direction it lies as seen from Earth. It is about 13 kpc (42,000 ly) from the center of the Milky Way Galaxy and a mere 8 kpc (25,000 ly) from Earth (about the same as the distance from Earth to the center of the Milky Way). Tidal forces exerted by the Milky Way on the Canis Major Dwarf are causing this dwarf galaxy to gradually disintegrate and leave a trail of debris behind it (Figure 24-20).

We may never know the total number of galaxies in the Local Group, because dust in the plane of the Milky Way obscures our view over a considerable region of the sky. Nevertheless, we can be certain that no additional large spiral galaxies are hidden by the Milky Way. Radio astronomers would have detected 21-cm radiation from them, even though their visible light is completely absorbed by interstellar dust.

The nearest fairly rich cluster is the Virgo cluster, a collection of more than 2000 galaxies covering a 10° × 12° area of the sky. Figure 24-8 shows a portion of this cluster. One member of this

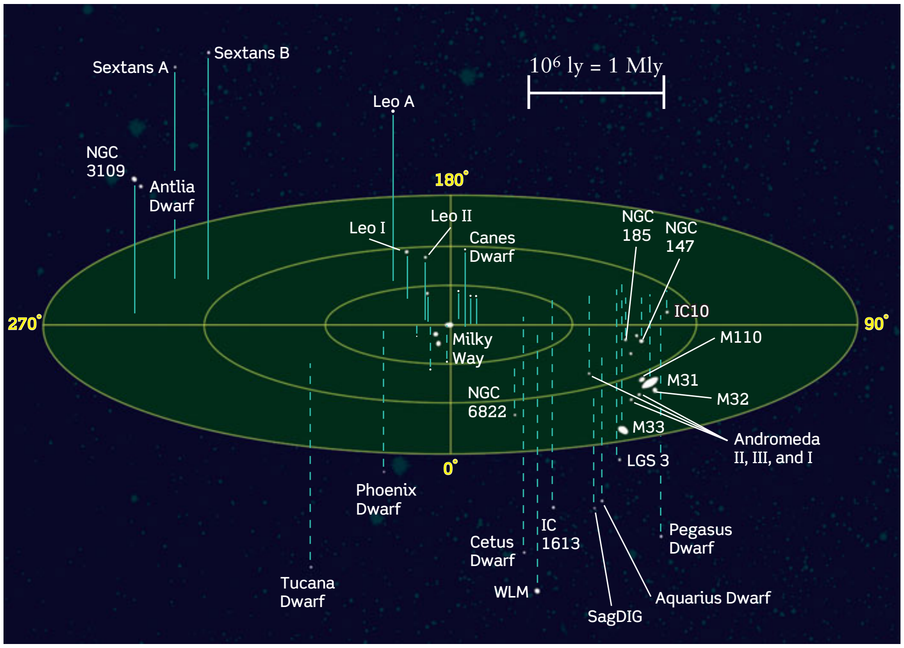
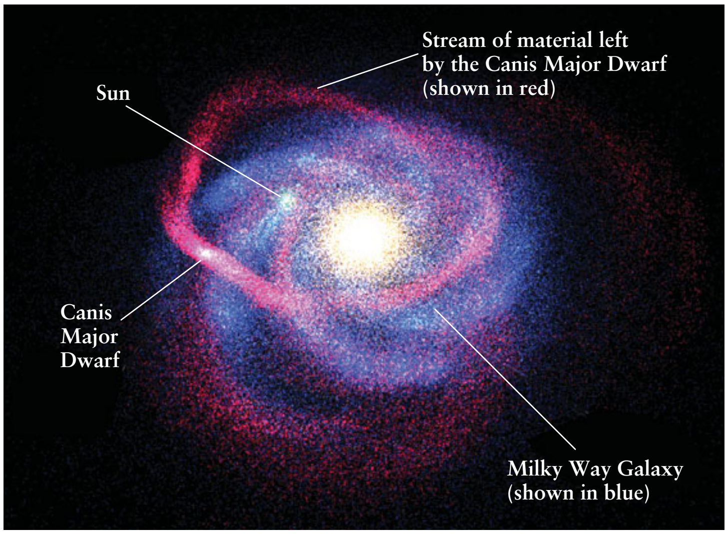


Figure 24-19

The Local Group This illustration shows the relative positions of the galaxies that comprise the Local Group, a poor, irregular cluster of which our Galaxy is part. (The blue rings represent the plane of the Milky Way's disk; 0° is the direction from Earth toward the Milky Way's center. Solid and dashed lines point to galaxies above and below the plane, respectively.) The largest and most massive galaxy in the Local Group is M31, the Andromeda Galaxy; in second place is the Milky Way, followed by the spiral galaxy M33. Both the Milky Way and M31 are surrounded by a number of small satellite galaxies. (© Richard Powell, www.atlasoftheuniverse.com)

cluster not shown in Figure 24-8 is the spiral galaxy M100; measurements of Cepheid variables in M100 (see Figure 24-4) give a distance of about 17 Mpc (56 million ly). The Tully-Fisher relation and observations of Type Ia supernovae (see Figure 24-13) give similar distances to this cluster. The overall diameter of the cluster is about 3 Mpc (9 million ly).

The center of the Virgo cluster is dominated by three giant elliptical galaxies. You can see two of these, M84 and M86, in Figure 24-8. The diameter of each of these enormous galaxies is comparable to the 750-kpc distance between the Milky Way and M31. In other words, one giant elliptical is approximately the same size as the entire Local Group!



ANIMATION 24.1 WEB LINK 24.13

Figure 24-20

The Canis Major Dwarf Discovered in 2003, this dwarf elliptical galaxy is actually slightly closer to Earth than is the center of the Milky Way Galaxy. This illustration shows the stream of material left behind by the Canis Major Dwarf as it orbits the Milky Way. This material is torn away by the Milky Way's tidal forces (see Section 4-8). (R. Ibata et al., Observatoire de Strasbourg/Université Louis Pasteur; 2MASS; and NASA)

Astronomers also categorize clusters of galaxies as regular or irregular, depending on the overall shape of the cluster. The Virgo cluster, for example, is called an **irregular cluster,** because its galaxies are scattered throughout a sprawling region of the sky. Our own Local Group is also an irregular cluster. In contrast, a **regular cluster** has a distinctly spherical appearance, with a marked concentration of galaxies at its center.

The nearest example of a rich, regular cluster is the Coma cluster, located about 90 Mpc (300 million light-years) from us toward the constellation Coma Berenices (Berenice's Hair) (Figure 24-21). Despite its great distance, telescopic images of this cluster show more than 1000 galaxies. The Coma cluster almost certainly contains many thousands of dwarf ellipticals, so the total membership of the cluster may be as many as 10,000 galaxies. The core of the Coma cluster is dominated by two giant ellipticals surrounded by many normal-sized galaxies.

The overall shape of a cluster is related to the dominant types of galaxies it contains. Rich, regular clusters contain mostly elliptical and lenticular galaxies. For example, about 80% of the brightest galaxies in the Coma cluster (see Figure 24-21) are ellipticals; only a few spiral galaxies are scattered around the cluster's outer regions. Irregular clusters, such as the Virgo cluster and the Hercules cluster shown in Figure 24-18, have a more even mixture of galaxy types.

Superclusters: Clusters of Clusters of Galaxies

Clusters of galaxies are themselves grouped together in huge associations called **superclusters.** A typical supercluster contains dozens of individual clusters spread over a region of space up to 45 Mpc (150 million ly) across. Figure 24-22 shows the distribution of clusters in our part of the universe. The nearer ones out to the Virgo cluster, including our own Local Group, are

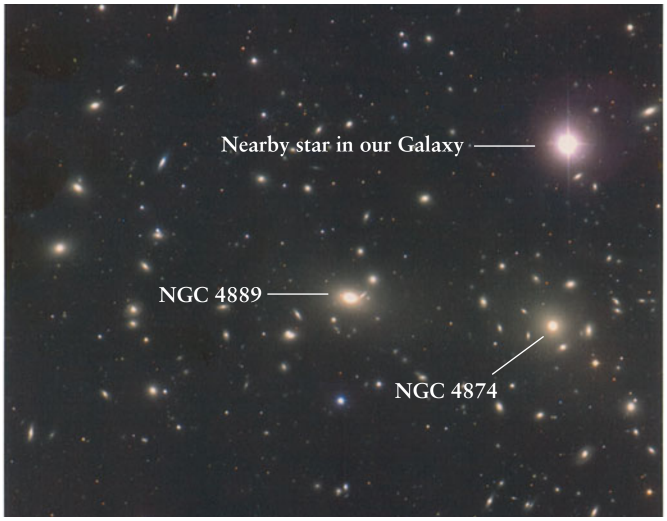


Figure 24-21 R I V U X G

The Coma Cluster This rich, regular cluster is about 90 Mpc (300 million light-years) from the Earth. Almost all of the spots of light in this image are individual galaxies of the cluster. Two giant elliptical galaxies, NGC 4889 and NGC 4874, dominate the center of the cluster. The bright star at the upper right is within our own Milky Way Galaxy, a million times closer than any of the galaxies shown here. (O. Lopez-Cruz and I. K. Shelton, University of Toronto; KPNO)

members of the *Local Supercluster*. The other clusters shown in Figure 24-22 belong to other superclusters. The most massive cluster in the local universe is called the *Great Attractor;* its gravity is so great that the Milky Way as well as the rest of the Local Supercluster is moving toward it at speeds of several hundred kilometers per second.

Observations indicate that unlike clusters, superclusters are not bound together by gravity. That is, most clusters in each supercluster are drifting away from most of the other clusters in that same supercluster. Furthermore, the superclusters are all moving away from each other due to the Hubble flow.

Cosmic Voids and Sheets: The Distribution of Superclusters

Since the 1980s, astronomers have been working to understand how superclusters are distributed in space. Some of this structure is revealed by maps such as that shown in Figure 24-23, which displays the positions on the sky of 1.6 million galaxies. Such maps reveal that superclusters are not randomly distributed, but seem to lie along filaments. But to comprehend more fully the distribution of superclusters, it is necessary to map their positions in three dimensions. This is done by measuring both the position of a galaxy on the sky as well as the galaxy's redshift. Using the Hubble law (see Section 24-5), astronomers can use each galaxy's redshift to estimate its distance from Earth and thus its position in three-dimensional space.

Superclusters of galaxies are not spread uniformly across the universe, but are found in vast sheets separated by truly immense voids

The first three-dimensional maps of this kind were made in the 1970s and included measurements of a few hundred galaxies. Collecting the data for such maps required many months or years of telescopic observations. Technology for astronomy has advanced tremendously since then, and it is now possible to measure the redshifts of 400 galaxies in a single hour!

WEB LINK 24.14

The most extensive galaxy maps available at this writing (early 2007) are those from the Sloan Digital Sky Survey, a joint project of astronomers from the United States, Japan, and Germany, and from the Two Degree Field Galactic Redshift Survey (2dFGRS), a collaboration between Australian, British, and U.S. astronomers. (The name refers to the 2° field of view of the telescope used for the observations, which is unusually wide for a research telescope.) Figure 24-24*a* shows a map made from 2dFGRS measurements of more than 60,000 galaxies. This particular map encompasses two wedge-shaped slices of the universe, one on either side of the plane of the Milky Way (Figure 24-24*b*). The Earth (in the Milky Way) is at the apex of the wedge-shaped map; each dot represents a galaxy. The measurements used to create this map included galaxies with redshifts as large as $z = 0.25$, corresponding to a recessional velocity of 66,000 km/s. With a Hubble constant of 73 km/s/Mpc, this means that the map in Figure 24-24*a* extends out to a distance of nearly 1000 Mpc, or 3 *billion* light-years from Earth.

Maps such as that shown in Figure 24-24*a* reveal enormous **voids** where exceptionally few galaxies are found. (These voids

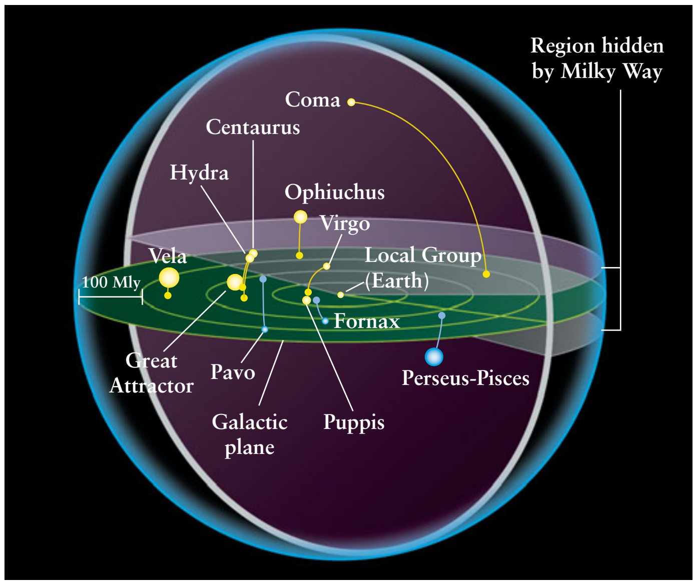


Figure 24-22

Nearby Clusters of Galaxies This illustration shows a sphere of space 800 million ly (250 Mpc) in diameter centered on the Earth in the Local Group. Each spherical dot represents a cluster of galaxies. To better see the three-dimensionality of this figure, colored arcs are drawn from each cluster to the green plane, which is an extension of the plane of the Milky Way outward into the universe. Note that clusters of galaxies are unevenly distributed here, as they are elsewhere in the universe.

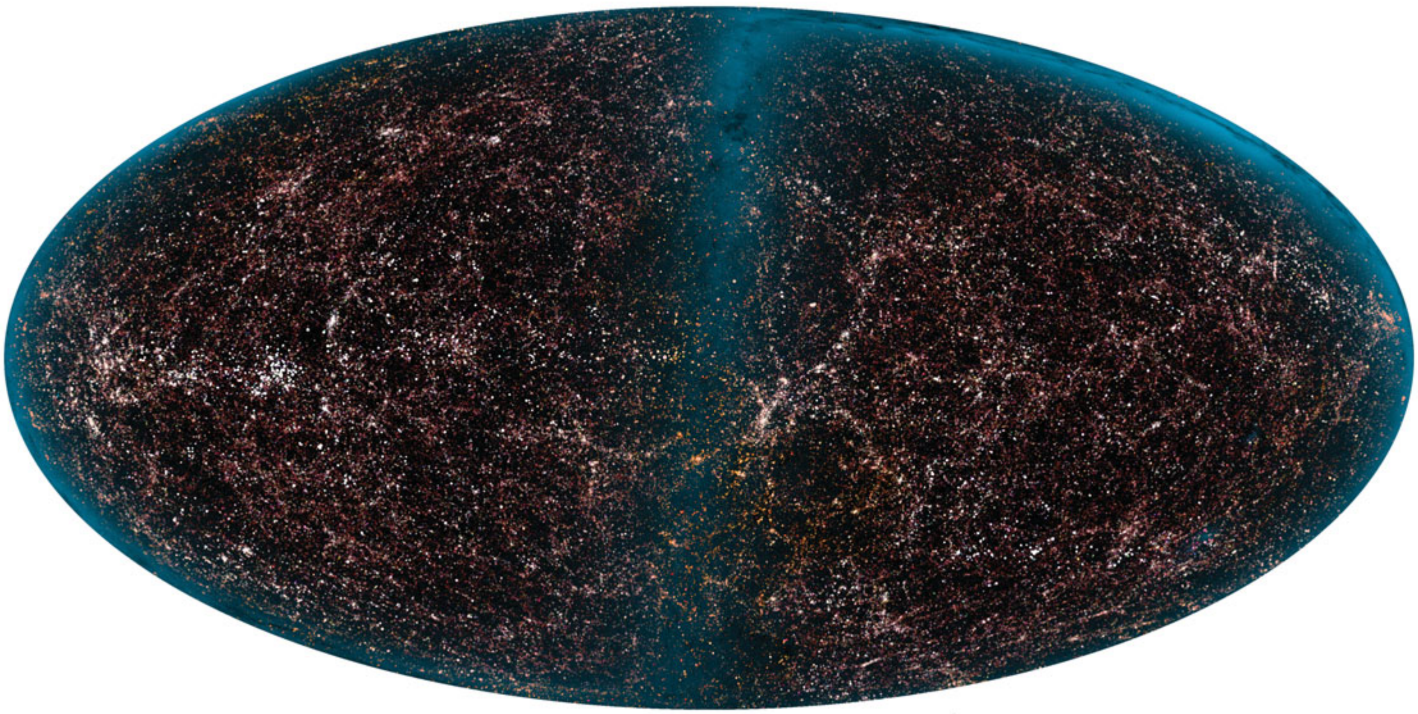
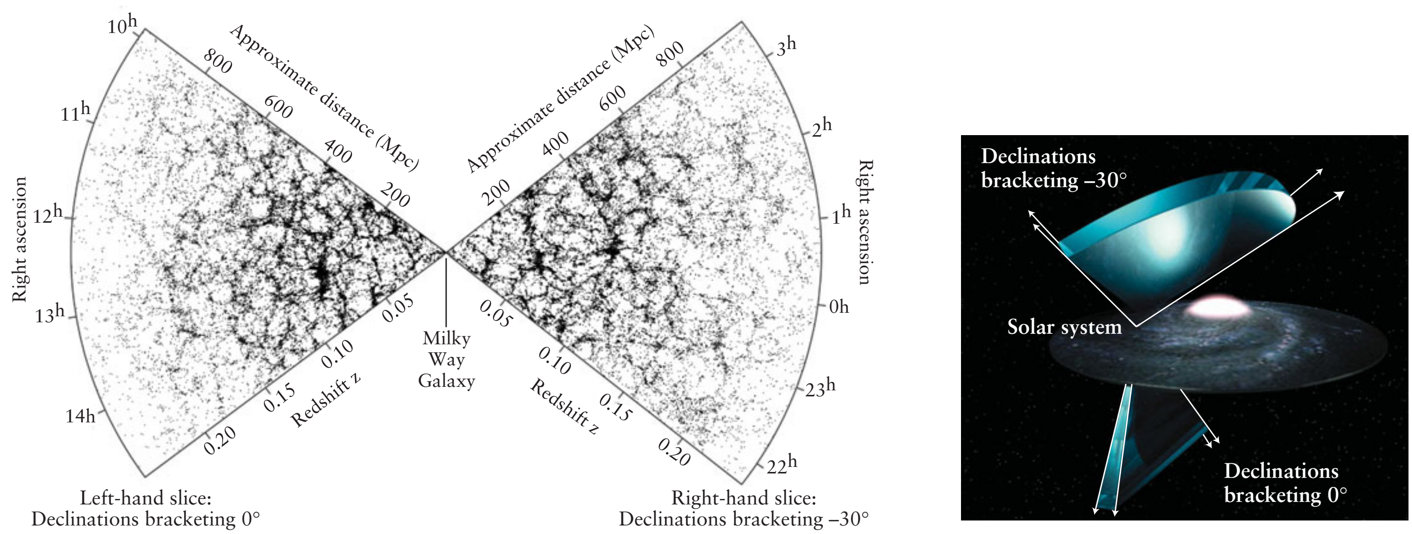


Figure 24-23 R I V U X G

Structure in the Nearby Universe This composite infrared image from the 2MASS (Two-Micron All-Sky Survey) project shows the light from 1.6 million galaxies. In this image, the entire sky is projected onto an oval; the blue band running vertically across the center of the image is light from the plane of the Milky Way. Note that galaxies form a lacy, filamentary structure. Note also the large, dark voids that contain few galaxies. (2MASS; IPAC/Caltech; and University of Massachusetts)



(a) The 2dF galaxy survey

(b) Fields of view in the 2dF survey

ANIMATION 24.2

Figure 24-24

The Large-Scale Distribution of Galaxies **(a)** This map shows the distribution of 62,559 galaxies in two wedges extending out to redshift $z = 0.25$. (For an explanation of right ascension and declination, see Box 2-1.) Note the prominent voids surrounded by thin regions full of galaxies. **(b)** The two wedges shown in (a) lie roughly perpendicular to the plane of the Milky Way. These were chosen to avoid the obscuring dust that lies in our Galaxy's plane. (Courtesy the 2dF Galaxy Redshift Survey Team/Anglo-Australian Observatory) were first discovered in 1978 in a pioneering study by Stephen Gregory and Laird Thompson at the Kitt Peak National Observatory.) These voids are roughly spherical and measure 30 to 120 Mpc (100 million to 400 million ly) in diameter. They are not entirely empty, however. There is evidence for hydrogen clouds in some voids, while others may be subdivided by strings of dim galaxies.

Figure 24-24*a* shows that most galaxies are concentrated in sheets on the surfaces between voids. These sheets can be more than 100 Mpc long and several megaparsecs thick. This pattern is similar to that of soapsuds in a kitchen sink, with sheets of soap film (analogous to galaxies) surrounding air bubbles (analogous to voids). These titanic sheets of galaxies are the largest structures known in the universe: On scales much larger than 100 Mpc, the distribution of galaxies in the universe appears to be roughly uniform. As we will see in Chapter 27, this pattern of sheets and voids contains important clues about how clusters of galaxies formed in the early universe.

24-7 Colliding galaxies produce starbursts, spiral arms, and other spectacular phenomena

Occasionally, two galaxies within a cluster or from adjacent clusters can collide with each other. Past collisions have hurled vast numbers of stars into intergalactic space. In some cases, we can even observe a collision in progress, a cosmic catastrophe that gives birth to new stars. And astronomers can predict collisions that will not take place for billions of years, such as the collision that is fated to occur between the galaxy M31 and our own Milky Way Galaxy.

High-Speed Galaxy Collisions: Shredding Gas and Dust

When two galaxies collide at high speed, the huge clouds of interstellar gas and dust in the galaxies slam into each other and can be completely stopped in their tracks. In this way, two colliding galaxies can be stripped of their interstellar gas and dust.

The best evidence that such collisions take place is that many rich clusters of galaxies are strong sources of X rays (Figure 24-25). This emission reveals the presence of substantial amounts of hot **intracluster gas** (that is, gas within the cluster) at temperatures between 10^7 and 10^8 K. The only way that such large amounts of gas could be heated to such extremely high temperatures is in violent collisions between galaxies.

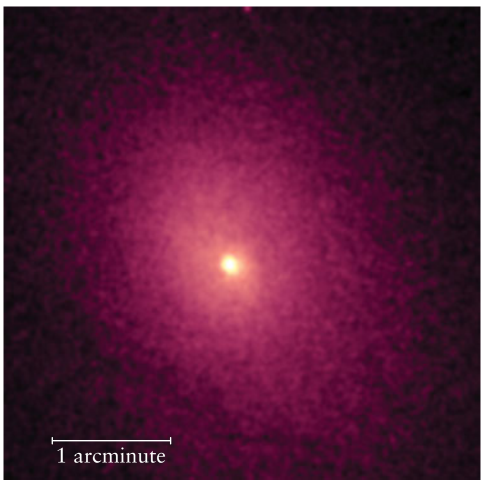
CAUTION! Although galaxies can and do collide, it is highly unlikely that the *stars* from two colliding galaxies actually run into each other. The reason is that the stars within a galaxy are very widely separated from one another, with a tremendous amount of space between them.

Gentle Galactic Collisions and Starbursts

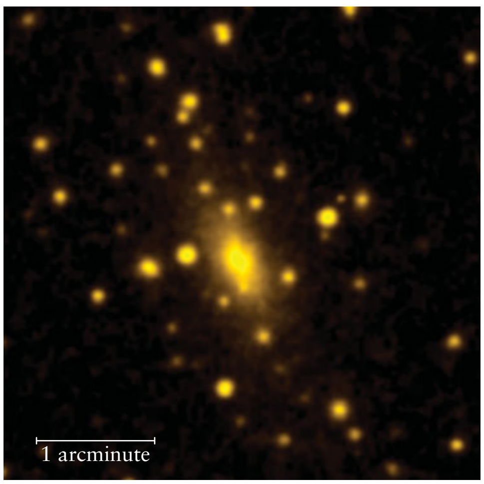
In a less violent collision or a near-miss between two galaxies, the compressed interstellar gas may have more time to cool, allowing many protostars to form. Such collisions may account for **starburst galaxies** such as M82 (Figure 24-26), which blaze with the light of numerous newborn stars. These galaxies have bright centers surrounded by clouds of warm interstellar dust, indicating recent, vigorous star birth. Their warm dust is so abundant that starburst galaxies are among the most luminous objects in the universe at infrared wavelengths. (The right-hand image on the first page of Chapter 6 shows the infrared emission from M82's warm dust.)

The starburst galaxy M82 shown in Figure 24-26 also shows the effects of strong winds from young, luminous stars. It also contains a number of luminous globular clusters. Unlike the globular clusters in our Galaxy, whose stars are about 12.5 billion years old, those in M82 are no more than 600 million years old. These young globular clusters are another sign of recent star formation.

M82 is one member of a nearby cluster of galaxies that includes the beautiful spiral galaxy M81 and a fainter elliptical companion called NGC 3077 (Figure 24-27*a*). Radio surveys of that region of the sky reveal enormous streams of hydrogen gas connecting the three galaxies (Figure 24-27*b*). The loops and



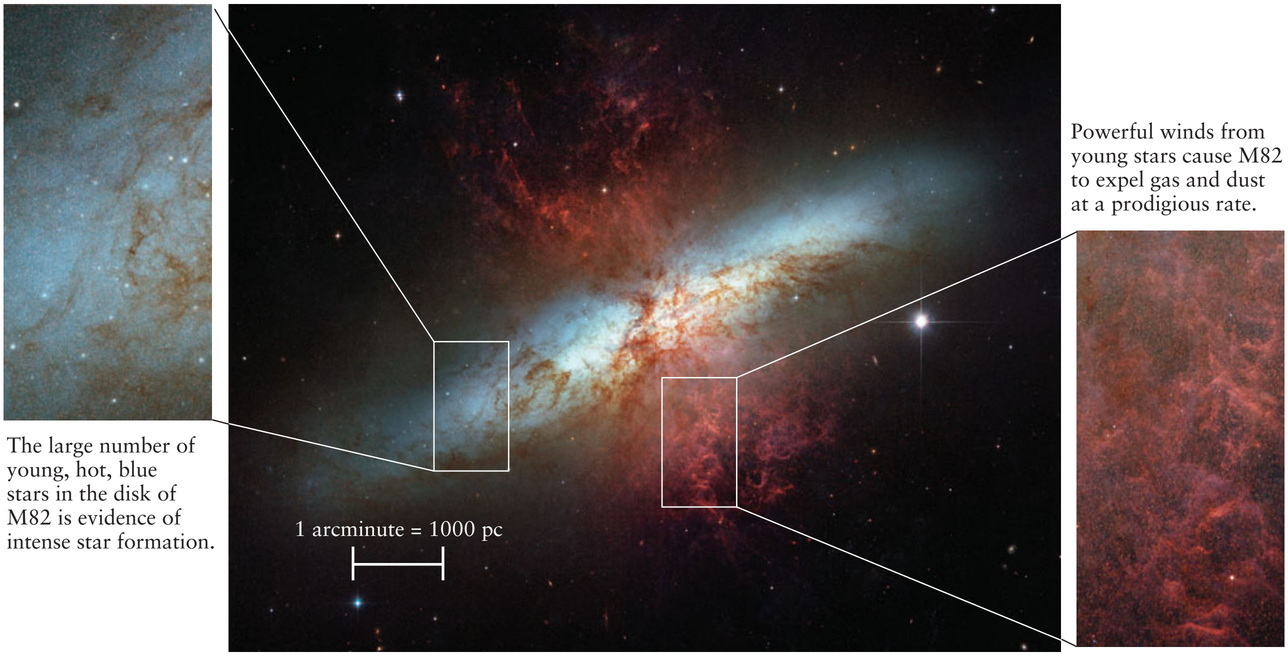
(a) An X-ray image of Abell 2029 shows emission from hot gas.



(b) A visible-light image of Abell 2029 shows the cluster's galaxies.

WEB LINK 24.15

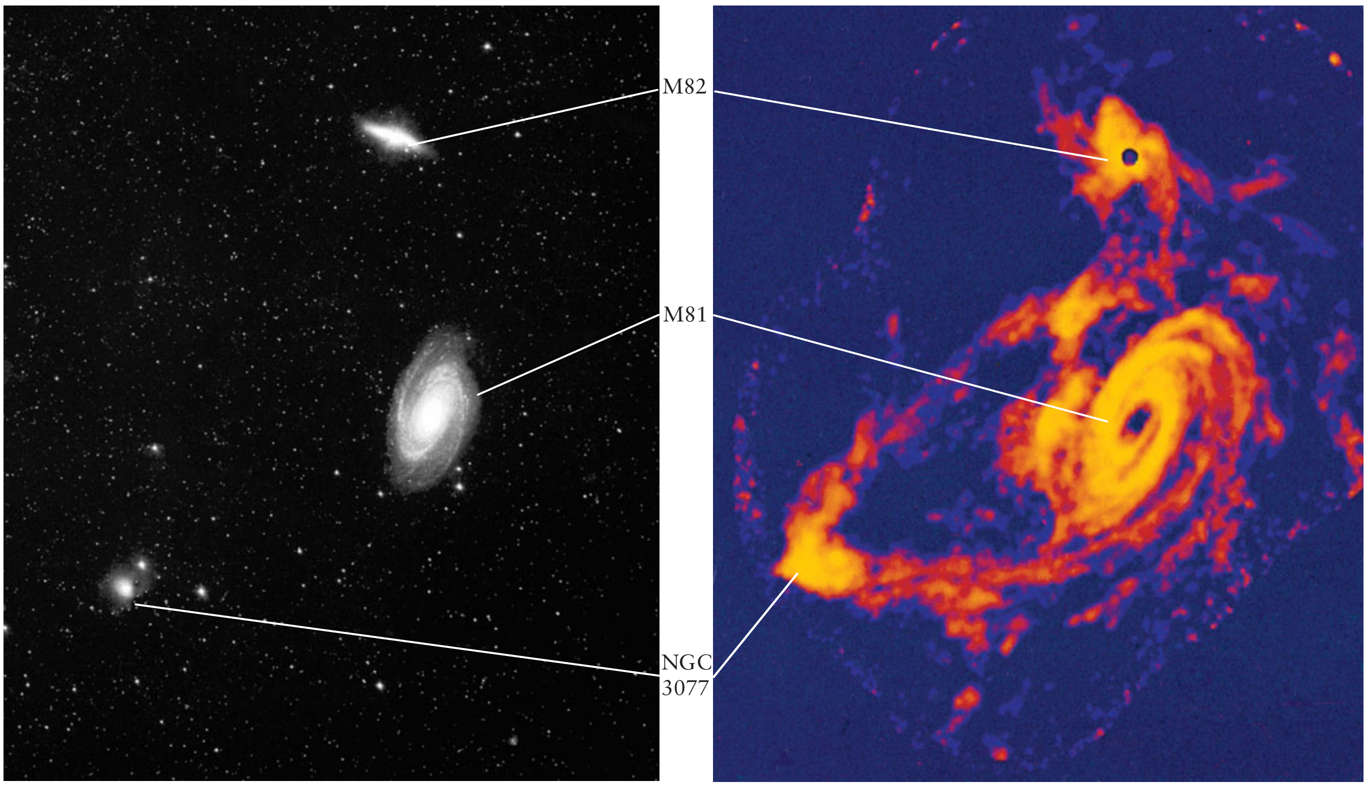
Figure 24-25
X-ray Emission from a Cluster of Galaxies **(a)** An X-ray image of this cluster of galaxies shows emission from hot gas between the galaxies. The gas was heated by collisions between galaxies within the cluster. **(b)** The galaxies themselves are too dim at X-ray wavelengths to be seen in (a), but are apparent at visible wavelengths. This cluster, one of many cataloged by the UCLA astronomer George O. Abell, is about 300 Mpc (1 billion ly) from Earth in the constellation Serpens. (a: NASA, CXC, and University of California, Irvine/A. Lewis et al. b: Palomar Observatory DSS)



VIDEO 24.3

Figure 24-26 R I V U X G

A Starburst Galaxy Prolific star formation is occurring at the center of the irregular galaxy M82, which lies about 3.6 Mpc (12 million ly) from Earth in the constellation Ursa Major. M82 also contains an unusual X-ray source, shown in Figure 22-17. (NASA; ESA; and the Hubble Heritage Team, STScI/AURA)



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

Figure 24-27

The M81 Group **(a)** The starburst galaxy M82 (see Figure 24-26) is part of a cluster of about a dozen galaxies. This wide-angle visible-light photograph shows the three brightest galaxies of the cluster. The area shown is about 1° across. **(b)** This false-color radio image of the same region, created from data taken by the Very Large Array, shows streamers of hydrogen gas that connect the three bright galaxies as well as several dim ones. (a: Palomar Sky Survey; b: M. S. Yun, VLA and Harvard) twists in these streamers suggest that the three galaxies have had several close encounters over the ages. A similar stream of hydrogen gas connects our Galaxy with its second nearest neighbor, the Large Magellanic Cloud (see Figure 24-12), suggesting a history of close encounters between our Galaxy and the LMC.

Tidal Forces and Galaxy Mergers

Tidal forces between colliding galaxies can deform the galaxies from their original shapes, just as the tidal forces of the Moon on the Earth deform the oceans and help give rise to the tides (see Section 4-8, especially Figure 4-27). The galactic deformation is so great that thousands of stars can be hurled into intergalactic space along huge, arching streams. (This same effect has stripped material away from the Canis Major Dwarf galaxy as it orbits the Milky Way, as shown in Figure 24-20.) Supercomputer simulations of such collisions show that while some of the stars are flung far and wide, other stars slow down and the galaxies may merge.

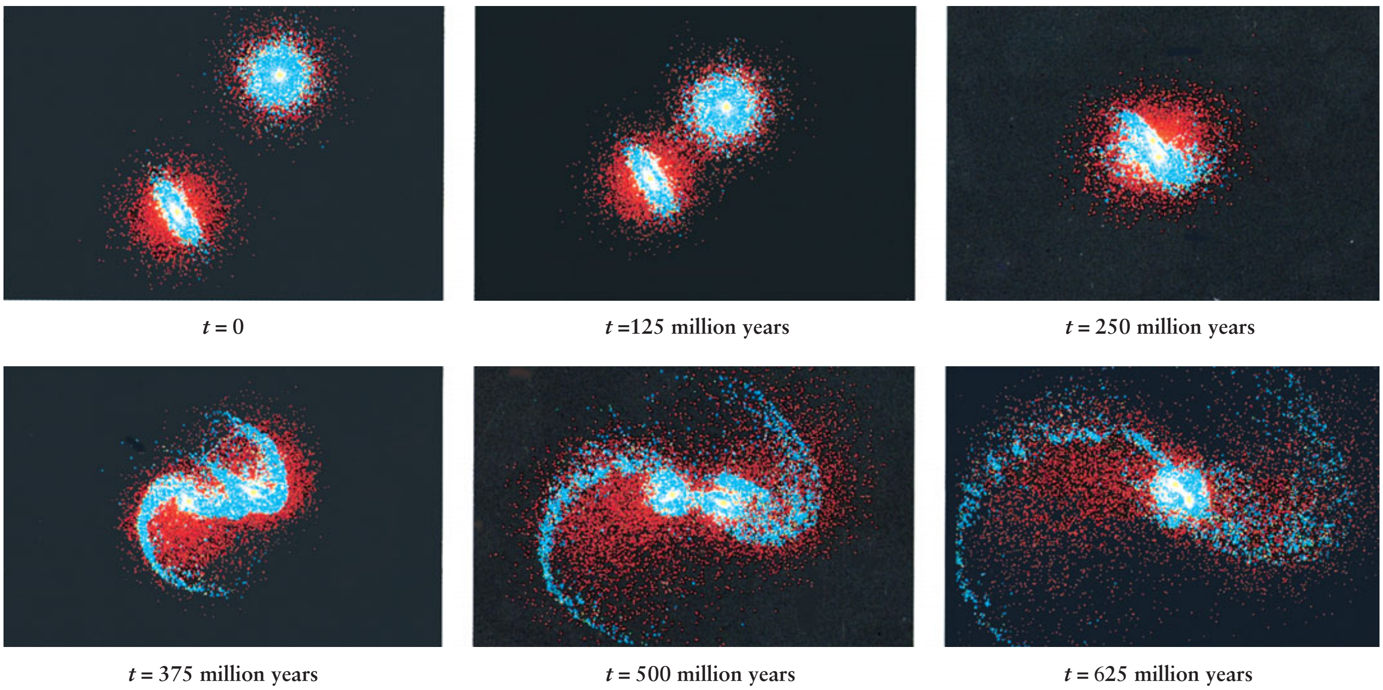
Galaxies need not actually collide to exert strong forces on each other

WEB LINK 24.16 Figure 24-28 shows one such simulation. As the two galaxies pass through each other, they are severely distorted by gravitational interactions and throw out a pair of extended tails. The interaction also prevents the galaxies in the simulation from continuing on their original paths. Instead, they fall back together for a second encounter (at 625 million years). The simulated galaxies merge soon thereafter, leaving a single object. The *Cosmic Connections* figure explores a real-life example of two galaxies that are colliding in just this way.

WEB LINK 24.17 Our own Milky Way Galaxy is expected to undergo a galactic collision like that shown in Figure 24-28. The Milky Way and the Andromeda Galaxy, shown in Figure 24-3, are actually approaching each other and should collide in another 6 billion years or so. (Recall that our solar system is only 4.56 billion years old.) When this happens, the sky will light up with a plethora of newly formed stars, followed in rapid succession by a string of supernovae, as the most massive of these stars complete their life spans. Any inhabitants of either galaxy will see a night sky far more dramatic and tempestuous than ours.

ANIMATION 24.5 When two galaxies merge, the result is a bigger galaxy. If this new galaxy is located in a rich cluster, it may capture and devour additional galaxies, growing to enormous dimensions by **galactic cannibalism.** Cannibalism differs from mergers in that the galaxy that does the devouring is bigger than its "meal," whereas merging galaxies are about the same size.

Many astronomers suspect that galactic cannibalism is the reason that giant ellipticals are so huge. As we have seen, giant galaxies typically occupy the centers of rich clusters. In many cases, smaller galaxies are located around these giants (see Figure 24-8



ANIMATION 24.3 ANIMATION 24.4

Figure 24-28

A Simulated Collision Between Two Galaxies These frames from a supercomputer simulation show the collision and merger of two galaxies accompanied by an ejection of stars into intergalactic space. Stars in the disk of each galaxy are colored blue, while stars in their central bulges are yellow-white. Red indicates dark matter that surrounds each galaxy. The frames progress at 125-million-year intervals. Compare the bottom frames with the image of the Antennae galaxies in the *Cosmic Connections* figure. (Courtesy of J. Barnes, University of Hawaii)

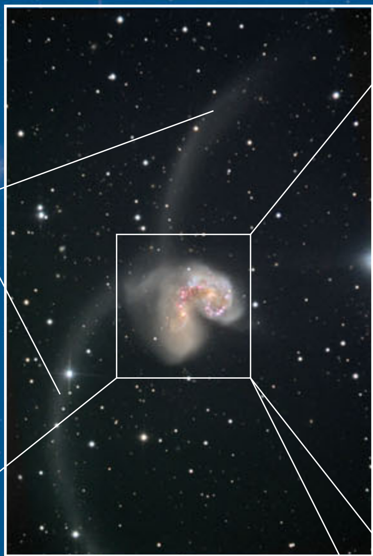
COSMIC CONNECTIONS

When Galaxies Collide

Although galaxies can collide at very high speeds by Earth standards, they are so vast that a collision can last hundreds of millions of years. Understanding what happens during a galactic collision requires ideas about tidal forces (Chapter 4), star formation (Chapter 18), and stellar evolution (Chapter 19).

1. One example of a galactic collision is the pair of galaxies called the Antennae, which lie 19 Mpc (16 million ly) from Earth in the constellation Corvus (the Crow). They probably began to interact several hundred million years ago.

Tidal forces between the galaxies pulled out these long "tidal tails" 200 to 300 million years ago.



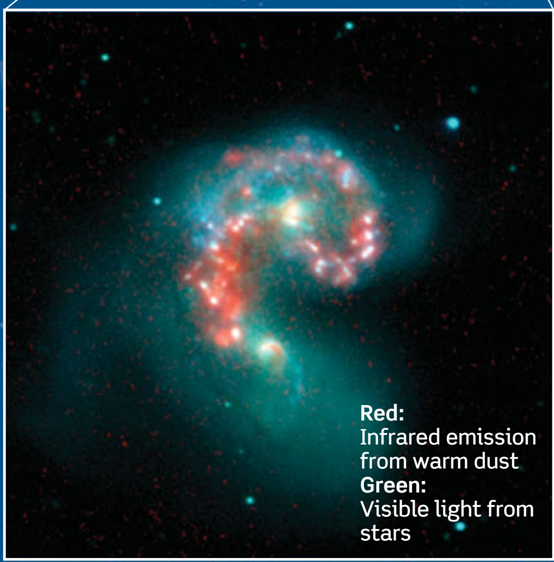
R I V U X G

(Bob and Bill Twardy/Adam Block/NOAO/AURA/NSF)

2. As the gas and dust clouds of the two galaxies collide with each other, they are greatly compressed. This causes stars to form in tremendous numbers.

R I V U X G

(NASA; ESA; and the Hubble Heritage Team, STScI/AURA-ESA/ Hubble Collaboration)

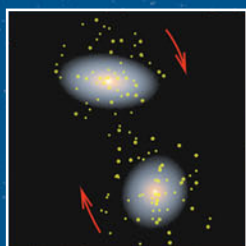


R I V U X G

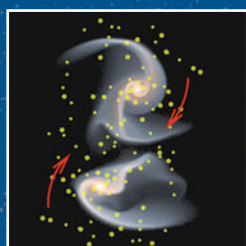
(NASA/JPL-Caltech/Z. Wang, Harvard-Smithsonian CfA)

3. This composite infrared and visible-light image of the Antennae allows us to see inside the two galaxies and reveals clouds of dust warmed by the light of hot young stars.

4. The globular clusters that orbit our Milky Way Galaxy contain only old stars; all of the short-lived blue stars have long since died. But some of the globular clusters that orbit the Antennae galaxies *do* have hot blue stars. Hence these clusters must be young. These, too, are a result of the compression of gas and dust that takes place in a collision between galaxies.

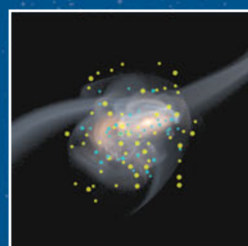


a. Two galaxies, each with old globular clusters (yellow), begin to interact.



b. The two galaxies swoop past each other before finally settling down as a single merged galaxy. Gas and dust is compressed in both galaxies in the process, creating new star clusters.

(John Dubinski, University of Toronto)



c. The combined galaxy has the original globular clusters (shown in yellow) as well as new ones (shown in blue).

VIDEO 24.4 WEB LINK 24.18

and Figure 24-21). As they pass through the extended halo of a giant elliptical, these smaller galaxies slow down and are eventually devoured by the larger galaxy.

Galaxy Interactions and Spiral Arms

Close encounters between galaxies provide a third way of forming spiral arms (in addition to density waves and self-propagating star formation, discussed in Section 23-5). Supercomputer simulations clearly demonstrate that spiral arms can be created during a collision, either by drawing out long streamers of stars or by compressing clouds of interstellar gas. For example, the spiral arms of M51 (examine Figure 24-2) may have been produced by a close encounter with a second galaxy. The disruptive galaxy, NGC 5195, is now located at the end of one of the spiral arms created by the collision. The two galaxies shown in the image that opens this chapter are thought to be interacting in the same way.

The very fact of our existence may be intimately related to interactions between galaxies. Some astronomers argue that the spiral arms of our Milky Way Galaxy were produced by a close encounter with the Large Magellanic Cloud. As we saw in Section 23-5, spiral arms compress the interstellar medium in the Milky Way's disk to form Population I stars like our own Sun, which have enough heavy elements to produce Earthlike planets. Thus, the chain of events that led to the formation of our Sun, our solar system, and life on our planet may have been initiated by a long-ago interaction between two galaxies.

24-8 Most of the matter in the universe is mysterious dark matter

A cluster of galaxies must be held together by gravity. In other words, there must be enough matter in the cluster to prevent the galaxies from wandering away. Nevertheless, careful examination of a rich cluster, like the Coma cluster shown in Figure 24-21, reveals that the mass of the visually luminous matter (principally the stars in the galaxies) is not at all sufficient to bind the cluster gravitationally. The observed line-of-sight speeds of the galaxies, measured by Doppler shifts, are so large that the cluster should have broken apart long ago. Considerably more mass than is visible is needed to keep the galaxies bound in orbit about the center of the cluster.

We encountered a similar situation in studying our own Milky Way Galaxy in Section 25-4: The total mass of our Galaxy is more than the amount of visible mass. As for our Galaxy, we conclude that clusters of galaxies must contain significant amounts of nonluminous *dark matter*. If this dark matter were not there, the galaxies would have long ago dispersed in random directions and the cluster would no longer exist today. Analyses demonstrate that the total mass needed to bind a typical rich cluster is about 10 times greater than the mass of material that shows up on visible-light images.

The Dark-Matter Problem and Rotation Curves

As for our Galaxy, the problem is to determine what form the invisible mass takes. A partial solution to this **dark-matter problem,** which dates from the 1930s, was provided by the discovery in the late 1970s of hot, X-ray–emitting gas within clusters of galaxies (see Figure 24-25*a*). By measuring the amount of X-ray emission, astronomers find that the total mass of intracluster gas in a typical rich cluster can be greater than the combined mass of all the stars in all the cluster's galaxies. This is sufficient to account for only about 10% of the invisible mass, however. The remainder is dark matter of unknown composition.

Vast assemblages of dark matter reveal their presence by bending passing rays of light

Although we do not know what dark matter is made of, it is possible to investigate how dark matter is distributed in galaxies and clusters of galaxies. It appears that dark matter lies within and immediately surrounding galaxies, not in the vast spaces between galaxies. The evidence for this comes principally from observations of the rotation curves of galaxies and of the gravitational bending of light by clusters of galaxies.

As we saw in Section 23-4, a rotation curve is a graph that shows how fast stars in a galaxy are moving at different distances from that galaxy's center. For example, Figure 23-18 is the rotation curve for our Galaxy. As Figure 24-29 illustrates, many other spiral galaxies have similar rotation curves that remain remarkably flat out to surprisingly great distances from each galaxy's center. In other words, the orbital speed of the stars remains roughly constant out to the visible edges of these galaxies.

This observation tells us that we still have not detected the *true* edges of these galaxies (and many similar ones). Near the true edge of a galaxy we should see a decline in orbital speed, in accordance with Kepler's third law (see Figure 23-18). Because this decline has not been observed, astronomers conclude that there must be a considerable amount of dark material that extends well beyond the visible portion of the disk.

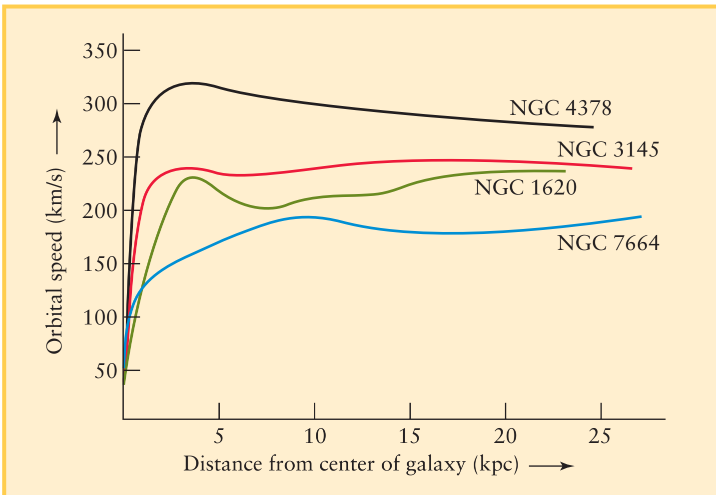


Figure 24-29

The Rotation Curves of Four Spiral Galaxies This graph shows how the orbital speed of material in the disks of four spiral galaxies varies with the distance from the center of each galaxy. If most of each galaxy's mass were concentrated near its center, these curves would fall off at large distances. But these and many other galaxies have flat rotation curves that do not fall off. This indicates the presence of extended halos of dark matter. (Adapted from V. Rubin and K. Ford)

"Seeing" Dark Matter with Gravitational Lensing

Further evidence about how dark matter is distributed comes from the gravitational bending of light rays, which we described in Section 22-2. As Figure 22-5 shows, the gravity of a single star like the Sun can deflect light by only a few arcseconds. But a more massive object such as a galaxy can produce much greater deflections, and the amount of this deflection can be used to determine the galaxy's mass. For example, suppose that the Earth, a massive galaxy, and a background light source (such as a more distant galaxy) are in nearly perfect alignment, as sketched in Figure 24-30*a*. Because of the warped space around the massive galaxy, light from the background source curves around the galaxy as it heads toward us. As a result, light rays can travel along two paths from the background source to us here on the Earth. Thus, we should see two images of the background source.

A powerful source of gravity that distorts background images is called a **gravitational lens.** For gravitational lensing to work, the alignment between the Earth, the massive galaxy, and a remote background light source must be almost perfect. Without nearly perfect alignment, the second image of the background star is too faint to be noticeable.

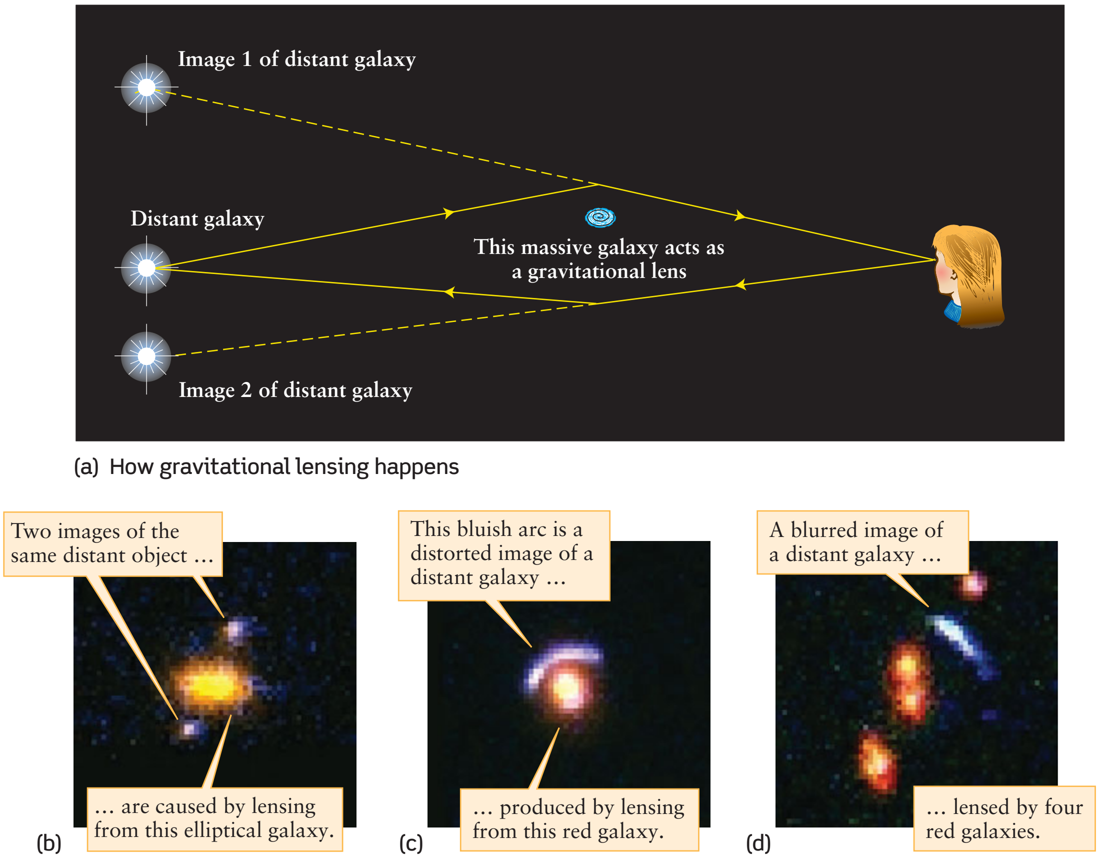
Beginning in 1979, astronomers have discovered a great number of examples of gravitational lensing. The example shown in Figure 24-30*b* is almost exactly like the ideal situation depicted in Figure 24-30*a*. If the alignment is very slightly off, the image of the distant galaxy is distorted into an arc as shown in Figure 24-30*c*. Figure 24-30*d* shows a more complicated example of lensing that results when the gravitational lens is not one but several massive galaxies.

Figure 24-31 shows a situation in which an entire cluster of galaxies acts as a gravitational lens. The image shows an ordinary-looking rich cluster of yellowish elliptical and spiral galaxies, but with a number of curious blue arcs. Reconstruction of the light paths through the cluster shows that all these blue arcs are actually distorted images of a single galaxy that lies billions of light-years beyond the cluster.

By measuring the distortion of the images of such background galaxies, J. Anthony Tyson of Bell Laboratories and his colleagues have determined that dark matter, which constitutes about 90% of the cluster's mass, is distributed much like the visible matter in the cluster. In other words, the overall arrangement of visible galaxies seems to trace the location of dark matter.

The Nature of Dark Matter

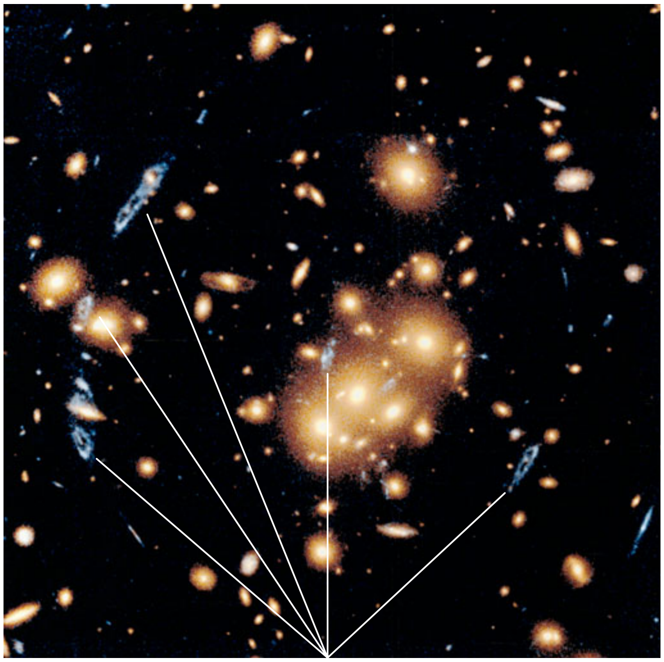
Many proposals have been made to explain the nature of dark matter. One reasonable suggestion was that clusters might contain



VIDEO 24.5

Figure 24-30 R I V U X G

Gravitational Lensing (a) A massive object such as a galaxy can deflect light rays like a lens so that an observer sees more than one image of a more distant galaxy. (Compare with Figure 25-17, which shows the same effect on a much smaller scale.) **(b), (c), (d)** Three examples of gravitational lensing. In each case a single, distant blue galaxy is "lensed" by a closer red galaxy or galaxies. (b, c, d: Kavan Ratnatunga, Carnegie Mellon University; ESA; and NASA)



All of these blue arcs are images of the same distant galaxy.

WEB LINK 24.19

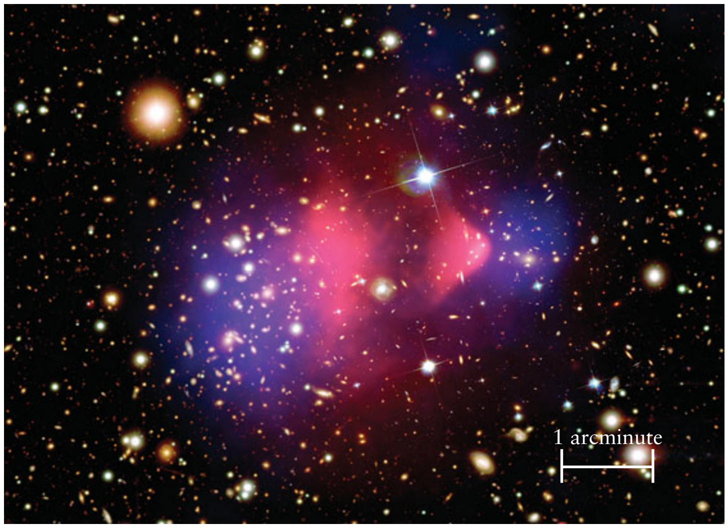
Figure 24-31 R I V U X G

Gravitational Lensing by a Cluster of Galaxies The blue arcs in this image of the rich cluster CL0024+1654 are distorted multiple images of a single more distant galaxy. These images are the result of gravitational lensing by the matter in CL0024+1654. The cluster is about 1600 Mpc (5 billion ly) from Earth; the blue galaxy is about twice as distant. The blue color of the remote galaxy suggests that it is very young and is actively forming stars. (W. N. Colley and E. Turner, Princeton University; J. A. Tyson, Bell Labs, Lucent Technologies; NASA)

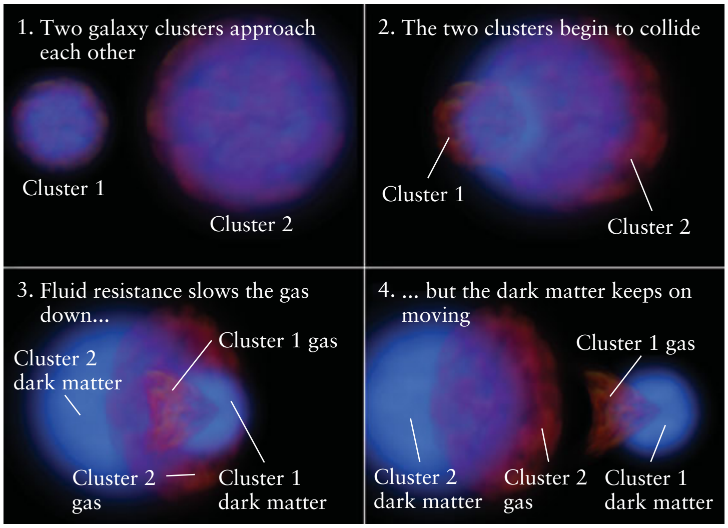
a large number of faint, red, low-mass (0.2 $M_\odot$ or less) stars. These faint stars could be located in extended halos surrounding individual galaxies or scattered throughout the spaces between the galaxies of a cluster. They would have escaped detection because their luminosity and hence apparent brightness would be very low. (The mass-luminosity relation for main-sequence stars, which we discussed in Section 17-9, tells us that low-mass stars are intrinsically very faint.) Searches for these stars around other galaxies as well as around the Milky Way have been carried out using the Hubble Space Telescope. None has yet been detected, so it is thought that faint stars are unlikely to constitute the majority of the dark matter in the universe.

As we described in Section 23-4, more exotic dark matter candidates include massive neutrinos, subatomic particles called WIMPs (weakly interacting massive particles), and MACHOs (massive compact halo objects, such as small black holes or brown dwarfs). To date, however, the true nature of dark matter remains unknown.

An important clue about the nature of dark matter was discovered in 2006 by examining a rich cluster of galaxies called 1E0657-56. Remarkably, the visible matter and dark matter in this cluster do *not* have the same distribution (Figure 24-32*a*).



(a) Composite image of galaxy cluster 1E0657-56 showing visible galaxies, X-ray-emitting gas (red) and dark matter (blue) R I V U X G



(b) A model of how the gas and dark matter in 1E0657-56 could have become separated

WEB LINK 24.20 ANIMATION 24.8

Figure 24-32

Isolated Dark Matter in a Cluster of Galaxies **(a)** This visible-light image of the galaxy cluster 1E0657-56 shows more than a thousand galaxies. The superimposed image in red shows the distribution of the cluster's hot, X-ray emitting gas, and the blue image shows the distribution of dark matter as determined by gravitational lensing (see Figure 24-30). **(b)** We can understand the separation of dark matter and gas in this cluster if we assume that dark matter does not feel any force of fluid resistance. This is what we would expect if dark matter responds to gravitational forces only. (a: X-ray: NASA/CXC/CfA/M. Markevitch et al.; Optical: NASA/STScI; Magellan/U. Arizona/D. Clew et al.; Lensing Map: NASA/STScI; SO FW; Magellan/U. Arizona/D. Clowe et al. b: NASA/CXC/M. Weiss)

The best explanation for how this could have come about is that 1E0657-56 is the result of a collision between two galaxy clusters, one larger than the other (Figure 24-32*b*). (In Section 24-7 we asked you to visualize collisions between entire galaxies; now you must imagine a vastly more immense collision between entire clusters of galaxies!) During the collision, the gas from one cluster slams into the gas from the other cluster and slows down due to fluid resistance. (You feel the force of fluid resistance pushing against you whenever you try to move through a liquid or gas—for example, when you swim in a lake or put your hand outside the window of a fast-moving car.) Fluid resistance is a consequence of the electric forces between adjacent atoms and molecules in a fluid. But if dark matter is made up of some curious material that responds only to *gravitational* forces, it is unaffected by fluid resistance. As a result, during the collision sketched in Figure 24-32*b* the gas is slowed by fluid resistance but the dark matter is not. The agreement of the simulation in Figure 24-32*b* with the observations in Figure 24-32*a* strongly reinforces the idea that dark matter, though mysterious, is quite real.

The study of dark matter forces us to redefine what we mean by "ordinary." As we will see in Chapter 26, we now know that there is about 5 times as much dark matter in the observable universe as there is visible matter. Because dark matter is so dominant, "ordinary" visible matter—including this book, the air that you breathe, all the planets and stars, and your own body—is in fact relatively rare. Thus, "ordinary" matter is rather *extraordinary* in the universe as a whole.

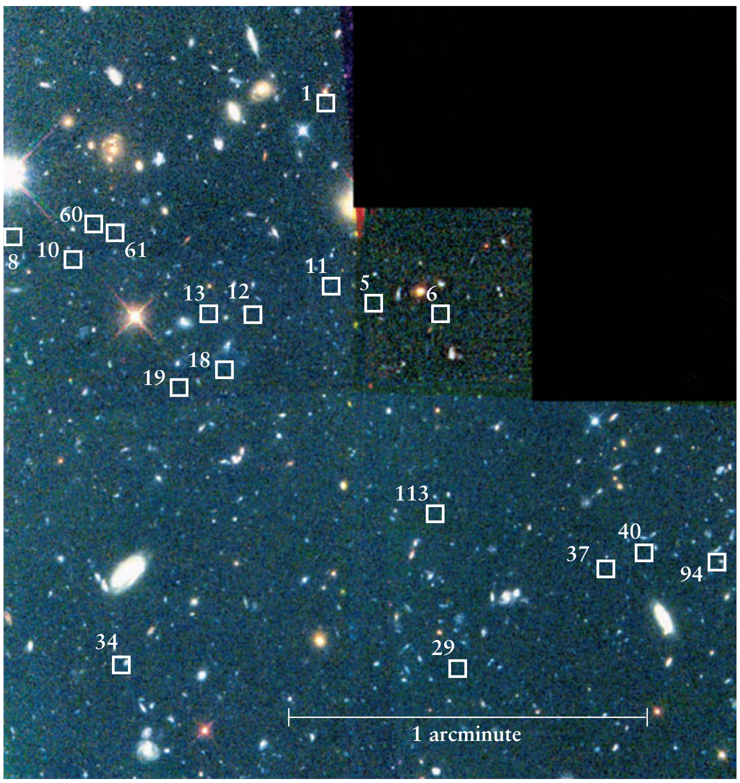
24-9 Galaxies formed from the merger of smaller objects

How do galaxies form and how do they evolve? Astronomers can gain important clues about galactic evolution simply by looking deep into space. The more distant a galaxy is, the longer its light takes to reach us. As we examine galaxies that are at increasing distances from the Earth, we are actually looking further and further back in time. By looking into the past, we can see galaxies in their earliest stages.

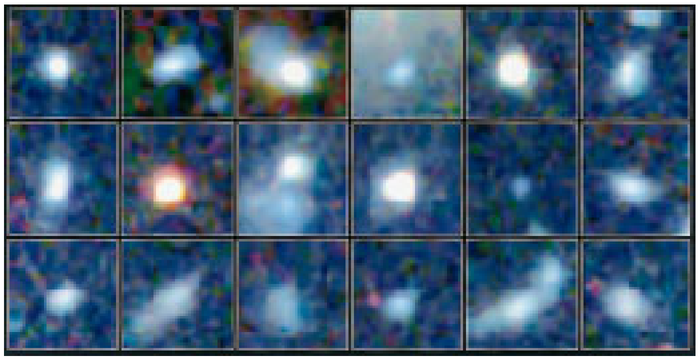
Over the eons, collisions and mergers have dramatically altered the population of galaxies

Building Galaxies from the "Bottom Up"

The Hubble Space Telescope images in Figure 24-33 provide a glimpse of galaxy formation in the early universe. Figure 24-33*a* shows a number of galaxylike objects some 11 billion light-years (3400 Mpc) away and are thus seen as they were 11 billion years ago. These objects are smaller than even the smallest galaxies we see in the present-day universe and have unusual, irregular shapes (Figure 24-33*b*). Furthermore, these objects are scattered over an area only 600 kpc (2 million light-years) across—less than the distance between the Milky Way Galaxy and M31—making it quite probable that they would collide and merge with each other. These collisions would be aided by the dark matter associated with each subgalactic object, which increases the object's mass



(a) A portion of the constellation Hercules



(b) Closeup images of the numbered objects in (a)

WEB LINK 24.21

Figure 24-33 R I V U X G

The Building Blocks of Galaxies (a) In this Hubble Space Telescope image, the objects outlined by boxes are about 3400 Mpc (11 billion ly) from Earth and are only 600 to 900 pc (2000 to 3000 ly) across—larger than a star cluster but smaller than even dwarf elliptical galaxies like that shown in Figure 24-9. **(b)** If these objects were to coalesce, the result would be a full-sized galaxy such as we see in the nearby universe today. (Rogier Windhorst and Sam Pascarelle, Arizona State University; NASA) and, hence, the gravitational forces pulling the objects together. Such mergers would eventually give rise to a normal-sized galaxy.

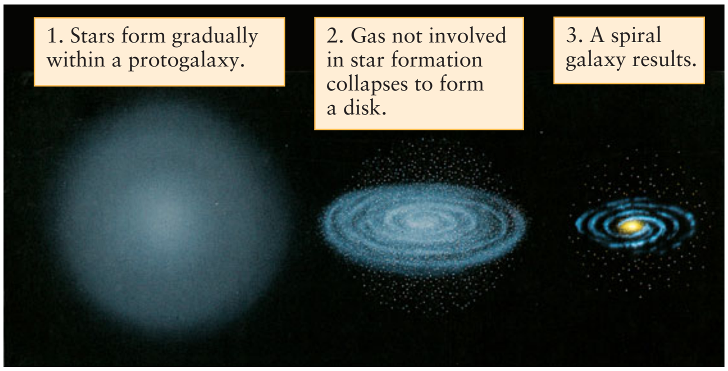
Images such as those in Figure 24-33 lead astronomers to conclude that galaxies formed "from the bottom up"—that is, by the merger of smaller objects like those in Figure 24-33*b* to form full-size galaxies. (These same images rule out an older idea that galaxies formed "from the top down"—that is, directly from immense, galaxy-sized clouds of material.) The blue color of the objects in Figure 24-33*b* indicates the presence of young stars. Observations indicate that the very first stars formed about 13.5 billion years ago, when the universe was only about 200 million years old. We will see evidence in Chapter 26 that the matter in the universe formed "clumps" even earlier than this. These clumps evolved into objects like those shown in Figure 24-33*b*, which in turn merged to form the population of galaxies that we see today.

Forming Spirals, Lenticulars, and Ellipticals

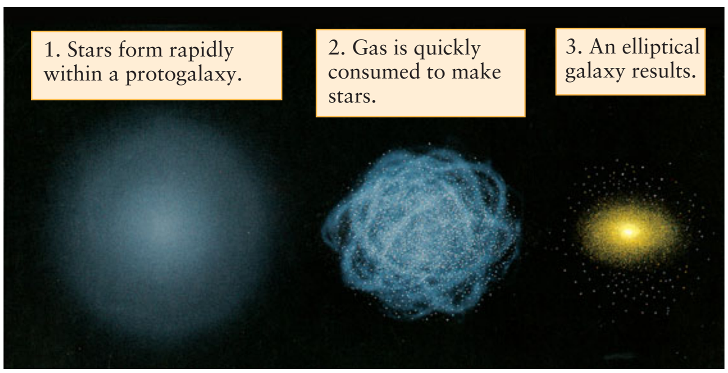
Once a number of subgalactic units combine, they make an object called a *protogalaxy*. The rate at which stars form within a protogalaxy may determine whether this protogalaxy becomes a spiral or an elliptical. If stars form relatively slowly, the gas surrounding them has enough time to settle by collisions to form a flattened disk, much as happened on a much smaller scale in the solar nebula (see Section 8-4). Star formation continues because the disk contains an ample amount of hydrogen from which to make new stars. The result is a spiral or lenticular galaxy (Figure 24-34*a*). But if stars initially form in the protogalaxy at a rapid rate, virtually all of the available gas is used up to make stars before a disk can form. In this case what results is an elliptical galaxy (Figure 24-34*b*).

Figure 24-34*c* compares the stellar birthrate in the two types of galaxies. This graph helps us understand some of the differences between spiral and elliptical galaxies that we described in Section 24-3. Protogalaxies are thought to have been composed almost exclusively of hydrogen and helium gas, so the first stars were Population II stars with hardly any metals (that is, heavy elements). As stars die and form planetary nebulae or supernovae, they eject gases enriched in metals into the interstellar medium. In a spiral galaxy there is ongoing star formation in the disk, so these metals are incorporated into new generations of stars, making relatively metal-rich Population I stars like the Sun. By contrast, an elliptical galaxy has a single flurry of star formation when it is young, after which star formation ceases. Elliptical galaxies therefore contain only metal-poor Population II stars.

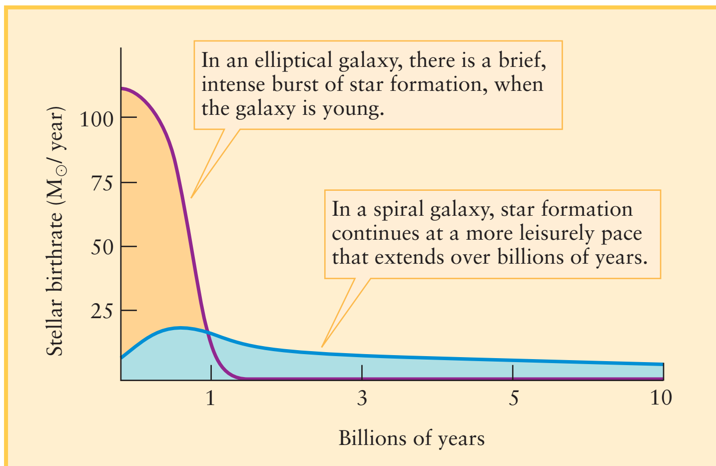
Figure 24-34*c* shows that both elliptical and spiral galaxies form stars most rapidly when they are young. This idea is borne out by the observation that very distant galaxies tend to be blue, which means that galaxies were bluer in the distant past than they are today. (Note the very blue colors of the distant, gravitationally lensed galaxies shown in Figure 24-30 and Figure 24-31, as well as those of the subgalactic objects shown in Figure 24-33*b*.) Spectroscopic studies of such galaxies in the 1980s by James Gunn and Alan Dressler demonstrated that most owe their blue color to vigorous star formation, often occurring in intense, episodic bursts. The hot, luminous, and short-lived O and B stars produced in these bursts of star formation give blue galaxies their characteristic color.



(a) Formation of a spiral galaxy



(b) Formation of an elliptical galaxy



(c) The stellar birthrate in galaxies

Figure 24-34

The Formation of Spiral and Elliptical Galaxies **(a)** If the initial star formation rate in a protogalaxy is low, it can evolve into a spiral galaxy with a disk. **(b)** If the initial star formation rate is rapid, no gas is left to form a disk. The result is an elliptical galaxy. **(c)** This graph shows how the rate of star birth (in solar masses per year) varies with age in spiral and elliptical galaxies.

An Evolving Universe of Galaxies

In addition to changes in galaxy colors, the character of the galactic population has also changed over the past several billion years. In nearby rich clusters, only about 5% of the galaxies are spirals. But observations of rich clusters at a redshift of $z = 0.4$—which corresponds to looking about 4 billion years into the past—show that about 30% of their galaxies were spirals.

Why were spiral galaxies more common in rich clusters in the distant past? Galactic collisions and mergers are probably responsible. During a collision, interstellar gas in the colliding galaxies is vigorously compressed, triggering a burst of star formation (see the *Cosmic Connections* figure in Section 24-7). A succession of collisions produces a series of star-forming episodes that create numerous bright, hot O and B stars that become dispersed along arching spiral arms by the galaxy's rotation. Eventually, however, the gas is used up; star formation then ceases and the spiral arms become less visible. Furthermore, tidal forces tend to disrupt colliding galaxies, strewing their stars across intergalactic space until the galaxies are completely disrupted (see Figure 24-28).

A full description of galaxy formation and evolution must include the effects of dark matter. As we have seen, only about 10% of the mass of a galaxy—its stars, gas, and dust—emits electromagnetic radiation of any kind. As yet we have no idea what the remaining 90% looks like or what it is made of. The dilemma of dark matter is one of the most challenging problems facing astronomers today.

Key Words

anisotropic, p. 642
barred spiral galaxy, p. 640
clusters (of galaxies), p. 650
dark-matter problem, p. 658
distance ladder, p. 645
dwarf elliptical galaxy, p. 641
elliptical galaxy, p. 641
fundamental plane, p. 645
galactic cannibalism, p. 656
giant elliptical galaxy, p. 641
gravitational lens, p. 659
groups (of galaxies), p. 650
Hubble classification, p. 639
Hubble constant, p. 647
Hubble flow, p. 647
Hubble law, p. 647
intracluster gas, p. 654
irregular cluster, p. 651
irregular galaxy, p. 642
isotropic, p. 642
lenticular galaxy, p. 642
Local Group, p. 650
maser, p. 646
poor cluster, p. 650
redshift, p. 647
regular cluster, p. 651
rich cluster, p. 650
spiral galaxy, p. 639
standard candle, p. 643
starburst galaxy, p. 654
supercluster, p. 651
Tully-Fisher relation, p. 645
tuning fork diagram, p. 642
void, p. 652

Key Ideas

INTERACTIVE EXERCISE 24.2

The Hubble Classification: Galaxies can be grouped into four major categories: spirals, barred spirals, ellipticals, and irregulars.

• The disks of spiral and barred spiral galaxies are sites of active star formation.

• Elliptical galaxies are nearly devoid of interstellar gas and dust, and so star formation is severely inhibited.

• Lenticular galaxies are intermediate between spiral and elliptical galaxies.

• Irregular galaxies have ill-defined, asymmetrical shapes. They are often found associated with other galaxies.

Distance to Galaxies: Standard candles, such as Cepheid variables and the most luminous supergiants, globular clusters, H II regions, and supernovae in a galaxy, are used in estimating intergalactic distances.

• The Tully-Fisher relation, which correlates the width of the 21-cm line of hydrogen in a spiral galaxy with its luminosity, can also be used for determining distance. A method that can be used for elliptical galaxies is the fundamental plane, which relates the galaxy's size to its surface brightness distribution and to the motions of its stars.

The Hubble Law: There is a simple linear relationship between the distance from the Earth to a remote galaxy and the redshift of that galaxy (which is a measure of the speed with which it is receding from us). This relationship is the Hubble law, $v = H_0 d$.

• The value of the Hubble constant, H_0, is not known with certainty but is close to 73 km/s/Mpc.

Clusters and Superclusters: Galaxies are grouped into clusters rather than being scattered randomly throughout the universe.

• A rich cluster contains hundreds or even thousands of galaxies; a poor cluster, often called a group, may contain only a few dozen.

• A regular cluster has a nearly spherical shape with a central concentration of galaxies; in an irregular cluster, galaxies are distributed asymmetrically.

• Our Galaxy is a member of a poor, irregular cluster called the Local Group.

• Rich, regular clusters contain mostly elliptical and lenticular galaxies; irregular clusters contain spiral, barred spiral, and irregular galaxies along with ellipticals.

• Giant elliptical galaxies are often found near the centers of rich clusters.

Galactic Collisions and Mergers: When two galaxies collide, their stars pass each other, but their interstellar media collide violently, either stripping the gas and dust from the galaxies or triggering prolific star formation.

• The gravitational effects during a galactic collision can throw stars out of their galaxies into intergalactic space.

• Galactic mergers may occur; a large galaxy in a rich cluster may tend to grow steadily through galactic cannibalism, perhaps producing in the process a giant elliptical galaxy.

The Dark-Matter Problem: The luminous mass of a cluster of galaxies is not large enough to account for the observed motions of the galaxies; a large amount of unobserved mass must also be present. This situation is called the dark-matter problem.

• Hot intergalactic gases in rich clusters account for a small part of the unobserved mass. These gases are detected by their X-ray emission. The remaining unobserved mass is probably in the form of dark-matter halos that surround the galaxies in these clusters.

• Gravitational lensing of remote galaxies by a foreground cluster enables astronomers to glean information about the distribution of dark matter in the foreground cluster.

Formation and Evolution of Galaxies: Observations indicate that galaxies arose from mergers of several smaller gas clouds.

• Whether a protogalaxy evolves into a spiral galaxy or an elliptical galaxy depends on its initial rate of star formation.

Questions

Review Questions

1. Why did many nineteenth-century astronomers think that the "spiral nebulae" are part of the Milky Way?
2. What was the Shapley-Curtis "debate" all about? Was a winner declared at the end of the "debate"? Whose ideas turned out to be correct?
3. How did Edwin Hubble prove that the Andromeda "Nebula" is not a nebula within our Milky Way Galaxy?
4. Are any galaxies besides our own visible with the naked eye from Earth? If so, which one(s)?
5. An educational publication for children included the following statement: "The Sun is in fact the only star in our galaxy. All of the other stars in the sky are located in other galaxies." How would you correct this statement?
6. What is the Hubble classification scheme? Which category includes the largest galaxies? Which includes the smallest? Which category of galaxy is the most common?
7. Which is more likely to have a blue color, a spiral galaxy or an elliptical galaxy? Explain why.
8. Which types of galaxies are most likely to have new stars forming? Describe the observational evidence that supports your answer.
9. Explain why the apparent shape of an elliptical galaxy may be quite different from its real shape.
10. Why do astronomers suspect that the Hubble tuning fork diagram does not depict the evolutionary sequence of galaxies?
11. Why are Cepheid variable stars useful for finding the distances to galaxies? Are there any limitations on their use for this purpose?
12. Why are Type Ia supernovae useful for finding the distances to very remote galaxies? Can they be used to find the distance to any galaxy you might choose? Explain.
13. What is the Tully-Fisher relation? How is it used for measuring distances? Can it be used for galaxies of all kinds? Why or why not?
14. What are masers? How can they be used to measure the distance to a galaxy?
15. What is the Hubble law? How can it be used to determine distances?
16. How did the discovery of the Hubble Law reinforce the idea that the spiral "nebulae" could not be part of the Milky Way?
17. Why do you suppose it has been so difficult to determine the value of H_0?
18. Some galaxies in the Local Group exhibit blueshifted spectral lines. Why aren't these blueshifts violations of the Hubble law?
19. What are the differences between regular and irregular clusters?
20. What is the difference between a cluster and a supercluster? Are both clusters and superclusters held together by their gravity?
21. What measurements do astronomers make to construct three-dimensional maps of the positions of galaxies in space?
22. Describe what voids are and what they tell us about the large-scale structure of the universe.
23. Why is the intracluster gas in galaxy clusters at such high temperatures?
24. What are starburst galaxies? How can they be produced by collisions between galaxies?
25. Why do giant elliptical galaxies dominate rich clusters but not poor clusters?
26. What evidence is there for the existence of dark matter in clusters of galaxies?
27. What is gravitational lensing? Why don't we notice the gravitational lensing of light by ordinary objects on Earth?
28. How do observations of galaxy cluster 1E0657-56 help constrain the nature of dark matter?
29. What observations suggest that present-day galaxies formed from smaller assemblages of matter?
30. On what grounds do astronomers think that in the past, spiral galaxies were more numerous in rich clusters than they are today? What could account for this excess of spiral galaxies in the past?

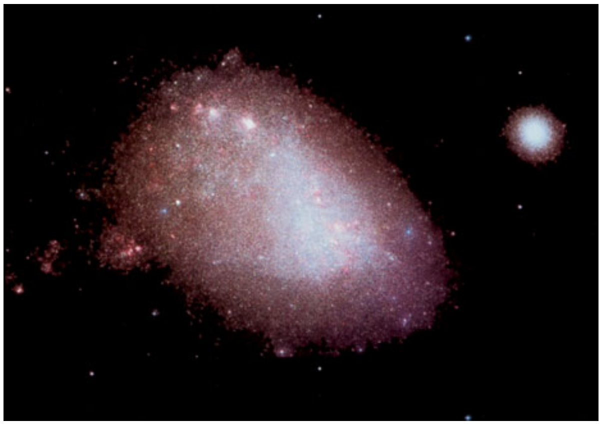
Advanced Questions

Questions preceded by an asterisk (*) involve topics discussed in the Boxes.

Problem-solving tips and tools

Box 1-1 explains the small-angle formula, and Box 17-3 discusses the relationship among apparent magnitude, absolute magnitude, and distance. As Box 23-2 explains, a useful form of Kepler's third law is $M = rv^2/G$, where M is the mass within an orbit of radius r, v is the orbital speed, and G is the gravitational constant. Another form of Kepler's third law, particularly useful for two stars or two galaxies orbiting each other, is given in Section 17-9. The volume of a sphere of radius r is $4\pi r^3/3$. The mass of a hydrogen atom (^{1}H) is given in Appendix 7.

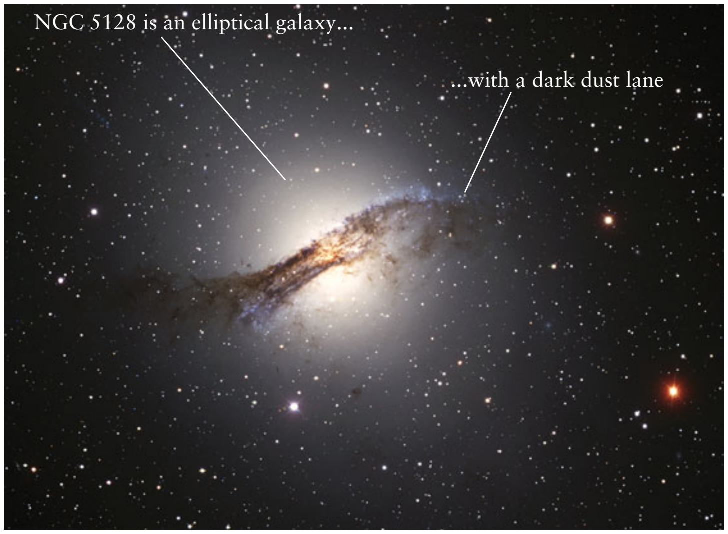
31. Hubble made his observations of Cepheids in M31 using the 100-inch (2.5-meter) telescope on Mount Wilson. Completed in 1917, this was the largest telescope in the world when Hubble carried out his observations in 1923. Why was it helpful to use such a large telescope?
32. The image on page 665 shows the Small Magellanic Cloud (SMC), an irregular galaxy that orbits the Milky Way. The SMC is 63 kpc (200,000 ly) from Earth and 8 kpc (26,000 ly) across, and can be seen with the naked eye from southern latitudes. What features of this image indicate that there has been recent star formation in the SMC? Explain.



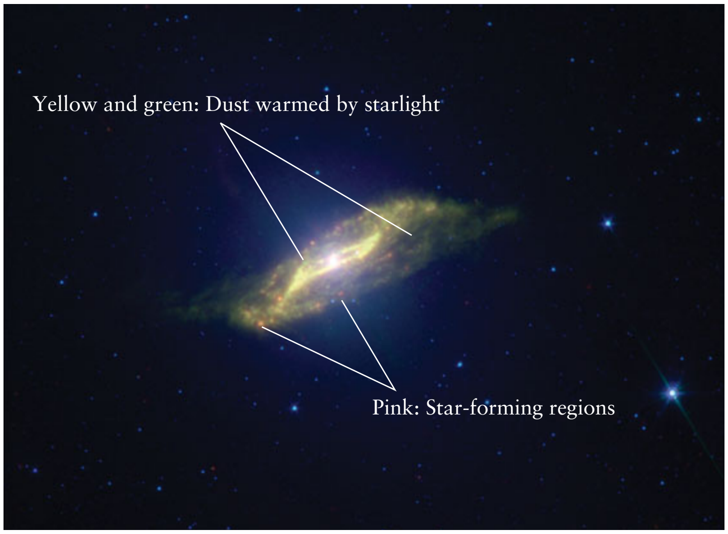
(Anglo-Australian Observatory)

33. When the results from the Hipparcos mission were released, with new and improved measurements of the parallaxes of *nearby* stars within 500 pc, astronomers had to revise the distances to many *remote* galaxies millions of parsecs away. Explain why.
34. As Figure 19-19 shows, there are two types of Cepheid variables. Type I Cepheids are metal-rich stars of Population I, while Type II Cepheids are metal-poor stars of Population II. (**a**) Which type of Cepheid variables would you expect to be found in globular clusters? Which type would you expect to be found in the disk of a spiral galaxy? Explain your reasoning. (**b**) When Hubble discovered Cepheid variables in M31, the distinction between Type I and Type II Cepheids was not yet known. Hence, Hubble thought that the Cepheids seen in the disk of M31 were identical to those seen in globular clusters in our own Galaxy. As a result, his calculations of the distance to M31 were in error. Using Figure 19-19, explain whether Hubble's calculated distance was too small or too large.
*35. Astronomers often state the distance to a remote galaxy in terms of its distance modulus, which is the difference between the apparent magnitude m and the absolute magnitude M (see Box 17-3). (**a**) By measuring the brightness of supernova 1994I in the galaxy M51 (see Figure 24-2), the distance modulus for this galaxy was determined to be $m - M = 29.2$. Find the distance to M51 in megaparsecs (Mpc). (**b**) A separate distance determination, which involved measuring the brightnesses of planetary nebulae in M51, found $m - M = 29.6$. What is the distance to M51 that you calculate from this information? (**c**) What is the difference between your answers to parts (a) and (b)? Compare this difference with the 750-kpc distance from the Earth to M31, the Andromeda Galaxy. The difference between your answers illustrates the uncertainties involved in determining the distances to galaxies!
*36. Suppose you discover a Type Ia supernova in a distant galaxy. At maximum brilliance, the supernova reaches an apparent magnitude of +10. How far away is the galaxy? (*Hint:* See Box 24-1.)
37. The masers that orbit the center of the spiral galaxy M106 travel at an orbital speed of about 1000 km/s. Astronomers observed these masers at intervals of 4 months. (**a**) What distance does a single maser move during a 4-month period? Give your answer in kilometers and in AU. (**b**) During this period, a maser moving across the line of sight (like the maser shown in green in Figure 24-15) appeared to move through an angle of only 10^{-5} arcsec. Calculate the distance to the galaxy.
38. The average radial velocity of galaxies in the Hercules cluster pictured in Figure 24-18 is 10,800 km/s. (**a**) Using $H_0 = 73$ km/s/Mpc, find the distance to this cluster. Give your answer in megaparsecs and in light-years. (**b**) How would your answer to (a) differ if the Hubble constant had a smaller value? A larger value? Explain.
39. A certain galaxy is observed to be receding from the Sun at a rate of 7500 km/s. The distance to this galaxy is measured independently and found to be 1.4×10^8 pc. From these data, what is the value of the Hubble constant?
*40. In the spectrum of the galaxy NGC 4839, the K line of singly ionized calcium has a wavelength 403.2 nm. (**a**) What is the redshift of this galaxy? (*Hint:* See Box 24-2.) (**b**) Determine the distance to this galaxy using the Hubble law with $H_0 = 73$ km/s/Mpc.
*41. The galaxy RD1 has a redshift of $z = 5.34$. (**a**) Determine its recessional velocity v in km/s and as a fraction of the speed of light. (**b**) What recessional velocity would you have calculated if you had erroneously used the low-speed formula relating z and v? Would using this formula have been a small or large error? (**c**) According to the Hubble law, what is the distance from Earth to RD1? Use $H_0 = 73$ km/s/Mpc for the Hubble constant, and give your answer in both megaparsecs and light-years.
42. It is estimated that the Coma cluster (see Figure 24-21) contains about 10^{13} $M_\odot$ of intracluster gas. (**a**) Assuming that this gas is made of hydrogen atoms, calculate the total number of intracluster gas atoms in the Coma cluster. (**b**) The Coma cluster is roughly spherical in shape, with a radius of about 3 Mpc. Calculate the number of intracluster gas atoms per cubic centimeter in the Coma cluster. Assume that the gas fills the cluster uniformly. (**c**) Compare the intracluster gas in the Coma cluster with the gas in our atmosphere (3×10^{19} molecules per cubic centimeter, temperature 300 K); a typical gas cloud within our own Galaxy (a few hundred molecules per cubic centimeter, temperature 50 K or less); and the corona of the Sun (10^5 atoms per cubic centimeter, temperature 10^6 K).
*43. Two galaxies separated by 600 kpc are orbiting each other with a period of 40 billion years. What is the total mass of the two galaxies?
*44. Figure 24-29 shows the rotation curve of the Sa galaxy NGC 4378. Using data from that graph, calculate the orbital period of stars 20 kpc from the galaxy's center. How much mass lies within 20 kpc from the center of NGC 4378?
45. How might you determine what part of a galaxy's redshift is caused by the galaxy's orbital motion about the center of mass of its cluster?

46. The accompanying image shows the unusual elliptical galaxy NGC 5128. Explain how the properties of this galaxy seen in the infrared image can be explained if NGC 5128 is the result of a merger of an elliptical galaxy and a spiral galaxy.



R I **V** U X G



R **I** V U X G

(visible: Eric Peng, Herzberg Institute of Astrophysics and NOAO/AURA/NSF; infrared: Jocelyn Keene, NASA/JPL and Caltech)

47. Explain why the dark matter in galaxy clusters could not be neutral hydrogen.
48. According to Figure 24-34*c*, elliptical galaxies continue to form stars for about a billion years after they form. Give an argument why we might expect to find some Population I stars in an elliptical galaxy. (*Hint:* Table 19-1 gives the main-sequence lifetimes for stars of different masses.)

Discussion Questions

49. The Earth is composed principally of heavy elements, such as silicon, nickel, and iron. Would you be likely to find such planets orbiting stars in the disk of a spiral galaxy? In the nucleus of a spiral galaxy? In an elliptical galaxy? In an irregular galaxy? Explain your answers.
50. Discuss what observations you might make to determine whether or not the various Hubble types of galaxies represent some sort of evolutionary sequence.
51. Discuss the advantages and disadvantages of using the various standard candle distance indicators to obtain extragalactic distances.
52. How would you distinguish star images from unresolved images of remote galaxies on a CCD?
53. Describe what sorts of observations you might make to search for as-yet-undiscovered galaxies in our Local Group. How is it possible that such galaxies might still remain to be discovered? In what part of the sky would these galaxies be located? What sorts of observations might reveal these galaxies?

Web/eBook Questions

54. When galaxies pass close to one another, as should happen frequently in a rich cluster, tidal forces between the galaxies can strip away their outlying stars. The result should be a loosely dispersed sea of "intergalactic stars" populating the space between galaxies in a cluster. Search the World Wide Web for information about intergalactic stars. Have they been observed? If so, where are they found? What would our nighttime sky look like if our Sun were an intergalactic star?

WEB LINK 24.22

55. The Hubble Space Telescope (HST) has made extensive observations of very distant galaxies. Visit the HST Web site to learn about these investigations. How far back in time has HST been able to look? What sorts of early galaxies are observed? What is the current thinking about how galaxies formed and evolved?

ANIMATION 24.4

56. **The Formation of "The Mice."** Access and view the animation "The Formation of 'The Mice'" in Chapter 24 of the *Universe* Web site or eBook. Based on this computer simulation, what will be the final fate of the two galaxies that make up the Antennae? (See the *Cosmic Connections* figure in Section 24-7.) Speculate on what might have happened if the two galaxies had collided at a much greater speed.

ANIMATION 24.7

57. **Radiation from a Rotating Galaxy.** Access and view the animation "Radiation from a Rotating Galaxy" in Chapter 24 of the *Universe* Web site or eBook. Describe how the animation would have to be changed for a spiral galaxy of the same size but (**a**) greater mass and (**b**) smaller mass.

Activities

Observing Projects

WEB LINK 24.23

58. Using a telescope with an aperture of at least 30 cm (12 in.), observe as many of the spiral galaxies listed in the following table as you can. If your copy of the book comes with the *Starry Night Enthusiast™* program, use it to help determine when these galaxies can best be viewed. Many of these galaxies are members of the Virgo cluster, which can best be seen from March through June. Because all galaxies are quite faint, be sure to schedule your observations for a moonless night. The best

view is obtained when a galaxy is near the meridian. While at the eyepiece, make a sketch of what you see. Can you distinguish any spiral structure? After completing your observations, compare your sketches with photographs found in an online catalog of Messier objects (the "M" in the galaxy designations stands for Messier).

Spiral galaxy	Right ascension	Declination	Hubble type
M31 (NGC 224)	$0^h\ 42.7^m$	$+41°\ 16'$	Sb
M58 (NGC 4579)	12 37.7	+11 49	Sb
M61 (NGC 4303)	12 21.9	+4 28	Sc
M63 (NGC 5055)	13 15.8	+42 02	Sb
M64 (NGC 4826)	12 56.7	+21 41	Sb
M74 (NGC 628)	1 36.7	+15 47	Sc
M83 (NGC 5236)	13 37.0	−29 52	Sc
M88 (NGC 4501)	12 32.0	+14 25	Sb
M90 (NGC 4569)	12 36.8	+13 10	Sb
M91 (NGC 4548)	12 35.4	+14 30	SBb
M94 (NGC 4736)	12 50.9	+41 07	Sb
M98 (NGC 4192)	12 13.8	+14 54	Sb
M99 (NGC 4254)	12 18.8	+14 25	Sc
M100 (NGC 4321)	12 22.9	+15 49	Sc
M101 (NGC 5457)	14 03.2	+54 21	Sc
M104 (NGC 4594)	12 40.0	−11 37	Sa
M108 (NGC 3556)	11 11.5	+55 40	Sc

Note: The right ascensions and declinations are given for epoch 2000.

WEB LINK 24.23

59. Using a telescope with an aperture of at least 30 cm (12 in.), observe as many of the following elliptical galaxies as you can. Six of these galaxies are in the Virgo cluster, which is conveniently located in the evening sky from March through June. If your copy of the book comes with the *Starry Night Enthusiast*™ program, use it to help determine when these galaxies can best be viewed. As in the previous exercise, be sure to schedule your observations for a moonless night, when the galaxies you wish to observe will be near the meridian. Do these elliptical galaxies differ in appearance from spiral galaxies?

Elliptical galaxy	Right ascension	Declination	Hubble type
M49 (NGC 4472)	$12^h\ 29.8^m$	$+8°\ 00'$	E4
M59 (NGC 4621)	12 42.0	+11 39	E3
M60 (NGC 4649)	12 43.7	+11 33	E1
M84 (NGC 4374)	12 25.1	+12 53	E1
M86 (NGC 4406)	12 26.2	+12 57	E3
M89 (NGC 4552)	12 35.7	+12 33	E0
M110 (NGC 205)	00 40.4	+41 41	E6

Note: The right ascensions and declinations are given for epoch 2000.

WEB LINK 24.23

60. Using a telescope with an aperture of at least 30 cm (12 in.), observe as many of the following interacting galaxies as you can. If your copy of the book comes with the *Starry Night Enthusiast*™ program, use it to help determine when these galaxies can best be viewed. As in the previous exercises, be sure to schedule your observations for a moonless night, when the galaxies you wish to observe will be near the meridian. While at the eyepiece, make a sketch of each galaxy. Can you distinguish hints of interplay among the galaxies? After completing your observations, compare your sketches with photographs found in an online catalog of Messier objects (the "M" in the galaxy designations stands for Messier).

Interacting galaxies	Right ascension	Declination
M51 (NGC 5194)	$13^h\ 29.9^m$	$+47°\ 12'$
NGC 5195	13 30.0	+47 16
M65 (NGC 3623)	11 18.9	+13 05
M66 (NGC 3627)	11 20.2	+12 59
NGC 3628	11 20.3	+13 36
M81 (NGC 3031)	9 55.6	+69 04
M82 (NGC 3034)	9 55.8	+69 41
M95 (NGC 3351)	10 44.0	+11 42
M96 (NGC 3368)	10 46.8	+11 49
M105 (NGC 3379)	10 47.8	+12 35

Note: The right ascensions and declinations are given for epoch 2000.

STARRY NIGHT

61. Use the *Starry Night Enthusiast*™ program to observe other galaxies. Select **Favourites > Guides > Atlas** to display the entire celestial sphere. Open the **Options** pane, expand the **Deep Space** list, and click the box for **Messier Objects** to display these objects. Open the **Find** pane and select each of the following galaxies in turn. Zoom in as necessary to see the shape of the galaxy to determine its Hubble classification as well as you can, and explain your reasoning: (i) Virgo A (M87); (ii) M105; (iii) M102; (iv) M104; (v) M109.

STARRY NIGHT

62. Use the *Starry Night Enthusiast*™ program to examine clusters of galaxies. Select **Favourites > Deep Space > Virgo Cluster** to center this collection of galaxies in the view, as seen from a distance of about 53 Mly from the Sun. You are looking at a three-dimensional view of the Tully Database. Open the **Find** pane and locate **Virgo A,** one of the galaxies examined in the previous question, which is close to the center of this cluster of galaxies. Right-click on this galaxy to open the contextual menu (Macintosh users Ctrl-click on this galaxy) and click on **Highlight "GA Virgo Cluster" Filament** to highlight this cluster in yellow. Click on the "**up**" arrow in the Viewing Location to move to about 30 Mly from the Sun. Hold down the Shift key while holding down the mouse button and move the mouse to use the location scroller to rotate this rich group of galaxies. (**a**) Describe the general shape of the Virgo cluster. (**b**) As you rotate the Virgo cluster, you should notice other groupings of galaxies. Stop this rotation at some position and make a sketch of the screen, circling what you believe are other groups on your sketch. Right-click (Macintosh users Ctrl-click) on one of the other clusters (and clouds and extensions) near to the Virgo Cluster to open the contextual menu and use the **Highlight** option to see how astronomers have grouped these other galaxies. Repeat this process until you have identified all of the clusters around Virgo. Outline and label these clusters on your drawing. (**c**) Choose three of these clusters, center on each in turn and right-click (Ctrl-click on a Macintosh) to open the contextual menu and use the **Centre** command. Use the **Zoom** facility and **location scroller** to move around these collections of galaxies, and describe their distribution compared to the Virgo cluster. For example, what are their shapes and relative sizes compared to Virgo and to each other? Are they rich spherical concentrations or walls of galaxies?

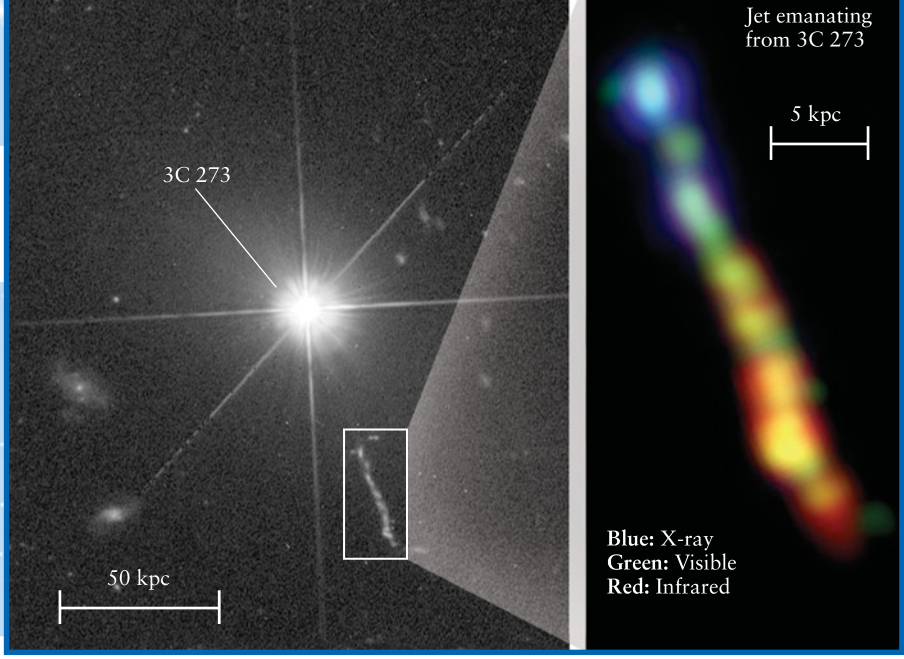
Collaborative Exercises

63. In the early twentieth century, there was considerable debate about the nature of spiral nebulae and their distance from us, but the debate was resolved by improvements in technology. As a group, list three issues that we, as a culture, did not understand in the past but understand today, and explain why we now have that understanding.

64. Even though there are billions of galaxies, there are not billions of different kinds. In fact, galaxies are classified according to their appearance. As a group, dig into your book bags and put all of the writing implements (pens, pencils, highlighters, and so on) you have in a central pile. Remember which ones are yours! Determine a classification scheme that sorts the writing implements into at least three to six piles. Write down the scheme and the number of items in each pile. Ask the group next to you to use your scheme and sort your materials. Correct any ambiguities before submitting your classification scheme.

65. Imagine your company, Astronomical Artistry, has been contracted by the local marching band to create a football halftime show about spiral galaxies. How exactly would you design the positions of the band members on the field to represent the different spiral galaxies of classes Sa, Sb, and Sc? Create two columns on your paper by drawing a line from top to bottom, drawing sketches in the left-hand column and writing a description of each sketch in the right-hand column. Also include what the band's opening formation and final formation should be.

25 Quasars and Active Galaxies



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

WEB LINK 25.1

A powerful jet tens of thousands of parsecs in length blasts outward from quasar 3C 273. (Left: NASA and J. Bahcall, Institute for Advanced Study; right: NASA/JPL-Caltech/Yale University)

An ordinary star emits radiation primarily at ultraviolet, visible, and infrared wavelengths, in proportions that reflect the star's surface temperature. Ordinary galaxies, too, emit most strongly in these wavelength regions. But the object shown here, called 3C 273, is outrageously different: It emits strongly over an immense range of wavelengths from radio to X-ray. 3C 273 is also intensely luminous, with thousands of times the radiation output of an ordinary galaxy. And, as this X-ray image shows, some source of energy within 3C 273 causes it to eject a glowing, high-speed jet that extends outward for hundreds of thousands of light-years.

3C 273 is a *quasar*—one of many thousands of distant objects whose luminosity is far too great to come from starlight alone. Quasars are intimately related to *active galaxies,* whose tremendous luminosity can fluctuate substantially over a period of months, weeks, or even days.

What makes quasars and active galaxies so tremendously energetic? The quest to answer this question was one of the greatest scientific detective stories of the twentieth century. Like a detective story, in this chapter we will begin with the often startling evidence about quasars and active galaxies that confronted astronomers. The plot will thicken as we reveal the prime suspects—supermassive black holes that are thought to lie at the center of every quasar and active galaxy. As matter accretes onto such a black hole, it releases fantastic amounts of energy and can produce powerful jets. We will even catch the culprit in the act, with an image that shows a quasar about to devour material to feed its central black hole.

25-1 Quasars look like stars but have huge redshifts

Many of the most revolutionary discoveries in astronomy have been the result of advances in technology. So it was in 1610 when Galileo used a new device called the telescope to scan the sky, and in the process found evidence that shattered the ancient geocentric model of the universe (see Section 4-3). And so it was during the middle years of the twentieth century, when the new technique of radio astronomy revealed the existence of remote, dazzlingly luminous objects called *quasi-stellar radio sources* or *quasars*.

The Discovery of Quasars

WEB LINK 25.2

Grote Reber, a radio engineer and ham radio enthusiast, built the first true radio telescope in 1936 in his backyard in Illinois (see Section 6-6). By 1944 he had detected strong radio emissions from sources in the constellations Cassiopeia, Sagittarius, and Cygnus.

Two of these sources, named Cassiopeia A and Sagittarius A, happen to lie in our own Galaxy; they are a supernova remnant (see Figure 20-25) and the center of the Galaxy (see Section 23-6). The nature of the

Examining the spectra of quasars revealed that they are immensely distant

Learning Goals

By reading the sections of this chapter, you will learn

25-1 The distinctive features of quasars

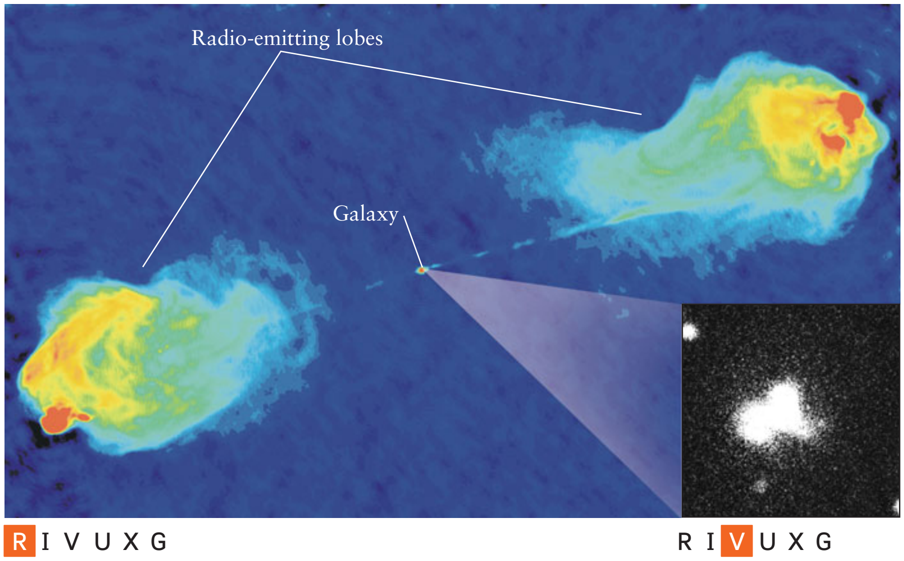
25-2 The connection between quasars and distant galaxies

25-3 What Seyfert galaxies and radio galaxies are and how they compare to quasars

25-4 The properties of active galactic nuclei

25-5 How supermassive black holes can power active galactic nuclei

25-6 Why many active galaxies emit ultrafast jets of material



(a) Radio image of Cygnus A

(b) Visible-light close-up of the central galaxy

Figure 25-1

Cygnus A (3C 405) **(a)** This false-color radio image from the Very Large Array shows that most of the emission from Cygnus A comes from luminous radio lobes located on either side of a peculiar galaxy. (Red indicates the strongest radio emission, while blue indicates the faintest). Each lobe extends about 70 kpc (230,000 ly) from the galaxy. **(b)** The galaxy at the heart of Cygnus A has a substantial redshift, so it must be extremely far from Earth (about 230 Mpc, or 740 million ly). To be so distant and yet be one of the brightest radio sources in the sky, Cygnus A must have an enormous energy output. (a: R. A. Perley, J. W. Dreher, and J. J. Cowan, NRAO/AUI; b: Palomar Observatory)

third source, called Cygnus A (Figure 25-1*a*), proved more elusive. The mystery only deepened in 1951, when Walter Baade and Rudolph Minkowski used the 200-inch (5-meter) optical telescope on Palomar Mountain to discover a dim, strange-looking galaxy at the position of Cygnus A (see Figure 25-1*b*).

When Baade and Minkowski photographed the spectrum of Cygnus A, they were surprised to find a number of bright *emission* lines. By contrast, a normal galaxy has *absorption* lines in its spectrum (see Section 24-3 and Figure 24-16). This absorption takes place in the atmospheres of the stars that make up the galaxy, as we described in Section 17-5. In order for Cygnus A to have emission lines, something must be exciting and ionizing its atoms. Furthermore, the wavelengths of Cygnus A's emission lines are all shifted by 5.6% toward the red end of the spectrum. Astronomers use the letter z to denote redshift (see Section 24-5), and thus this object has $z = 0.056$, corresponding to a recessional velocity of 16,000 km/s.

If Cygnus A participates in the same Hubble flow as clusters of galaxies (described in Section 24-5), then this recessional velocity corresponds to a tremendous distance from the Earth—about 230 Mpc (740 million ly) if the Hubble constant equals 73 km/s/Mpc. Yet despite its tremendous distance from the Earth, radio waves from Cygnus A can be picked up by amateur astronomers with backyard equipment. This means that Cygnus A must be one of the most luminous radio sources in the sky. In fact, its radio luminosity is 10^7 times as great as that of an ordinary galaxy like the Milky Way. The object that creates the Cygnus A radio emission must be something quite extraordinary.

Cygnus A was not the only curious radio source to draw the attention of astronomers. In 1960, Allan Sandage used the 200-inch telescope to discover a "star" at the location of a radio source designated 3C 48 (Figure 25-2). (The "3C" refers to the *Third Cambridge Catalogue,* a compendium of radio sources.) Ordinary stars are not strong sources of radio emission, so 3C 48 must be something unusual. Like Cygnus A, its spectrum showed a series of emission lines, but astronomers were unable to identify the chemical elements that produced these lines.

Two years later, astronomers discovered a similar starlike optical counterpart to the radio source 3C 273. This "star" is even

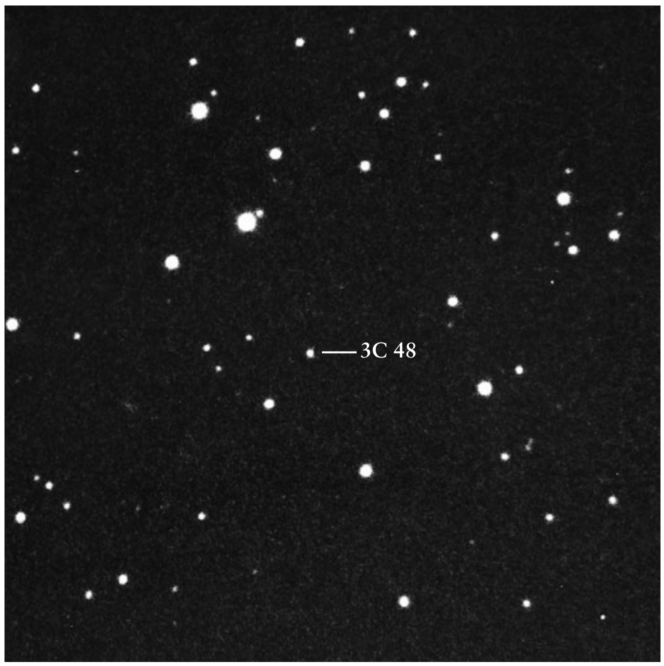


Figure 25-2 R I V U X G

The Quasar 3C 48 This object in the constellation Triangulum was thought at first to be a nearby star that happened to emit radio waves. In fact, the redshift of 3C 48 is so great ($z = 0.367$) that, according to the Hubble law, it must presently be approximately 1400 Mpc (4.6 billion ly) away. (Alex G. Smith, Rosemary Hill Observatory, University of Florida)

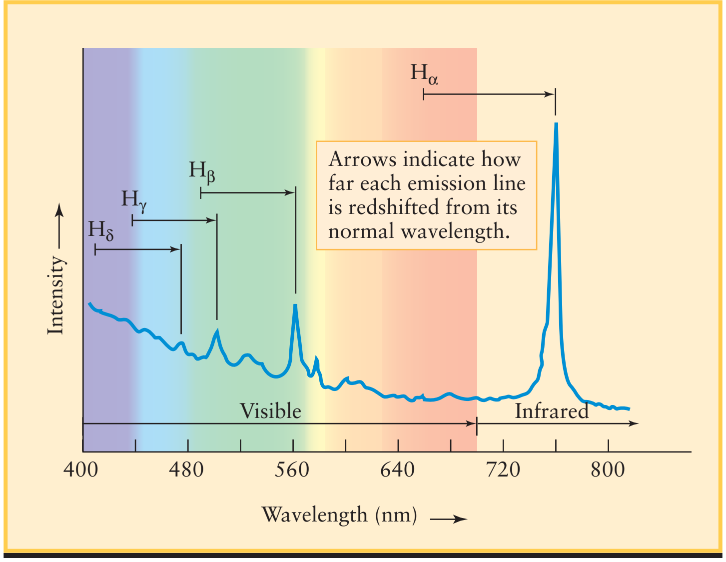


Figure 25-3

The Spectrum of 3C 273 The visible-light and infrared spectrum of this quasar is dominated by four bright emission lines of hydrogen (see Section 5-8). The redshift is $z = 0.158$, so the wavelength of each line is 15.8% greater than for a sample of hydrogen on Earth. For example, the wavelength of H_β is shifted from 486 nm (a blue-green wavelength) to $1.158 \times$ (486 nm) = 563 nm (a yellow wavelength).

more unusual, with a luminous jet protruding from one side. (The image that opens this chapter shows 3C 273 and its accompanying jet.) Like 3C 48, its visible spectrum contains a series of emission lines that no one could explain.

Clues from Spectra

Although 3C 48 and 3C 273 were clearly oddballs, many astronomers thought they were just strange stars in our own Galaxy. A breakthrough occurred in 1963, when Maarten Schmidt at Caltech took another look at the spectrum of 3C 273. He realized that four of its brightest emission lines are positioned relative to one another in precisely the same way as four of the Balmer lines of hydrogen. However, these emission lines from 3C 273 were all shifted to much longer wavelengths than the usual wavelengths of the Balmer lines. Schmidt determined that 3C 273 has a redshift of $z = 0.158$, corresponding to a recessional velocity of 44,000 km/s (15% of the speed of light). No star could be moving this fast and remain within our Galaxy. Hence, Schmidt concluded that 3C 273 could not be a nearby star, but must lie outside the Milky Way.

According to the Hubble law, the recessional velocity of 3C 273 implies that its present distance from us is 630 Mpc (2 billion ly). To be detected at such distances, 3C 273 must be an extraordinarily powerful source of both visible light and radio radiation.

Figure 25-3 shows the visible and near-infrared spectrum of 3C 273. As we saw in Section 6-5 (review Figure 6-21), such a spectrum is a graph of intensity versus wavelength on which emission lines appear as peaks and absorption lines appear as valleys. Emission lines are caused by excited atoms, which emit radiation at specific wavelengths.

Upon learning how Schmidt deciphered the spectrum of 3C 273, two other Caltech astronomers, Jesse Greenstein and Thomas Matthews, found they could identify the spectral lines of 3C 48 as having suffered an even larger redshift of $z = 0.367$. That shift corresponds to a speed three-tenths that of light, which means that 3C 48 is presently twice as far away as 3C 273, or approximately 1400 Mpc (4.6 billion ly) from Earth.

Quasars: High Redshifts, Extreme Distances

WEB LINK 25.3

Because of their strong radio emission and starlike appearance, 3C 48 and 3C 273 were dubbed *quasi-stellar radio sources,* a term soon shortened to **quasars.** After the first quasars were discovered by their radio emission, many similar, high-redshift, starlike objects were found that emit little or no radio radiation. These "radio-quiet" quasars were originally called *quasi-stellar objects,* or *QSOs*, to distinguish them from radio emitters. Today, however, the term "quasar" is often used to include both types. Only about 10% of quasars are "radio-loud" like 3C 48 and 3C 273.

Thanks to sky surveys like those we discussed in Section 24-6, more than 100,000 quasars are now known. They all look rather like stars (see Figure 1-10), and all have large redshifts, ranging from 0.06 to at least 5.8 (Figure 25-4). Most quasars have redshifts of 0.3 or more, which implies that they are more

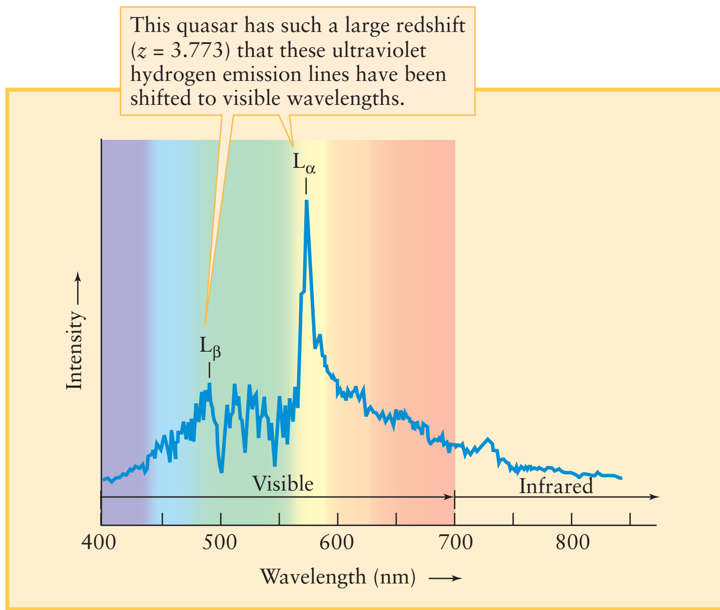


Figure 25-4

The Spectrum of a High-Redshift Quasar This quasar, known as PKS 2000-030, has a redshift $z = 3.773$. This is so large that the Lyman spectral lines L_α and L_β of hydrogen, which are normally at ultraviolet wavelengths, have been shifted into the visible part of the spectrum. The designation of the quasar refers to its position on the celestial sphere and to the Parkes Observatory in Australia, where it was first discovered (see Figure 6-22).

than 1000 Mpc (3 billion light-years) from Earth. For example, the quasar PC 1247+3406 has a redshift of $z = 4.897$, which corresponds to a recessional velocity of more than 94% of the speed of light and (from the Hubble law) a distance of roughly 7950 Mpc (25.9 billion ly) from Earth.

CAUTION! A value of z greater than 1 does *not* mean that a quasar is receding from us faster than the speed of light. At high speeds, the relationship between redshift and recessional velocity must be modified by the special theory of relativity, as explained in Box 24-2. As the speed of a receding source approaches the speed of light, its redshift can become very much greater than 1. An infinite redshift ($z = \infty$) corresponds to a recessional velocity equal to the speed of light.

Light takes time to travel across space, so when we observe a very remote object, we are seeing it in the remote past. This means that for very remote objects, the relationship between redshift and distance from the Earth depends on how the universe has evolved over time. As we will see in Chapter 26, the Hubble law reveals that the universe is expanding. In other words, if you could watch the motions of widely separated clusters of galaxies over millions of years, you would see them gradually moving away from one another. Furthermore, in Chapter 26 we will see evidence that the universe has expanded at different rates at different times in the past. Table 25-1 relates the redshift to the recessional velocity and distance in a model consistent with our present understanding of the expansion of the universe.

CAUTION! Since the universe is expanding, there is more than one way to state the distance between us and a distant quasar. The light from quasar PC 1247+3406 has taken 12.5 billion years to reach us, so we see this quasar as it was 12.5 billion years ago. This elapsed time is called the *light travel time*. But the universe has continued to expand during that time. Hence, PC 1247+3406 is now much more than 12.5 billion light-years away, about 25.9 billion ly or 7950 Mpc. This is called the *comoving radial distance* to the quasar, and is actually the correct distance d to use in the Hubble law $v = H_0d$. (The distances we cited above for Cygnus A, 3C 273, 3C 48, and PKS 1247+3406 are all comoving radial distances.) From this point forward, when we give the distance to a remote galaxy or quasar, we mean the light travel time multiplied by the speed of light—that is, the distance that light traveled to reach us from that galaxy or quasar. Then the distance to a remote object expressed in light-years tells you how many years into the past you are looking when you view that object. (Table 25-1 lists both the distance calculated in this way and the comoving radial distance.) To avoid these issues about how best to define distances, many astronomers prefer simply to state the redshift z of a distant object; they then use the simple rule that the greater the redshift, the more distant the object (see Section 24-5).

WEB LINK 25.4

Table 25-1 Redshift and Distance

Redshift z	Recessional velocity v/c	Distance at which we see the object (Mpc)	Distance at which we see the object (10^9 ly)	Present distance to the object (comoving radial distance) (Mpc)	Present distance to the object (comoving radial distance) (10^9 ly)
0	0	0	0	0	0
0.1	0.095	384	1.25	403	1.32
0.2	0.180	721	2.35	790	2.58
0.3	0.257	1020	3.32	1160	3.79
0.4	0.324	1280	4.17	1510	4.94
0.5	0.385	1510	4.93	1850	6.04
0.75	0.508	2070	6.48	2620	8.54
1	0.600	2350	7.65	3290	10.7
1.5	0.724	2840	9.26	4390	14.3
2	0.800	3160	10.3	5250	17.1
3	0.882	3520	11.5	6500	21.2
4	0.923	3710	12.1	7370	24.0
5	0.946	3830	12.5	8010	26.1
10	0.984	4060	13.2	9790	31.9
Infinite	1	4210	13.7	14500	47.4

This table assumes a Hubble constant $H_0 = 73$ km/s/Mpc, a matter density parameter $\Omega_m = 0.24$, and a dark energy density parameter $\Omega_\Lambda = 0.76$ (see Chapter 26). The distance at which we see the object (in light-years) is equal to the light travel time in years. The present distance to the object is the distance d to be used in the Hubble law, $v = H_0d$.

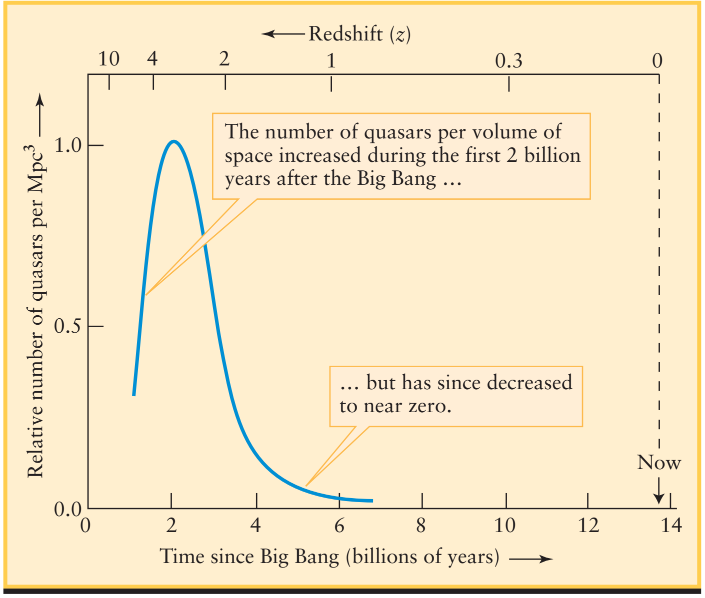


Figure 25-5

Quasars Are Extinct The greater the redshift of a quasar, the farther it is from Earth and the farther back in time we are seeing it. By observing the number of quasars found at different redshifts, astronomers can calculate how the density of quasars in the universe has changed over the history of the universe. The history of quasars is reminiscent of the history of the dinosaurs, which once populated the entire Earth but today are extinct. (Peter Shaver, European Southern Observatory)

Because there are no quasars with small redshifts, it follows that *there are no nearby quasars*. The nearest one is some 250 Mpc (800 million ly) from the Earth. Hence, the absence of nearby quasars means that there have been no quasars for 800 million years. Indeed, the number of quasars began to decline precipitously roughly 10 billion years ago (Figure 25-5). Quasars were a common feature of the universe in the distant past, but there are none in the present-day universe.

25-2 Quasars are the ultraluminous centers of distant galaxies

If quasars are as distant as their redshifts show them to be, they must be extraordinarily luminous to be visible from Earth. What, then, are these strange objects? Quasars cannot simply be very large and luminous galaxies, because their spectra are totally different: The visible-light and infrared spectrum of a galaxy is dominated by absorption lines from the galaxy's stars, while the corresponding spectrum of a quasar is dominated by emission lines. Quasars also emit much more of their light at ultraviolet wavelengths than do stars or galaxies. Hence, whatever the origin of a quasar's radiation, it is not simply starlight. While quasars are not galaxies, we will see compelling evidence that the two are intimately related: Quasars turn out to be ultraluminous objects located at the centers of remote galaxies.

Quasars and Their Luminosities

A quasar's luminosity can be calculated from its apparent brightness and distance using the inverse-square law (see Section 17-2). For example, 3C 273 has a luminosity of about 10^{40} watts, which is equivalent to 2.5×10^{13} (25 trillion) Suns. (The Sun's luminosity is $L_{\odot} = 3.90 \times 10^{26}$ watts.) Generally, quasar luminosities range from about 10^{38} watts up to nearly 10^{42} watts. For comparison, a typical large galaxy, like our own Milky Way, shines with a luminosity of 10^{37} watts, which equals 2.5×10^{10} (25 billion) Suns. Thus, a bright quasar can be many thousands of times more luminous than the entire Milky Way Galaxy.

The most luminous quasars emit 100,000 times more radiation than the entire Milky Way Galaxy

WEB LINK 25.5

When quasars were first discovered, their energy output seemed so absurdly huge that a few astronomers began to question long-held ideas, such as the Hubble law. Traditionally, astronomers had used the Hubble law to calculate the distance to a galaxy from its redshift: the higher the redshift, the greater the distance. In the 1960s the American astronomer Halton C. Arp suggested that the Hubble law might not apply to quasars. If part of a quasar's redshift was caused by some yet-undiscovered phenomenon, then the quasar could be much closer to Earth than the Hubble law would have us believe. If so, a quasar's luminosity would not be so incredibly large.

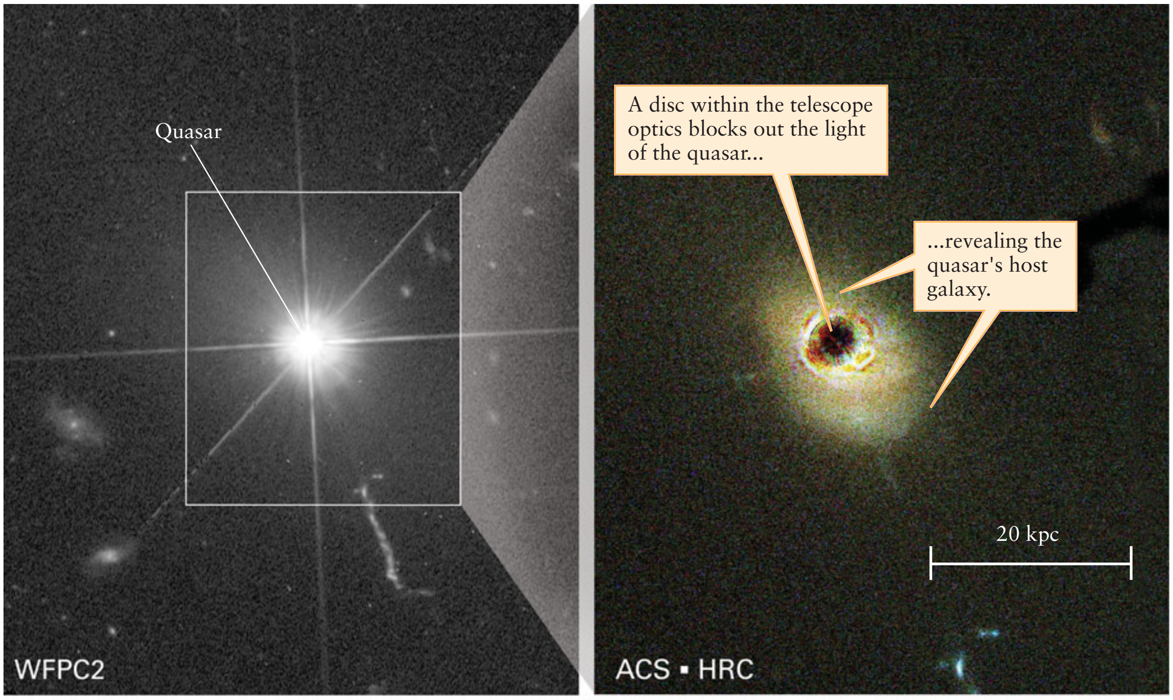
In support of this theory, Arp and his collaborators drew attention to certain high-redshift quasars that seem to be located in or associated with low-redshift galaxies. These examples of "discordant redshifts" fueled a heated debate during the 1970s that was reminiscent of the Shapley-Curtis "debate" 50 years earlier (see Section 24-1). By the 1980s, however, the preponderance of evidence clearly favored the standard interpretation of redshifts according to the Hubble law. Most astronomers today regard Arp's discordant redshifts as simply a *projection effect,* whereby a distant quasar just happens to be in the same part of the sky as a nearby galaxy.

ANALOGY An analogy to the projection effect sometimes happens in photography on Earth. If a photographer poses a person with a telephone pole in the background, the photo can give the misleading impression that the unlucky person has a telephone pole growing out of his head.

What largely ended the redshift debate of the 1970s were observations showing that quasars are associated with remote galaxies. Long-exposure images of quasars showed that they are often found in groups or clusters of galaxies, and that the redshift of each of these quasars is essentially the same as the redshifts of the galaxies that surround it. Thus, the quasar's redshift indicates its distance from Earth, just as do the redshifts of galaxies.

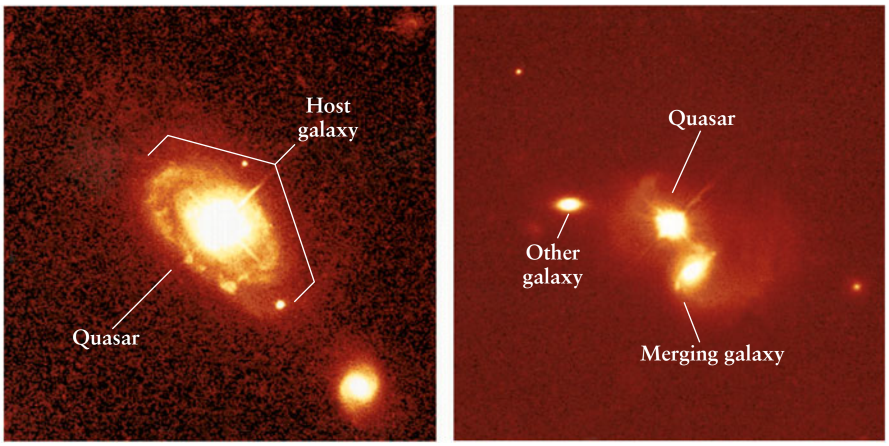
Quasar "Fuzz" and Host Galaxies

A second important link between quasars and galaxies was established in the 1980s. In that decade astronomers first had the



(a) 3C 273

(b) The host galaxy of 3C 273



(c) PG 0052+251 and its host galaxy

(d) PG 1012+008 and its host galaxy

WEB LINK 25.6

Figure 25-6 R I V U X G

Quasars and Their Host Galaxies (a) In this 1994 image from the Hubble Space Telescope (HST), the glare of quasar 3C 273 hides its host galaxy. **(b)** This 2002 image that reveals the host galaxy was made using an upgraded camera aboard HST. **(c)** Quasar PG 0052+251 is located at the center of an apparently normal spiral galaxy at redshift $z = 0.155$. Other quasars are found at the centers of ordinary-looking elliptical galaxies. **(d)** The galaxy that hosts quasar PG 1012+008 (redshift $z = 0.185$) is in the process of merging with a second luminous galaxy. The wispy material surrounding the quasar may have been pulled out of the galaxies by tidal forces (see Figure 24-28). The two merging galaxies are just 9500 pc (31,000 ly) apart. Another small galaxy to the left of the quasar may also be merging with the others. (a: NASA and J. Bahcall (Institute for Advanced Study); b: NASA, A. Martel (JHU), H. Ford (JHU), M. Clampin (STScI), G. Hartig (STScI), G. Illingworth (UCO/Lick Observatory), the ACS Science Team and ESA; c, d: J. Bahcall (Institute for Advanced Study), M. Disney (University of Wales), and NASA)

technology to examine the faint "fuzz" seen around the images of some quasars. The spectrum of this fuzz shows stellar absorption lines, indicating that each of these quasars is embedded in a galaxy whose light is far fainter than that of the quasar itself. In each case, the absorption lines of the stars have the same redshift as the quasar's emission lines, further supporting the idea that quasars are at distances indicated by their redshifts and the Hubble law.

It is very difficult to observe the "host galaxy" in which a quasar is located because the quasar's light overwhelms light from the galaxy's stars (Figure 25-6*a*). Nevertheless, improved technology and painstaking observations have revealed some basic properties of these host galaxies (Figure 25-6*b*). Relatively nearby radio-quiet quasars (with redshifts of less than 0.2, corresponding to distances of less than about 750 Mpc) tend to be located in spiral galaxies, whereas radio-loud quasars as well as more distant radio-quiet quasars tend to be located in ellipticals (Figure 25-6*c*). However, a large percentage of these host galaxies have distorted shapes or are otherwise peculiar. Many have nearby companion galaxies, suggesting a link between collisions or mergers and the quasar itself (Figure 25-6*d*). We will see in Section 25-6 how the supermassive black hole model explains this link, and also helps explain why quasars are found only at great distances from Earth—and hence only in the distant past of the universe.

While observations like those in Figure 25-6 show that quasars lie at the centers of galaxies, they do not really explain what quasars *are*. As we will see, important clues come from a class of galaxies that are intermediate in luminosity between quasars and normal galaxies.

25-3 Seyferts and radio galaxies bridge the gap between normal galaxies and quasars

We have seen that some scientists in the 1970s preferred to challenge the Hubble law rather than accept the existence of highly luminous objects. One reason for their skepticism was the huge gap in energy output between normal galaxies and quasars. The gap was bridged, however, when astronomers realized that the centers of certain galaxies look like low-luminosity quasars.

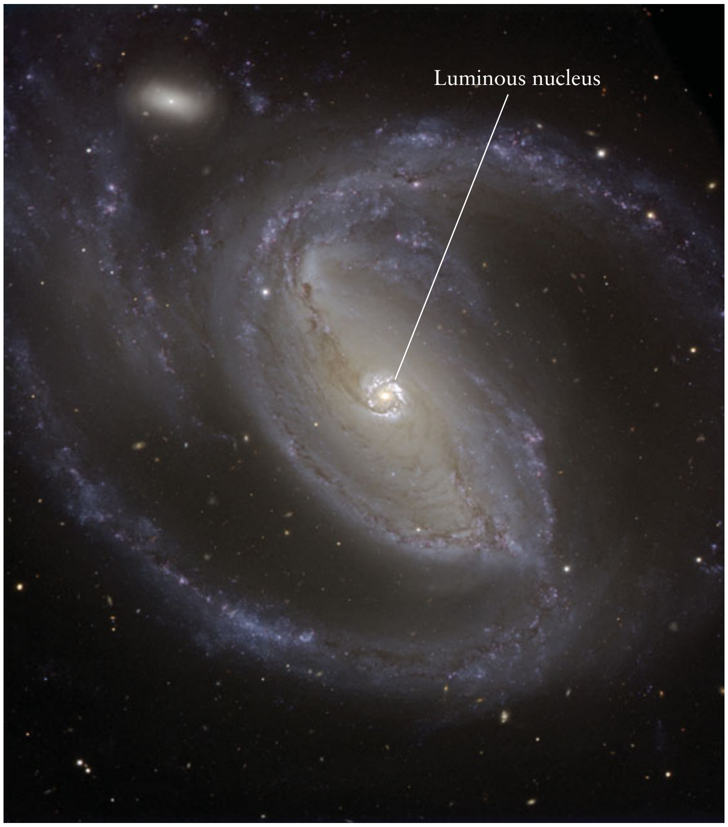
Seyfert Galaxies and Radio Galaxies

The first "missing links" between quasars and ordinary galaxies were actually discovered before quasars themselves. In 1943, Carl Seyfert at the Mount Wilson Observatory made a systematic study of spiral galaxies with bright, compact nuclei that seem to show signs of intense and violent activity (Figure 25-7). Like quasars, the nuclei of these galaxies have strong emission lines in their spectra. These galaxies are now referred to as **Seyfert galaxies.**

Fast-moving jets emanate from quasars, radio galaxies, and Seyfert galaxies

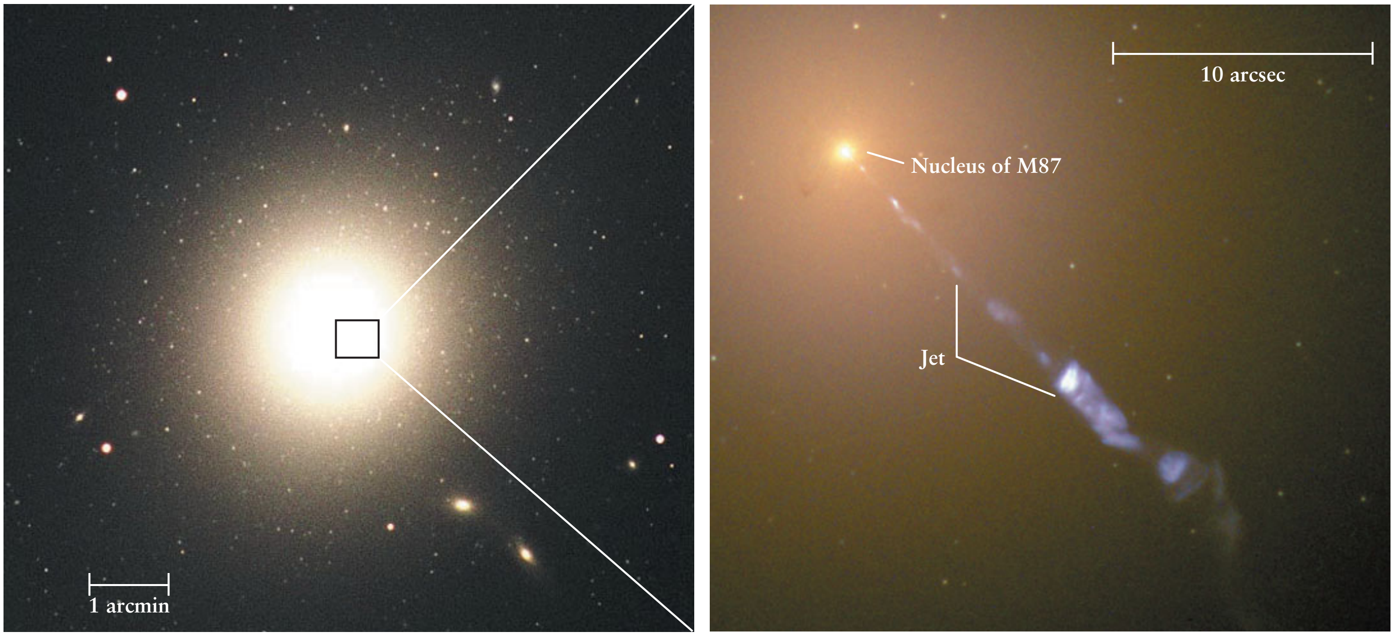
A few percent of the most luminous spiral galaxies are Seyfert galaxies. These galaxies range in luminosity from about 10^{36} to 10^{38} watts, which makes the brightest Seyferts as luminous as faint quasars. Indeed, there is no sharp dividing line between the properties of Seyferts and those of quasars. Like radio-quiet quasars, Seyferts tend to have only weak radio emissions. And like quasars, some Seyferts are members of interacting pairs or exhibit the vestiges of mergers and collisions.

While Seyfert galaxies resemble dim, radio-quiet quasars, certain elliptical galaxies, called **radio galaxies** because of their strong radio emission, are like dim, radio-loud quasars. The energy output of radio galaxies covers approximately the same range as that of Seyferts. The first of these peculiar galaxies was discovered in 1918 by Heber D. Curtis. His short-exposure photograph of the giant elliptical galaxy M87 revealed a bright, starlike nucleus with a protruding jet like that of quasar 3C 273 (see the image that opens this chapter). Figure 25-8*a* shows an overall view of the entire galaxy. The Hubble Space Telescope image in Figure 25-8*b* shows the jet extending outward from the nucleus of M87.



WEB LINK 25.7 **Figure 25-7** R I V U X G
A Seyfert Galaxy This barred spiral galaxy, called NGC 1097, is a Seyfert galaxy that lies some 14 Mpc (45 million ly) from Earth in the southern constellation Fornax (the Furnace). Its nucleus glows much more brightly than that of a normal barred spiral galaxy. The small elliptical galaxy at upper left is a satellite of NGC 1097. (European Southern Observatory)

Most of the light from the central regions of M87 is **thermal radiation,** with a spectrum like that of a blackbody. This is radiation that is caused by the random thermal motion of the atoms and molecules that make up the emitting object. The spectrum of thermal radiation depends only on the object's temperature. For a given temperature, an object emits the maximum amount of thermal radiation at one particular frequency and emits less at frequencies above or below that value. There are also absorption lines in the spectrum of thermal radiation from M87's nucleus, which indicates that this radiation is due to a profusion of stars crowded around the galaxy's center.



(a) The giant elliptical galaxy M87

(b) A shorter exposure reveals M87's jet

WEB LINK 25.8

Figure 25-8 R I V U X G

The Radio Galaxy M87 **(a)** M87 measures about 100 kpc (300,000 ly) across and lies in the Virgo cluster some 15 Mpc (50 million ly) from Earth. The numerous fuzzy dots that surround the galaxy are globular clusters. **(b)** From the galaxy's tiny, bright nucleus—less than 2 pc (6.5 ly) in diameter—a jet extends outward some 1500 pc (5000 ly). (a: David Malin/Anglo-Australian Observatory; b: NASA and the Hubble Heritage Team, STScI/AURA)

By contrast, the light from M87's jet is **nonthermal radiation.** This radiation is *not* due to random thermal motion, and has a very different spectrum than does thermal radiation. In this case of M87's jet, the spectrum extends from radio to X-ray frequencies, which is a much broader range than thermal radiation.

The particular type of nonthermal radiation emitted from the jet is *synchrotron radiation*. As we saw in the discussion of the Crab Nebula in Section 21-4, synchrotron radiation is produced by relativistic electrons traveling in a strong magnetic field. (Recall that "relativistic" means "traveling near the speed of light.") As the electrons spiral around the magnetic field, they emit electromagnetic radiation with a distinctive spectrum very different from that of blackbody radiation (see Figure 21-6). The presence of synchrotron radiation coming from M87's jet indicates that relativistic particles are being ejected from the galaxy's nucleus and encountering a magnetic field.

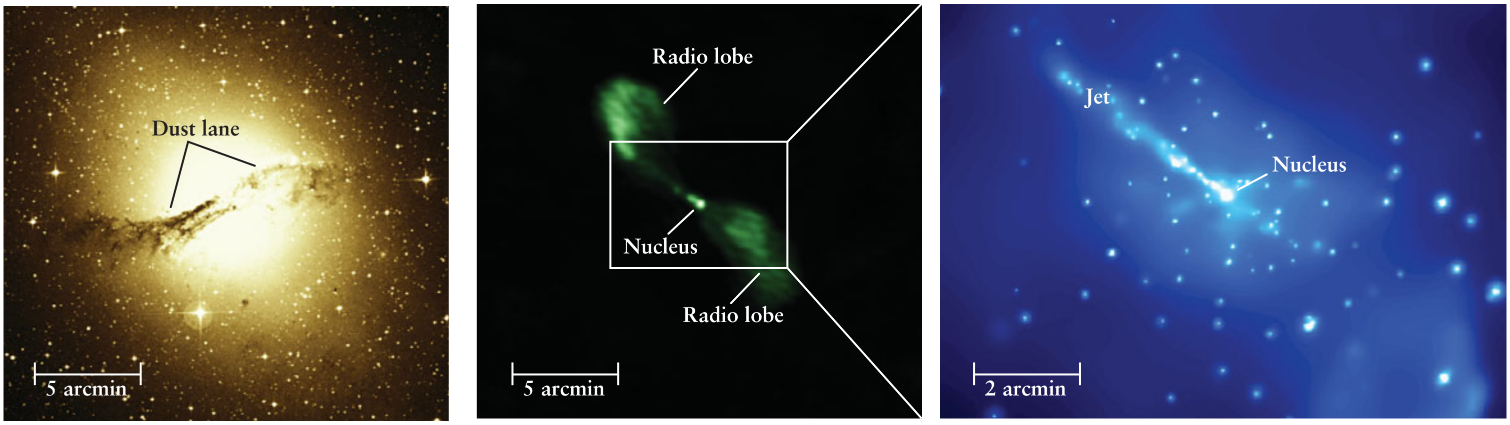
One piece of evidence that the radiation from M87's jet is nonthermal is that some of it is **polarized radiation,** which means that the electric fields of the waves are oriented in a specific direction. By contrast, thermal radiation—like light from the Sun, stars, or lightbulbs—is *unpolarized* radiation. It consists of waves whose electric fields are oriented at random angles. (Most radio stations transmit signals that are polarized in a vertical plane. In order to detect these signals, the radio antennas on cars are oriented vertically rather than horizontally.)

Jets like those found with quasars and radio galaxies are also observed with Seyfert galaxies. One example is NGC 1097, which has jets that extend away from the galaxy's nucleus for more than 50 pc. However, these are too faint to appear in Figure 25-7.

Jets and Radio Lobes

During the 1960s and 1970s, astronomers found that many radio galaxies have a set of two **radio lobes,** one on each side of a parent galaxy. The radio lobes generally span a distance that is 5 to 10 times the size of the parent galaxy, although smaller and much larger examples are known. The parent galaxy—almost always a giant elliptical—sits midway between the radio lobes. Because of their overall shape, radio galaxies (and many similar quasars) are sometimes called **double radio sources.** Cygnus A, shown in Figure 25-1, is a fine example.

The spectrum and polarization of the radio-frequency emission from the lobes of a double radio source bear all the characteristics of synchrotron radiation. This suggests that radio galaxies should have jets of relativistic particles leading from the galaxy out to the radio lobes. A spectacular example of this is Centaurus A, one of the brightest radio sources in the southern sky (Figure 25-9). The peculiar parent galaxy of Centaurus A, shown in Figure 25-9*a,* has a broad dust lane studded with young, hot, massive stars. This galaxy, called NGC 5128, is thought to be an elliptical galaxy that has collided with a spiral galaxy. Radio waves pour from two lobes on opposite sides of the galaxy (Figure 25-9*b*). A second set of radio lobes farther from the galaxy (not shown in Figure 25-9) spans a volume 2 million light-years across. Figure 25-9*c* shows an X-ray–emitting jet emanating from



(a) Centaurus A: light from stars
R I **V** U X G

(b) Centaurus A: radio lobes
R I V U X G

(c) An X-ray-emitting jet emanates from the nucleus
R I V U **X** G

WEB LINK 25.9

Figure 25-9

The Radio Galaxy Centaurus A **(a)** This elliptical galaxy, called NGC 5128, lies about 4 Mpc (13 million ly) from Earth at the location of the radio source Centaurus A. The dust lane is evidence of a collision with another galaxy. **(b)** At radio wavelengths we see synchrotron radiation from two radio lobes centered on the galaxy's nucleus. **(c)** A luminous jet extends from the nucleus directly toward one of the lobes. (a: Digital Sky Survey/U. K. Schmidt Image/STScI; b: NRAO/VLA/J. Condon et al.; c: NASA/SAO/R. Kraft et al.)

the galaxy's nucleus. This jet, which is perpendicular to the galaxy's dust lane, is aimed toward one of the radio lobes.

How does a jet give rise to a radio lobe? To see the answer, note that Centaurus A is a member of a cluster of galaxies. We saw in Section 24-7 that the space between galaxies in such a cluster is filled with a substantial amount of hot intracluster gas whose presence can be detected in X-ray images (see Figure 24-25). As the particles in a jet plow through this gas, they are subjected to fluid resistance that slows them down. The particles lose energy in the process, and this energy is converted to electromagnetic radiation at radio wavelengths.

The idea that jets of particles are found in double radio sources is reinforced by the existence of **head-tail sources.** These are radio galaxies with a concentrated source of radio emission (the "head") to which are attached two long radio-emitting streams (the "tails"). Figure 25-10 shows one such head-tail source. The "head" coincides with the position of the elliptical galaxy NGC 1265, which is a member of the Perseus cluster of galaxies. NGC 1265 is known to be moving through the cluster at 2500 km/s. Figure 25-10 shows that the two "tails" are swept backward, like smoke emerging from a fast-moving steam locomotive. This is just what we would expect if the "tails" are caused by jets emerging from the galaxy; as NGC 1265 moves through the intergalactic gas of the cluster, the jets feel a 2500-km/s wind blowing them back.

Like NGC 1265, many radio galaxies are found near the centers of rich clusters of galaxies. Hence, they are probably subjected to collisions and mergers like the one that NGC 5128 has experienced (see Figure 25-9*a*).

Seyfert galaxies and radio galaxies share many properties with the more remote, more luminous quasars. A key difference

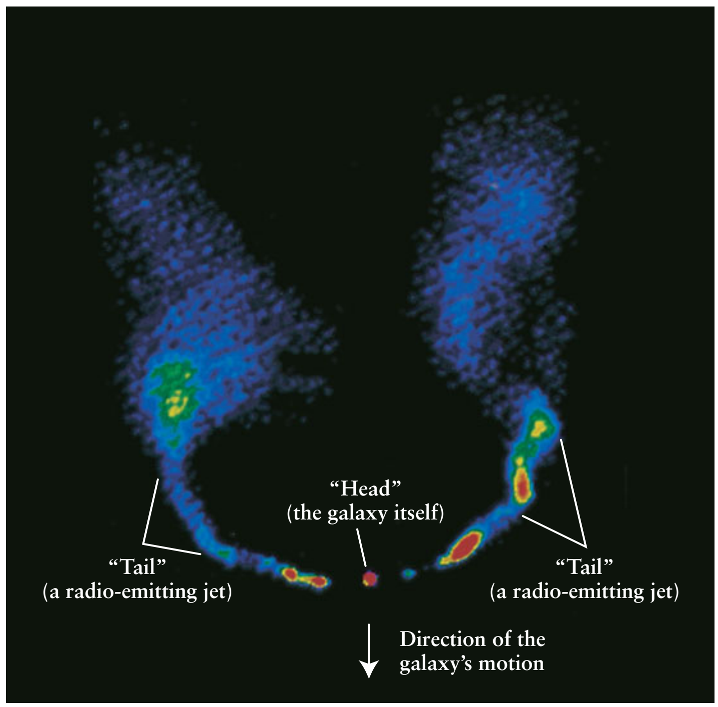
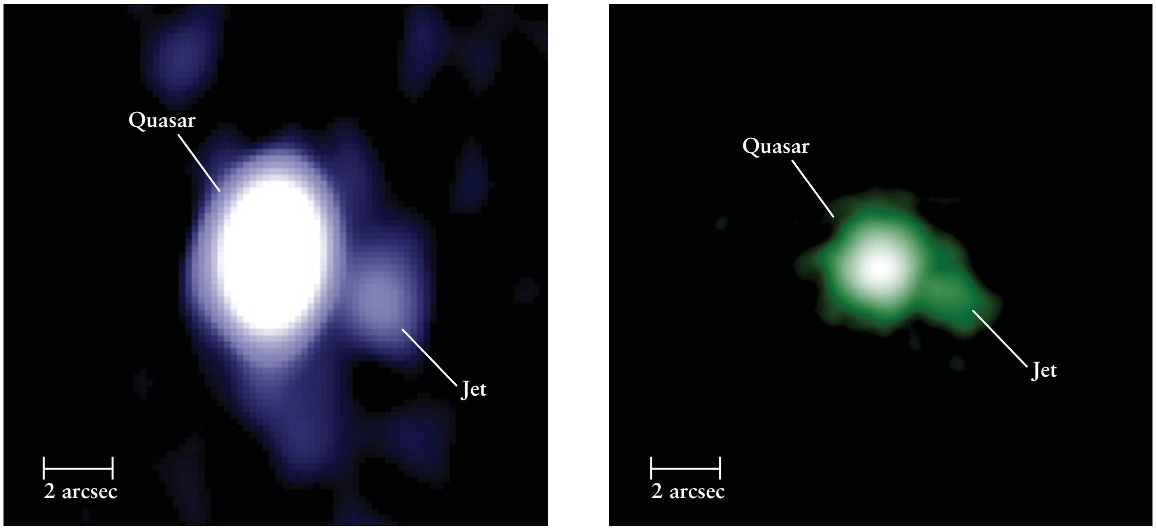


Figure 25-10 **R** I V U X G

The Head-Tail Source NGC 1265 The elliptical galaxy NGC 1265 would probably be an ordinary double radio source, except that it is moving at a high speed through the Perseus cluster of galaxies. Because of this motion, its two jets trail behind the galaxy, giving this radio source a distinctly windswept appearance. (NRAO)



(a) A quasar jet at radio wavelengths... R I V U X G

(b) ...and at X-ray wavelengths R I V U X G

Figure 25-11

A Jet Emanating from a High-Redshift Quasar
Designated GB1508+5714, this quasar has a redshift $z = 4.3$, corresponding to a distance of 3800 Mpc (12.2 billion ly). We see this quasar as it was when the universe was only 10% of its present age. Quasars like this may have evolved into radio galaxies like those shown in Figures 25-8, 25-9, and 25-10. (a: NRAO/VLA/Brandeis/T. Cheung; b: NASA/CXC/SAO/A. Siemiginowska et al.)

is that we see many Seyferts and radio galaxies that are relatively close to us and thus in the recent past, whereas we only see quasars at great distances and hence in the remote past (Figure 25-11). Thus, quasarlike objects have not completely disappeared from the universe, but those that remain in the present-day universe—Seyferts and radio galaxies—are only a pale shadow of the intensely luminous quasars that populated the heavens when the universe was young.

Whatever lies at the centers of quasars, Seyferts, and radio galaxies, it must deliver energy at a prodigious rate to power not only these objects' tremendous luminosities but also their high-speed jets. As we will see in the next section, these energy sources are not only powerful; they are also highly variable.

25-4 Quasars, blazars, Seyferts, and radio galaxies are active galaxies

Inspired by the discovery of quasars, Seyferts, and luminous radio galaxies, astronomers during the 1960s and 1970s searched for clues to these powerful energy sources. One important clue turned out to be a new class of unusual objects called *blazars*. Like quasars, blazars are extraordinarily luminous objects that look like stars but prove to be the nuclei of distant galaxies. The difference is that unlike a quasar, the spectrum of a blazar is almost featureless, with hardly any absorption or emission lines at all!

Blazars and the Strange Case of Superluminal Motion

The first of this new class of objects to be discovered was BL Lacertae, or BL Lac for short, in the constellation Lacerta (the Lizard). When discovered in 1929, it was mistaken for a variable star, largely because its brightness varies by a factor of 15 within only a few months. Careful examination, however, revealed some fuzz (visible in Figure 25-12) around its bright, starlike core. In the early 1970s, Joseph Miller at the University of California's Lick Observatory blocked out the light from the bright center of BL Lac and managed to obtain a spectrum of the fuzz. This spectrum, which contains stellar absorption lines, strongly resembles the spectrum of an elliptical galaxy. In other words, BL Lac is an elliptical galaxy with an intensely luminous center.

Objects such as BL Lac are called **blazars.** They emit polarized light with a featureless, nonthermal spectrum typical of synchrotron radiation, as shown in Figure 21-6*b*. (Careful observations have revealed emission and absorption lines in blazar spectra, but these are swamped by the intense synchrotron radiation.) Detailed radio observations show that blazars are probably double radio sources—but oriented so that we see their jets end-on. For example, high-resolution studies of blazars with the Very Large Array have revealed diffuse radio emission around a bright core in almost all cases. A faint radio halo around a bright radio core is exactly what you would expect if you were looking straight down the jet from a radio galaxy. As seen from this angle, the intense synchrotron radiation from the galaxy's bright nucleus is surrounded by weaker emission from its radio lobes.

The idea that blazars are radio galaxies with their jets aimed toward the Earth was supported by the surprising discovery of movement that appeared to be faster than the speed of light. Such **superluminal motion** is also observed in some quasars, where very-long-baseline interferometry (described in Section 6-6) reveals a lumpy structure that changes with time. For example, Figure 25-13 shows four high-resolution

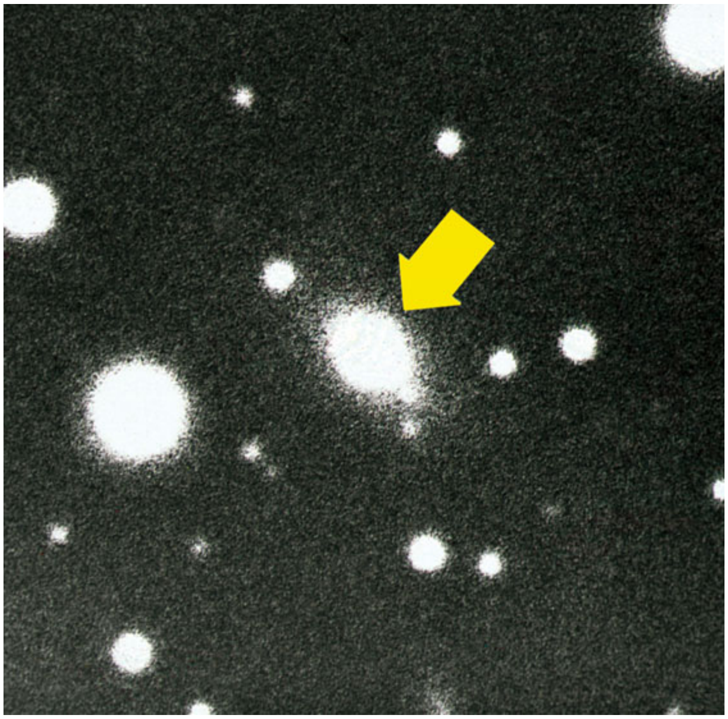


Figure 25-12 R I V U X G

BL Lac This photograph shows fuzz around the blazar BL Lac. The redshift of this fuzz indicates that BL Lac is about 280 Mpc (900 million ly) from Earth. Blazars are giant elliptical galaxies with bright, starlike nuclei that have many quasarlike properties. (Courtesy of T. D. Kinman; Kitt Peak National Observatory)

images of the quasar 3C 273 spanning three years. During this interval, a "blob" moved away from the quasar at a rate of almost 0.001 arcsec per year. Taking into account the distance to the source, this rate of angular separation corresponds to a speed 10 times that of light!

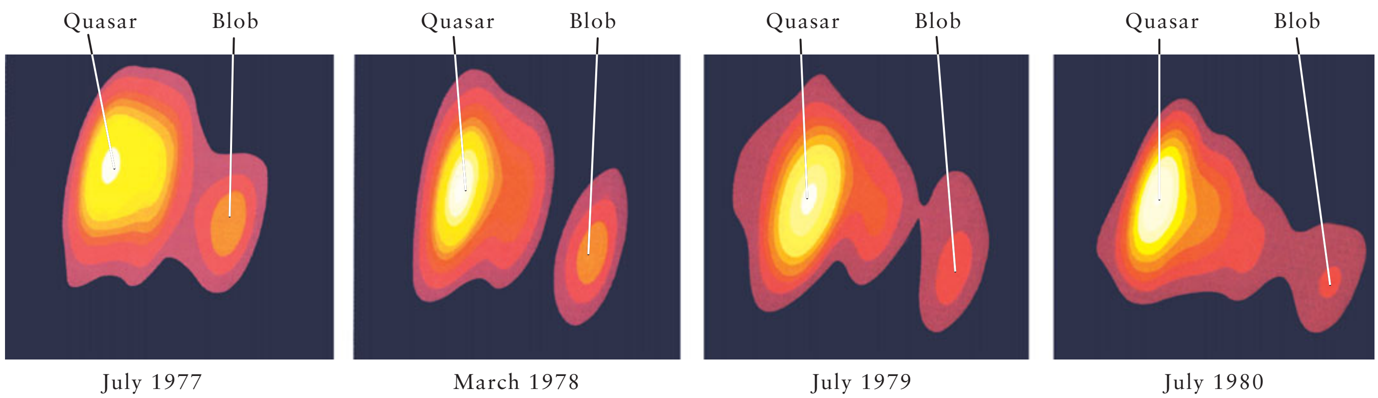
Superluminal motion was puzzling when it was first discovered, because it seemed to violate one of the basic tenets of the special theory of relativity: Nothing can move faster than light. Astronomers soon realized, however, that superluminal motion can be explained as movement slower than light—once we take into account the angle at which we view the radio source. If a relativistic beam of material is aimed close to your line of sight, it can appear to be moving faster than the speed of light. Figure 25-14 shows an example. Quasars generally exhibit lower superluminal speeds (from c to $5c$) than blazars (from $5c$ to $10c$), probably because the relativistic jets from quasars are not aimed as close to our line of sight as are blazar jets.

Active Galaxies and AGNs

Because of the many properties they share, quasars, blazars, Seyfert galaxies, and radio galaxies are now collectively called **active galaxies.** The activity of such a galaxy comes from an energy source at its center. Hence, astronomers say that these galaxies possess **active galactic nuclei,** or **AGNs.**

The rapid variability of active galactic nuclei tells us that they must be very small

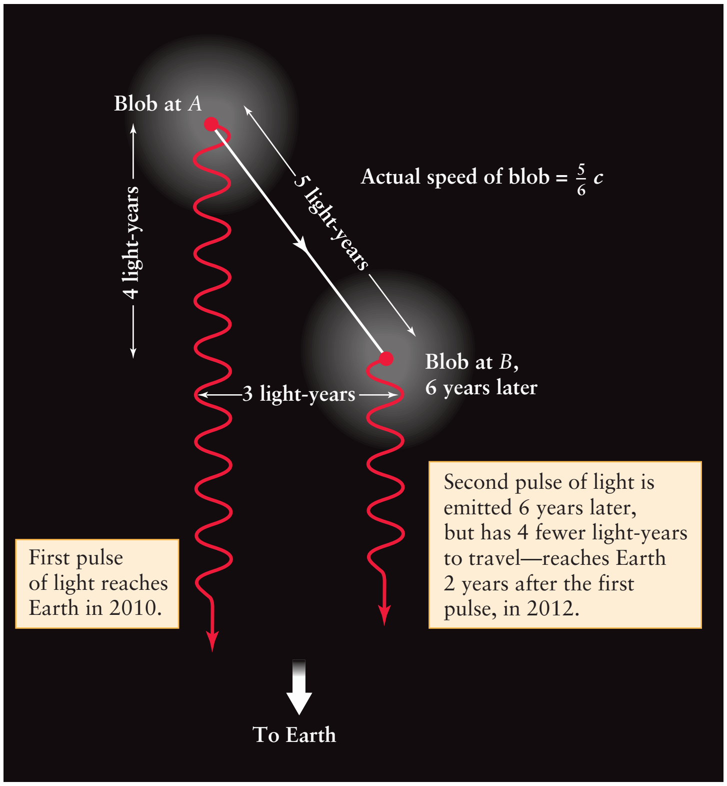
Table 25-2 summarizes the properties of different kinds of active galactic nuclei. As this table shows, one of the features that distinguishes different types of AGNs from each other is the width of the emission lines in their spectra. Figures 25-3 and 25-4 show the spectra of typical quasars in the visible and infrared parts of the spectrum. The "spikes" in these spectra are emission lines from hydrogen gas as well as other elements. The widths of these emission lines indicate that individual light-emitting gas clouds are moving within the quasar at very high speeds (10,000 km/s or so). Some of the clouds are moving toward us, causing the emitted light to have a shorter wavelength and higher frequency, while other clouds are moving away from us and emit light with longer wavelength and lower frequency. This explains why the emission lines of quasars are quite broad. By contrast, the emission lines of radio galaxies do not show the same kind of broadening. If there



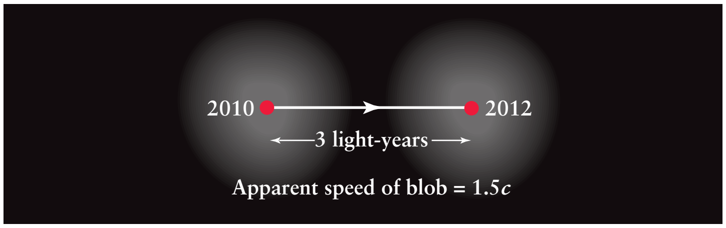
VIDEO 25.1

Figure 25-13 R I V U X G

Superluminal Motion in 3C 273 These four images are false-color, high-resolution radio maps of the quasar 3C 273 (shown in the image that opens this chapter). They show a blob that seems to move away from the quasar at 10 times the speed of light. In fact, a beam of relativistic particles from 3C 273 is aimed almost directly at the Earth, giving the illusion of faster-than-light motion. Each image is about 7 milliarcseconds (0.007 arcsec) across. (NRAO)



(a) View from above



(b) View from Earth

INTERACTIVE EXERCISE 25.1

Figure 25-14

An Explanation of Superluminal Motion **(a)** If a blob of material ejected from a quasar moves at five-sixths of the speed of light, it covers the 5 ly from point *A* to point *B* in 6 years. In the case shown here, it moves 4 ly toward the Earth and 3 ly in a transverse direction. The light emitted by the blob at *A* reaches us in 2020. The light emitted by the blob at *B* reaches us in 2022. The light left the blob at *B* 6 years later than the light from *A* but had 4 fewer light-years to travel to reach us. **(b)** From Earth we can see only the blob's transverse motion across the sky, as in Figure 25-13. It appears that the blob has traveled 3 ly in just 2 years, so its apparent speed is 3/2 of the speed of light, or 1.5*c*.

are fast-moving gas clouds in these galaxies, they are hidden from our sight. (In Section 25-6 we will see why this is.) In a similar way, astronomers distinguish between different subtypes of Seyfert galaxies (called *Seyfert 1* and *Seyfert 2*), depending on whether their dominant emission lines are broad or narrow.

AGN Variability and the Size of the Light Source

One characteristic that is common to *all* types of active galactic nuclei is variability. For example, Figure 25-15 shows brightness fluctuations of the quasar 3C 273 as determined from 29 years of observations. The brightness of 3C 273 increased by 60% from the beginning to the end of 1982, then declined to the starting value in just five months. Other AGNs undergo greater fluctuations in brightness (by a factor of 25 or more) that occur even more rapidly (X-ray observations reveal that some blazars vary in brightness over time intervals as short as 3 hours).

These fluctuations in brightness allow astronomers to place strict limits on the maximum size of a light source. An object cannot vary in brightness faster than light can travel across that object. For example, an object that is 1 light-year in diameter cannot vary significantly in brightness over a period of less than 1 year.

Table 25-2 Properties of Active Galactic Nuclei (AGNs)

Object	Found in which type of galaxy	Strength of radio emission	Type of emission lines in spectrum	Luminosity (watts)	Luminosity (Milky Way Galaxy = 1)
Blazar	Elliptical	Strong	None	10^{38} to 10^{42}	10 to 10^5
Radio-loud quasar	Elliptical	Strong	Broad	10^{38} to 10^{42}	10 to 10^5
Radio galaxy	Elliptical	Strong	Narrow	10^{36} to 10^{38}	0.1 to 10
Radio-quiet quasar	Spiral or elliptical	Weak	Broad	10^{38} to 10^{42}	10 to 10^5
Seyfert 1	Spiral	Weak	Broad	10^{36} to 10^{38}	0.1 to 10
Seyfert 2	Spiral	Weak	Narrow	10^{36} to 10^{38}	0.1 to 10

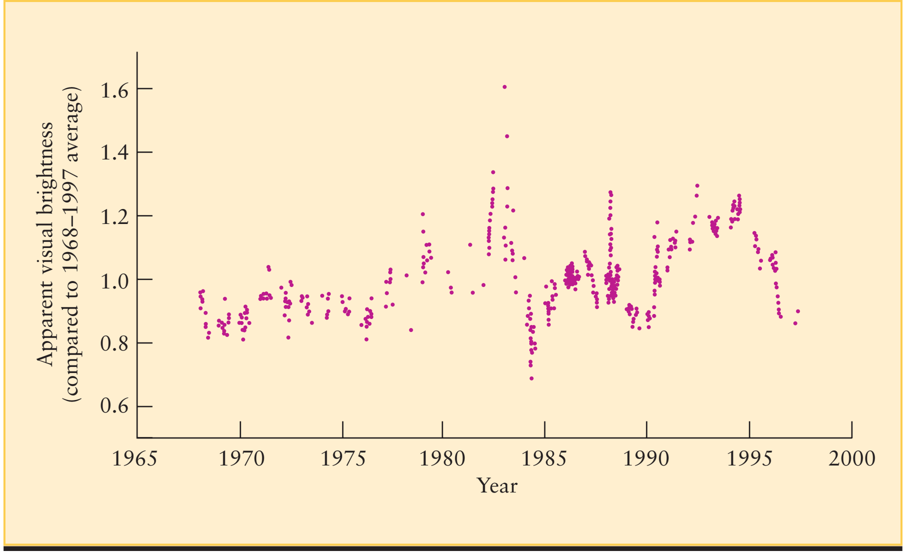


Figure 25-15

Brightness Variations of an AGN This graph shows variations over a 29-year period in the apparent brightness of the quasar 3C 273 (see Figure 25-6*a* and Figure 25-6*b*). Note the large outburst in 1982–1983 and the somewhat smaller ones in 1988 and 1992. (Adapted from M. Türler, S. Paltani, and T. J.-L. Courvoisier)

To understand this limitation, imagine an object that measures 1 light-year across, as in Figure 25-16. Suppose the entire object emits a brief flash of light. Photons from that part of the object nearest the Earth arrive at our telescopes first. Photons from the middle of the object arrive at Earth 6 months later. Finally, light from the far side of the object arrives a year after the first photons. Although the object emitted a sudden flash of light, we observe a gradual increase in brightness that lasts a full year. In other words, the flash is stretched out over an interval equal to the difference in the light travel time between the nearest and farthest observable regions of the object.

The rapid flickering exhibited by active galactic nuclei means that they emit their energy from a small volume, possibly less than 1 light-day across. In other words, a region no larger than our solar system can emit more energy per second than a thousand galaxies! Astrophysicists therefore face the challenge of explaining how so much energy can be produced in such a very small volume.

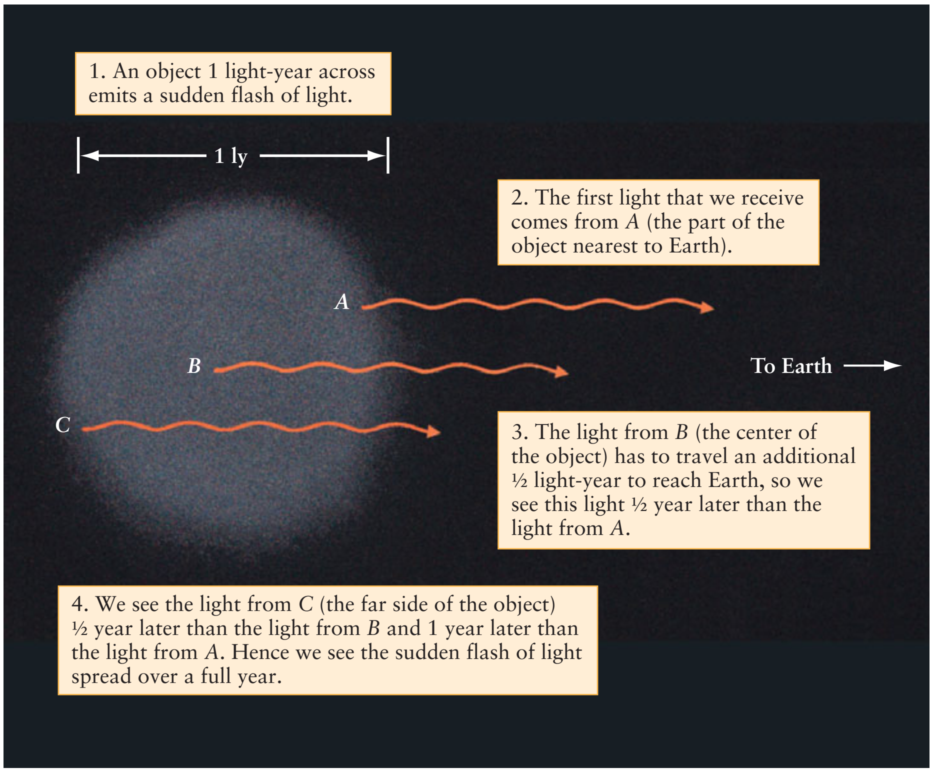


Figure 25-16

A Limit on the Speed of Variations in Brightness The rapidity with which the brightness of an object can vary significantly is limited by the time it takes light to travel across the object. If the object 1 light-year in size emits a sudden flash of light, the flash will be observed from Earth to last a full year. If the object is 2 light-years in size, brightness variations will last at least 2 years as seen from Earth, and so on.

25-5 Supermassive black holes are the "central engines" that power active galactic nuclei

As long ago as 1968, the British astronomer Donald Lynden-Bell pointed out that a black hole lurking at the center of a galaxy could be the "central engine" that powers an active galactic nucleus. Lynden-Bell theorized that as gases fall onto a black hole, their gravitational energy would be converted into radiation. (As we saw in Section 22-3, a similar process produces radiation from black holes in close binary star systems.) To produce as much radiation as is seen from active galactic nuclei, the black hole would have to be very massive. But even a gigantic black hole would occupy a volume much smaller than our solar system—exactly what is needed to explain how active galactic nuclei can vary so rapidly in brightness.

The Eddington Limit and Black Hole Sizes

How large a black hole would be needed to power an active galactic nucleus? You might think that what really matters is not the size of the black hole, but rather the amount of gas that falls onto it and releases energy. However, there is a natural limit to the luminosity that can be radiated by accretion onto a compact object like a black hole. This is called the **Eddington limit,** after the British astrophysicist Sir Arthur Eddington.

If the luminosity exceeds the Eddington limit, there is so much *radiation pressure*—the pressure produced by photons streaming outward from the infalling material—that the surrounding gas is pushed outward rather than falling inward onto the black hole. Without a source of gas to provide energy, the luminosity naturally decreases to below the Eddington limit, at which point gas can again fall inward. This limit allows us to calculate the minimum mass of an active galactic nucleus.

Numerically, the Eddington limit is:

The Eddington limit

$$L_{\mathrm{Edd}} = 30{,}000\left(\frac{M}{\mathrm{M}_\odot}\right)\mathrm{L}_\odot$$

L_{Edd} = maximum luminosity that can be radiated by accretion onto a compact object

M = mass of the compact object

$\mathrm{M}_\odot$ = mass of the Sun

$\mathrm{L}_\odot$ = luminosity of the Sun

The tremendous luminosity of an active galactic nucleus must be less than or equal to its Eddington limit, so this limit must be very high indeed. Hence, the mass of the black hole must also be quite large. For example, consider the quasar 3C 273 (shown in the image that opens this chapter and in Figure 25-6*a*), which has a luminosity of about 3×10^{13} $\mathrm{L}_\odot$. To calculate the minimum mass of a black hole that could continue to attract gas to power the quasar, assume that the quasar's luminosity equals the Eddington limit. Inserting $L_{\mathrm{Edd}} = 3 \times 10^{13}$ $\mathrm{L}_\odot$ into the above equation, we find that $M = 10^9$ $\mathrm{M}_\odot$. Therefore, if a black hole is responsible for the energy output of 3C 273, its mass must be greater than a billion Suns.

Astronomers have indeed found evidence for such **supermassive black holes** at the centers of many nearby normal galaxies (see Section 22-4). As we saw in Section 23-6, at the center of our own Milky Way Galaxy lies what is almost certainly a black hole of about 3.7×10^6 solar masses—supermassive in comparison to a star, but less than 1% the mass of the behemoth black hole at the center of 3C 273.

Most or all large galaxies have supermassive black holes at their centers

Theory suggests that unlike stellar-mass black holes, which require a supernova to produce them (see Figure 20-26), supermassive black holes can be produced without extreme pressures or densities. This may help to explain why they appear to be a feature of so many galaxies.

Measuring Black Hole Masses in Galaxies

One galaxy that probably has a black hole at its center is the Andromeda Galaxy (M31), shown in Figure 24-3. M31 is only 750 kpc (2.5 million ly) from Earth, close enough that details in its core as small as 1 parsec across can be resolved under the best seeing conditions.

In the mid-1980s, astronomers made high-resolution spectroscopic observations of M31's core. By measuring the Doppler shifts of spectral lines at various locations in the core, we can determine the orbital speeds of the stars about the galaxy's nucleus.

Figure 25-17 plots the results for the innermost 80 arcseconds of M31. (At M31's distance, this corresponds to a linear distance of 290 pc or 950 ly.) Note that the rotation curve in the galaxy's nucleus does not follow the trend set in the outer core. Rather, there are sharp peaks—one on the approaching side of the galaxy and the other on the receding side—within 5 arcsec of the galaxy's center.

The most straightforward interpretation is that the peaks are caused by the orbital motions of stars around M31's center. Stars on one side of the galaxy's center are approaching us while stars on the other side are receding from us.

The high speeds of stars orbiting close to M31's center indicate the presence of a massive central object. We can use Newton's form of Kepler's third law (described in Section 4-7 and Box 4-4) and our knowledge of these stars' orbital speeds to calculate the mass of this object. (As described in Section 23-6, much the same method is used to calculate the mass of the supermassive black hole at the center of our own Milky Way Galaxy. The difference is that we can track individual stars at the center of our Galaxy, while the data in Figure 25-17 comes from the combined light of many stars in M31.) Such calculations show that there must be about 3×10^7 solar masses within 5 pc (16 ly) of the center of M31. That much matter confined to such a small volume strongly suggests the presence of a supermassive black hole. Observations of M31 with the Chandra X-ray Observatory are consistent with this picture.

By applying high-resolution spectroscopy to the cores of other nearby galaxies, astronomers have discovered a number of

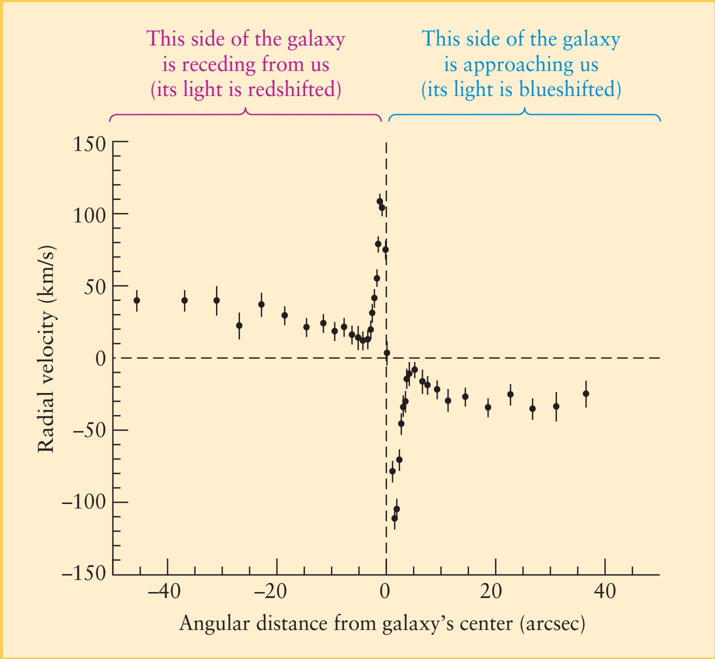


Figure 25-17

The Rotation Curve of the Core of M31 This graph plots radial velocity of matter in the core of M31 versus the angular distance from the galaxy's center. Note the sharp peaks, one blueshifted and one redshifted, within 5 arcsec of the galaxy's center. This indicates the presence of a compact, very massive object at the center of the galaxy. At the distance of M31, 1 arcsec corresponds to 3.6 pc (12 ly). (Adapted from J. Kormendy)

supermassive black holes like the one in M31. Unfortunately, this technique for identifying black holes is difficult to apply to quasars, which are very distant and have small angular sizes. The evidence for supermassive black holes in quasars is therefore circumstantial, yet compelling: No other known energy source could provide enough power to sustain a quasar's intense light output.

If most galaxies have supermassive black holes at their centers, and supermassive black holes are the "central engines" of active galactic nuclei, why aren't all galaxies active? Why are there no nearby quasars? Why do radio galaxies and some quasars have jets? As we will see in the next section, the answers to all these questions may be provided by a model of what happens in an active galactic nucleus.

25-6 Quasars, blazars, and radio galaxies are probably the same kind of object seen from different angles

Accretion onto a supermassive black hole is the most likely explanation of the immense energy output of active galactic nuclei. The challenge to astrophysicists is to understand how that accretion takes place. A successful model of the accretion process must also explain other properties of active galactic nuclei, including their unusual spectra, variable light output, and energetic jets.

Accretion Disks around Supermassive Black Holes

In the leading model of this process, at the heart of an active galaxy is a supermassive black hole surrounded by an **accretion disk**, an immense disk of matter captured by the hole's gravity and spiraling into it. We saw in Section 18-5 that accretion disks are found around stars in the process of formation. The accretion disks that astrophysicists envision around supermassive black holes are similar but far larger and far more dynamic. The *Cosmic Connections* figure depicts the physics of such accretion disks.

The power source for active galaxies is gravitational energy released by material falling toward a central black hole

Imagine a billion-solar-mass black hole sitting at the center of a galaxy, surrounded by a rotating accretion disk. According to Kepler's third law, the inner regions of this accretion disk would orbit the hole more rapidly than would the outer parts. Thus, the rapidly spinning inner regions would constantly rub against the slower moving gases in the outer regions. This friction, aided by magnetic forces within the disk, would cause the gases to lose energy and spiral inward toward the black hole.

As the gases move inward within the accretion disk, they are compressed and heated to very high temperatures. This causes the accretion disk to glow, thus producing the brilliant luminosity of an active galactic nucleus. The temperature of material in the accretion disk reaches 100,000 K or more, which helps explain why many active galactic nuclei emit far more X-ray and gamma-ray radiation than do ordinary galaxies.

In this model, the fundamental source of the energy output by an active galactic nucleus is *gravitational* energy released by infalling material in the accretion disk. We saw in Section 20-6 that gravitational energy is also what powers the immense light output of a core-collapse supernova. The difference between such a supernova and an active galactic nucleus is that the infall of a supernova's outer layers is a one-time event, while gas in an AGN's accretion disk falls inward continuously. Any variations in the density of this infalling gas will cause the disk's luminosity to fluctuate, giving rise to the brightness variations observed by astronomers (see Figure 25-15).

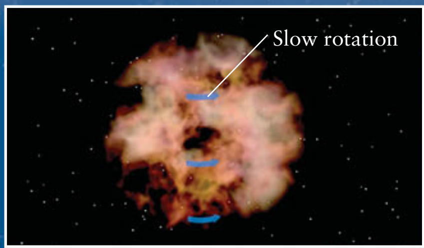
Supercomputer simulations can help us visualize the flow of matter in such an accretion disk. These simulations combine general relativity with equations describing gas flow. At first, matter accelerates to supersonic speeds as it spirals inward toward a black hole. But because the matter is rotating around the hole as well as moving toward it, this inward rush stops abruptly near the hole. The reason is the law of conservation of angular momentum (described in Section 8-4 and Section 21-3). Thanks to this law, rotating objects of all kinds tend to expand (like a spinning lump of dough that expands to make a pizza) rather than contract. The inward motion of gases stops where this tendency to expand outward balances the pull of the black hole's gravity. As a result, not all of the infalling gas reaches the black hole. Instead, part of the gas can become concentrated in high-speed orbits quite close to the hole. The sudden halt in the supersonic inflow creates a shock wave, which defines the inner edge of the accretion disk (see the *Cosmic Connections* figure). Only a

COSMIC CONNECTIONS

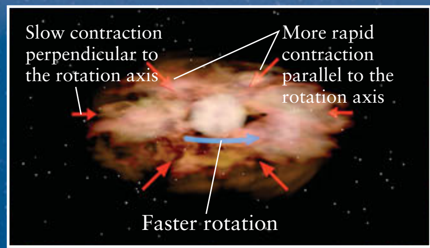
Accretion Disks

Quasars and active galaxies are thought to be powered by gravitational energy released when gas falls toward a central supermassive black hole. Due to conservation of angular momentum, the gas does not fall straight into the black hole. Instead, it must negotiate its way through an accretion disk.

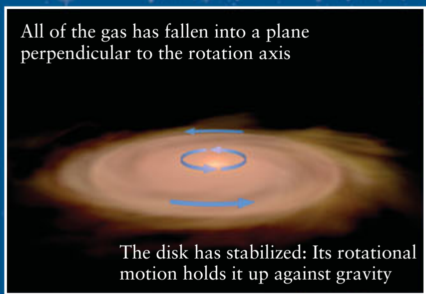
How an accretion disk forms around a black hole



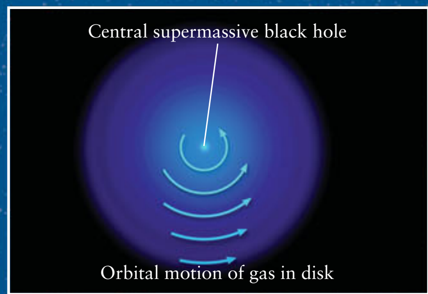
1. A slowly rotating cloud of gas coalesces around the black hole.



2. Due to conservation of angular momentum, the gas moves inward at different rates in different directions.

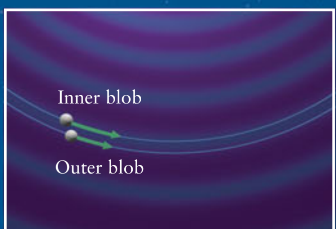


3. The final configuration of the gas is a flattened disk.

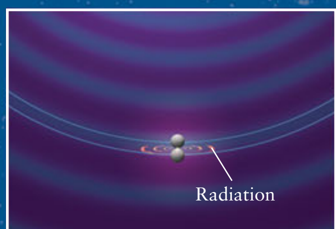


4. Like planets orbiting around the Sun, the closer a blob of gas is to the center of the disk, the faster its orbital speed.

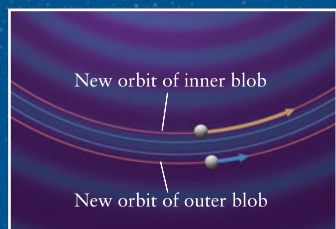
Why gas in an accretion disk spirals in and radiates



5. Consider two adjacent blobs of gas in the accretion disk. The inner blob moves a bit faster than the outer blob (see 4).

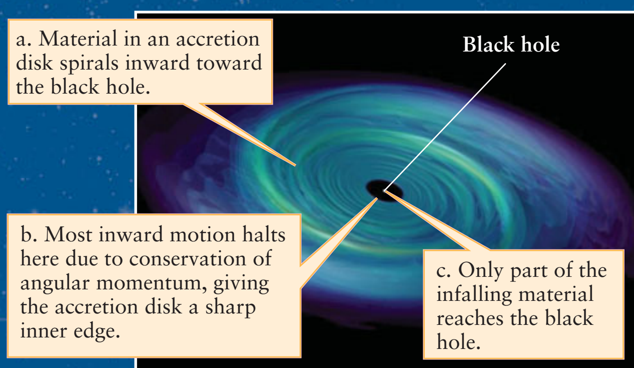


6. When the two blobs collide, energy is transferred from the inner blob to the outer blob. The heated gas emits radiation.



7. Having lost energy, the inner blob moves inward toward the center of the disk and speeds up. The outer blob moves outward and slows down.

8. Conservation of angular momentum slows the inward motion of the blobs of gas.



Upper images: Alfred T. Kamajian, from Omer Blaes, "A Universe of Disks," *Scientific American*, October 2004. Lower right image: Courtesy of Michael Owen and John Blondin, North Carolina State University.

fraction of the material in the accretion disk can cross this inner edge and fall into the black hole.

Because of the constant inward crowding of hot gases, pressures climb rapidly in the inner accretion disk. These pressures relieve the congestion by expelling matter at extremely high speeds. This ejected material escapes moving at right angles to the plane of the accretion disk.

ANIMATION 25.1

Magnetic forces play a crucial role in steering these fast-moving particles. These forces arise because the hot gases in the accretion disk are ionized, forming a plasma, and the motions of this plasma generate a magnetic field (see Section 16-9). As the plasma in the disk rotates around the black hole, it pulls the magnetic field along with it. But because the disk rotates faster in its inner regions than at its outer rim, the magnetic field becomes severely twisted. This twisted field forms two helix shapes, one on either side of the plane of the disk. Relativistic particles flowing outward from the accretion disk tend to follow these magnetic field lines. The result is that the outflowing beams of particles are focused into two jets oriented perpendicular to the plane of the accretion disk, as Figure 25-18 shows. Figure 25-19 is an artist's conception of the accretion disk and jets as they might appear from a nearby planet.

Observations of active galaxies provide evidence in support of this model. One example is the radio galaxy NGC 4261, shown in Figure 22-16, whose radio lobes appear to emanate from an accretion disk around a supermassive black hole. While the accretion disk itself is too small to see, we can observe a dusty ring, or torus, about 250 pc (800 ly) in diameter orbiting the central black hole (Figure 25-20). The motions of this torus reveal the mass of the central black hole to be $1.2 \times 10^9\ M_\odot$.

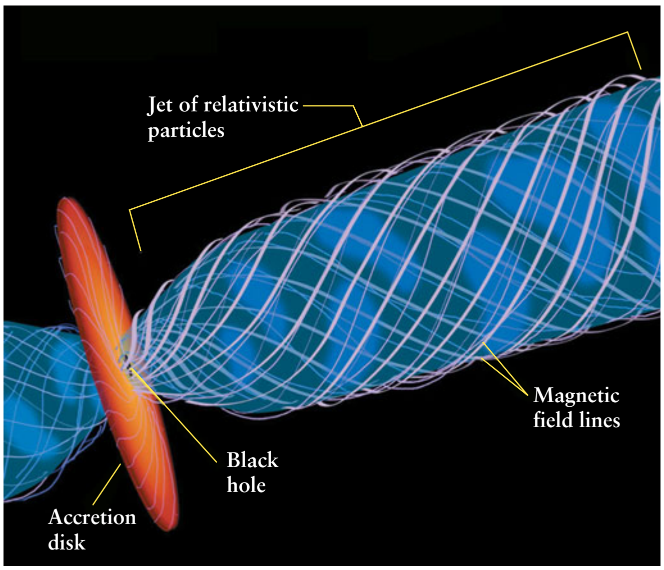


Figure 25-18

Jets from a Supermassive Black Hole The rotation of the accretion disk surrounding a supermassive black hole twists the disk's magnetic field lines into a helix. The field then channels the flow of subatomic particles pouring out of the disk. Over a distance of about a light-year, this channeling changes a broad flow into a pair of tightly focused jets, one on each side of the disk. Figure 18-16 shows a similar process that takes place on a much smaller scale in protostars. (NASA and Ann Feild, Space Telescope Science Institute)

ANIMATION 25.2

Figure 25-19

At the Core of an Active Galaxy This artist's rendering shows a supermassive black hole surrounded by an accretion disk. In the inner regions of the accretion disk, matter crowding toward the hole is diverted outward along two oppositely directed jets. As the relativistic particles move in spiral paths within the jets, they emit synchrotron radiation. (*Astronomy*)

The Unified Model of AGNs

If there were no dusty torus around an accreting supermassive black hole, an observer could view such a black hole from any angle and see the intense radiation from the accretion disk. But the presence of such a torus seems to be a natural result of the accretion process. As a result, from certain angles the torus blocks the view of the innermost part of the active galactic nucleus. This idea offers a single explanation for several types of active galaxies. Many astronomers suspect that the main difference between blazars, quasars, and radio galaxies is only the angle at which the black hole "central engine" is viewed. This idea is called the *unified model* of active galactic nuclei.

Figure 25-21 illustrates the unified model for a luminous active galactic nucleus with jets. If an observer looks straight down the axis of the jet, the observed radiation is dominated by synchrotron radiation from the jet. This has a continuous spectrum with no emission or absorption lines (see Figure 21-6*b*). Hence, the observer sees a blazar, with a nearly featureless spectrum.

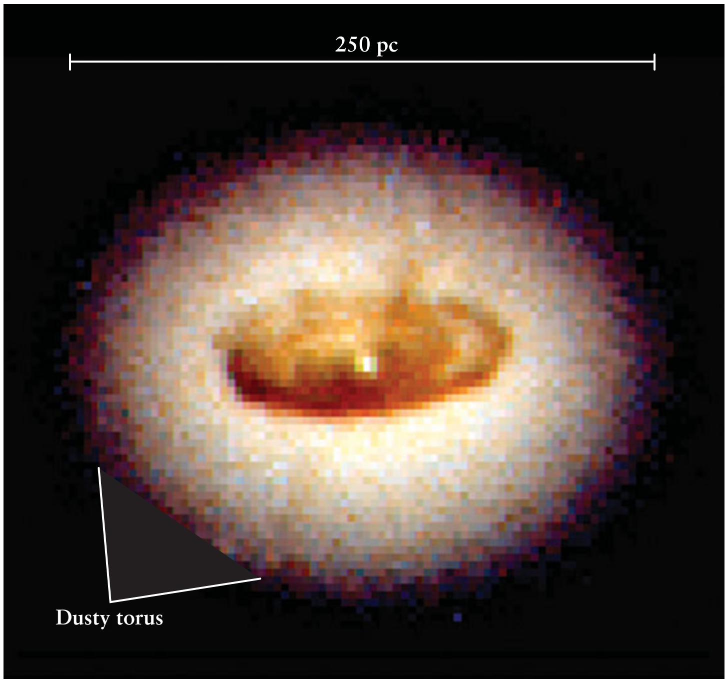


Figure 25-20 R I V U X G

A Dusty Torus Around a Supermassive Black Hole This immense doughnut of dust and gas orbits the black hole at the center of the radio galaxy NGC 4261. The radio lobes of this galaxy appear to be the endpoints of jets emerging parallel to the plane of this torus (see Figure 22-16). (L. Ferrarese/Johns Hopkins University, NASA)

At a more oblique angle, the observer gets a clear view of the luminous accretion disk and the turbulent region around the black hole. Because gases move at many different velocities in this region, the observer sees spectral lines that have been broadened by the Doppler effect. The observer also sees intense thermal radiation from the accretion disk and synchrotron radiation from the jets. From this angle, what the observer sees is a radio-loud quasar.

If the observer looks nearly edge-on at the torus, the accretion disk will be completely hidden. Some of the light reaching the observer comes from hot gas flowing out of the accretion disk, and this light has an emission-line spectrum. But this gas is not moving rapidly either toward or away from the observer, so there is little Doppler shift and the emission lines are narrow. The synchrotron emission from the jets is still visible, and so our observer reports seeing a radio galaxy.

It may happen that no jets are present, which means that the active galactic nucleus will lack the powerful synchrotron radiation that particles in a jet produce. In this case, an observer viewing the "central engine" either face-on or at an oblique angle will see a radio-quiet quasar.

Unlike our imaginary observer, we cannot move the vast distances through space that would be needed to view a given active galaxy from different angles. Instead, we may see a given active galaxy as either a blazar, a quasar, or a radio galaxy, depending on how that galaxy's accretion disk and torus are oriented to our line of sight.

CAUTION! Note that we are really making use of *two* different but complementary models here. The unified model says that different types of active galactic nuclei are really different views of the same type of "central engine," while the accretion-disk model explains how the "central engine" works.

The accretion-disk idea helps to explain why there are no nearby quasars (see Figure 25-5). Over time, most of the available gas and dust surrounding a quasar's "central engine" is accreted onto the black hole. The "central engine" has less and less infalling matter to act as "fuel," and the quasar becomes much less active. The result is a relatively less luminous radio galaxy or

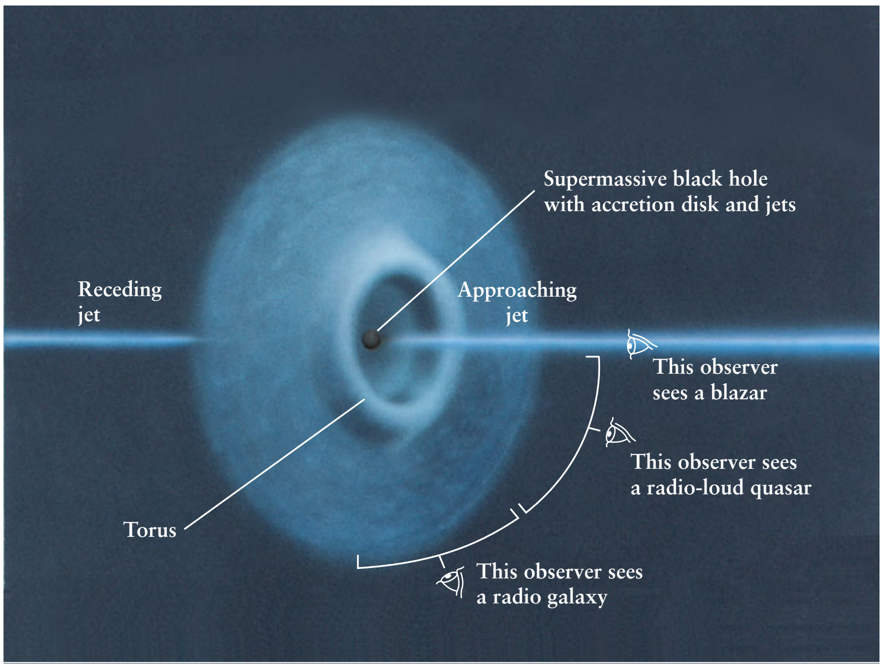


Figure 25-21

A Unified Model of Active Galaxies Blazars, radio-loud quasars, and radio galaxies may be the same type of object—a supermassive black hole, its accretion disk, and its relativistic jets—viewed at different angles. (Compare Figure 22-12, which shows the similar environment of a stellar-mass black hole.)

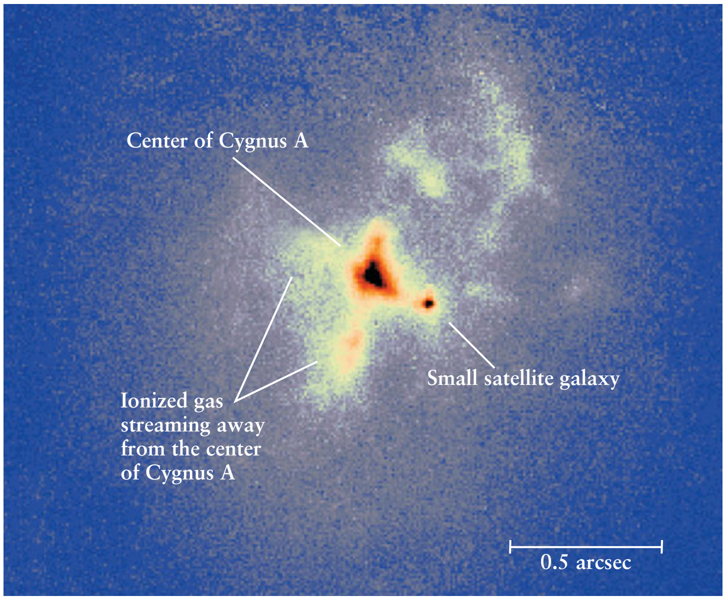


Figure 25-22 R I V U X G

Providing Fresh "Fuel" for a Supermassive Black Hole A small satellite galaxy has fallen into the central regions of the radio galaxy Cygnus A (see Figure 25-1). Its material may eventually accrete onto the supermassive black hole at Cygnus A's very center. This false-color extreme close-up was made using adaptive optics, described in Section 6-3. (Courtesy G. Canalizo, C. Max, D. Whysong, R. Antonucci, and S. E. Dahm)

Seyfert galaxy. (A Seyfert galaxy is essentially a low-luminosity, radio-quiet quasar. Whether it is a Seyfert 1 with predominantly broad emission lines or a Seyfert 2 with predominantly narrow lines may depend on whether we view its "central engine" more nearly face-on or edge-on.)

Finally, the accretion-disk idea explains what happens when active galaxies collide or merge with other galaxies. Such collisions and mergers can transfer gas and dust from one galaxy to another, providing more "fuel" for a supermassive black hole. For example, the image in Figure 25-22 shows a close-up view of a supermassive black hole about to take in more "fuel" from a merger. Thus, it is not surprising that many of the most luminous active galaxies are also those that have recently undergone collisions, such as the radio galaxy Centaurus A shown in Figure 25-9.

Eventually an active galaxy will run out of "fuel" altogether, and its central supermassive black hole will become inactive. Most galaxies, including our own Milky Way, have supermassive black holes at their centers that are not presently active. Perhaps many of these galaxies once led much more dramatic lives as quasars or radio galaxies.

Key Words

accretion disk, p. 683
active galactic nucleus (AGN), p. 679
active galaxy, p. 679
blazar, p. 678
double radio source, p. 676
Eddington limit, p. 682
head-tail source, p. 677
nonthermal radiation, p. 676
polarized radiation, p. 676
quasar, p. 671
radio galaxy, p. 675
radio lobes, p. 676
Seyfert galaxy, p. 675
superluminal motion, p. 678
supermassive black hole, p. 682
thermal radiation, p. 675

Key Ideas

Quasars: A quasar looks like a star but has a huge redshift. These redshifts show that quasars are several hundred megaparsecs or more from the Earth, according to the Hubble law.

• To be seen at such large distances, quasars must be very luminous, typically about 1000 times brighter than an ordinary galaxy.

• About 10% of all quasars are strong sources of radio emission and are therefore called "radio-loud"; the remaining 90% are "radio-quiet."

• Some of the energy emitted by quasars is synchrotron radiation produced by high-speed particles traveling in a strong magnetic field.

Seyfert Galaxies: Seyfert galaxies are spiral galaxies with bright nuclei that are strong sources of radiation. Seyfert galaxies seem to be nearby, low-luminosity, radio-quiet quasars.

Radio Galaxies: Radio galaxies are elliptical galaxies located midway between the lobes of a double radio source.

• Relativistic particles are ejected from the nucleus of a radio galaxy along two oppositely directed beams.

Blazars: Blazars are bright, starlike objects that can vary rapidly in their luminosity. They are probably radio galaxies or quasars seen end-on, with a jet of relativistic particles aimed toward the Earth.

Active Galaxies: Quasars, blazars, and Seyfert and radio galaxies are examples of active galaxies. The energy source at the center of an active galaxy is called an active galactic nucleus.

• Rapid fluctuations in the brightness of active galaxies indicate that the region that emits radiation is quite small.

Black Holes and Active Galactic Nuclei: The preponderance of evidence suggests that an active galactic nucleus consists of a supermassive black hole onto which matter accretes.

• As gases spiral in toward the supermassive black hole, some of the gas may be redirected to become two jets of high-speed particles that are aligned perpendicularly to the accretion disk.

• An observer sees a radio galaxy when the accretion disk is viewed nearly edge-on, so that its light is blocked by a surrounding torus. At a steeper angle, the observer sees a quasar. If one of the jets is aimed almost directly at the Earth, a blazar is observed.

Questions

Review Questions

1. When quasi-stellar radio sources were first discovered and named, why were they called "quasi-stellar"?
2. How were quasars first discovered? How was it discovered that they are very distant objects?
3. Explain why astronomers cannot use any of the standard candles described in Section 24-4 to determine the distances to quasars.

4. Suppose you saw an object in the sky that you suspected might be a quasar. What sort of observations might you perform to test your suspicion?
5. Quasar PC 1247+3406 is presently about 25.9 billion light-years from Earth. Explain how it is possible for astronomers to see this quasar, even though light travels at a speed of 1 light-year per year.
6. How does the spectrum of a quasar differ from that of an ordinary galaxy? How do spectral lines help astronomers determine the distances to quasars?
7. If quasars lie at the centers of galaxies, why don't we see strong absorption lines from the galaxy's stars when we look at the spectrum of a quasar (like those shown in Figures 25-3 and 25-4)?
8. It was suggested in the 1960s that quasars might be compact objects ejected at high speeds from the centers of nearby ordinary galaxies. Explain why the absence of blueshifted quasars disproves this hypothesis.
9. Why were some astronomers skeptical that the redshifts of quasars gave a true indication of their distance?
10. How do astronomers know that quasars are located in galaxies? In what sorts of galaxies are they found?
11. What is a Seyfert galaxy? Why do astronomers think that Seyfert galaxies may be related to radio-quiet quasars?
12. What is a radio galaxy? What is a double radio source? Why do astronomers think these objects may be related to radio-loud quasars?
13. How would you distinguish between thermal and nonthermal radiation?
14. What is the difference between polarized and unpolarized radiation? How does the polarization of radiation from M87's jet show that the radiation is nonthermal radiation?
15. What are head-tail sources? How do they provide evidence that double radio sources include jets of fast-moving particles?
16. What is a blazar? What is unique about its spectrum? How is it related to other active galaxies?
17. Some blazars or quasars appear to be ejecting material at speeds faster than light. Is the material really moving that fast? If so, how is this possible? If not, why does the material appear to be traveling so fast?
18. What do astronomers learn from the widths of the spectral lines of quasars?
19. What do the brightness fluctuations of a particular active galaxy tell us about the size of the energy-emitting region within that galaxy?
20. How could a supermassive black hole, from which nothing—not even light—can escape, be responsible for the extraordinary luminosity of a quasar?
21. What is the Eddington limit? Explain how it can be used to set a limit on the mass of a supermassive black hole, and explain why this limit represents a minimum mass for the black hole.
22. Explain how the rotation curve of a galactic nucleus can help determine whether a supermassive black hole is present.
23. How does matter falling inward toward a central black hole find itself being ejected outward in a high-speed jet?
24. Explain how the unified model of active galaxies suggests that quasars, blazars, and radio galaxies are the same kind of object viewed from different angles.
25. Why do you suppose there are no quasars relatively near our Galaxy?

Advanced Questions

Questions preceded by an asterisk (*) involve topics discussed in the Boxes in Chapters 4, 22, 24.

> **Problem-solving tips and tools**
>
> Table 25-1 gives the connection between redshift and distance, Box 24-2 discusses relativistic redshift, and Box 22-1 describes time dilation in the special theory of relativity. You may find it useful to recall from Section 1-7 that 1 light-year = 63,240 AU. A speed of 1 km/s is the same as 0.211 AU/yr. You can find Newton's form of Kepler's third law in Box 4-4 and the formula for the Schwarzschild radius in Box 22-2.

26. When we observe a quasar with redshift $z = 0.75$, how far into its past are we looking? If we could see that quasar as it really is right now (that is, if the light from the quasar could somehow reach us instantaneously), would it still look like a quasar? Explain why or why not.

*27. The quasar SDSS 1044–0125 has a redshift $z = 5.80$. At what speed does this quasar seem to be receding from us? Give your answer in km/s and as a fraction of the speed of light c.

*28. Suppose that an astronomer discovers a quasar with a redshift of 8.0. With what speed would this quasar seem to be receding from us? Give your answer in km/s and as a fraction of the speed of light.

29. In the quasar spectrum shown in Figure 25-4, there are many deep absorption lines to the left of the L_α emission line (that is, at shorter wavelengths). These lines, collectively called the *Lyman-alpha forest,* are due to remote gas clouds along our line of sight to the quasar. Hydrogen atoms in these clouds absorb L_α photons from the quasar. Explain why the L_α absorption lines due to these clouds are at shorter wavelengths than the L_α emission line from the quasar itself.
30. Explain how the existence of gravitational lenses involving quasars (review Section 24-8) constitutes evidence that quasars are located at the great distances inferred from the Hubble law.
31. The Seyfert galaxy NGC 1275 is actually two galaxies that are colliding. Images of NGC 1275 show a number of globular clusters with a distinctive blue color. Explain how this color shows that these clusters formed relatively recently, perhaps as a result of the collision.
32. In the image that opens this chapter, the close-up view of the Cygnus A jet shows that different wavelengths are preferentially emitted at different locations along the jet's length. Explain why, using the following principle: As an individual particle moves in a magnetic field, the greater its speed, the shorter the wavelength of the synchrotron radiation that it emits.
33. Suppose the distance from point *A* to point *B* in Figure 25-14*a* is 26 light-years and the blob moves at 13/15 of the speed of light. As the blob moves from *A* to *B*, it moves

24 light-years toward the Earth and 10 light-years in a transverse direction. (**a**) How long does it take the blob to travel from *A* to *B*? (**b**) If the light from the blob at *A* reaches Earth in 2020, in what year does the light from *B* reach Earth? (**c**) As seen from Earth, at what speed does the blob appear to move across the sky?

*34. Suppose a blazar at $z = 1.00$ goes through a fluctuation in brightness that lasts one week (168 hours) as seen from Earth. (**a**) At what speed does the blazar seem to be moving away from us? (**b**) Using the idea of time dilation, determine how long this fluctuation lasted as measured by an astronomer within the blazar's host galaxy. (**c**) What is the maximum size (in AU) of the region from which this blazar emits energy?

35. (**a**) Calculate the maximum luminosity that could be generated by accretion onto a black hole of 3.7×10^6 solar masses. (This is the size of the black hole found at the center of the Milky Way, as described in Section 25-6.) Compare this to the total luminosity of the Milky Way, about $2.5 \times 10^{10}\ L_\odot$. (**b**) Speculate on what we might see if the center of our Galaxy became an active galactic nucleus with the luminosity you calculated in (a).

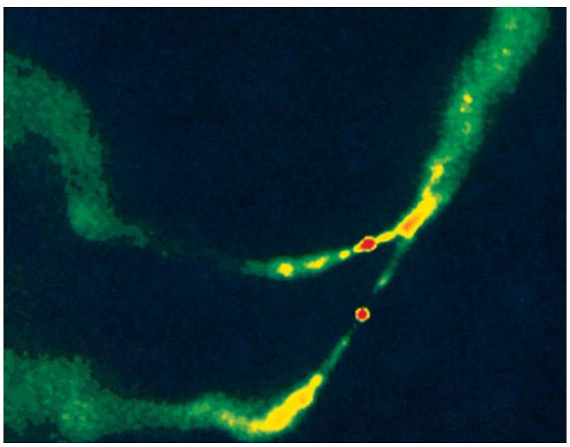
*36. Observations of a certain galaxy show that stars at a distance of 16 pc from the center of the galaxy orbit the center at a speed of 200 km/s. Use Newton's form of Kepler's third law to determine the mass of the central black hole.

*37. Calculate the Schwarzschild radius of a 10^9-solar-mass black hole. How does your answer compare with the size of our solar system (given by the diameter of Pluto's orbit)?

38. Figure 25-9 shows the double radio source Centaurus A. Is it possible that somewhere in the universe there is an alien astronomer who observes this same object as a blazar? Explain your answer with a drawing showing the relative positions of the Earth, the alien astronomer, and Centaurus A.

Discussion Questions

39. The accompanying image from the Very Large Array (VLA) shows the radio galaxy 3C 75 in the constellation Cetus. This galaxy has several radio-emitting jets. High-resolution optical photographs reveal that the galaxy has two nuclei, which are the two red spots near the center of the VLA image. Propose a scenario that might explain the appearance of 3C 75.



R I V U X G

(NRAO)

40. Some quasars show several sets of absorption lines whose redshifts are less than the redshifts of their emission lines. For example, the quasar PKS 0237-23 has five sets of absorption lines with redshifts in the range from 1.364 to 2.202, whereas the quasar's emission lines have a redshift of 2.223. Propose an explanation for these sets of absorption lines.

41. The Milky Way Galaxy is in the process of absorbing the satellite galaxy called the Canis Major Dwarf (see Section 24-6). Discuss whether this process could cause the Milky Way to someday become an active galaxy.

42. Figure 25-22 shows ionized gas streaming away from the "central engine" of the radio galaxy Cygnus A. Instead of spreading outward equally in all directions, the gas appears to be funneled into two oppositely directed cones. Discuss how this could be caused by a dusty torus surrounding a supermassive black hole at the center of Cygnus A.

Web/eBook Questions

43. Search the World Wide Web for information about "microquasars." These are objects that are found *within* the Milky Way Galaxy. How are they detected? What are the similarities and differences between these objects and true quasars? Are they long-lasting or short-lived?

44. The *Lockman hole* is a region in the constellation Ursa Minor where the Milky Way's interstellar hydrogen is the thinnest. By observing in this part of the sky, astronomers get the clearest possible view of distant galaxies and quasars. Search the World Wide Web for information about observations of the Lockman hole. What has been learned through X-ray observations made by spacecraft such as ROSAT and XMM-Newton?

AIMM 25.1

45. **Relativistic Redshift.** Access the Active Integrated Media Module "Relativistic Redshift" in Chapter 25 of the *Universe* Web site or eBook. Use this to calculate the redshift and recessional velocity of a quasar in whose spectrum the H_α emission line of hydrogen (unshifted wavelength 656 nm) appears at a wavelength of (**a**) 937 nm and (**b**) 5000 nm.

Activities

Observing Projects

46. Use a telescope with an aperture of at least 20 cm (8 in.) to observe the Seyfert Galaxy NGC 1068 (also known as M77). Located in the constellation Cetus (the Whale), this galaxy is most easily seen from September through January. The epoch 2000 coordinates are R.A. = $2^h\ 2.7^m$ and Decl. = $-0°\ 01'$. Sketch what you see. Is the galaxy's nucleus diffuse or starlike? How does this compare with other galaxies you have observed?

47. Use a telescope with an aperture of at least 20 cm (8 in.) to observe the two companions of the Andromeda Galaxy, M32 and M110. Both are small elliptical galaxies, but only M32 is suspected of harboring a supermassive black hole. They are located on opposite sides of the Andromeda Galaxy and are

most easily seen from September through January. The epoch 2000 coordinates are:

Galaxy	Right ascension	Declination
M32 (NGC 221)	0^{h} 42.7^{m}	+40° 52′
M110 (NGC 205)	0 40.4	+41 41

Make a sketch of each galaxy. Can you see any obvious difference in the appearance of these galaxies? How does this difference correlate with what you know about these galaxies?

48. If you have access to a telescope with an aperture of at least 30 cm (12 in.), you should definitely observe the Sombrero Galaxy, M104, which is suspected to have a supermassive black hole at its center. It is located in Virgo and can most easily be seen from March through July. The epoch 2000 coordinates are R.A. = 12^{h} 40.0^{m} and Decl. = −11° 37′. Make a sketch of the galaxy. Can you see the dust lane? How does the nucleus of M104 compare with the centers of other galaxies you have observed?

49. If you have access to a telescope with an aperture of at least 30 cm (12 in.), observe M87 and compare it with the other two giant elliptical galaxies, M84 and M86, that dominate the central regions of the Virgo cluster. If your copy of the book comes with the *Starry Night Enthusiast*™ program, use it to help you plan your observations. The epoch 2000 coordinates are:

Galaxy	Right ascension	Declination
M84 (NGC 4374)	12^{h} 25.1^{m}	+12° 53′
M86 (NGC 4406)	12 26.2	+12 57
M87 (NGC 4486)	12 30.8	+12 24

50. If you have access to a telescope with an aperture of at least 40 cm (16 in.), you might try to observe the brightest-appearing quasar, 3C 273, which has an apparent magnitude of nearly +13. It is located in Virgo at coordinates R.A. = 12^{h} 29^{m} 07^{s} and Decl. = +2° 03′ 07″.

51. If the Milky Way had an active galactic nucleus, with an accretion disk around its central black hole, there might be a pair of relativistic jets emanating from its center. Use the *Starry Night Enthusiast*™ program to investigate how these jets might appear from Earth. On the toolbar, set the date and time to June 15 of this year at 12:00:00 A.M. (midnight), when the center of the Milky Way is prominent in the sky. Open the **Find** pane and center the field of view on the star HIP86919. The position of this star on the celestial sphere is less than 1° from the black hole at the center of the Milky Way. Select **View > Stars** and ensure that **Milky Way** is being displayed. Select **Options > Stars > Milky Way** to open the **Milky Way Options** dialog window, move the slidebar to the right to brighten the galaxy, and click the **OK** button. Close any open panes to ensure that that the entire window is again devoted to a view of the sky. Make a sketch of the Milky Way Galaxy and attempt to show how the night sky might appear on June 15 at 12:00:00 A.M. if our Galaxy had an active galactic nucleus. Label the Milky Way, the jets, the central black hole, and the accretion disk. Assume that the plane of the accretion disk is aligned with the plane of the Milky Way. Zoom in to a field of view of about 6°. An X-ray image of the Milky Way center is superimposed upon the galaxy. (If not, select **View > Deep Space** and click on **Chandra Images** to display this image.) Open the object contextual menu over this image and click on **Magnify** to enlarge this image to see the high temperature features of this violently active region of our galaxy.

52. Use the *Starry Night Enthusiast*™ program to examine the vicinity of the galaxy M87, shown in Figure 25-8. Select **Favourites > Deep Space > Virgo Cluster** to display this large cluster of galaxies. (**a**) You can use the upward and downward pointing triangles in the **Viewing Location** panel of the toolbar to move toward or away from the cluster. You can also rotate the Virgo Cluster by putting the mouse cursor over the image and, while holding down both the Shift key and the mouse button, move the mouse. (On a two-button mouse, hold down the left mouse button.) Use these controls to get a sense of the extent of the Virgo Cluster. Use the **Viewing Location** controls to move to a distance of about 30 Mly from the Sun. Open the **Find** pane and enter Virgo A. Click the menu button associated with Virgo A in the Find pane and select **Highlight "GA Virgo Cluster" Filament** to highlight the members of this cluster in yellow to see the extent of this huge grouping of galaxies. Describe where this active galaxy, also known as M87, is located in the cluster. (**b**) Discuss how the position of M87 in the Virgo Cluster might relate to its being an active galaxy.

Collaborative Exercises

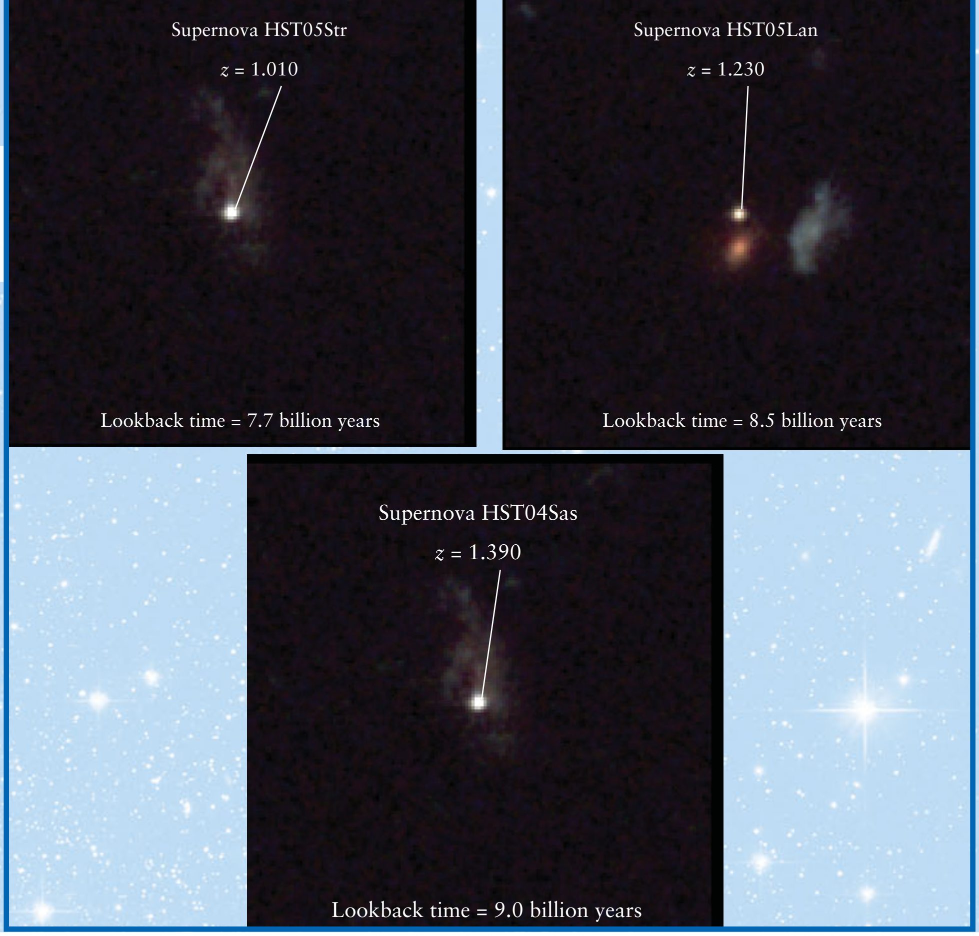
53. Make a labeled sketch clearly showing how the spectrum of 3C 273, shown in Figure 25-3, would be different if the object was moving toward us at the same velocity. Compare your sketch to that of another group and resolve any inconsistencies.

54. Consider the relationship between redshift and distance as shown in Table 25-1. If an object has a recessional velocity (v/c) of 0.902, how many light-years away was the object when it emitted the light that we see today? How many light-years away is it now? Are these closer or farther than an object with a z of 3?

55. Two dramatic images of the radio galaxy M87 are shown in Figure 25-8. Invent an imaginary scenario that would be analogous to a local dance club where two pictures would show different aspects of the event and write a description of each. Be sure to include a description of how this is analogous to the images in Figure 25-8.

26

Cosmology: The Origin and Evolution of the Universe



R I V U X G

Very distant supernovae—which we see as they were billions of years ago—help us understand the evolution of the universe. (NASA; ESA; and A. Riess, STScI)

So far in this book we have cataloged the contents of the universe. Our scope has ranged from submicroscopic objects, such as atomic nuclei, to superclusters of galaxies hundreds of millions of light-years across. In between, we have studied planets, moons, and stars.

But now we turn our focus beyond what we find in the universe to the nature of the universe itself—the subject of the science called *cosmology*. How large is the universe? What is its structure? How long has it existed, and how has it changed over time?

In this chapter we will see that the universe is expanding. This expansion began with an event at the beginning of time called the Big Bang. We will see direct evidence of the Big Bang in the form of microwave radiation from space. This radiation is the faint afterglow of a primordial fireball that filled all space shortly after the beginning of the universe.

Will the universe continue to expand forever, or will it eventually collapse back on itself? We will find that to predict the future of the universe, we must first understand what happened in the remote past. To this end, astronomers study luminous supernovae like the example shown in the accompanying image. These can be seen across billions of light-years, and so can tell us about conditions in the universe billions of years ago. We will see how recent results from such supernovae, as well as from studies of the Big Bang's afterglow, have revolutionized our understanding of cosmology and given us new insights into our place in the cosmos.

26-1 The darkness of the night sky tells us about the nature of the universe

Cosmology is the science concerned with the structure and evolution of the universe as a whole. One of the most profound and basic questions in cosmology may at first seem foolish: Why is the sky dark at night? This question,

Learning Goals

By reading the sections of this chapter, you will learn

26-1 What the darkness of the night sky tells us about the nature of the universe

26-2 What it means to say that the universe is expanding

26-3 How to estimate the age of the universe from its expansion rate

26-4 How astronomers detect the afterglow of the Big Bang

26-5 What the universe was like during its first 380,000 years

26-6 How the curvature of the universe reveals its matter and energy content

26-7 What distant supernovae tell us about the expansion history of the universe

26-8 How our understanding of the universe continues to evolve which haunted Johannes Kepler as long ago as 1610, was brought to public attention in the early 1800s by the German amateur astronomer Heinrich Olbers.

Olbers' Paradox and Newton's Static Universe

Olbers and his contemporaries pictured a universe of stars scattered more or less randomly throughout infinite space. Isaac Newton himself thought that no other model made sense. The gravitational forces between any *finite* number of stars, he argued, would in time cause them all to fall together, and the universe would soon be a compact blob.

Obviously, this has not happened. Newton concluded that we must be living amid a static, infinite expanse of stars. In this model, the universe is infinitely old, and it will exist forever without major changes in its structure. Olbers noticed, however, that a static, infinite universe presents a major puzzle.

If space goes on forever, with stars scattered throughout it, then any line of sight must eventually hit a star. In this case, no matter where you look in the night sky, you should ultimately see a star. The entire sky should be as bright as an average star, so, even at night, the sky should blaze like the surface of the Sun. **Olbers's paradox** is that the night sky is actually *dark* (Figure 26-1).

Olbers's paradox suggests that something is wrong with Newton's infinite, static universe. According to the classical, Newtonian picture of reality, space is like a gigantic flat sheet of inflexible, rectangular graph paper. (Space is actually three-dimensional, but it is easier to visualize just two of its three dimensions. In a similar way, an ordinary map represents the three-dimensional surface of the Earth, with its hills and valleys, as a flat, two-dimensional surface.)

Figure 26-1 R I V U X G

The Dark Night Sky If the universe were infinitely old and filled uniformly with stars that were fixed in place, the night sky would be ablaze with light. In fact the night sky is dark, punctuated only by the light from isolated stars and galaxies, Hence, this simple picture of an infinite, static universe cannot be correct. (NASA; ESA; and the Hubble Heritage Team, STScI/AURA)

This rigid, flat, Newtonian space stretches on and on, totally independent of stars or galaxies or anything else. The same is true of time in Newton's view of the universe; a Newtonian clock ticks steadily and monotonously forever, never slowing down or speeding up. Furthermore, Newtonian space and time are unrelated, in that a clock runs at the same rate no matter where in the universe it is located.

Einstein's Revolution and His "Greatest Blunder"

WEB LINK 26.3

Albert Einstein overturned this view of space and time. His special theory of relativity (recall Section 22-1 and Box 22-1) shows that measurements with clocks and rulers depend on the motion of the observer. What is more, Einstein's general theory of relativity (Section 22-2) tells us that gravity curves the fabric of space. As a result, the matter that occupies the universe influences the overall shape of space throughout the universe.

If we represent the universe as a sheet of graph paper, the sheet is not perfectly flat but has a dip wherever there is a concentration of mass, such as a person, a planet, or a star (see Figure 22-4). Because of gravitational effects, clocks run at different rates depending on whether they are close to or far from a massive object, as Figure 22-7*a* shows.

What does the general theory of relativity, with its many differences from the Newtonian picture, have to say about the structure of the universe as a whole? Einstein attacked this problem shortly after formulating his general theory in 1915. At that time, the prevailing view was that the universe was static, just as Newton had thought.

Einstein narrowly missed predicting that our universe is not static

Einstein was therefore dismayed to find that his calculations could not produce a truly static universe. According to general relativity, the universe must be either expanding or contracting. In a desperate move to force his theory to predict a static universe, he added to the equations of general relativity a term called the **cosmological constant** (denoted by Λ, the capital Greek letter lambda). The cosmological constant was intended to represent a pressure that tends to make the universe expand as a whole. Einstein's idea was that this pressure would just exactly balance gravitational attraction, so that the universe would be static and not collapse.

ANALOGY Einstein's cosmological constant is analogous to the pressure of gas inside a bicycle tire. This pressure exactly balances the inward force exerted by the stretched rubber of the tire itself, so the tire maintains the same size.

Unlike other aspects of Einstein's theories, the cosmological constant did not have a firm basis in physics. He just added it to make the general theory of relativity agree with the prejudice that the universe is static.

Because Einstein doubted his original equations, he missed an incredible opportunity: He could have postulated that we live in an expanding universe. Einstein has been quoted as saying in his later years that the cosmological constant was "the greatest blunder of my life." (In fact, the cosmological constant plays an important role in modern cosmology, although a very different one from what Einstein proposed. We will explore this in Section 26-7.)

Instead, the first hint that we live in an expanding universe came more than a decade later from the observations of Edwin Hubble. As we will see in Section 26-3, Hubble's discovery provides the resolution of Olbers's paradox.

26-2 The universe is expanding

WEB LINK 26.4

Hubble is usually credited with discovering that our universe is expanding. He found a simple linear relationship between the distances to remote galaxies and the redshifts of the spectral lines of those galaxies (review Section 24-5, especially Figures 24-16 and 24-17). This relationship, now called the *Hubble law,* states that the greater the distance to a galaxy, the greater is the galaxy's redshift. Thus, remote galaxies are moving away from us with speeds proportional to their distances.

The Hubble Law and the Expanding Universe

The Hubble law can be stated as a very simple equation that relates the recessional velocity v of a galaxy to its distance d from the Earth:

$$v = H_0 d$$

where H_0 is the Hubble constant. Because remote galaxies are getting farther and farther apart as time goes on, astronomers say that the universe is expanding.

ANALOGY What does it actually mean to say that the universe is expanding? According to general relativity, space itself is not rigid. The amount of space between widely separated locations changes over time. A good analogy is that of baking a chocolate chip cake, as in Figure 26-2. As the cake expands during baking, the amount of space between the chocolate chips gets larger and larger. In the same way, as the universe expands, the amount of space between widely separated galaxies increases. The expansion of the universe *is* the expansion of space.

CAUTION! It is important to realize that the expansion of the universe occurs primarily in the vast spaces that separate clusters of galaxies. Just as the chocolate chips in Figure 26-2 do not expand as the cake expands during baking, galaxies themselves do not expand. Einstein and others have established that an object that is held together by its own gravity, such as a galaxy or a cluster of galaxies, is always contained within a patch of nonexpanding space. A galaxy's gravitational field produces this nonexpanding region, which is indistinguishable from the rigid space described by Newton. Thus, the Earth and your body, for example, are not getting any bigger. Only the distance between widely separated galaxies increases with time. The *Cosmic Connections* figure has more to say about several misconceptions concerning the expanding universe.

The Hubble law is a direct proportionality—that is, a galaxy twice as far away is receding from us twice as fast. This is just what we would expect in an expanding universe. To see why this is so, imagine a grid of parallel lines (as on a piece of graph paper) crisscrossing the universe. Figure 26-3*a* shows a series of such gridlines 100 Mpc apart, along with five galaxies labeled A, B, C, D, and E that happen to lie where gridlines cross. As the universe expands in all directions, the gridlines and the attached galaxies spread apart. (This is just what would happen if the universe were a two-dimensional rubber sheet that was being pulled equally on all sides. Alternatively, you can imagine that Figure 26-3*a* depicts a very small portion of the chocolate chip cake in Figure 26-2, with galaxies taking the place of chocolate chips.)

Figure 26-3*b* shows the universe at a later time, when the gridlines are 50% farther apart (150 Mpc) and all the distances between galaxies are 50% greater than in Figure 26-3*a*. Imagine that A represents our galaxy, the Milky Way. The table accompanying Figure 26-3 shows how far each of the other galaxies have moved away from us during the expansion: Galaxies A and B and galaxies A and E were originally 100 Mpc apart, and have moved away from each other by an additional 50 Mpc; A and C, which were originally 200 Mpc apart, have increased their separation by an additional 100 Mpc; and the distance between A and D, originally 300 Mpc, has increased by an additional 150 Mpc. In other words, the increase in distance between any pair of galaxies is in direct proportion to the original distance; if the original

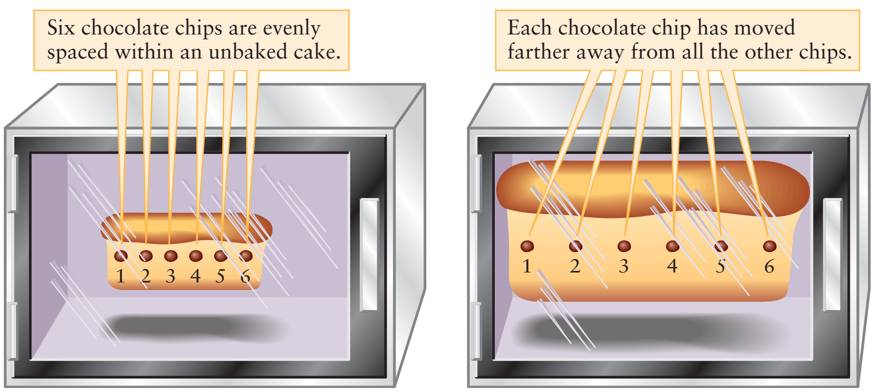
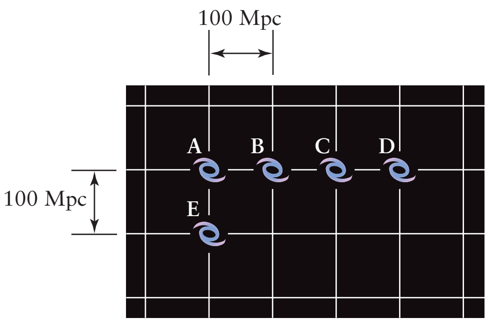
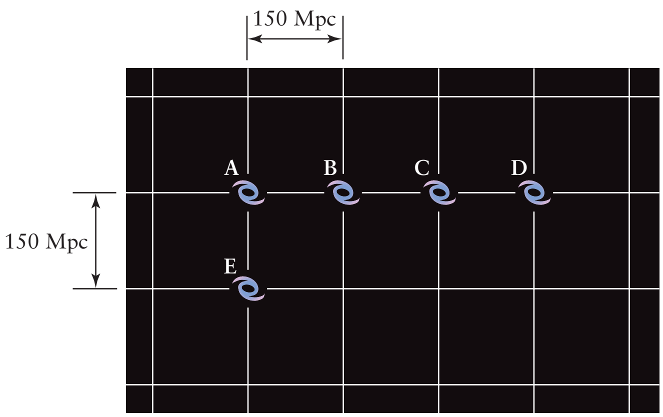


Figure 26-2

The Expanding Chocolate Chip Cake Analogy The expanding universe can be compared to what happens inside a chocolate chip cake as the cake expands during baking. (The cake is floating weightlessly inside the oven of an orbiting spacecraft crewed by hungry astronauts.) All of the chocolate chips in the cake recede from one another as the cake expands, just as all the galaxies recede from one another as the universe expands.



(a) Five galaxies spaced 100 Mpc apart



(b) The expansion of the universe spreads the galaxies apart

	Original distance (Mpc)	Later distance (Mpc)	Change in distance (Mpc)
A–B	100	150	50
A–C	200	300	100
A–D	300	450	150
A–E	100	150	50

ANIMATION 26.1

Figure 26-3

The Expanding Universe and the Hubble Law (a) Imagine five galaxies labeled A, B, C, D, and E. At the time shown here, adjacent galaxies are 100 Mpc apart. **(b)** As the universe expands, by some later time the spacing between adjacent galaxies has increased to 150 Mpc. The table shows that the greater the original distance between galaxies, the greater the amount that distance has increased. This agrees with the Hubble law.

distance is twice as great, the increase in distance is also twice as great.

To see what this tells us about the recessional velocities of galaxies, remember that velocity is equal to the distance moved divided by the elapsed time. (For example, if you traveled in a straight line for 360 kilometers in 4 hours, your velocity was (360 km)/(4 hours) = 90 kilometers per hour.) Because the distance that each galaxy moves away from A during the expansion is directly proportional to its original distance from A, it follows that the velocity v at which each galaxy moves away from A is also directly proportional to the original distance d. This is just the Hubble law, $v = H_0 d$.

CAUTION! It may seem that if the universe is expanding, and if we see all the distant galaxies rushing away from us, then we must be in a special position at the very center of the universe. In fact, the expansion of the universe looks the same from the vantage point of *any* galaxy. For example, as seen from galaxy D in Figure 26-3, the initial distances to galaxies A, B, and C are 300 Mpc, 200 Mpc, and 100 Mpc, respectively. Between parts *a* and *b* of the figure, these distances increase by 150 Mpc, 100 Mpc, and 50 Mpc, respectively. So, as seen from D as well, the recessional velocity increases in direct proportion with the distance, and in the same proportion as seen from A. In other words, no matter which galaxy you call home, you will see all the other galaxies receding from you in accordance with the same Hubble law (and the same Hubble constant) that we observe from Earth.

WEB LINK 26.5

Figure 26-2 also shows that the expansion of the universe looks the same from one galaxy as from any other. An insect sitting on any one of the chocolate chips would see all the other chips moving away. If the cake were infinitely long, it would not actually have a center; as seen from any chocolate chip within such a cake, the cake would extend off to infinity to the left and to the right, and the expansion of the cake would appear to be centered on that chip. Likewise, because every point in the universe appears to be at the center of the expansion, it follows that our universe has no center at all. (Later in this chapter we will see evidence that the universe, like our imaginary cake, is indeed infinite.)

CAUTION! "If the universe is expanding, what is it expanding into?" This commonly asked question arises only if we take our chocolate chip cake analogy too literally. In Figure 26-2, the cake (representing the universe) expands in three-dimensional space into the surrounding air. But the actual universe includes *all* space; there is nothing "beyond" it, because there is no "beyond." Asking "What lies beyond the universe?" is as meaningless as asking "Where on the Earth is north of the North Pole?"

The ongoing expansion of space explains why the light from remote galaxies is redshifted. Imagine a photon coming toward us from a distant galaxy. As the photon travels through space, the space is expanding, so the photon's wavelength becomes stretched. When the photon reaches our eyes, we see an increased wavelength: The photon has been redshifted. The longer the photon's journey, the more its wavelength will have been stretched. Thus, photons from distant galaxies have larger redshifts than those of photons from nearby galaxies, as expressed by the Hubble law.

The redshifts of distant galaxies are not Doppler shifts; they are caused by the expansion of space itself

A redshift caused by the expansion of the universe is properly called a **cosmological redshift.** It is *not* the same as a Doppler shift. Doppler shifts are caused by an object's *motion through*

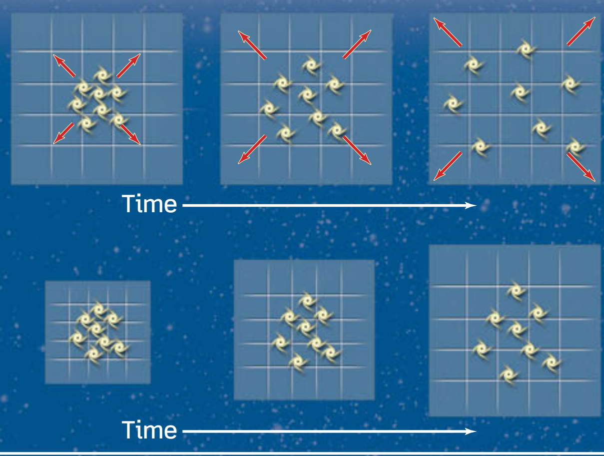
COSMIC CONNECTIONS

"Urban Legends" about the Expanding Universe

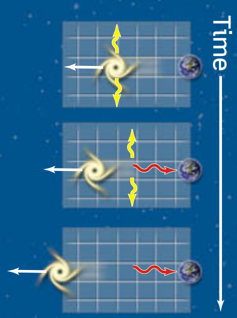
There are a number of common misconceptions or "urban legends" about what happens as the universe expands. The illustrations below depict the myth and the reality for three of these "urban legends."

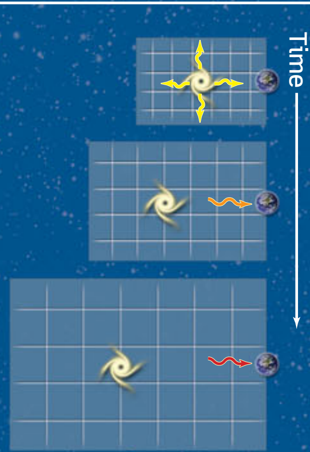
Urban Legend #1:
The expansion of the universe means that as time goes by, galaxies move away from each other through empty space. In this picture, space is simply a background upon which the galaxies act out their parts .

Reality:
The expansion of the universe means that as time goes by, *space itself* expands. As it expands, it carries the galaxies along with it.



Urban Legend #2:
The redshift of light from distant galaxies is a Doppler shift. It occurs because these galaxies are moving away from us rapidly.

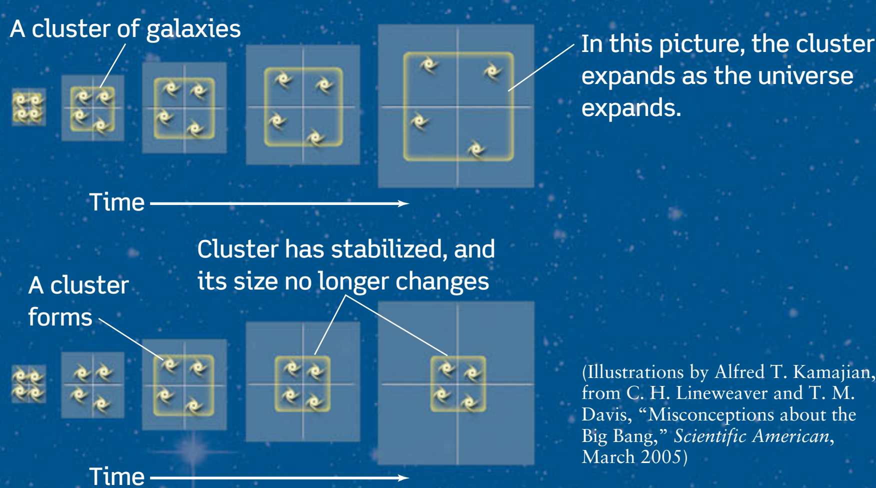




Reality:
As a photon travels through intergalactic space, its wavelength expands as the space through which it is traveling expands. This is called a *cosmological* redshift.

Urban Legend #3:
As the universe expands, so do objects within the universe. Hence galaxies within a cluster are now more spread out than they were billions of years ago.

Reality:
At first the expansion of the universe tends to pull the galaxies of a cluster away from each other. But the force of gravitational attraction binds the members of the clusters together, so the cluster stabilizes at a certain size.



(Illustrations by Alfred T. Kamajian, from C. H. Lineweaver and T. M. Davis, "Misconceptions about the Big Bang," *Scientific American*, March 2005)

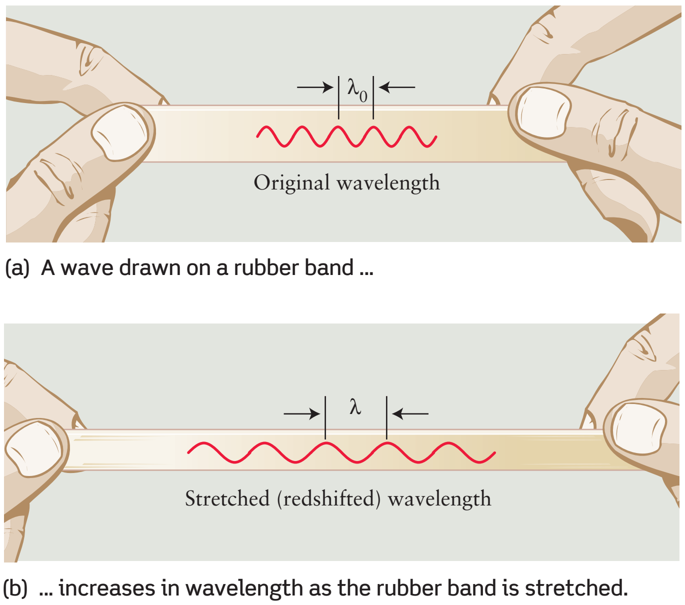


Figure 26-4

Cosmological Redshift A wave drawn on a rubber band stretches along with the rubber band. In an analogous way, a light wave traveling through an expanding universe "stretches," that is, its wavelength increases.

space, whereas a cosmological redshift is caused by the *expansion of space* (Figure 26-4).

Cosmological Redshift and Lookback Time

We can calculate the factor by which the universe has expanded since some ancient time from the redshift of light emitted by objects at that time. As we saw in Section 24-5, redshift (z) is defined as

$$z = \frac{\lambda - \lambda_0}{\lambda_0}$$

In this equation λ_0 is the unshifted wavelength of a photon and λ is the wavelength we observe. For example, λ_0 could be the wavelength of a particular emission line in the spectrum of light leaving a remote galaxy. As the galaxy's light travels through space, its wavelength is stretched by the expansion of the universe. Thus, at our telescopes we observe the spectral line to have a wavelength λ. The ratio λ/λ_0 is a measure of the amount of stretching. By rearranging terms in the preceding equation, we can solve for this ratio to obtain

$$\frac{\lambda}{\lambda_0} = 1 + z$$

For example, consider a galaxy with a redshift $z = 3$. Since the time that light left that galaxy, the universe has expanded by a factor of $1 + z = 1 + 3 = 4$. In other words, when the light left that galaxy, representative distances between widely separated galaxies were only one-quarter as large as they are today. A representative volume of space, which is proportional to the cube of its dimensions, was only $(1/4)^3 = 1/64$ as large as it is today. Thus, the density of matter in such a volume was 64 times greater than it is today.

If you know the redshift z of a distant object such as a remote galaxy, you can calculate that object's recessional velocity v (see Box 24-2 for how this is done). Then, using the Hubble law, you can determine the distance d to that object if you know the value of the Hubble constant H_0. This also tells you the **lookback time** of that object, that is, how far into the past you are looking when you see that object. For example, if the lookback time for a distant galaxy is a billion years, that means the light from the galaxy took a billion years to reach us and so we are seeing it as it was a billion years ago. The images that open this chapter show supernovae in distant galaxies with lookback times from 7.7 to 9.0 billion years. (Table 25-1 lists the distance at which we see objects with different redshifts. This distance measured in light-years is equal to the lookback time in years.)

Distances and lookback times determined in this way are somewhat uncertain, because there is some uncertainty in the value of the Hubble constant. Furthermore, as we will see in Section 26-8, the universe has not always expanded at the same rate. This means that the value of the Hubble constant H_0 was different in the distant past. (In other words, the Hubble "constant" is not actually constant in time.) Furthermore, the correct distance d to use in the Hubble law is not the distance at which we *see* the object, but rather the distance between us and the object *now*. The latter distance (which we called the *commoving radial distance* in Section 25-1) is larger because during the time that it takes to reach us from a distant object, that object has moved farther away due to the expansion of the universe.

To avoid dealing with these uncertainties and complications, astronomers commonly refer to times in the distant past in terms of redshift rather than years. For example, instead of asking, "How common were quasars 5 billion years ago?," an astronomer might ask, "How common were quasars at $z = 1.0$?" In this question, "at $z = 1.0$" is a shorthand way of saying "at the lookback time that corresponds to objects at $z = 1.0$." We will use this terminology later in this chapter. Remember that the greater the redshift, the greater the lookback time and hence the further back into time we are peering.

The Cosmological Principle

These ideas about the expanding universe demonstrate the central philosophy of cosmology. In cosmology, unlike other sciences, we cannot carry out controlled experiments or even make comparisons: There is only one universe that we can observe. To make progress in cosmology, we must accept certain philosophical assumptions or abandon hope of understanding the nature of the universe. The Hubble law provides a classic example. It could be interpreted to mean that we are at the center of the universe. We reject this interpretation, however, because it violates a cosmological extension of Copernicus's belief that we do not occupy a special location in space.

When Einstein began applying his general theory of relativity to cosmology, he made a daring assumption: Over very large distances the universe is **homogeneous,** meaning that every region is the same as every other region, and **isotropic,** meaning that the universe looks the same in every direction. In other words, if you could stand back and look at a very large region of space, any one part of the universe would look basically the same as any other part, with the same kinds of galaxies distributed through space in the same way. The assumption that the universe is homogeneous and isotropic constitutes the **cosmological principle.** It gives precise meaning to the idea that we do not occupy a special location in space.

Models of the universe based on the cosmological principle have proven remarkably successful in describing the structure and evolution of the universe and in interpreting observational data. All the discussion about the universe in this chapter and the next assumes that the universe is homogeneous and isotropic on the largest scale.

26-3 The expanding universe emerged from a cataclysmic event called the Big Bang

The Hubble flow shows that the universe has been expanding for billions of years. This means that in the past the matter in the universe must have been closer together and therefore denser than it is today. If we look far enough into the very distant past, there must have been a time when the density of matter was almost inconceivably high. This leads us to conclude that some sort of tremendous event caused ultradense matter to begin the expansion that continues to the present day. This event, called the **Big Bang,** marks the creation of the universe.

CAUTION! It is not correct to think of the Big Bang as an explosion. When a bomb explodes, pieces of debris fly off *into space* from a central location. If you could trace all the pieces back to their origin, you could find out exactly where the bomb had been. This process is not possible with the universe, however, because the universe itself always has and always will consist of all space. As we have seen, the universe logically cannot have an edge (see the discussion of the expanding chocolate chip cake in Section 26-2).

Estimating the Age of the Universe

How long ago did the Big Bang take place? To estimate an answer, imagine two galaxies that today are separated by a distance d and receding from each other with a velocity v. A movie of these galaxies would show them flying apart. If you run the movie backward, you would see the two galaxies approaching each other as time runs in reverse. We can calculate the time T_0 it will take for the galaxies to collide by using the equation

$$T_0 = \frac{d}{v}$$

This says that the time to travel a distance d at velocity v is equal to the ratio d/v. (As an example, to travel a distance of 360 km at a velocity of 90 km/h takes (360 km)/(90 km/h) = 4 hours.) If we use the Hubble law, $v = H_0 d$, to replace the velocity v in this equation, we get

$$T_0 = \frac{d}{H_0 d} = \frac{1}{H_0}$$

Note that the distance of separation, d, has canceled out and does not appear in the final expression. This means that T_0 is the same for *all* galaxies. This is the time in the past when all galaxies were crushed together, the time back to the Big Bang. In other words, the reciprocal of the Hubble constant H_0 gives us an estimate of the age of the universe, which is one reason why H_0 is such an important quantity in cosmology.

Observations suggest that H_0 = 73 km/s/Mpc to within a few percent, and this is the value we choose as our standard (see Section 24-5). Using this value, our estimate for the age of the universe is

$$T_0 = \frac{1}{73 \text{ km/s/Mpc}}$$

To convert this into units of time, we simply need to remember that 1 Mpc equals 3.09×10^{19} km and 1 year equals 3.156×10^7 s. Using the technique we discussed in Box 1-3 for converting units, we get

$$\begin{aligned} T_0 &= \frac{1}{73}\;\frac{\text{Mpc s}}{\text{km}} \times \frac{3.09 \times 10^{19} \text{ km}}{1 \text{ Mpc}} \times \frac{1 \text{ year}}{3.156 \times 10^7 \text{s}} \\ &= 1.34 \times 10^{10} \text{ years} = 13.4 \text{ billion years} \end{aligned}$$

By comparison, the age of our solar system is only 4.56 billion years, or about one-third the age of the universe. Thus, the formation of our home planet is a relatively recent event in the history of the cosmos.

The value of H_0 has an uncertainty of about 5%, so our simple estimate of the age of the universe is likewise uncertain by at least 5%. Furthermore, the formula $T_0 = 1/H_0$ is at best an approximation, because in deriving it we assumed that the universe expands at a constant rate. In Section 26-8 we will discuss how the expansion rate of the universe has changed over its history. When these factors are taken into consideration, we find that the age of the universe is 13.7 billion years, with an uncertainty of about 0.2 billion years. This is remarkably close to our simple estimate.

WEB LINK 26.6 Whatever the true age of the universe, it must be at least as old as the oldest stars. The oldest stars that we can observe readily lie in the Milky Way's globular clusters (see Section 19-4 and Section 23-1). The most recent observations, combined with calculations based on the theory of stellar evolution, indicate that these stars are about 13.4 billion years old (with an uncertainty of about 6%). Encouragingly enough, this is less than the calculated age of the universe: The oldest stars in our universe are younger than the universe itself!

Our Observable Universe and the Dark Night Sky

The Big Bang helps resolve Olbers's paradox, which we discussed in Section 26-1. We know that the universe had a definite beginning, and thus its age is finite (as opposed to infinite). If the universe is 13.7 billion years old, then the most distant objects that we can see are those whose light has traveled 13.7 billion years to reach us. (Due to the expansion of the universe, these objects are now more than 13.7 billion light-years away.) As a result, we

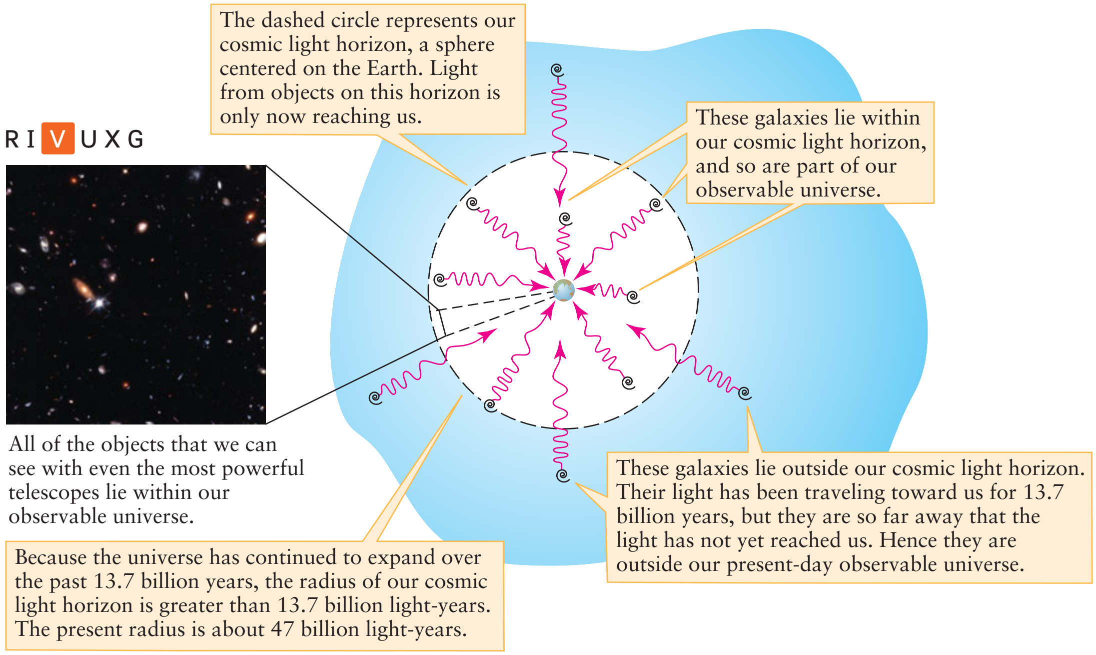


Figure 26-5

Our Observable Universe The part of the universe that we can observe lies within our cosmic light horizon. The galaxies that we can just barely make out with our most powerful telescopes lie inside our cosmic light horizon; we see them as they were less than a billion years after the Big Bang. We cannot see objects beyond our cosmic light horizon, because in the 13.7 billion years since the Big Bang their light has not had enough time to reach us. (Inset: Robert Williams and the Hubble Deep Field Team, STScI; NASA)

can only see objects that lie within an immense sphere centered on the Earth (Figure 26-5). This is true even if the universe is infinite, with galaxies scattered throughout its limitless expanse.

The surface of the sphere depicted in Figure 26-5 is called our **cosmic light horizon.** Our entire **observable universe** is located inside this sphere. We cannot see anything beyond our cosmic light horizon, because the time required for light to reach us from these incredibly remote distances is greater than the present age of the universe. As time goes by, light from more distant parts of the universe reaches us for the first time, and the size of the cosmic light horizon (and hence the size of our observable universe) increases. Galaxies are distributed sparsely enough in our observable universe that there are no stars along most of our lines of sight. This helps explain why the night sky is dark.

Measuring the recessional velocities of galaxies allows us to estimate the age of the universe

Besides the finite age of the universe, a second effect also contributes significantly to the darkness of the night sky—the redshift. According to the Hubble law, the greater the distance to a galaxy, the greater the redshift. When a photon is redshifted, its wavelength becomes longer, and its energy—which is inversely proportional to its wavelength (see Section 5-5)—decreases. Consequently, even though there are many galaxies far from the Earth, they have large redshifts and their light does not carry much energy. A galaxy nearly at the cosmic light horizon has a nearly infinite redshift, meaning that the light we receive from that galaxy carries practically no energy at all. This decrease in photon energy because of the expansion of the universe decreases the brilliance of remote galaxies, helping to make the night sky dark.

The concept of a Big Bang origin for the universe is a straightforward, logical consequence of having an expanding universe. If you can just imagine far enough back into the past, you can arrive at a time 13.7 billion years ago, when the density throughout the universe was infinite. As a result, throughout the universe space and time were completely jumbled up in a condition of infinite curvature similar to that at the singularity found at the center of a black hole (see Section 22-5). For this reason, a better name for the Big Bang is the **cosmic singularity.** Thanks to the infinite curvature, the usual laws of physics do not tell us exactly what happened at the moment of the Big Bang.

A very short time after the Big Bang, space and time began to behave in the way we think of them today. This short time interval, called the **Planck time** (t_P), is given by the following expression:

The Planck time

$$t_P = \sqrt{\frac{Gh}{c^5}} = 1.35 \times 10^{-43} \text{ s}$$

t_P = Planck time

G = universal constant of gravitation

h = Planck's constant

c = speed of light

LOOKING DEEPER 26.1

We do not yet understand how space, time, and matter behaved in that brief but important interval from the beginning of the Big Bang to the Planck time, about 10^{-43} seconds later. (Indeed, the laws of physics suggest that it may be impossible ever to know what happened during this extremely short time interval.) Hence, the Planck time represents a limit to our knowledge of conditions at the very beginning of the universe.

26-4 The microwave radiation that fills all space is evidence of a hot Big Bang

One of the major advances in twentieth-century astronomy was the discovery of the origin of the heavy elements. We know today that essentially all the heavy elements are created by thermonuclear reactions at the centers of stars and in supernovae (see Chapter 20). The starting point of all these reactions is the fusion of hydrogen into helium, which we described in Section 16-1. But as astronomers began to understand the details of thermonuclear synthesis in the 1960s, they were faced with a dilemma: There is far more helium in the universe than could have been created by hydrogen fusion in stars.

For example, the Sun consists of about 74% hydrogen and 25% helium by mass, leaving only 1% for all the remaining heavier elements combined. This 1% can be understood as material produced inside earlier generations of massive stars that long ago cast these heavy elements out into space when they became supernovae. Some freshly made helium, produced by the thermonuclear fusion of hydrogen within the stars, certainly accompanied these heavy elements. But calculations showed that the amount of helium produced in this way was not nearly enough to account for one-quarter of the Sun's mass. Because it was thought that the universe originally contained only hydrogen—the simplest of all the chemical elements—the presence of so much helium posed a major dilemma.

A Hot Big Bang and the Cosmic Microwave Background

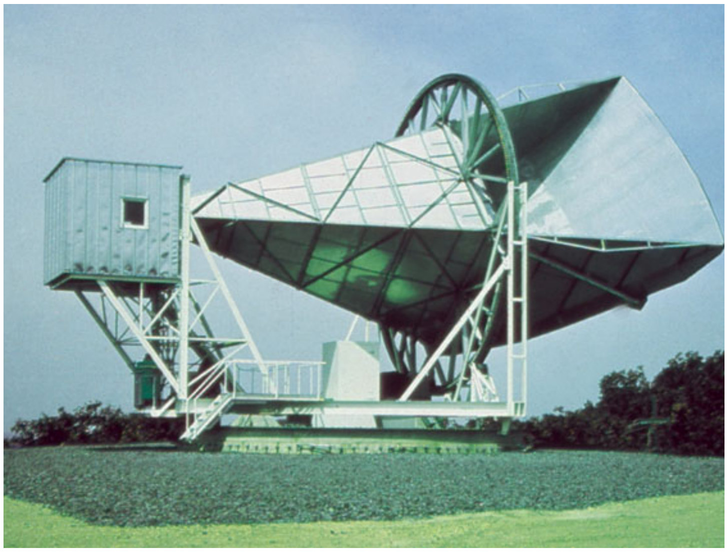
Shortly after World War II, Ralph Alpher and Robert Hermann proposed that the universe immediately following the Big Bang must have been so incredibly hot that thermonuclear reactions occurred everywhere throughout space. Following up this idea in 1960, Princeton University physicists Robert Dicke and P. J. E. Peebles discovered that they could indeed account for today's high abundance of helium by assuming that the early universe had been at least as hot as the Sun's center, where helium is currently being produced. The hot early universe must therefore have been filled with many high-energy, short-wavelength photons. The properties of this radiation field depended on its temperature, as described by Planck's blackbody law (review Figure 5-11).

The universe has expanded so much since those ancient times that all those short-wavelength photons have had their wavelengths stretched by a tremendous factor. As a result, they have become low-energy, long-wavelength photons. The temperature of this cosmic radiation field is now only a few degrees above absolute zero. By Wien's law, radiation at such a low temperature should have its peak intensity at microwave wavelengths of approximately 1 millimeter. Hence, this radiation field, which fills all of space, is called the **cosmic microwave background** or **cosmic background radiation.** In the early 1960s, Dicke and his colleagues began designing an antenna to detect this microwave radiation.

Meanwhile, just a few miles from Princeton University, Arno Penzias and Robert Wilson of Bell Telephone Laboratories in New Jersey were working on a new microwave horn antenna designed to relay telephone calls to Earth-orbiting communications satellites (Figure 26-6). Penzias and Wilson were deeply puzzled when, no matter where in the sky they pointed their antenna, they detected faint background noise. Thanks to a colleague, they happened to learn about the work of Dicke and Peebles and came to realize that they had discovered the cooled-down cosmic background radiation left over from the hot Big Bang. Penzias and Wilson shared the 1978 Nobel Prize in Physics for their discovery.

The afterglow of the Big Bang was first discovered by a happy coincidence—and can be detected with an ordinary television

You can actually detect cosmic background radiation using an ordinary television set. This radiation is responsible for about 1% of the random noise or "hash" that appears on the screen when you tune a television to a station that is off the air. Using far more sophisticated detectors than TV sets, scientists have made many measurements of the intensity of the background radiation at a variety of wavelengths. Unfortunately, the Earth's atmosphere is almost totally opaque to wavelengths between about 10 μm and 1 cm (see Figure 6-25), which is just the wavelength

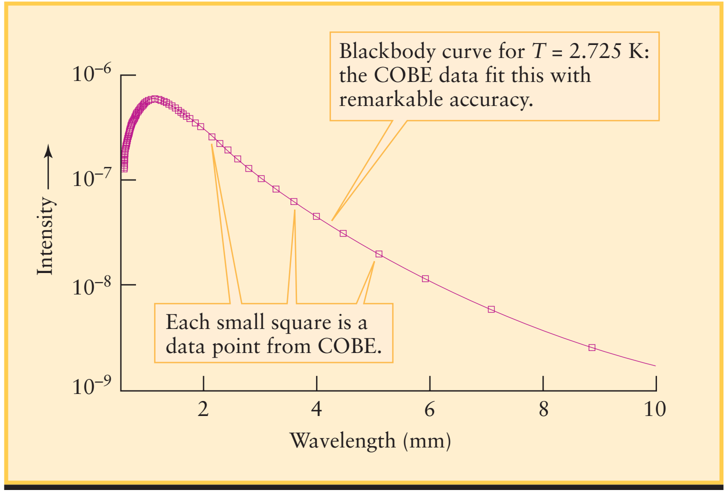


WEB LINK 26.7

Figure 26-6 R I V U X G

The Bell Labs Horn Antenna Using this microwave horn antenna, originally built for communications purposes, Arno Penzias and Robert Wilson detected a signal that seemed to come from all parts of the sky. After carefully removing all potential sources of electronic "noise" (including bird droppings inside the antenna) that could create a false signal, Penzias and Wilson realized that they were actually detecting radiation from space. This radiation is the afterglow of the Big Bang. (Bell Labs)

(a) The COBE spacecraft



(b) The spectrum of the cosmic microwave background

WEB LINK 26.8

Figure 26-7

COBE and the Spectrum of the Cosmic Microwave Background **(a)** The Cosmic Background Explorer satellite (COBE), launched in 1989, measured the spectrum and angular distribution of the cosmic microwave background over a wavelength range from 1 μm to 1 cm. **(b)** A blackbody curve gives an excellent match to the COBE data. (a: Courtesy of J. Mather/NASA; b: Courtesy of E. Cheng/NASA COBE Science Team)

range in which the background radiation is most intense. As a result, scientists have had to place detectors either on high-altitude balloons (which can fly above the majority of the obscuring atmosphere) or, even better, on board orbiting spacecraft.

A Detailed Look at the Cosmic Microwave Background

The first high-precision measurements of the cosmic microwave background came from the Cosmic Background Explorer (COBE, pronounced "coe-bee") satellite, which was placed in Earth orbit in 1989 (Figure 26-7*a*). Data from COBE's spectrometer, shown in Figure 26-7*b*, demonstrate that this ancient radiation has the spectrum of a blackbody with a temperature of 2.725 K. In recognition of this discovery, as well as others that we will discuss in Section 26-5, COBE team leaders John Mather of NASA and George Smoot of the University of California, Berkeley, were awarded the 2006 Nobel Prize in Physics.

An important feature of the microwave background is that its intensity is almost perfectly isotropic, that is, the same in all directions. In other words, we detect nearly the same background intensity from all parts of the sky. This is a striking confirmation of Einstein's assumption that the universe is isotropic (see Section 26-2). However, extremely accurate measurements first made from high-flying airplanes, and later from high-altitude balloons and from COBE, reveal a very slight variation in temperature across the sky. The microwave background appears slightly warmer than average toward the constellation of Leo and slightly cooler than average in the opposite direction toward Aquarius. Between the warm spot in Leo and the cool spot in Aquarius, the background temperature declines smoothly across the sky. Figure 26-8 is a map of the microwave sky showing this variation.

This apparent variation in temperature is caused by the Earth's motion through the cosmos. If we were at rest with respect to the microwave background, the radiation would be even more nearly isotropic. Because we are moving through this radiation field, however, we see a Doppler shift. Specifically, we see shorter-than-average wavelengths in the direction toward which we are moving, as drawn in Figure 26-9. A decrease in wavelength corresponds to an increase in photon energy and thus an

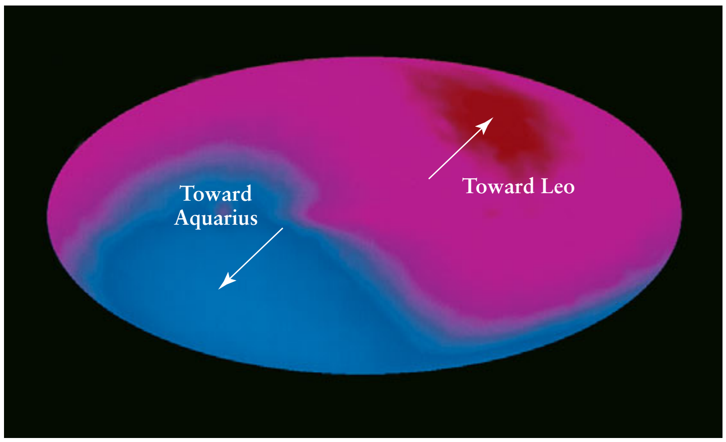


Figure 26-8 R I V U X G

The Microwave Sky In this map of the entire sky made from COBE data, the plane of the Milky Way runs horizontally across the map, with the galactic center in the middle. Color indicates temperature—red is warm and blue is cool. The small temperature variation across the sky—only 0.0033 K above or below the average radiation temperature of 2.725 K—is caused by the Earth's motion through the microwave background. (NASA)

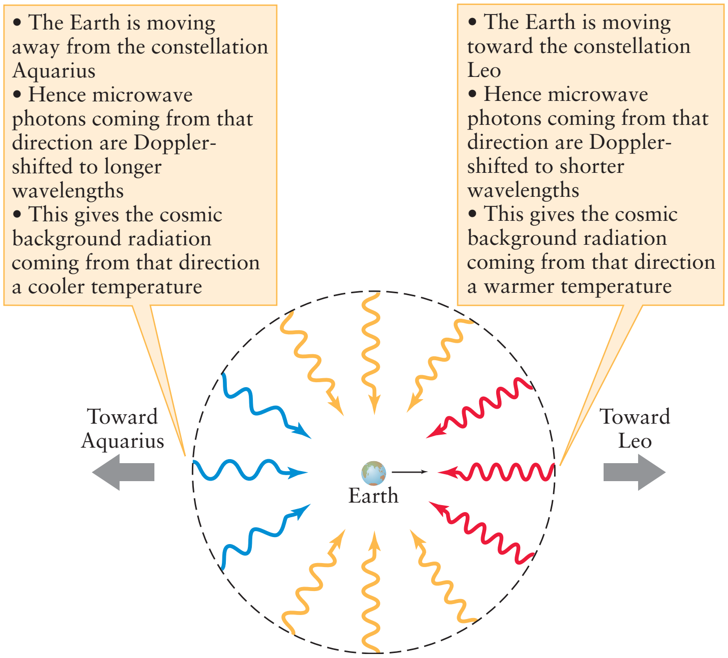


Figure 26-9

Our Motion Through the Microwave Background Because of the Doppler effect, we detect shorter wavelengths in the microwave background and a higher temperature of radiation in that part of the sky toward which we are moving. This part of the sky is the area shown in red in Figure 26-8. In the opposite part of the sky, shown in blue in Figure 26-8, the microwave radiation has longer wavelengths and a cooler temperature.

increase in the temperature of the radiation. The slight temperature excess observed, about 0.00337 K, corresponds to a speed of 371 km/s. Conversely, we see longer-than-average wavelengths in that part of the sky from which we are receding. An increase in wavelength corresponds to a decline in photon energy and, hence, a decline in radiation temperature.

Our solar system is thus traveling away from Aquarius and toward Leo at a speed of 371 km/s. Taking into account the known velocity of the Sun around the center of our Galaxy, we find that the entire Local Group of galaxies, including our Milky Way Galaxy, is moving at about 620 km/s toward the Hydra-Centaurus supercluster. Observations show that thousands of other galaxies are being carried in this direction, as is the Hydra-Centaurus supercluster itself. This tremendous flow of matter is thought to be due to the gravitational pull of an enormous collection of visible galaxies and dark matter lying in that direction. This immense object, dubbed the *Great Attractor,* lies about 50 Mpc (150 million ly) from Earth (see Figure 24-22).

The existence of such concentrations of mass, as well as the existence of superclusters of galaxies (see Section 22-6), shows that the universe is rather "lumpy" on scales of 100 Mpc or smaller. It is only on larger scales that the universe is homogeneous and isotropic.

26-5 The universe was a hot, opaque plasma during its first 380,000 years

Everything in the universe falls into one of two categories—energy or matter. One form of energy is radiation, that is, photons. (We will encounter another type of energy in Section 26-7.) There are many photons of starlight traveling across space, but the vast majority of photons in the universe belong to the cosmic microwave background. The matter in the universe is contained in such luminous objects as stars, planets, and galaxies, as well as in nonluminous dark matter. A natural question to ask is this: Which plays a more important role in the universe, radiation or matter? As we will see, the answer to this question is different for the early universe from the answer for our universe today.

Radiation and Matter in the Universe

To make a comparison between radiation and matter, recall Einstein's famous equation $E = mc^2$ (see Section 16-1). We can think of the photon energy in the universe (E) as being equal to a quantity of mass (m) multiplied by the square of the speed of light (c). The amount of this equivalent mass in a volume V, divided by that volume, is the **mass density of radiation** (ρ_{rad}; say "rho sub rad").

We can combine $E = mc^2$ with the Stefan-Boltzmann law (see Section 5-4) to give the following formula:

Mass density of radiation

$$\rho_{rad} = \frac{4\sigma T^4}{c^3}$$

ρ_{rad} = mass density of radiation

T = temperature of radiation

σ = Stefan-Boltzmann constant
$= 5.67 \times 10^{-8}\ \mathrm{W\ m^{-2}\ K^{-4}}$

c = speed of light = 3.00×10^8 m/s

For the present-day temperature of the cosmic background radiation, $T = 2.725$ K, this equation yields

$$\rho_{rad} = 4.6 \times 10^{-31}\ \mathrm{kg/m^3}$$

The **average density of matter** (ρ_m; say "rho sub em") in the universe is harder to determine. To find this density, we look at a large volume (V) of space, determine the total mass (M) of all the stars, galaxies, and dark matter in that volume, and divide the volume into the mass: $\rho_m = M/V$. (We emphasize that this is the *average* density of matter. It would be the actual density if all the matter in the universe were spread out uniformly rather than being clumped into galaxies and clusters of galaxies.) Determining how much mass is in a large volume of space is a challenging task. A major part of the challenge involves dark matter, which emits no electromagnetic radiation and can be detected only by its gravitational influence (see Section 23-4 and Section 24-8).

One method that appears to deal successfully with this challenge is to observe clusters of galaxies, within which most of the mass in the universe is concentrated. Rich clusters are surrounded by halos of hot, X-ray–emitting gas, typically at temperatures of 10^7 to 10^8 K (see Figure 24-25). Such a halo should be in hydrostatic equilibrium, so that it neither expands nor contracts but remains the same size (see Section 16-2, especially Figure 16-2). The outward gas pressure associated with the halo's high temperature would then be balanced by the inward gravitational pull due to the total mass of the cluster. Thus, by measuring the temperature of the halo—which can be determined from the properties of the halo's X-ray emission—astronomers can infer the cluster's mass, including the contribution from dark matter.

From galaxy clusters and other measurements, the present-day average density of matter in the universe is thought to be

$$\rho_m = 2.4 \times 10^{-27}\ \text{kg/m}^3$$

with an uncertainty of about 15%. The mass of a single hydrogen atom is 1.67×10^{-27} kg. Hence, if the mass of the universe were spread uniformly over space, there would be the equivalent of $1\frac{1}{2}$ hydrogen atoms per cubic meter of space. By contrast, there are 5×10^{25} atoms in a cubic meter of the air you breathe! The very small value of ρ_m shows that our universe has very little matter in it.

Furthermore, by counting galaxies and other measurements, astronomers determine that the average density of *luminous* matter (that is, the stars and gas within clusters of galaxies) is about 4.2×10^{-28} kg/m^3. This is only about 17% of the average density of matter of all forms. Thus, nonluminous dark matter is actually the predominant form of matter in our universe. The "ordinary" matter of which the stars, the planets, and ourselves are made is actually very exceptional!

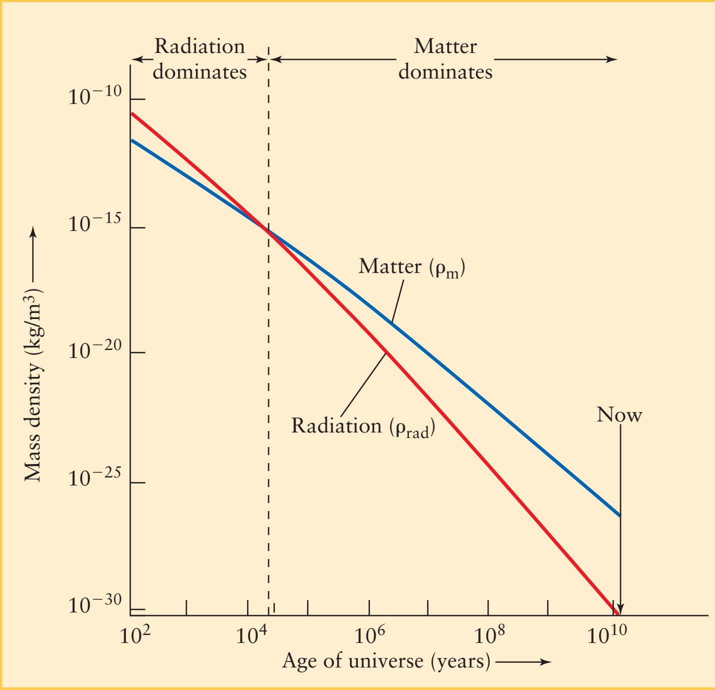


Figure 26-10

The Evolution of Density For approximately 24,000 years after the Big Bang, the mass density of radiation (ρ_{rad}, shown in red) exceeded the matter density (ρ_m, shown in blue), and the universe was radiation-dominated. Later, however, continued expansion of the universe caused ρ_{rad} to become less than ρ_m, at which point the universe became matter-dominated. (Graph courtesy of Clem Pryke, University of Chicago)

When Radiation Held Sway Over Matter

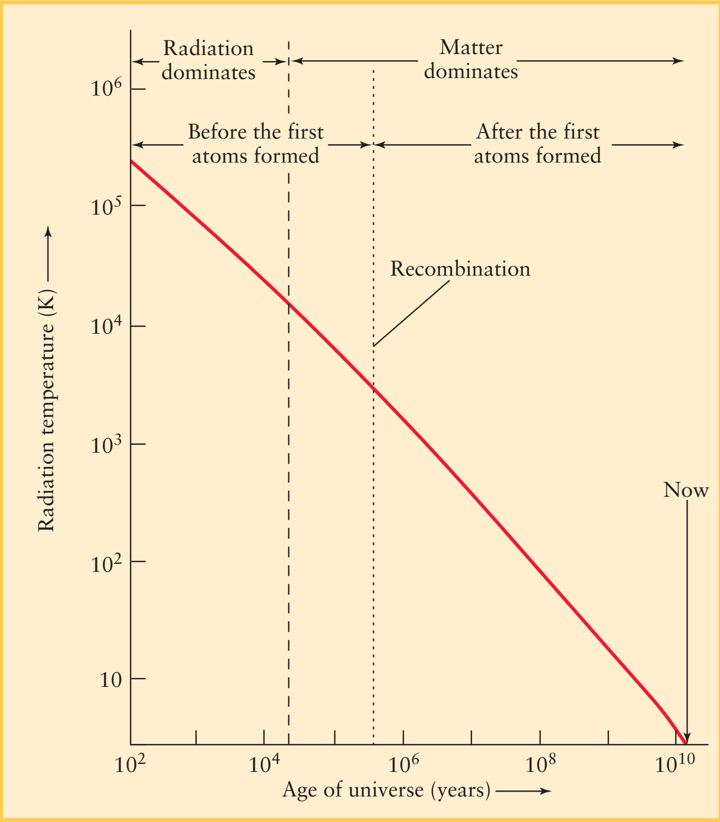
Although the average density of matter in the universe is tiny by Earth standards, it is thousands of times larger than ρ_{rad}, the mass density of radiation. However, this was not always the case. Matter prevails over radiation today only because the energy now carried by microwave photons is so small. Nevertheless, the number of photons in the microwave background is astounding. From the physics of blackbody radiation it can be demonstrated that there are today 410 million (4.1×10^8) photons in every cubic meter of space. In other words, the photons in space outnumber atoms by roughly a billion (10^9) to one. In terms of total number of particles, the universe thus consists almost entirely of microwave photons. This radiation field no longer has much "clout," however, because its photons have been redshifted to long wavelengths and low energies after 13.7 billion years of being stretched by the expansion of the universe.

In contrast, think back toward the Big Bang. The universe becomes increasingly compressed, and thus the density of matter increases as we go back in time. The photons in the background radiation also become more crowded together as we go back in time. But, in addition, the photons become less redshifted and thus have shorter wavelengths and higher energy than they do today. Because of this added energy, the mass density of radiation (ρ_{rad}) increases more quickly as we go back in time than does the average density of matter (ρ_m). In fact, as Figure 26-10 shows, there was a time in the ancient past when ρ_{rad} equaled ρ_m. Before this time, ρ_{rad} was greater than ρ_m, and radiation thus held sway over matter. Astronomers call this state a **radiation-dominated universe.** After ρ_m became greater than ρ_{rad}, so that matter prevailed over radiation, our universe became a **matter-dominated universe.**

This transition from a radiation-dominated universe to a matter-dominated universe occurred about 24,000 years after the Big Bang, at a time that corresponds to a redshift of about $z = 5200$. Since that time the wavelengths of photons have been stretched by a factor of $1 + z$, or about 5200. Today these microwave photons typically have wavelengths of about 1 mm. But when the universe was about 24,000 years old, they had wavelengths of about 190 nm, in the ultraviolet part of the spectrum.

The temperature of the background radiation has declined over the eons thanks to the expansion of the universe

To calculate the temperature of the cosmic background radiation at the time of this transition from a radiation-dominated universe to a matter-dominated one, we use Wien's law (see Section 5-4). This law says that the wavelength of maximum emission (λ_{max}) of a blackbody is inversely proportional to its



WEB LINK 26.10

Figure 26-11

The Evolution of Radiation Temperature As the universe expanded, the photons in the radiation background became increasingly redshifted and the temperature of the radiation fell. Approximately 380,000 years after the Big Bang, when the temperature fell below 3000 K, hydrogen atoms formed and the radiation field "decoupled" from the matter in the universe. After that point, the temperature of matter in the universe was not the same as the temperature of radiation. The time when the first atoms formed is called the era of recombination (see Figure 26-12). (Graph courtesy of Clem Pryke, University of Chicago)

temperature (T): A *decrease* of λ_{max} by a factor of 2 corresponds to an *increase* of T by a factor of 2.

The present-day peak wavelength of the cosmic background radiation corresponds to a blackbody temperature of 2.725 K. Hence, a peak wavelength 5200 times smaller corresponds to a temperature 5200 times greater: $T = 5200 \times 2.725\text{ K} = 14{,}000\text{ K}$. In other words, the radiation temperature at redshift z was greater than the present-day radiation temperature by a factor of $1 + z$. This means that the temperature of the radiation background was once much greater and has been declining over the ages, as Figure 26-11 shows.

When the First Atoms Formed

The nature of the universe changed again in a fundamental way about 380,000 years after the Big Bang, when z was roughly 1100 and the temperature of the radiation background was about $1100 \times 2.725\text{ K} = 3000\text{ K}$. To see the significance of this moment in cosmic history, recall that hydrogen is by far the most abundant element in the universe—hydrogen atoms outnumber helium atoms by about 12 to 1. A hydrogen atom consists of a single proton orbited by a single electron, and it takes relatively little energy to knock the proton and electron apart. In fact, ultraviolet radiation warmer than about 3000 K easily ionizes hydrogen. Thus, hydrogen atoms could not survive in the universe that existed before $z = 1100$. That is, in the first 380,000 years after the Big Bang, the background photons had energies great enough to prevent electrons and protons from binding to form hydrogen atoms (Figure 26-12*a*). Only since $z = 1100$ (that is,

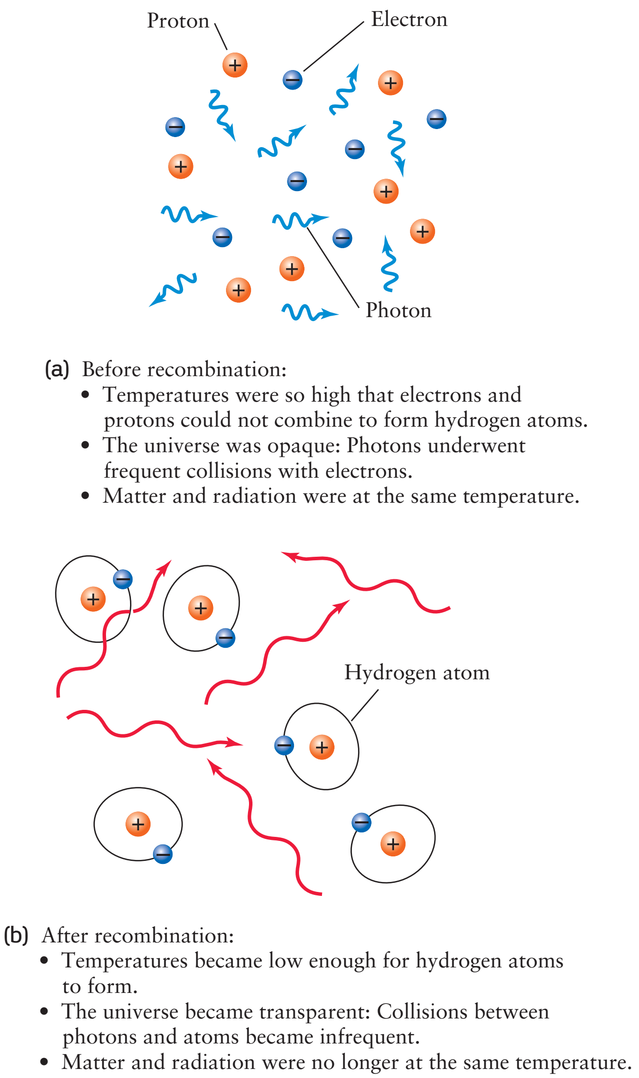


Figure 26-12

The Era of Recombination **(a)** Before recombination, the energy of photons in the cosmic background was high enough to prevent protons and electrons from forming hydrogen atoms. **(b)** Some 380,000 years after the Big Bang, the energy of the cosmic background radiation became low enough that hydrogen atoms could survive.

since $t = 380{,}000$ years) have the energies of these photons been low enough to permit hydrogen atoms to exist (Figure 26-12*b*).

The epoch when atoms first formed at $t = 380{,}000$ years is called the **era of recombination.** This refers to electrons "recombining" to form atoms. (The name is a bit misleading, because the electrons and protons had never before combined into atoms.)

Prior to $t = 380{,}000$ years, the universe was completely filled with a shimmering expanse of high-energy photons colliding vigorously with protons and electrons. This state of matter, called a **plasma,** is opaque, just like the glowing gases inside a discharge tube (like a neon advertising sign). The surface and interior of the Sun are also a hot, glowing, opaque plasma (see Section 16-9). P. J. E. Peebles coined the term **primordial fireball** to describe the universe during its first 380,000 years of existence.

After $t = 380{,}000$ years, the photons no longer had enough energy to keep the protons and electrons apart. As soon as the temperature of the radiation field fell below about 3000 K, protons and electrons began combining to form hydrogen atoms. These atoms do not absorb low-energy photons, and so space became transparent! All the photons that heretofore had been vigorously colliding with charged particles could now stream unimpeded across space. Today, these same photons constitute the microwave background.

Before recombination, matter and the radiation field had the same temperature, because photons, electrons, and protons were all in continuous interaction with one another. After recombination, photons and atoms hardly interacted at all, and thus the temperature of matter in the universe was no longer the same as the temperature of the background radiation. Thus, $T = 2.725$ K is the temperature of the present-day background radiation field, *not* the temperature of the matter in the universe. Note that while the temperature of the background radiation is very uniform, the temperature of matter in the universe is anything but: It ranges from hundreds of millions of kelvins in the interiors of giant stars to a few tens of kelvins in the interstellar medium.

ANALOGY A good analogy is the behavior of a glass of cold water. If you hold the glass in your hand, the water will get warmer and your hand will get colder until both the water and your hand are at the same temperature. But if you set the glass down and do not touch it, so that the glass and your hand do not interact, their temperatures are decoupled: The water will stay cold and your hand will stay warm for much longer.

Because the universe was opaque prior to $t = 380{,}000$ years, we cannot see any further into the past than the era of recombination. In particular, we cannot see back to the era when the universe was radiation-dominated. The microwave background, whose photons have suffered a redshift of $z = 1100$, contains the most ancient photons we will ever be able to observe.

Nonuniformities in the Early Universe and the Origin of Galaxies

Careful analysis of COBE data showed that the cosmic background radiation is not completely isotropic. Even when the effects of the Earth's motion are accounted for, there remain variations in the temperature of the radiation field of about 100 μK (100 microkelvins, or 10^{-4} K) above or below the average 2.725 K temperature. These tiny temperature variations indicate that the matter and radiation in the universe were not totally uniform at the moment of recombination. Regions that were slightly denser than average were also slightly cooler than average; less dense regions were slightly warmer. When radiation decoupled from matter at the time of recombination, the radiation preserved a record of these variations in temperature and density.

Astronomers place great importance on studying temperature variations in the cosmic background radiation. The reason is that concentrations of mass in our present-day universe, such as superclusters of galaxies, are thought to have formed from the denser regions in the early universe. Within these immense concentrations formed the galaxies, stars, and planets. Thus, by studying these nonuniformities, we are really studying our origins.

WEB LINK 26.11

Unfortunately, the detectors on board COBE had a relatively coarse angular resolution of 7° and thus could not give a detailed picture of these temperature variations. In 1998 two balloon-borne experiments, BOOMERANG and MAXIMA, carried new, high-resolution telescopes aloft to study the cosmic background radiation with unprecedented precision. The best all-sky coverage of the background radiation has come from the state-of-the-art detectors on board the Wilkinson Microwave Anisotropy Probe (WMAP for short), a NASA spacecraft that was launched in 2001. Shown in Figure 26-13, the spacecraft is named for the late David Wilkinson of Princeton University, who was a pioneer in studies of the cosmic background

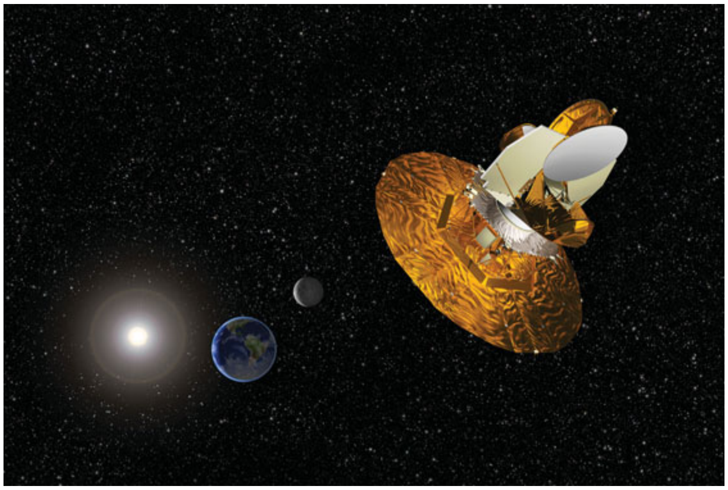
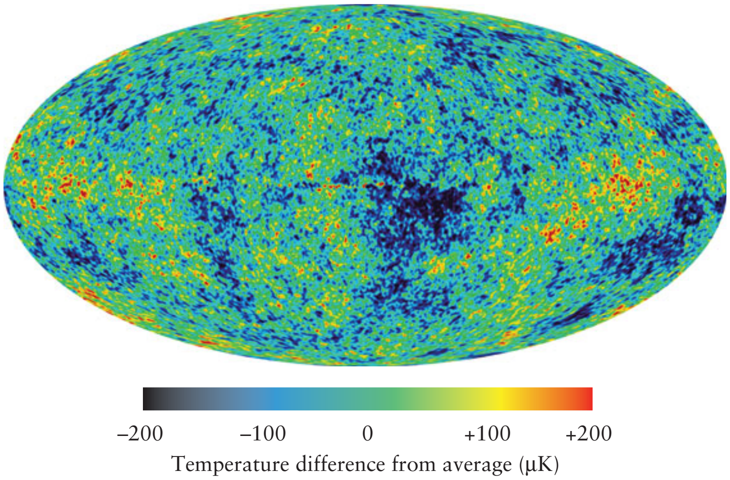


Figure 26-13

In Search of Ancient Photons This artist's impression shows the Wilkinson Microwave Anisotropy Probe (WMAP) en route to a location in space called L2, which lies about 1.5 million kilometers from Earth on the side opposite the Sun. At this position WMAP takes one year to orbit the Sun. The solar panels continually shade WMAP's detectors from sunlight, keeping them cold so that they can accurately measure the low-temperature photons ($T = 2.725$ K) of the cosmic background radiation. (NASA/WMAP Science Team)



ANIMATION 26.2 ANIMATION 26.3

Figure 26-14 R I V U X G

Temperature Variations in the Cosmic Microwave Background This map from WMAP data shows small variations in the temperature of the cosmic background radiation across the entire sky. (The variations due to the Earth's motions through space, shown in Figure 26-8, have been factored out.) Lower-temperature regions (shown in blue) show where the early universe was slightly denser than average. Note that the variations in temperature are no more than 200 μK, or 2×10^{-4} K. (NASA/WMAP Science Team)

radiation. Figure 26-14 shows a map of the sky based on data taken by the WMAP detectors. This map shows us the state of the universe when it was less than 0.003% of its present age.

Temperature variations in the cosmic background radiation do more than show us the origins of large-scale structure in the universe. As we will see in the next two sections, they actually reveal the shape of the universe as a whole.

26-6 The shape of the universe indicates its matter and energy content

We have seen that by following the mass densities of radiation (ρ_{rad}) and of matter (ρ_m), we can learn about the evolution of the universe. But it is equally important to know the combined mass density of *all* forms of matter and energy. (In an analogous way, an accountant needs to know the overall financial status of a company, not just individual profits or losses.) Remarkably, we can do this by investigating the overall shape of the universe.

The Curvature of the Universe

Einstein's general theory of relativity explains that gravity curves the fabric of space. Furthermore, the equivalence between matter and energy, expressed by Einstein's equation $E = mc^2$, tells that either matter or energy produces gravity. Thus, the matter and energy scattered across space should give the universe an overall curvature. The degree of curvature depends on the **combined average mass density** of all forms of matter and energy. This quantity, which we call ρ_0 (say "rho sub zero"), is the sum of the average mass densities of matter, radiation, and any other form of energy. Thus, by measuring the curvature of space, we should be able to determine the value of ρ_0 and, hence, learn about the content of the universe as a whole.

To see what astronomers mean by the curvature of the universe, imagine shining two powerful laser beams out into space so that they are perfectly parallel as they leave the Earth. Furthermore, suppose that nothing gets in the way of these two beams, so that we can follow them for billions of light-years as they travel across the universe and across the space whose curvature we wish to detect. There are only three possibilities:

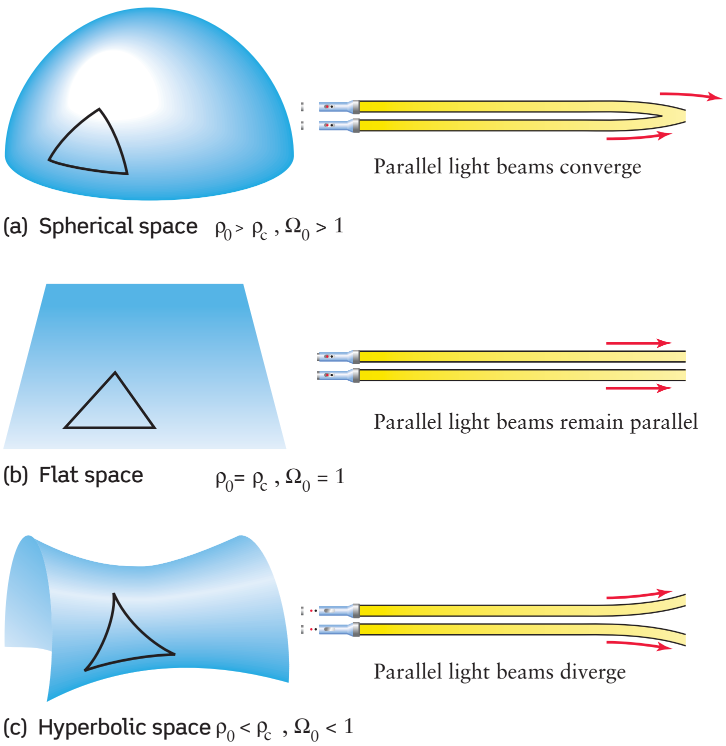
1. We might find that our two beams of light remain perfectly parallel, even after traversing billions of light-years. In this case, space would not be curved: The universe would have **zero curvature,** and space would be **flat.**

2. Alternatively, we might find that our two beams of light gradually converge. In such a case, space would not be flat. Recall that lines of longitude on the Earth's surface are parallel at the equator but intersect at the poles. Thus, in this case the three-dimensional geometry of the universe would be analogous to the two-dimensional geometry of a spherical surface. We would then say that space is **spherical** and that the universe has **positive curvature.** Such a universe is also called **closed,** because if you travel in a straight line in any direction in such a universe, you will eventually return to your starting point.

3. Finally, we might find that the two initially parallel beams of light would gradually diverge, becoming farther and farther apart as they moved across the universe. In this case, the universe would still have to be curved, but in the opposite sense from the spherical model. We would then say that the universe has **negative curvature.** In the same way that a sphere is a positively curved two-dimensional surface, a saddle is a good example of a negatively curved two-dimensional surface. Parallel lines drawn on a sphere always converge, but parallel lines drawn on a saddle always diverge. Mathematicians say that saddle-shaped surfaces are hyperbolic. Thus, in a negatively curved universe, we would describe space as **hyperbolic.** Such a universe is also called **open** because if you were to travel in a straight line in any direction, you would never return to your starting point.

Figure 26-15 illustrates the three cases of positive curvature, zero curvature, and negative curvature. Real space is three-dimensional, but we have drawn the three cases as analogous, more easily visualized two-dimensional surfaces. Therefore, as you examine the drawings in Figure 26-15, remember that the real universe has one more dimension. For example, if the universe is in fact hyperbolic, then the geometry of space must be the (difficult to visualize) three-dimensional analog of the two-dimensional surface of a saddle.

Note that in accordance with the cosmological principle, none of these models of the universe has an "edge" or a "center." This is clearly the case for both the flat and hyperbolic universes, because they are infinite and extend forever in all directions. A spherical universe is finite, but it also lacks a center and an edge.



INTERACTIVE EXERCISE 26.1

Figure 26-15

The Geometry of the Universe The curvature of the universe is either **(a)** positive, **(b)** zero, or **(c)** negative. The curvature depends on whether the combined mass density is greater than, equal to, or less than the critical density, or, equivalently, on whether the density parameter Ω_0 is greater than, equal to, or less than 1. In theory, the curvature could be determined by seeing whether two laser beams initially parallel to each other would converge, remain parallel, or spread apart.

You could walk forever around the surface of a sphere (like the surface of the Earth) without finding a center or an edge.

The curvature of the universe is determined by the value of the combined mass density ρ_0. If ρ_0 is greater than a certain value ρ_c (say "rho sub cee"), the universe has positive curvature and is closed. If ρ_0 is less than ρ_c, the universe has negative curvature and is open. In the special case that ρ_0 is exactly equal to ρ_c, the universe is flat. Because of its crucial role in determining the geometry of the universe, ρ_c is called the **critical density.** It is given by the expression

Critical density of the universe

$$\rho_c = \frac{3H_0^2}{8\pi G}$$

ρ_c = critical density of the universe

H_0 = Hubble constant

G = universal constant of gravitation

Using a Hubble constant H_0 = 73 km/s/Mpc, we get

$$\rho_c = 1.0 \times 10^{-26} \text{ kg/m}^3$$

A sample of hydrogen gas with this density would contain just 6 hydrogen atoms per cubic meter.

Many astronomers prefer to characterize the combined average mass density of the universe in terms of the **density parameter** Ω_0 (say "omega sub zero"). This is just the ratio of the combined average mass density to the critical density:

Density parameter

$$\Omega_0 = \frac{\rho_0}{\rho_c}$$

Ω_0 = density parameter

ρ_0 = combined average mass density

ρ_c = critical density

An open universe has a density parameter Ω_0 between 0 and 1, and a closed universe has Ω_0 greater than 1. In a flat universe, Ω_0 is equal to 1. Thus, we can use the value of Ω_0 as a measure of the curvature of the universe (Table 26-1).

Measuring the Cosmic Curvature

How can we determine the curvature of space across the universe? In theory, if you drew an enormous triangle whose sides were each a billion light-years long (see Figure 26-15), you could determine the curvature of space by measuring the three angles of the triangle. If their sum equaled 180°, space would be flat. If the sum was greater than 180°, space would be spherical. And if the sum of the three angles was less than 180°, space would be hyperbolic. Unfortunately, this direct method for measuring the curvature of space is not practical.

INTERACTIVE EXERCISE 26.1

Table 26-1 The Geometry and Average Density of the Universe

Geometry of space	Curvature of space	Type of universe	Combined average mass density (ρ_0)	Density parameter (Ω_0)
Spherical	positive	closed	$\rho_0 > \rho_c$	$\Omega_0 > 1$
Flat	zero	flat	$\rho_0 = \rho_c$	$\Omega_0 = 1$
Hyperbolic	negative	open	$\rho_0 < \rho_c$	$\Omega_0 < 1$

A way to determine the curvature of the universe that is both practical and precise is to see if light rays bend toward or away from each other, as shown in Figure 26-15. The greater the distance a pair of light rays has traveled, and, hence, the longer the time the light has been in flight, the more pronounced any such bending should be. Therefore, astronomers test for the presence of such bending by examining the oldest radiation in the universe: the cosmic microwave background.

If the cosmic microwave background were truly isotropic, so that equal amounts of radiation reached us from all directions in the sky, it would be impossible to tell whether individual light rays have been bent. However, as we saw in Section 26-5, there are localized "hot spots" in the cosmic microwave background due to density variations in the early universe. The apparent size of these hot spots depends on the curvature of the universe (**Figure 26-16**). If the universe is closed, the bending of light rays from a hot spot will make the spot appear larger (Figure 26-16*a*); if the universe is open, the light rays will bend the other way and the hot spots will appear smaller (Figure 26-16*c*). Only in a flat universe will the light rays travel along straight lines, so that the hot spots appear with their true size (Figure 26-16*b*).

By calculating what conditions were like in the primordial fireball, astrophysicists find that in a flat universe, the dominant "hot spots" in the cosmic background radiation should have an angular size of about 1°. (In Section 26-9 we will learn how this is deduced.) This is just what the BOOMERANG and MAXIMA experiments observed, and what the WMAP observations have confirmed (see Figure 26-14). Hence, the curvature of the universe must be very close to zero, and the universe must either be flat or very nearly so.

As Table 26-1 shows, once we know the curvature of the universe, we can determine the density parameter Ω_0 and hence the combined average mass density ρ_0. By analyzing the data shown in Figure 26-14, astrophysicists find that $\Omega_0 = 1.0$ with an uncertainty of about 2%. In other words, ρ_0 is within 2% of the critical density ρ_c.

The flatness of the universe poses a major dilemma. We saw in Section 26-5 that the average mass density of matter in the universe, ρ_m, is 2.4×10^{-27} kg/m^3. This is only 0.24 of the critical density ρ_c. We can express this in terms of the **matter density parameter** Ω_m (say "omega sub em"), equal to the ratio of ρ_m to the critical density:

$$\Omega_m = \frac{\rho_m}{\rho_c} = 0.24$$

If matter and radiation were all there is in the universe, the combined average mass density ρ_0 would be equal to ρ_m (plus a tiny contribution from radiation, which we can neglect because the average mass density of radiation is only about 0.02% that of matter). Then the density parameter Ω_0 would be equal to Ω_m—that is, equal to 0.24—and the universe would be open. But the temperature variations in the cosmic microwave background clearly show that the universe is either flat or very nearly so. These variations also show that the density parameter Ω_0, which includes the effects of *all* kinds of matter and energy, is equal to 1.0. In other words, radiation and matter, including dark matter, together account for only 24% of the total density of the universe! The dilemma is this: What could account for the rest?

Dark Energy

WEB LINK 26.12

The source of the missing density must be some form of energy that we cannot detect from its gravitational effects (the technique astronomers use to detect dark matter). It must also not emit detectable radiation of any kind. We refer to this mysterious energy as **dark energy**.

The geometry of space reveals that the universe is filled with dark energy

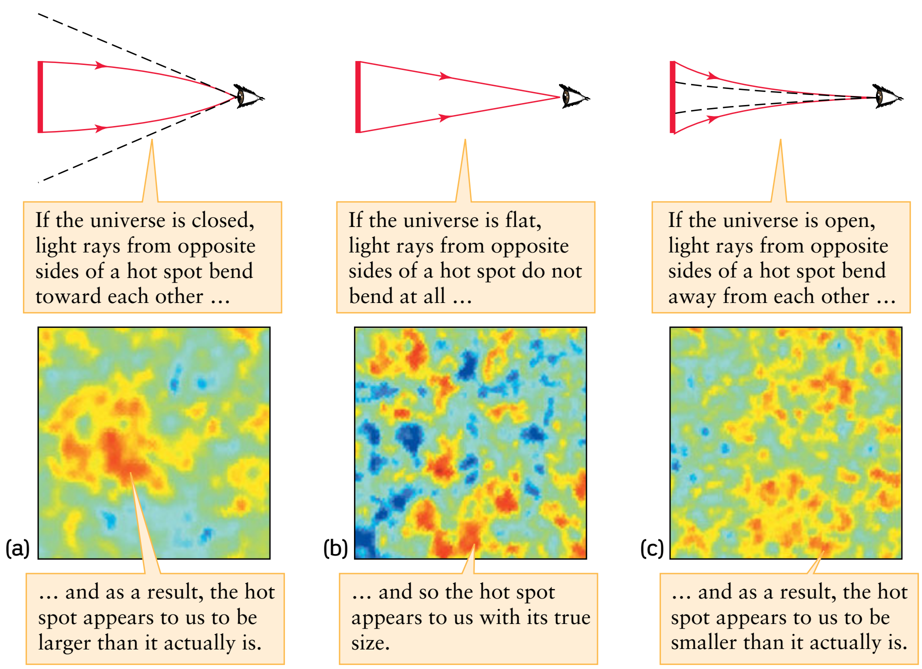


Figure 26-16 R I V U X G
The Cosmic Microwave Background and the Curvature of Space Temperature variations in the early universe appear as "hot spots" in the cosmic microwave background. The apparent size of these spots depends on the curvature of space. (The BOOMERANG Group, University of California, Santa Barbara)

Just as we express the average density of matter and radiation by the matter density parameter Ω_m, we can express the average density of dark energy in terms of the **dark energy density parameter** Ω_Λ (say "omega sub lambda"). This is equal to the average mass density of dark energy, ρ_Λ, divided by the critical density ρ_c:

$$\Omega_\Lambda = \frac{\rho_\Lambda}{\rho_c}$$

We can determine the value of Ω_Λ by noting that the combined average mass density ρ_0 must be the sum of the average mass densities of matter, radiation, and dark energy. As we have seen, the contribution of radiation is so small that we can ignore it, so we have

$$\rho_0 = \rho_m + \rho_\Lambda$$

If we divide this through by the critical density ρ_c, we obtain

$$\Omega_0 = \Omega_m + \Omega_\Lambda$$

That is, the density parameter Ω_0 is the sum of the matter density parameter Ω_m and the dark energy density parameter Ω_Λ. Solving for Ω_Λ, we find

$$\Omega_\Lambda = \Omega_0 - \Omega_m$$

Since Ω_0 is close to 1.0 and Ω_m is 0.24, we conclude that Ω_Λ must be $1.0 - 0.24 = 0.76$. Thus, whatever dark energy is, it accounts for 76% (about three-quarters) of the contents of the universe!

The concept of dark energy is actually due to Einstein. When he proposed the existence of a cosmological constant, he was suggesting that the universe is filled with a form of energy that by itself tends to make the universe expand (see Section 26-1). Unlike gravity, which tends to make objects attract, the energy associated with a cosmological constant would provide a form of "antigravity." Hence, it would not be detected in the same way as matter. (The subscript Λ in the symbol for the dark energy density parameter pays homage to the symbol that Einstein chose for the cosmological constant.)

If dark energy is in fact due to a cosmological constant, the value of this constant must be far larger than Einstein suggested. This is needed if we are to explain why Ω_Λ has a large value of 0.76. If Einstein felt he erred by introducing the idea of a cosmological constant, his error was giving it too small a value!

These ideas concerning dark energy are extraordinary, and extraordinary claims require extraordinary evidence to confirm them. As we will see in the next section, a crucial test is to examine how the rate of expansion of the universe has evolved over the eons.

26-7 Observations of distant supernovae reveal that we live in an accelerating universe

We have seen that the universe is expanding. But does the rate of expansion stay the same? Because there is matter in the universe, and because gravity tends to pull the bits of matter in the universe toward one another, we would expect that the expansion should slow down with time. (In the same way, a cannonball shot upward from the surface of the Earth will slow down as it ascends because of the Earth's gravitational pull.) If there is a cosmological constant, however, its associated dark energy will exert an outward pressure that tends to accelerate the expansion. Which of these effects is more important?

Modeling the Expansion History of the Universe

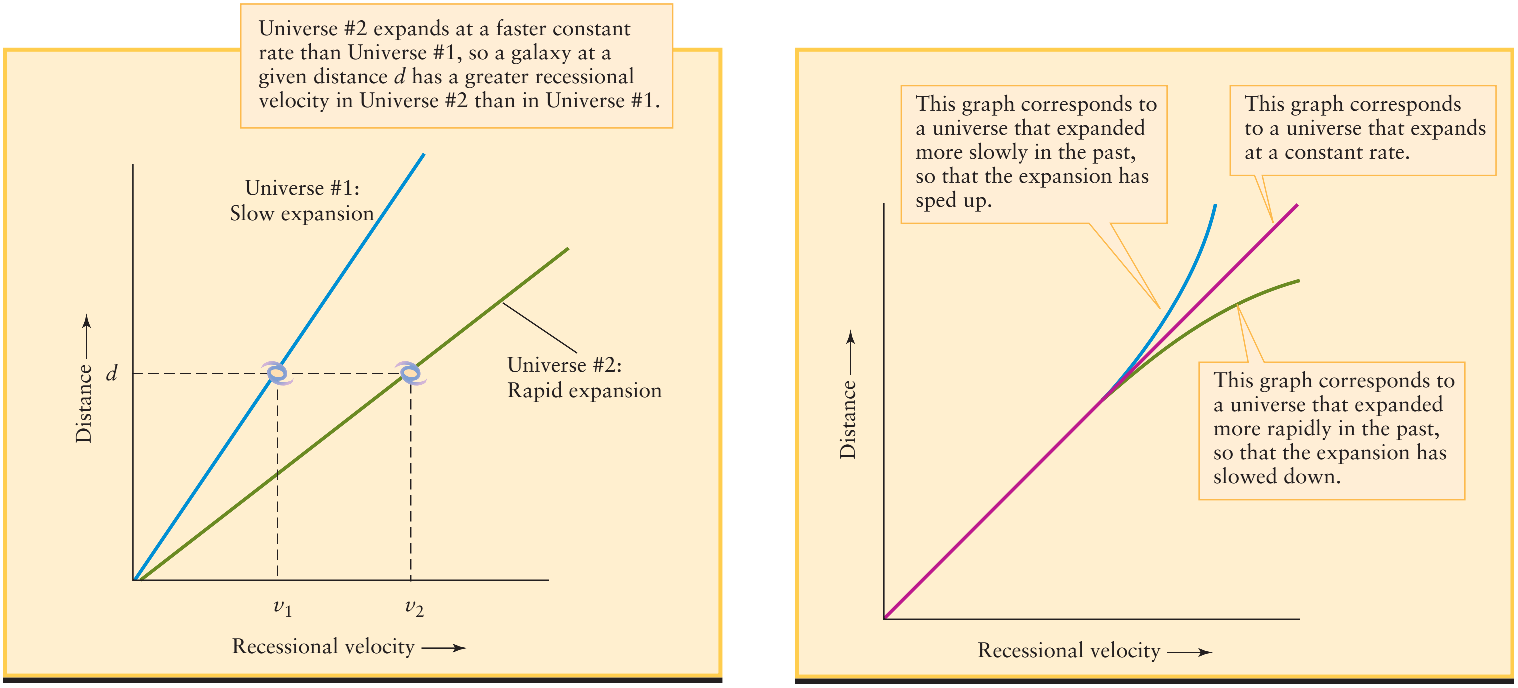
To answer this question, astronomers study the relationship between redshift and distance for extremely remote galaxies. We see these galaxies as they were billions of years ago. If the rate of expansion was the same in the distant past as it is now, the same Hubble law should apply to distant galaxies as to nearby ones. But if the rate of expansion has either increased or decreased, we will find important deviations from the Hubble law.

To see how astronomers approach this problem, first imagine two different parallel universes. Both Universe #1 and Universe #2 are expanding at constant rates, so for both universes there is a direct proportion between recessional velocity v and distance d as expressed by the Hubble law $v = H_0 d$. Hence, a graph of distance versus recessional velocity for either universe is a straight line, as Figure 26-17*a* shows. The only difference is that Universe #1 is expanding at a slower rate than Universe #2. Hence, a galaxy at a certain distance from Earth in Universe #1 will have a slower recessional velocity than a galaxy at the same distance from Earth in Universe #2. As a result, the graph of distance versus recessional velocity for slowly expanding Universe #1 (shown in blue) has a steeper slope than the graph for rapidly expanding Universe #2 (shown in green). Keep this observation in mind: A slower expansion means a steeper slope on a graph of distance versus recessional velocity.

Now consider *our* universe and allow for the possibility that the expansion rate may change over time. If we observe very remote galaxies, we are seeing them as they were in the remote past. If the expansion of the universe in the remote past was slower or faster than it is now, the slope of the graph of distance versus recessional velocity will be different for those remote galaxies. If the expansion was slower, then the slope will be steeper for distant galaxies (shown in blue in Figure 26-17*b*); if the expansion was faster, the slope will be shallower for distant galaxies (shown in green in Figure 26-17*b*). In either case, there will be a deviation from the straight-line Hubble law (shown in red in Figure 26-17*b*).

Measuring Ancient Expansion with Type Ia Supernovae

Which of the possibilities shown in Figure 26-17*b* represents the actual history of the expansion of our universe? In Section 24-5 we looked at the observed relationship between distance and recessional velocity for galaxies. Figure 24-17 is a plot of some representative data. The data points appear to lie along a straight line, suggesting that the rate of cosmological expansion has not changed. (Figure 24-17 is actually a graph of recessional velocity versus distance, not the other way around. But a straight line on one kind of graph will be a straight line on the other, because in either case there is a direct proportion between the two quantities being graphed.) However, the graph in Figure 24-17 was based on measurements of galaxies no farther than 400 Mpc



(a) Two universes with different expansion rates

(b) Possible expansion histories of the universe

Figure 26-17

Varying Rates of Cosmic Expansion **(a)** Imagine two universes, #1 and #2. Each expands at its own constant rate. For a galaxy at a given distance, the recessional velocity will be greater in the more rapidly expanding universe. Hence, the graph of distance *d* versus recessional velocity *v* will have a shallower slope for the rapidly expanding universe and a steeper slope for the slowly expanding one. **(b)** If the rate of expansion of our universe was more rapid in the distant past, corresponding to remote distances, the graph of *d* versus *v* will have a shallower slope for large distances (green curve). If the expansion rate was slower in the distant past, the graph will have a steeper slope for large distances (blue curve).

(1.3 billion ly) from the Earth, which means we are looking only 1.3 billion years into the past. The straightness of the line in Figure 24-17 means only that the expansion of the universe has been relatively constant over the past 1.3 billion years—only 10% of the age of the universe, and a relatively brief interval on the cosmic scale.

Now suppose that you were to measure the redshifts and distances of galaxies *several* billion light-years from the Earth. The light from these galaxies has taken billions of years to arrive at your telescope, so your measurements will reveal how fast the universe was expanding billions of years ago. To do this, we need a technique that will allow us to find the distances to these very remote galaxies. We saw in Section 24-4 that one way to do this is to identify Type Ia supernovae in such galaxies. These supernovae are among the most luminous objects in the universe, and hence can be detected even at extremely large distances (see Figure 24-14). The maximum brightness of a supernova tells astronomers its distance through the inverse-square law for light, and the redshift of the supernova's spectrum tells them its recessional velocity. As an example, the image that opens this chapter shows Type Ia supernovae with redshifts $z = 1.010$, 1.230, and 1.390, corresponding to recessional velocities of 60%, 67%, and 70% of the speed of light. We see these supernovae as they were 7.7 to 9.0 billion years ago, when the universe was less than half of its present age.

WEB LINK 26.13

In 1998, two large research groups—the Supernova Cosmology Project, led by Saul Perlmutter of Lawrence Berkeley National Laboratory, and the High-Z Supernova Search Team, led by Brian Schmidt of the Mount Stromlo and Siding Springs Observatories in Australia—reported their results from a survey of Type Ia supernovae in galaxies at redshifts of 0.2 or greater, corresponding to distances beyond 750 Mpc (2.4 billion ly). Figure 26-18 shows some of their data, along with more recent observations, on a graph of apparent magnitude versus redshift. A greater apparent magnitude corresponds to a dimmer supernova (see Section 17-3), which means that the supernova is more distant. A greater redshift implies a greater recessional velocity. Hence, this graph is basically the same as those in Figure 26-17.

To interpret these results, we need guidance from **relativistic cosmology.** This is a theoretical description of the universe and its expansion, based on Einstein's general theory of relativity and developed in the 1920s by Alexander Friedmann in Russia, Georges Lemaître in France, and Willem de Sitter in the Netherlands. Given values of the mass density parameter Ω_m and the dark energy density parameter Ω_Λ, cosmologists can use the equations of relativistic cosmology to predict how the expansion rate of the universe should change over time. Such predictions can be expressed as curves on a graph of distance versus redshift such as Figure 26-18.

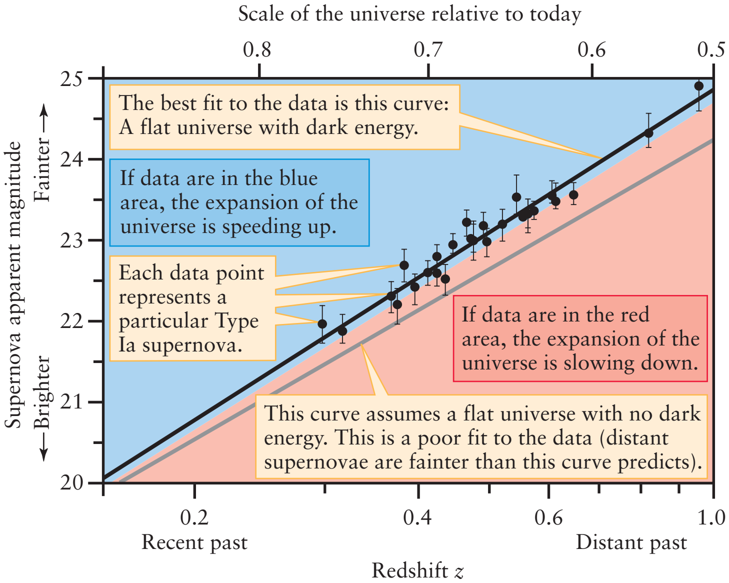


Figure 26-18

The Hubble Diagram for Distant Supernovae This graph shows apparent magnitude versus redshift for supernovae in distant galaxies. The greater the apparent magnitude, the dimmer the supernova and the greater the distance to it and its host galaxy. If the expansion of the universe is speeding up, the data will lie in the blue area; if it is slowing down, the data will lie in the red area. The data show that the expansion is in fact speeding up. (The Supernova Cosmology Project/R. A. Knop et al.)

The lower, gray curve in Figure 26-18 shows what would be expected in a flat universe with $\Omega_0 = 1.00$ but with no dark energy, so that $\Omega_\Lambda = 0$ and $\Omega_m = \Omega_0$ (that is, the density parameter is due to matter and radiation alone). In this model, and in fact in any model whose curve lies in the red area in Figure 26-18, the absence of dark energy means that gravitational attraction between galaxies would cause the expansion of the universe to slow down with time. Hence, the expansion rate would have been greater in the past (compare with the green curve in Figure 26-17*b*).

In fact, the data points in Figure 26-18 are almost all in the *blue* region of the graph, and agree very well with the curve shown in black. This curve also assumes a flat universe, but with an amount of dark energy consistent with the results from the cosmic microwave background ($\Omega_m = 0.24$, $\Omega_\Lambda = 0.76$, $\Omega_0 = \Omega_m + \Omega_\Lambda = 1.00$). In this model, and indeed in any model whose curve lies in the blue region of Figure 26-18, dark energy has made the expansion of the universe speed up over time. Hence, the expansion of the universe was slower in the distant past, which means that we live in an *accelerating* universe.

Just like the blue curve in Figure 26-17*b*, the data in Figure 26-18 show that supernovae of a certain brightness (and hence a given distance) have smaller redshifts (and hence smaller recessional velocities) than would be the case if the expansion rate had always been the same. These data provide compelling evidence of the existence of dark energy. (Professor Robert Kirshner's essay following this chapter discusses more about this remarkable discovery.)

Roughly speaking, the data in Figure 26-18 indicate the relative importance of dark energy (which tends to make the expansion speed up) and gravitational attraction between galaxies (which tends to make the attraction slow down). Thus, these data tell us about the *difference* between the values of the dark energy density parameter Ω_Λ and the matter density parameter Ω_m. By contrast, measurements of the cosmic microwave background (Section 26-6) give information about Ω_0, equal to the *sum* of Ω_Λ and Ω_m. Observations of galaxy clusters (Section 26-5) set limits on the value of Ω_m by itself. By combining these three very different kinds of observations as shown in Figure 26-19, we can set more stringent limits on both Ω_Λ and Ω_m.

Taken together and combined with other observations, all these data suggest the following values.

$$\Omega_m = 0.241 \pm 0.034$$

$$\Omega_\Lambda = 0.759 \pm 0.034$$

$$\Omega_0 = \Omega_m + \Omega_\Lambda = 1.02 \pm 0.02$$

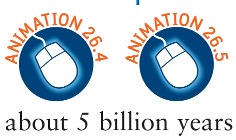
In each case, the number after the ± sign is the uncertainty in the value.

This collection of numbers points to a radically different model of the universe from what was suspected just a few years ago. In the 1980s there was no compelling evidence for an accelerating expansion of the universe, so it was widely assumed that $\Omega_\Lambda = 0$. Evidence from distant galaxies suggested a flat universe, so it was presumed that $\Omega_m = 1$. Figure 26-19 shows that modern data rule out this model.

Dark energy became the dominant form of energy in the universe about the same time that our solar system formed

The model we are left with is one in which the universe is suffused with a curious dark energy due to a cosmological constant. Unlike matter or radiation, whose average densities decrease as the universe expands and thins out, the average density of this dark energy remains constant throughout the history of the universe (Figure 26-20). The dark energy was relatively unimportant over most of the early history of the universe. Today, however, the density of dark energy is greater than that of matter (Ω_Λ is greater than Ω_m). In other words, we live in a **dark-energy-dominated universe.**

Cosmic Expansion: From Slowing Down to Speeding Up



As Figure 26-20 shows, the dominance of dark energy is a relatively recent development in the history of the universe. Prior to about 5 billion years ago, the density of matter should have been greater than that of dark energy. Hence, we would expect that up until about 5 billion years ago, the expansion of the universe should have been slowing down rather than speeding up. Recently, astronomers have found evidence of this picture by using the Hubble Space Telescope to observe extremely distant Type Ia supernovae with redshifts z greater than 1. (The image that opens this chapter shows three of these supernovae.) When astronomers compare the distance to these supernovae (determined from their brightness) to their redshifts, they find that the redshifts are *greater* than would be the case if the expansion of the universe had always been at the same rate or had always been speeding

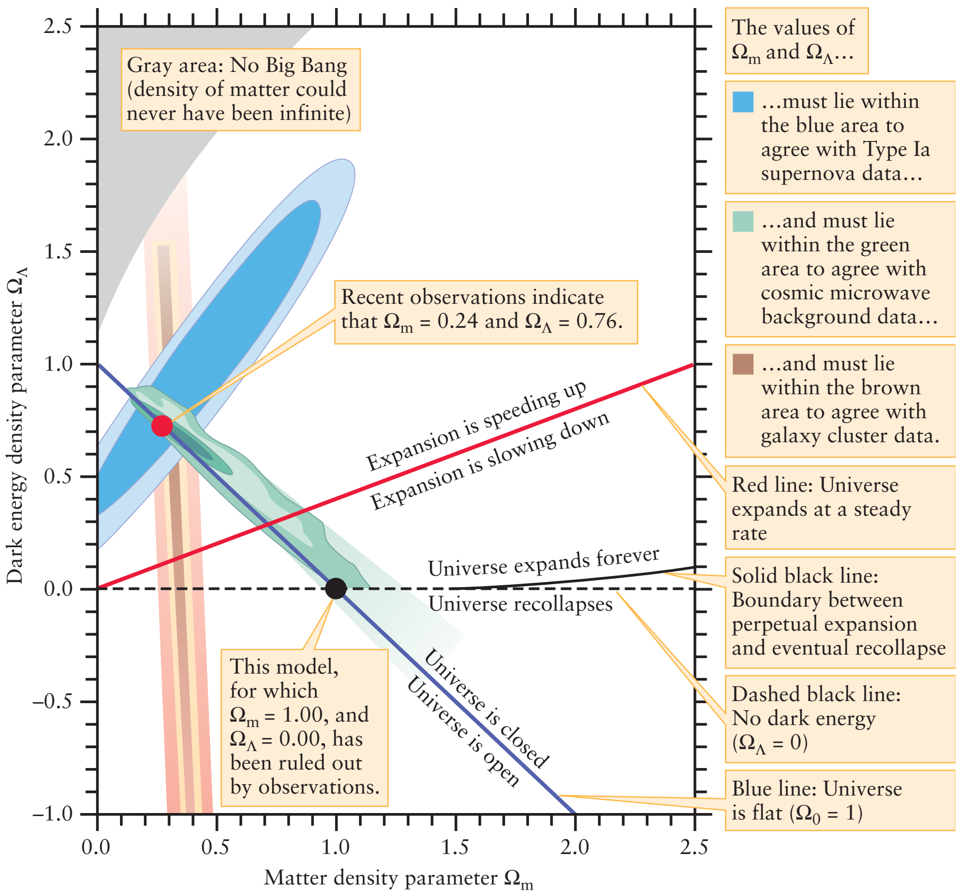
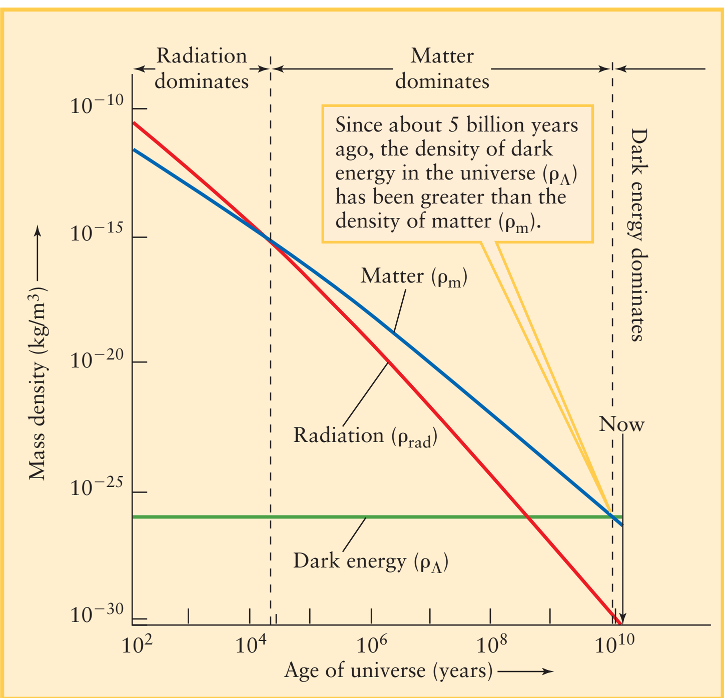


Figure 26-19

Limits on the Nature of the Universe The three regions on this graph show values of the mass density parameter Ω_m and the dark energy density parameter Ω_Λ that are consistent with various types of observations. Galaxy cluster measurements (in brown) set limits on Ω_m. Observations of the cosmic microwave background (in green) set limits on the sum of Ω_m and Ω_Λ: A larger value of Ω_m (to the right in the graph) implies a smaller value of Ω_Λ (downward in the graph) to keep the sum the same, which is why this band slopes downward. Observations of Type Ia supernovae (in blue) set limits on the difference between Ω_m and Ω_Λ; this band slopes upward since a larger value of Ω_m implies a larger value of Ω_Λ to keep the difference the same. The best agreement to all these observations is where all three regions overlap (the red dot). (The Supernova Cosmology Project/R. A. Knop et al.)

up (see Figure 26-17). This is just what would be expected if the expansion was slowing down in the very early universe. After about 5 billion years ago, the effects of dark energy became dominant and the expansion began to speed up.



ANALOGY If you see a red light up ahead while driving, you would probably apply the brakes to make the car slow down. But if the light then turns green before your car comes to a stop, and the road ahead is clear, you would step on the gas to make the car speed up again. The expansion of the universe has had a similar history. The mutual gravitational attraction of all the matter in the universe means that "the brakes were on" for about the first 9 billion years after the Big Bang, so that the expansion slowed down. But for about the past 5 billion years, dark energy has "had its foot on the gas," and the expansion has been speeding up.

If dark energy truly is a cosmological constant, the density of dark energy will continue to remain constant, as shown in Figure 26-20. Due to the effects of this dark energy, the universe will keep on expanding forever, and the rate of expansion will continue to accelerate. Eventually, some 30 billion years from now, the universe will have expanded so much that only a thousand or so of the nearest galaxies will still be visible. The billions of other

Figure 26-20

The Evolution of Density, Revisited The average mass density of matter, ρ_m, and the average mass density of radiation, ρ_{rad}, both decrease as the universe expands and becomes more tenuous. But if the dark energy is due to a cosmological constant, its average mass density ρ_Λ remains constant. In this model, our universe became dominated by dark energy about 5 billion years ago.

galaxies that we can observe today will have moved so far away from us that their light will have faded to invisibility. Furthermore, they will be moving away from us so rapidly that what light we do receive from them will have been redshifted out of the visible range.

There may be other explanations for dark energy besides a cosmological constant, however. Several physicists have proposed a type of dark energy whose density decreases slowly as the universe expands. Depending on how the density of dark energy evolves over time, the universe could continue to expand or could eventually recollapse on itself. Future observations, including space-based measurements of both the cosmic background radiation and of Type Ia supernovae, should help resolve the nature of the mysterious dark energy.

26-8 Primordial sound waves help reveal the character of the universe

We have seen how studying the “hot spots” in the cosmic background radiation reveals that we live in a flat universe. In fact, temperature variations reveal more: They give us a window on conditions in the early universe, and actually help us pin down the values of other important quantities such as the Hubble constant and the density of matter in the universe. The key to extracting this additional information from the cosmic background radiation is recognizing that the hot and cold spots in a map such as Figure 26-14 are actually sound waves.

Sound Waves in the Early Universe

Sound waves can travel in fluids of all kinds. Sound waves in air are used in human speech and hearing, while whales communicate using high-frequency underwater clicks and whistles. If you could take a snapshot of a sound wave, you would see that at any moment there are some regions, called **compressions**, where the fluid is squeezed together, and other regions, called **rarefactions**, where the fluid is thinned out or rarefied. (When a sound wave enters a human ear, the air next to the eardrum is alternately compressed and rarefied, which makes the eardrum flex back and forth. The ear detects this flexing and translates it into an electrical signal that is sent to the brain.) Even a seemingly quiet fluid, such as the still air inside a room, has sound waves in it that are triggered by random motions of the fluid. Such random sounds in still air are too faint to be detected by your ear. This is a good thing; it would be very annoying to hear a continuous background noise from the air!

Immense sound waves in the early universe left their imprint as variations in the cosmic microwave background

There would also have been sound waves in the early universe before recombination. During the first 380,000 years after the Big Bang, the universe was filled with a fluid composed primarily of photons, electrons, and protons, with a density more than 10^9 times greater than that of our present-day universe. Just as water molecules in a glass of water collide with each other, photons and particles collided frequently with each other in this primordial fluid, triggering random sound waves with compressions and rarefactions.

Because there was more mass in a compression than in a rarefaction, photons emerging from a compression experienced a greater gravitational redshift than did photons emerging from a rarefaction (see Section 22-2). As a result, while the light we see from either a compression or a rarefaction has a blackbody spectrum, the light from a compression is shifted to slightly longer wavelengths. We saw in Sections 5-3 and 5-4 that a blackbody spectrum dominated by longer wavelengths corresponds to a lower blackbody temperature (see Figure 5-11). Hence, we see compressions as the cold spots in Figure 26-14, and we see rarefactions as the hot spots. The overall pattern of cold and hot spots is thus a record of the sound waves that were present just as the universe became transparent, some 380,000 years after the Big Bang. As we discussed in Section 26-5, from the cold compressions arose our present-day population of galaxies.

The nature of a sound wave depends on the material through which it passes. For example, sound waves travel faster in helium than they do in air (because helium is less dense) and faster still in water (which, while denser than air, is much more resistant to compression). So by studying the primordial sound waves recorded in Figure 26-14, we can learn about the properties of the fluid that made up the early universe. These properties include the average densities of matter and dark energy in the fluid, as well as the value of the Hubble constant (which helps determine how rapidly the fluid was expanding and thinning out as the universe expanded). We can also determine the age of the universe at the time that the cosmic background radiation was emitted, since this determines the maximum size to which a hot spot (rarefaction) or cold spot (compression) could have grown since the Big Bang.

Figure 26-21 shows an important way in which astronomers systematize their data about hot and cold spots in the cosmic background radiation. This graph shows the number of observed hot or cold spots of different angular sizes, with larger spots on the left and smaller spots on the right. The presence of peaks in the graph shows that spots of certain sizes are more common than others. The largest peak tells us that the predominant angular size is about 1°, which corresponds to a region of compression or rarefaction that was about a million (10^6) light-years across at the time of recombination at $z = 1100$. (By contrast, the compressions and rarefactions in the sound waves most used in speech are a few meters across.) Since then the universe has expanded by a factor of about 1100, so that same region is now about a billion (10^9) light-years across.

Different cosmological models predict different shapes for the curve shown in Figure 26-21. Astronomers determine the best model by seeing which one gives a curve that best fits the data points. For example, the peak of the curve at an angular size of 1° is just what would be expected for a flat universe with $\Omega_0 = 1$. Table 26-2 summarizes the results of a flat-universe model that yields the particular curve shown in Figure 26-21.

Polarization of the Cosmic Microwave Background

Even more information can be obtained from the *polarization* of light in the cosmic microwave background. Ordinary light from

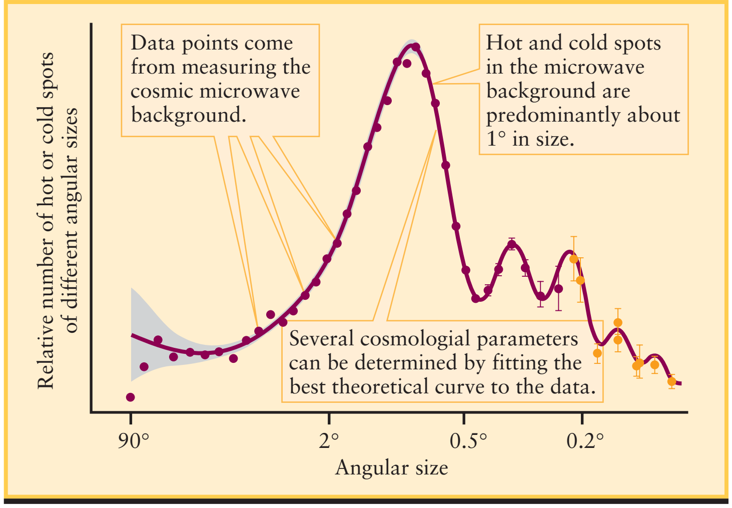


Figure 26-21

Sound Waves in the Early Universe Observations of the cosmic background radiation show that hot and cold spots of certain angular sizes are more common than others. A model that describes these observations helps to constrain the values of important cosmological parameters. Most of the data shown here is from the Wilkinson Microwave Anisotropy Probe; the data for the smallest angles (at the right of the graph) come from the CBI detector in the Chilean Andes and the ACBAR detector at the South Pole. (NASA/WMAP Science Team)

a lightbulb or from the Sun is *unpolarized,* which means that the electric fields of the light waves are oriented in random directions. But when light collides with and bounces off an object, it tends to become *polarized,* with its electric fields oriented in a specific direction. (As an example, the electric fields in sunlight reflected from the ground are mostly oriented horizontally. Polarizing sunglasses work by screening out electric fields of this orientation only, which helps eliminate distracting reflections.) In a similar way, the cosmic background radiation acquires a polarization when it scatters from material in a large hot spot. The amount of polarization turns out to be a very sensitive probe of conditions in the early universe, and, hence, of the nature of the universe itself.

In 2002 astronomers using the DASI microwave telescope at the South Pole reported the first detection of polarized light in the cosmic background radiation. Measurements made with the Wilkinson Microwave Anisotropy Probe have since provided additional information about polarization. Such measurements are very challenging to do. But the rewards for cosmology are very great, which is why researchers are devoting great effort to further studies of the polarization of the background radiation.

Our understanding of the universe as a whole has increased tremendously over the past several years. We have found compelling evidence that dark energy exists, and that it is the dominant form of energy in the universe. Studies of supernovae, galaxy clusters, and the cosmic background radiation have provided us with so much high-quality data that we can now express the key parameters of the universe (Table 26-2) with very high accuracy. When we look back to the situation in the 1980s, when the value of the Hubble constant was uncertain by at least 50%, it is no exaggeration to say that we have entered an age of *precision* cosmology.

Yet many questions remain unanswered. What is the nature of dark matter? What actually is dark energy? Can these mysterious entities be detected and studied in the laboratory? These and other questions will continue to occupy cosmologists for many years to come.

WEB LINK 26.16

Table 26-2 Some Key Properties of the Universe

Quantity	Significance	Value*
Hubble constant, H_0	Present-day expansion rate of the universe	$73.2^{+3.1}_{-3.2}$ km/s/Mpc
Density parameter, Ω_0	Combined mass density of all forms of matter *and* energy in the universe, divided by the critical density	1.02 ± 0.02
Matter density parameter, Ω_m	Combined mass density of all forms of matter in the universe, divided by the critical density	0.241 ± 0.034
Density parameter for ordinary matter, Ω_b	Mass density of ordinary atomic matter in the universe, divided by the critical density	0.0416 ± 0.001
Dark energy density parameter, Ω_Λ	Mass density of dark energy in the universe, divided by the critical density	0.759 ± 0.034
Age of the universe, T_0	Elapsed time from the Big Bang to the present day	$(1.373^{+0.016}_{-0.015}) \times 10^{10}$ years
Age of the universe at the time of recombination	Elapsed time from the Big Bang to when the universe became transparent, releasing the cosmic background radiation	$(3.79^{+0.08}_{-0.07}) \times 10^5$ years
Redshift z at the time of recombination	Since the cosmic background radiation was released, the universe has expanded by a factor $1 + z$	1089 ± 1

**Values for H_0, Ω_m, Ω_b, Ω_Λ, and T_0 are based on the three-year WMAP data with the assumption of a flat universe. Values for the time and redshift of recombination and for Ω_0 are from the first-year WMAP data. (NASA/WMAP Science Team)*

Key Words

average density of matter, p. 701
Big Bang, p. 697
closed universe, p. 705
combined average mass density, p. 705
compression, p. 712
cosmic background radiation, p. 699
cosmic microwave background, p. 699
cosmic light horizon, p. 698
cosmic singularity, p. 698
cosmological constant, p. 692
cosmological principle, p. 697
cosmological redshift, p. 694
cosmology, p. 691
critical density, p. 706
dark energy, p. 707
dark energy density parameter, p. 708
dark-energy-dominated universe, p. 710
density parameter, p. 706
era of recombination, p. 704
flat space, p. 705
homogeneous, p. 696
hyperbolic space, p. 705
isotropic, p. 696
lookback time, p. 696
mass density of radiation, p. 701
matter density parameter, p. 707
matter-dominated universe, p. 702
negative curvature, p. 705
observable universe, p. 698
Olbers's paradox, p. 692
open universe, p. 705
Planck time, p. 698
plasma, p. 704
positive curvature, p. 705
primordial fireball, p. 704
radiation-dominated universe, p. 702
rarefaction, p. 712
relativistic cosmology, p. 709
spherical space, p. 705
zero curvature, p. 705

Key Ideas

The Expansion of the Universe: The Hubble law describes the continuing expansion of space. The redshifts that we see from distant galaxies are caused by this expansion, not by the motions of galaxies through space.

• The redshift of a distant galaxy is a measure of the scale of the universe at the time the galaxy emitted its light.

• It is meaningless to speak of an edge or center to the universe or of what lies beyond the universe.

The Cosmological Principle: Cosmological theories are based on the idea that on large scales, the universe looks the same at all locations and in every direction.

The Big Bang: The universe began as an infinitely dense cosmic singularity that began its expansion in the event called the Big Bang, which can be described as the beginning of time.

• The observable universe extends about 14 billion light-years in every direction from the Earth. We cannot see objects beyond this distance because light from these objects has not had enough time to reach us.

• During the first 10^{-43} second after the Big Bang, the universe was too dense to be described by the known laws of physics.

Cosmic Background Radiation and the Evolution of the Universe: The cosmic microwave background radiation, corresponding to radiation from a blackbody at a temperature of nearly 3 K, is the greatly redshifted remnant of the hot universe as it existed about 380,000 years after the Big Bang.

• The background radiation was hotter and more intense in the past. During the first 380,000 years of the universe, radiation and matter formed an opaque plasma called the primordial fireball. When the temperature of the radiation fell below 3000 K, protons and electrons could combine to form hydrogen atoms and the universe became transparent.

• The abundance of helium in the universe is explained by the high temperatures in its early history.

The Geometry of the Universe: The curvature of the universe as a whole depends on how the combined average mass density ρ_0 compares to a critical density ρ_c.

• If ρ_0 is greater than ρ_c, the density parameter Ω_0 has a value greater than 1, the universe is closed, and space is spherical (with positive curvature).

• If ρ_0 is less than ρ_c, the density parameter Ω_0 has a value less than 1, the universe is open, and space is hyperbolic (with negative curvature).

• If ρ_0 is equal to ρ_c, the density parameter Ω_0 is equal to 1 and space is flat (with zero curvature).

Cosmological Parameters and Dark Energy: Observations of temperature variations in the cosmic microwave background indicate that the universe is flat or nearly so, with a combined average mass density equal to the critical density. Observations of galaxy clusters suggest that the average density of matter in the universe is about 0.24 of the critical density. The remaining contribution to the average density is called dark energy.

• Measurements of Type Ia supernovae in distant galaxies show that the expansion of the universe is speeding up. This may be due to the presence of dark energy in the form of a cosmological constant, which provides a pressure that pushes the universe outward.

Cosmological Parameters and Primordial Sound Waves: Temperature variations in the cosmic background radiation are a record of sound waves in the early universe. Studying the character of these sound waves, and the polarization of the background radiation that they produce, helps constrain models of the universe.

Questions

Review Questions

1. Why did Isaac Newton conclude that the universe was static? Was he correct?
2. What is Olbers's paradox? How can it be resolved?
3. What is a cosmological constant? Why did Einstein introduce it into cosmology?
4. What does it mean when astronomers say that we live in an expanding universe? What is actually expanding?
5. Describe how the expansion of the universe explains Hubble's law.
6. Would it be correct to say that due to the expansion of the universe, the Earth is larger today than it was 4.56 billion years ago? Why or why not?
7. Using a diagram, explain why the expansion of the universe as seen from a distant galaxy would look the same as seen from our Galaxy.

8. How does modern cosmology preclude the possibility of either a center or an edge to the universe?
9. Explain the difference between a Doppler shift and a cosmological redshift.
10. Explain how redshift can be used as a measure of lookback time. In what ways is it superior to time measured in years?
11. By what factor has the universe expanded since $z = 1$? Explain your reasoning.
12. What does it mean to say that the universe is homogeneous? That it is isotropic?
13. What is the cosmological principle? How is it justified?
14. How was the Big Bang different from an ordinary explosion? Where in the universe did it occur?
15. Some people refer to the Hubble constant as "the Hubble variable." In what sense is this justified?
16. What is meant by "the observable universe"?
17. (**a**) Explain why the radius of the observable universe is continually increasing. (**b**) Although the universe is 13.7 billion years old, the observable universe includes objects that are more than 13.7 billion light-years away from Earth. Explain why.
18. Imagine an astronomer living in a galaxy a billion light-years away. Is the observable universe for that astronomer the same as for an astronomer on Earth? Why or why not?
19. How did the abundance of helium in the universe suggest the existence of the cosmic background radiation?
20. Can you see the cosmic background radiation with the naked eye? With a visible-light telescope? Explain why or why not.
21. If the universe continues to expand forever, what will eventually become of the cosmic background radiation?
22. How can astronomers measure the average mass density of the universe?
23. What does it mean to say that the universe was once radiation-dominated? What happened when the universe changed from being radiation-dominated to being matter-dominated? When did this happen?
24. What was the era of recombination? What significant events occurred in the universe during this era? Was the universe matter-dominated or radiation-dominated during this era?
25. (**a**) Was there ever an era when the universe was radiation-dominated and matter and radiation were at the same temperature? If so, approximately when was this, and were there atoms during that era? If not, explain why not. (**b**) Was there ever an era when the universe was radiation-dominated and matter and radiation were *not* at the same temperature? If so, approximately when was this, and were there atoms during that era? If not, explain why not.
26. Describe two different ways in which the cosmic microwave background is not isotropic.
27. What is meant by the critical density of the universe? Why is this quantity important to cosmologists?
28. Describe how astronomers use the cosmic background radiation to determine the geometry of the universe.
29. Explain why it is important to measure how the expansion rate of the universe has changed over time. How is this rate measured?
30. What is dark energy? Describe two ways that we can infer its presence.
31. What does it mean to say that the universe is dark-energy–dominated? What happened when the universe changed from being matter-dominated to being dark-energy–dominated?
32. How can we detect the presence of sound waves in the early universe? What do these sound waves tell us?

Advanced Questions

Problem-solving tips and tools

We discussed Wien's law in Section 5-4. You may find it useful to recall that 1 parsec equals 3.26 light-years, that 1 Mpc equals 3.09×10^{19} km, and that a year contains 3.16×10^7 seconds.

33. (**a**) For what value of the redshift z were representative distances between galaxies only 20% as large as they are now? (**b**) Compared to representative distances between galaxies in the present-day universe, how large were such distances at $z = 8$? Compared to the density of matter in the present-day universe, what was the density of matter at $z = 8$? (**c**) If dark energy is in the form of a cosmological constant, how does its present-day density compare to the density of dark energy at $z = 2$? At $z = 5$? Explain your answers.
34. The host galaxy of the supernova HST04Sas (see the image that opens this chapter) has a redshift $z = 1.390$. The light from this galaxy includes the Lyman-alpha (L_α) spectral line of hydrogen, with an unshifted wavelength of 121.6 nm. Calculate the wavelength at which we detect the Lyman-alpha photons from this galaxy. In what part of the electromagnetic spectrum does this wavelength lie?
35. Estimate the age of the universe for a Hubble constant of (**a**) 50 km/s/Mpc, (**b**) 75 km/s/Mpc, and (**c**) 100 km/s/Mpc. On the basis of your answers, explain how the ages of globular clusters could be used to place a limit on the maximum value of the Hubble constant.
36. Some so-called "creation scientists" claim that the universe came into being about 6000 years ago. Find the value of the Hubble constant for such a universe. Is this a reasonable value for H_0? Explain.
37. The quasar HS 1946+7658 has a redshift $z = 3.02$. At the time when the light we see from HS 1946+7658 left the quasar, how many times more dense was the matter in the universe than it is today?
38. Use Wien's law (Section 5-4) to calculate the wavelength at which the cosmic microwave background ($T = 2.725$ K) is most intense.
39. If the mass density of radiation in the universe were 625 times larger than it is now, what would the background temperature be?
40. Suppose that the present-day temperature of the cosmic background radiation were somehow increased by a factor of 100, from 2.725 K to 272.5 K. (**a**) Calculate ρ_{rad} in this situation. (**b**) If the average density of matter (ρ_m) remained unchanged, would it be more accurate to describe our universe as matter-dominated or radiation-dominated? Explain.

41. Calculate the mass density of radiation (ρ_{rad}) in each of the following situations, and explain whether each situation is matter-dominated or radiation-dominated: (**a**) the photosphere of the Sun ($T = 5800$ K, $\rho_m = 3 \times 10^{-4}$ kg/m^3); (**b**) the center of the Sun ($T = 1.55 \times 10^7$ K, $\rho_m = 1.6 \times 10^5$ kg/m^3); (**c**) the solar corona ($T = 2 \times 10^6$ K, $\rho_m = 5 \times 10^{-13}$ kg/m^3).

42. If a photon from the cosmic microwave background had wavelength λ_0 when it was emitted at redshift z, its wavelength today is $\lambda = \lambda_0/(1 + z)$. (**a**) Let T be the symbol for the temperature of the cosmic microwave background today. Explain why the radiation temperature was $T_0 = T(1 + z)$ at redshift z. (**b**) What was the radiation temperature at $z = 1$? (**c**) At what redshift was the radiation temperature equal to 293 K (a typical room temperature)?

43. What would be the critical density of matter in the universe (ρ_c) if the value of the Hubble constant were (**a**) 50 km/s/Mpc? (**b**) 100 km/s/Mpc?

44. Consider the quasar HS 1946+7658 (see Advanced Question 37), which has $z = 3.02$. (**a**) Suppose that in the present-day universe, two clusters of galaxies are 500 Mpc apart. At the time that the light was emitted from HS 1946+7658 to produce an image on Earth tonight, how far apart were those two clusters? (**b**) What was the average density of matter (ρ_m) at that time? Assume that in today's universe, $\rho_m = 2.4 \times 10^{-27}$ kg/m^3. (**c**) What were the temperatures of the cosmic background radiation and the mass density of radiation (ρ_{rad}) at that time? (**d**) At this time in the remote past, was the universe matter-dominated, radiation-dominated, or dark-energy-dominated? Explain.

45. Whether the expansion of the universe is speeding up or slowing down can be expressed in term of a quantity called the *deceleration parameter,* denoted by q_0. The expansion is slowing down if q_0 is positive and speeding up if q_0 is negative; if $q_0 = 0$, the expansion proceeds at a constant rate. If we assume that the dark energy is due to a cosmological constant, the deceleration parameter can be calculated using the formula

$$q_0 = \frac{1}{2}\Omega_0 - \frac{3}{2}\Omega_\Lambda$$

(Recall that the density parameter Ω_0 is equal to $\Omega_m + \Omega_\Lambda$.) (**a**) Show that if there is no cosmological constant, the expansion of the universe must slow down. (**b**) Using the values of Ω_m and Ω_Λ given in Table 26-2, find the value of the deceleration parameter for our present-day universe. Based on this, is the expansion of the universe speeding up or slowing down? (**c**) Imagine a universe that has the same value of Ω_m as our universe but in which the expansion of the universe is neither speeding up nor slowing down. What would be the value of Ω_Λ in such a universe? Which would be dominant in such a universe, matter or dark energy? Explain.

46. In general, the deceleration parameter (see Advanced Question 45) is not constant but varies with time. For a flat universe, the deceleration parameter at a redshift z is given by the formula

$$q_z = \frac{1}{2} - \frac{3}{2}\left[\frac{\Omega_\Lambda}{\Omega_\Lambda + (1 - \Omega_\Lambda)(1 + z)^3}\right]$$

where Ω_Λ is the dark energy density parameter. Using the value of Ω_Λ given in Table 26-2, find the value of q_z for (**a**) $z = 0.5$ and (**b**) $z = 1.0$. (**c**) Explain how your results show that the expansion of the universe was actually decelerating at $z = 1.0$, but changed from deceleration to acceleration between $z = 1.0$ and $z = 0.5$.

47. The dark energy density parameter Ω_Λ is related to the value of the cosmological constant Λ by the formula

$$\Omega_\Lambda = \frac{\Lambda c^2}{3H_0^2}$$

where $c = 3.00 \times 10^8$ m/s is the speed of light. Determine the value of Λ if Ω_Λ and H_0 have the values given in Table 26-2. (*Hint:* You will need to convert units to eliminate kilometers and megaparsecs.)

Discussion Questions

48. Suppose we were living in a radiation-dominated universe. Discuss how such a universe would be different from what we now observe.

49. How can astronomers be certain that the cosmic microwave background fills the entire cosmos, not just the vicinity of the Earth?

50. Do you think there can be "other universes," regions of space and time that are not connected to our universe? Should astronomers be concerned with such possibilities? Why or why not?

Web/eBook Questions

51. Before the discovery of the cosmic microwave background, it seemed possible that we might be living in a "steady-state universe" with overall properties that do not change with time. The steady-state model, like the Big Bang model, assumes an expanding universe, but does not assume a "creation event." Instead, matter is assumed to be created continuously everywhere in space to ensure that the average density of the universe remains constant. Search the World Wide Web for information about the steady-state theory. Explain why the existence of the cosmic microwave background was a fatal blow to the steady-state theory.

52. Search the World Wide Web for information on a European Space Agency mission called Planck. In what ways is Planck an improvement over the WMAP mission? Has it been launched? If yes, what have scientists learned from Planck? If no, what do they hope plan to learn?

53. **Temperatures in the Early Universe.** Access the Active Integrated Media Module "Blackbody Curves" in Chapter 26 of the *Universe* Web site or eBook. (**a**) Use the module to determine by trial and error the temperature at which a blackbody spectrum has its peak at a wavelength of 1 μm. (**b**) At the time when the temperature of the cosmic background radiation was equal to the value you found in (a), was the universe matter-dominated or radiation-dominated? Explain your answer.

Activities

Observing Projects

54. Use the *Starry Night Enthusiast™* program to determine how the solar system moves through the cosmic microwave background. This motion appears to be taking us towards the constellation Leo. First, select **Favourites > Guides > Atlas** to display the entire celestial sphere from the center of a transparent Earth. Open the **Find** pane and click on the magnifying glass symbol to display the Find categories and click on **Constellation.** Double-click on Leo to center on this constellation and click again on the **Find** pane tab to close this pane and display the full screen. Select **View > Constellations > Astronomical** and **View > Constellations > Labels** to display and label the constellations. (**a**) Draw a sketch showing the Sun, the plane in which the Earth orbits the Sun, and the direction in which the solar system moves through the cosmic microwave background. (**b**) Use the date controls in the toolbar to step through the months of the year. In which month is the Sun placed most nearly in front of the Earth as the solar system travels through the cosmic background radiation?

55. Use *Starry Night Enthusiast™* to compare the distances of objects in the **Tully Database** with the radius of the Cosmic Light Horizon, the limit of our observable universe. As you will find, the most distant galaxies in this database are a long way away from the Earth and yet these distances are only a small fraction of the distances from which we can see light in our universe. Select **Favourites > Deep Space > Tully Database** to display this collection of galaxies in their correct 3-dimensional positions in space around our position. Stop **Time** and click on **View > Feet** to remove the image of the astronaut's suit from the view. Select **Preferences** from the **File** menu (Windows) or the **Starry Night Enthusiast** menu (Macintosh). In the Preferences dialog, select **Cursor Tracking (HUD)** in the drop-down box and ensure that **Distance from observer**, **Name** and **Object type** are selected. The view shows the boundaries of the Tully database as a cube. Use the **location scroller** (hold down the **Shift** key and mouse button while moving the mouse) to rotate the cube to allow you to choose galaxies on the outer fringes of this space. Use the **Hand Tool** to examine a selection of the furthest objects from the Earth, which is centered in the view, and write a list of 10–20 objects, noting the **Object type** and **Distance from observer.** (**a**) In your sample, is there a predominance of any one kind of galaxy? If so, what type of galaxy appears to be most common at these distances? (**b**) Select the furthest of these galaxies and compare their distances with the radius of the cosmic light horizon. What fraction of the radius of the observable universe is covered by the Tully database?

Collaborative Exercises

56. As a group, create a four- to six-panel cartoon strip showing a discussion between two individuals describing why the sky is dark at night.

57. Imagine your firm, Creative Cosmologists Coalition, has been hired to create a three-panel, folded brochure describing the principal observations that astronomers use to infer the existence of a Big Bang. Create this brochure on an $8\frac{1}{2} \times 11$ piece of paper. Be sure each member of your group supervises the development of a different portion of the brochure and that the small print acknowledges who in your group was primarily responsible for which portion.

58. The three potential geometries of the universe are shown in Figure 26-15. To demonstrate this, ask one member of your group to hold a piece of paper in one of the positions while another member draws two parallel lines that never change in one geometry, eventually cross in another geometry, and eventually diverge in another.

The Extravagant Universe

Robert Kirshner

Exploding stars halfway across the observable universe reveal a surprising fact. Judging the distances to distant supernovae from their apparent brightness, the rate of cosmic expansion has been speeding up in the last 5 billion years. While gravitation acts to slow cosmic expansion, these observations require something else to make the universe accelerate. We call this the "dark energy," though, in truth, we do not know what it is. Perhaps it is the modern version of Einstein's notorious cosmological constant.

This result was a big surprise to people working on the problem. Early in 1998, I wrote an e-mail to the members of our High-Z Supernova Team, saying, "In your heart you know this is wrong, though your head tells you [that] you don't care and you're just reporting the observations." One reason for our reluctance was a contrary result. The Supernova Cosmology Project, based at the Lawrence Berkeley Lab, had published a paper in the *Astrophysical Journal* in July 1997 claiming that distant supernovae showed the universe was slowing down. Our methods weren't that different, but our results were pointing the other way. We couldn't both be right.

Another reason to be wary was Einstein's bad experience with this idea, invented to make a static universe. Did we think we were smarter than Einstein? Einstein never liked the cosmological constant, as he wrote, "I am unable to believe that such an ugly thing should be realized in nature." It became a kind of theoretical poison ivy, touched only by the unwary for about 65 years. But the data were leading us to reconsider. As my fellow team member Adam Riess wrote to the rest of us, "Approach these results not with your heart or your head, but with your eyes. We are observers, after all!"

By 1998, both teams saw the signature of cosmic acceleration in supernovae at redshift 0.5: light that had been en route for about 5 billion years. Now we're using the Hubble Space Telescope to search for even fainter and more distant supernovae at redshift 1.5, roughly 9 billion light-years away, where we see signs of deceleration caused by denser dark matter that was winning the cosmic tug-of-war in the early history of cosmic expansion.

Robert Kirshner is Harvard College Professor and Clowes Professor of Science at Harvard University in Cambridge, Massachusetts. Born in Sudbury, Massachusetts, in 1949, he obtained his A.B. degree from Harvard College in 1970 and his Ph.D. from the California Institute of Technology in 1975. Kirshner was elected to the National Academy of Science in 1998 and elected president of the American Astronomical Society in 2003. At Harvard, he teaches a large course for students who are not science concentrators, entitled "The Energetic Universe." He is also master of Quincy House, one of Harvard's undergraduate residences. In 2003, Princeton University Press published Kirshner's popular-level book *The Extravagant Universe: Exploding Stars, Dark Energy, and the Accelerating Cosmos.*

Technology is a lot better now than it was just a decade ago: We've trained faster computers to scan bigger megapixel digital images of galaxies for us, to pick out the objects that change from one night to the next, and to pop out a list of supernova candidates just hours after we take the images. Speed is important because supernovae are like fish—after a few days, they lose their freshness. If you want to see the peak of a distant supernova's light curve, prompt action is essential.

Energetic supernova observations by us and by others have strengthened the case for an accelerating universe by measuring the distance and the redshift to hundreds of new objects, both near and far. Converging lines of evidence from other directions, including measurements of galaxy clustering and ripples in the glow from the Big Bang, all point to the same strange picture of a runaway universe that is ~70% dark energy, 25% dark matter, with just a few percent of the universe made of ordinary neutrons and protons, like the Galaxy, the Sun, the Earth, and ourselves.

In the future, telescopes scanning the whole sky and specialized satellites will give us a better chance to pin down the nature of the dark energy. Is it really Einstein's cosmological constant, retrieved from the dumpster of history, smoothed out and made new again? Or is it some more general "quintessence" whose energy density changes over time? We're doing better measurements to show whether the dark energy is constant, or changes subtly as the universe expands. Either way will be very interesting. So far, everything is consistent with a cosmological constant, but we should be ready for surprises ahead. The dark energy is important because it points to a deep mystery right at the heart of physics—the thorny problem of connecting gravitation with the quantum world. Will the solution come from string theory, extra dimensions of space, or from modifications to Einstein's general relativity? We don't know, and at this moment, only astronomical observations can shed light on the dark energy.

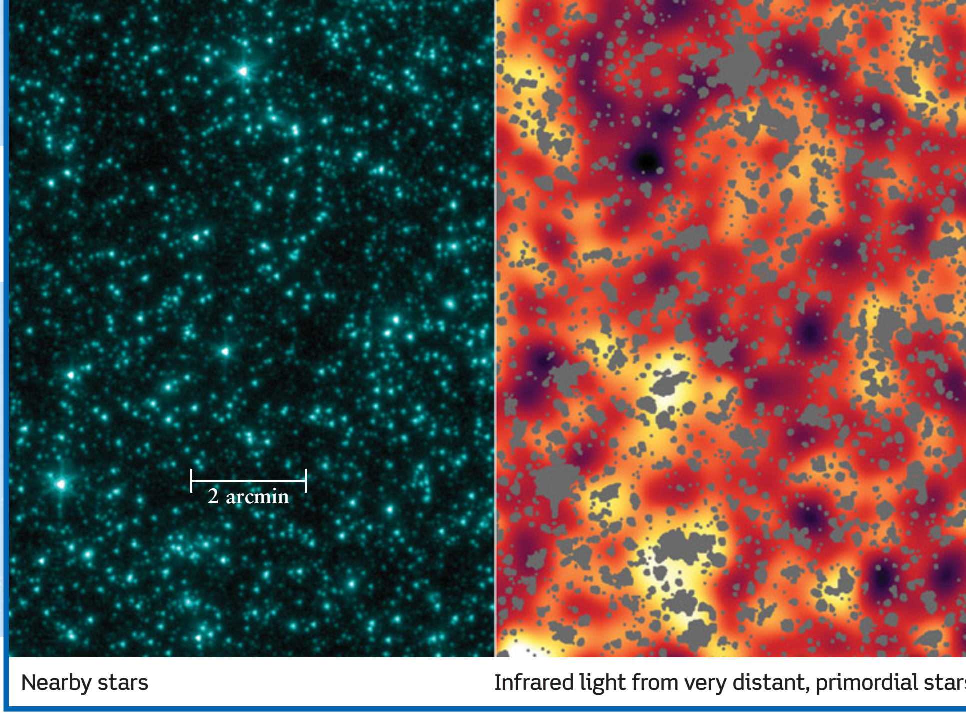
Faint light from distant stellar catastrophes traces the history of cosmic expansion. It is not what we expected to see. The universe contains more parts than the simplest universe we could imagine: atoms that glow, atoms that don't, neutrinos with mass, and another dark matter particle with more mass, something that made the universe expand exponentially in the era of inflation, and something more that is now making the universe accelerate. Perhaps some day in the future all of this will seem essential, but at the moment, it seems we live in a recklessly extravagant universe, with extra parts whose function we do not yet fathom.

When people ask what science is for, the answer is often framed in terms of economic development, or national security, or improved life span. Nobody is against being rich, or safe, or healthy, but there is something deeper at work when we study how the universe began and where we are going. This is research we do for the joy of finding out how the world works.

This page intentionally left blank

27

Exploring the Early Universe



R I V U X G

These Spitzer Space Telescope images show (left) light from nearby stars and (right) light from a remote population of ancient stars in the same patch of sky. (NASA; JPL-Caltech; and A. Kashlinsky, Goddard Space Flight Center)

The two images shown here are of a patch of sky near the Big Dipper in the constellation Ursa Major. The left-hand image is dominated by relatively nearby stars. But when these are removed digitally, what remains in the right-hand, false-color image is an intriguing pattern of highly redshifted light from objects much farther away. This is thought to be some of the oldest starlight in the universe: It was emitted by members of the very first generation of stars, born when our universe was less than a billion years old.

Only recently have new telescopes begun to reveal the story of the first stars and first galaxies. But no telescope can ever hope to directly observe events during the first 380,000 years after the Big Bang, when the universe was so opaque that it blocked the free passage of light. We can nonetheless reconstruct some of the events of that hidden epoch, because many aspects of today's universe are relics of the earliest events in the cosmos.

In this chapter we will see evidence that during the first miniscule fraction of a second after the Big Bang, the universe inflated in size by a stupendously large factor of about 10^{50}. During the next 15 minutes after inflation came to an end, the universe was so dense and hot that particles were constantly colliding at high speeds. As we will see, the events of those 15 violent minutes set the stage for all that came afterward—from the formation of atoms 380,000 years later, to the appearance of the first stars and galaxies some 400 million years after the Big Bang, down to the diverse present-day universe of which we are part.

27-1 The newborn universe underwent a brief period of vigorous expansion

With the discovery of the cosmic microwave background, astronomers had direct evidence that the universe began with a hot Big Bang (see Section 26-4). Remarkably, the microwave background is incredibly uniform, or *isotropic,* across the sky. If we subtract the effects of our own motion through the microwave background (see Figure 26-8), we find that the temperature of the microwave background is the same in all parts of the sky to an accuracy of 1 part in 10,000. The small nonuniformities in the microwave background are also remarkable, in part because their angular sizes help indicate that our universe has a flat geometry (see Section 26-6, especially Figure 26-16). As striking as these observations are, they pose substantial challenges to the standard theory of the expansion of the universe.

Learning Goals

By reading the sections of this chapter, you will learn

27-1 How the very young universe expanded enormously in a brief instant of time

27-2 How the fundamental forces of nature and the properties of empty space changed during the first second after the Big Bang

27-3 How the physics of subatomic particles affected the evolution of the early universe

27-4 Why antimatter was once common in the universe, but is very rare today

27-5 Which chemical elements in today's universe are remnants of the primordial fireball

27-6 How the first stars and galaxies formed in the early universe

27-7 What steps scientists are taking in the quest toward an all-encompassing "theory of everything"

The Isotropy Problem: Why Is the Microwave Background So Uniform?

To appreciate why the uniformity of the microwave background poses a problem, think about two opposite parts of the sky, labeled *A* and *B* in Figure 27-1. Both of these points lie on a sphere centered on the Earth, called our **cosmic light horizon** (see Section 26-3, especially Figure 26-5). We can see light coming from any object on or inside our cosmic light horizon, but we cannot yet see objects beyond this horizon. Even traveling at the speed of light—the ultimate speed limit—no information of any kind from objects beyond our cosmic light horizon has had time to reach us over the entire 13.7-billion-year history of the universe. Thus, the cosmic light horizon defines the limits of our observable universe. (As time passes, our cosmic light horizon expands, so eventually we will receive light from objects that are beyond the present-day horizon.)

Points *A* and *B* in Figure 27-1 lie just on our cosmic light horizon, so when we look at these points we are seeing as far back into the past as possible—that is, the light we see from these points is the cosmic background radiation. In order for the radiation that reaches us from *A* and *B* to be nearly the same, the material of the early universe at *A* must have had the same temperature as at *B*. But for two objects to be at the same temperature, they should have been in contact and able to exchange heat with each other. (A hot cup of coffee and a cold spoonful of cream reach the same temperature only after you pour the cream into the coffee.)

The problem is that the widely separated regions at *A* and *B* have absolutely no connection with each other. As Figure 27-1 shows, point *A* lies outside the cosmic light horizon for point *B*, and point *B* lies outside the cosmic light horizon for point *A*. (Put another way, point *A* lies outside the observable universe of point *B* and vice versa.) Hence, no information has had time to travel between points *A* and *B* over the entire history of the universe. How is it possible, then, that these unrelated parts of the universe have almost exactly the same temperature? This dilemma is called the **isotropy problem** or the **horizon problem.**

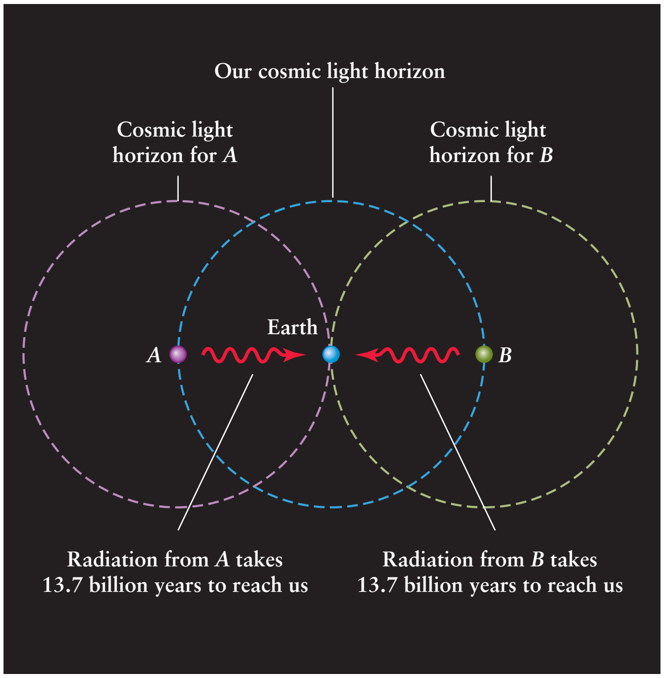


Figure 27-1

The Isotropy Problem Regions *A* and *B*, both of which lie on our cosmic light horizon, are so far apart that they seem never to have been in communication over the lifetime of the universe. Yet the cosmic background radiation from *A* and *B*, and from all other parts of the sky, shows that these disconnected regions have almost exactly the same temperature. The dilemma of why this should be is called the isotropy problem.

The Flatness Problem: Why Is $\Omega_0 = 1$?

The flatness of our universe presents us with a second enigma. Recall that the geometry of our universe depends on the density parameter Ω_0, which is the ratio of the combined average mass density in the universe (ρ_0) to the critical density (ρ_c) (see Section 26-6). Observations of temperature variations in the cosmic microwave background indicate that Ω_0 is very close to 1, which corresponds to a flat universe.

For the density parameter Ω_0 to be close to 1 today, it must have been *extremely* close to 1 during the Big Bang. In other words, the density of the early universe was almost exactly equal to the critical density. (The density was much higher than it is today, but the value of the critical density was also much higher. In a flat universe, the average mass density and the critical density decrease together as the universe expands, so that they remain equal at all times.)

The equations for an expanding universe show that any deviation from exact equality would have mushroomed within a fraction of a second. Had the average mass density been slightly less than ρ_c, the universe would have expanded so rapidly that matter could never have clumped together to form galaxies. Without galaxies, there would be no stars or planets, and humans would never have evolved. If, on the other hand, the density had been slightly greater than ρ_c, the universe would soon have become tightly packed with matter. Had this been the case, the gravitational attraction of this matter would long ago have collapsed the entire cosmos in a reversed Big Bang or "Big Crunch," and again humans would never have evolved.

In other words, immediately after the Big Bang the fate of the universe hung in the balance. The tiniest deviation from the precise equality $\rho_0 = \rho_c$ would have rapidly propelled the universe away from the special case of $\Omega_0 = 1$. Had there been such a deviation, we would not be here to contemplate the nature of the universe.

Happily, our universe is one in which galaxies, stars, planets, and humans do exist, the Big Crunch has not taken place, and Ω_0 is close to 1. This tells us that the density of the universe immediately after the Big Bang must have been equal to the critical density to an incredibly high order of precision. In order for space to be as flat as it is today, the value of ρ_0 right after the Big Bang must have been equal to ρ_c to more than 50 decimal places!

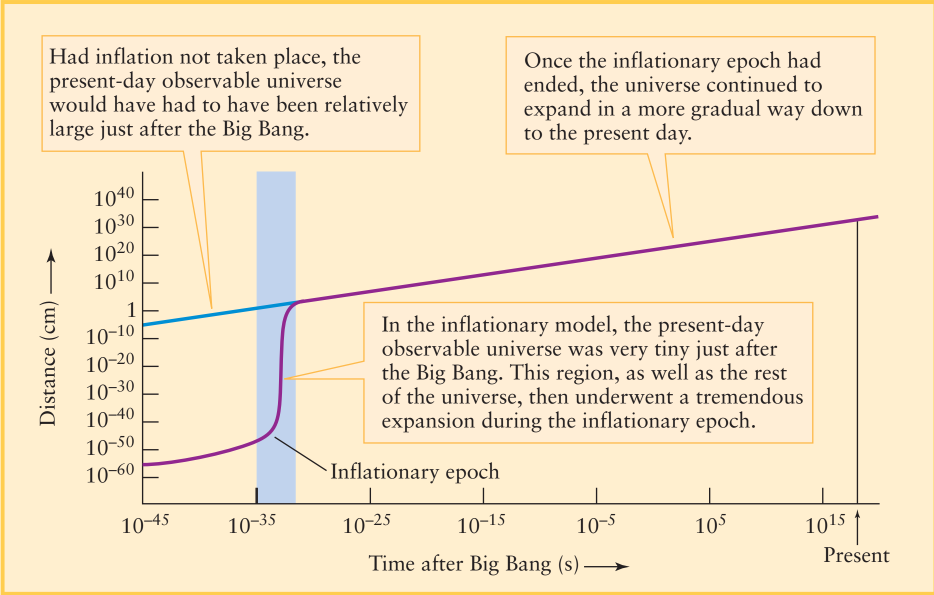


Figure 27-2

The Observable Universe With and Without Inflation According to the inflationary model (shown in red), the universe expanded by a factor of about 10^{50} shortly after the Big Bang. This growth in the size of the present-day observable universe—that portion of the universe that lies within our present cosmic light horizon—occurred during a very brief interval, as indicated by the vertical shaded area on the graph. The blue line shows the projected size of the present-day observable universe soon after the Big Bang if inflation had not taken place. (Adapted from A. Guth)

What could have happened during the first few moments of the universe to ensure that $\rho_0 = \rho_c$ to such an astounding degree of accuracy? Because $\rho_0 = \rho_c$ means that space is flat, this enigma is called the **flatness problem.**

Solving the Problems: The Inflationary Model

WEB LINK 27.2

In the early 1980s, a remarkable solution was proposed to both the isotropy problem and the flatness problem. Working independently, Alexei Starobinsky at the L. D. Landau Institute of Theoretical Physics in Moscow and Alan Guth at Stanford University suggested that the universe might have experienced a brief period of **inflation,** or extremely rapid expansion, shortly after the Planck time. (As we saw in Section 26-3, before the Planck time the normal laws of physics do not properly describe the behavior of space, time, and matter.) During this **inflationary epoch,** the universe expanded outward in all directions by a factor of about 10^{50}. This epoch of dramatic expansion may have lasted only about 10^{-32} s (Figure 27-2).

The inflationary model explains not only why the cosmic microwave background is so uniform, but also how stars, galaxies, and humans can exist

Inflation accounts for the isotropy of the microwave background. During the inflationary epoch, much of the material that was originally near our location was moved out to tremendous

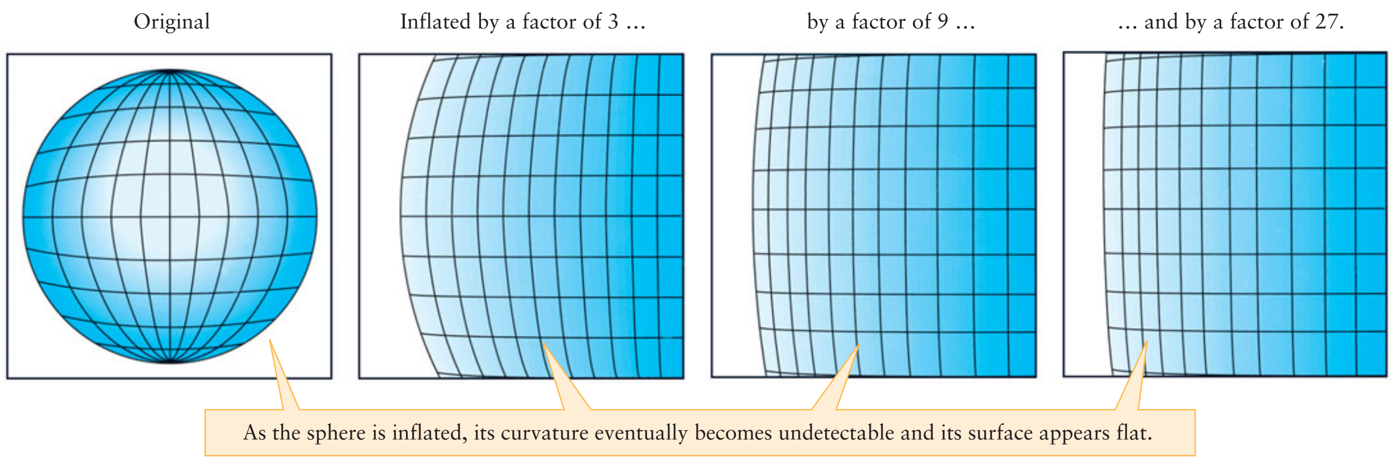


Figure 27-3

Inflation Solves the Flatness Problem This sequence of drawings shows how inflation can produce a locally flat geometry. In each successive frame, the sphere is inflated by a factor of 3 (and the number of grid lines on the sphere is increased by the same factor). Note how the curvature of the surface quickly becomes undetectable on the scale of the illustration. (Adapted from A. Guth and P. Steinhardt)

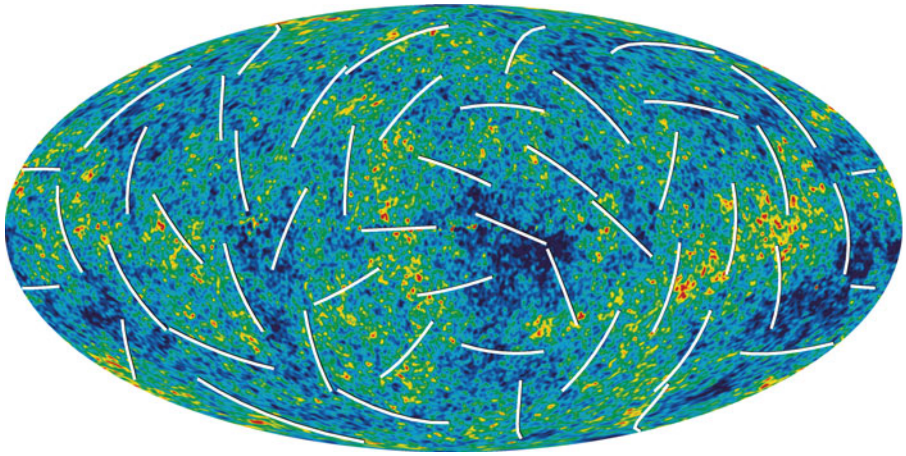


Figure 27-4 R I V U X G

Polarization of the Cosmic Microwave Background
The white lines on this map of the cosmic background radiation (compare Figure 26-14) indicate the directions in which radiation from different parts of the sky is polarized. This polarization was caused when photons in the background radiation scattered from electrons in the early universe. The distribution and density of these electrons depends on the earlier expansion history of the universe. Hence, polarization measurements can reveal whether inflation took place as well as the nature of the inflation. (NASA/WMAP Science Team)

distances. Over the past 13.7 billion years, our cosmic light horizon has expanded so that we can see radiation from these distant regions. Hence, when we examine microwaves from opposite parts of the sky, we are seeing radiation from parts of the universe that were originally in intimate contact with one another. This common origin is why all parts of the sky have almost exactly the same temperature.

ANALOGY An inflationary epoch can also account for the flatness of the universe. To see why, think about a small portion of the Earth's surface, such as your backyard. For all practical purposes, it is impossible to detect the Earth's curvature over such a small area, and your backyard looks flat. Similarly, the observable universe is such a tiny fraction of the entire inflated universe that any overall curvature in it is virtually undetectable (Figure 27-3). Like your backyard, the segment of space we can observe looks flat.

CAUTION! It is important to note that the concept of inflation does not violate Einstein's dictum that nothing can travel faster than the speed of light. Remember that the expansion of the universe is the expansion of space itself, not the motion of objects through space. During the inflationary epoch, the distances between particles increased enormously, but this was entirely the result of a sudden vigorous *expansion of space*. Particles did not move through space; space itself inflated.

As we will see later in this chapter, the inflationary model helps us to understand a number of other aspects of the universe. This is an important reason why many astronomers have confidence that inflation really did take place. However, other cosmological models have been proposed to explain the isotropy and flatness problems. It turns out that a stringent test on such models, as well as on different types of inflationary models, is what they predict about the polarization of the cosmic background radiation. (We discussed this polarization in Section 26-8.) Figure 27-4 shows measurements of the background radiation's polarization from the Wilkinson Microwave Anisotropy Probe. While such measurements are very difficult and still in their infancy, the data obtained to date are entirely consistent with the inflationary model.

27-2 Inflation was one of several profound changes that occurred in the very early universe

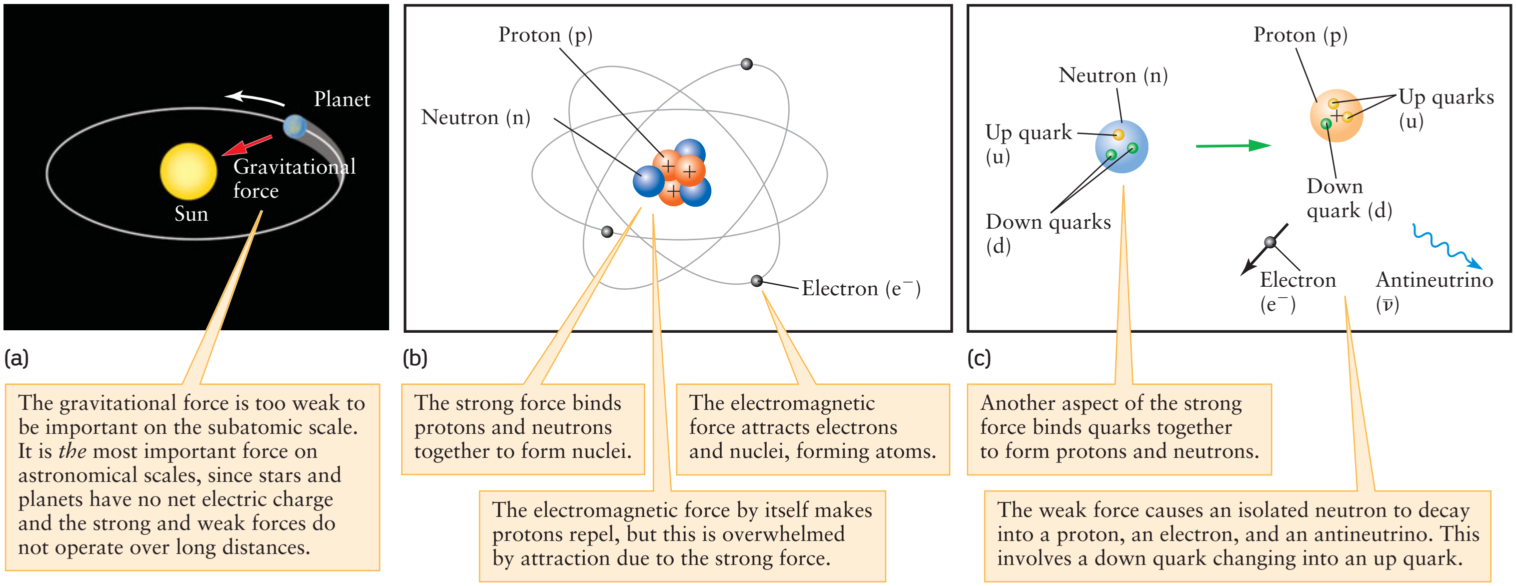
If the universe went through an episode of extreme inflation, what could have triggered it? Our understanding is that inflation was one of a sequence of remarkable events during the first 10^{-12} second after the Big Bang. In each of these events there was a fundamental transformation of the basic physical properties of the universe. To understand what happened during that brief moment of time, when the universe was a hot, dense sea of fast-moving particles and energetic photons colliding with each other, we must first understand how particles interact at very high energies.

The Fundamental Forces of Nature

Just *four* fundamental forces—gravitation, electromagnetism, and the strong and weak forces—explain the interactions of everything in the universe. Of these forces, gravitation is the most familiar (Figure 27-5*a*). It is a long-range force that dominates the universe over astronomical distances. The electromagnetic force is also a long-range force, but it is intrinsically much stronger than the gravitational force. For example, the electromagnetic force between an electron and a proton is about 10^{39} times stronger than the gravitational force between those two particles. That is why the electromagnetic force, not the gravitational force, is responsible for holding electrons in orbit about the nuclei in atoms.

We do not generally observe longer-range effects of the electromagnetic force, because there is usually a negative electric charge for every positive charge and a south magnetic pole for every north magnetic pole. Thus, over great volumes of space the effects of electromagnetism effectively cancel out. No similar canceling occurs with gravity because there is no equivalent "negative mass." This explains why the force that holds the Earth in orbit around the Sun is gravitational, not electromagnetic.

The **strong force** holds protons and neutrons together inside the nuclei of atoms (Figure 27-5*b*). It is said to be a *short-range force:* Its influence extends only over distances less than about 10^{-15} m, about the diameter of a proton. Without the strong force, nuclei would disintegrate because of the electromagnetic



Force	Relative strength	Particles exchanged	Particles on which the force can act	Range	Example
Strong	1	gluons	quarks	10^{-15} m	holding protons, neutrons, and nuclei together
Electromagnetic	1/137	photons	charged particles	infinite	holding atoms together
Weak	10^{-4}	intermediate vector bosons	quarks, electrons, neutrinos	10^{-16} m	radioactive decay
Gravitational	6×10^{-39}	gravitons	everything	infinite	holding the solar system together

INTERACTIVE EXERCISE 27.1

Figure 27-5

The Four Forces **(a)** Gravitation is dominant on the scales of planets, star systems, and galaxies, while **(b)**, **(c)** the strong, electromagnetic, and weak forces hold sway on the scale of atoms and nuclei.

repulsion of the positively charged protons. In fact, the strong force overpowers the electromagnetic forces inside nuclei.

The **weak force,** which also is a short-range force, is at work in certain kinds of radioactive decay (Figure 27-5*c*). An example is the transformation of a neutron (n) into a proton (p), in which an electron (e^-) is released along with a nearly massless particle called an antineutrino ($\bar{\nu}$):

$$n \rightarrow p + e^- + \bar{\nu}$$

Numerous experiments in nuclear physics demonstrate that protons and neutrons are composed of more basic particles called **quarks,** the most common varieties being "up" (u) quarks and "down" (d) quarks. A proton is composed of two up quarks and one down quark, a neutron of two down quarks and one up quark.

In the 1970s the concept of quarks led to a breakthrough in our understanding of the strong and weak forces. The strong force holds quarks together, while the weak force is at work whenever a quark changes from one variety to another. For example, when a neutron decays into a proton, one of the neutron's down quarks changes into an up quark as shown in Figure 27-5*c*. Thus, the weak force is responsible for transformations such as

$$d \rightarrow u + e^- + \bar{\nu}$$

WEB LINK 27.3

In the 1940s, physicists Richard P. Feynman and Julian S. Schwinger (working independently in the United States) and Sin-Itiro Tomonaga (in Japan) succeeded in developing a basic description of what we mean by force. Focusing their attention on the electromagnetic force, they tried to describe exactly what happens when two charged particles interact. According

Experimental and theoretical research into the basic forces of nature helps us understand the evolution of the cosmos

to their theory, now called **quantum electrodynamics,** charged particles interact by exchanging photons that cannot be observed directly, because they exist for extremely short time intervals. (We will explore such *virtual* photons in Section 27-3.)

Quantum electrodynamics has proved the most successful theory in modern physics. It describes with remarkable accuracy many details of the electromagnetic interaction between charged particles. Inspired by these successes, physicists have tried to develop similar theories for the other three forces. In these theories, the weak force occurs when particles exchange **intermediate vector bosons,** the gravitational force occurs when particles exchange **gravitons,** and quarks stick together by exchanging **gluons.** The table in Figure 27-5 summarizes these features of the four fundamental forces.

Unified Theories of the Fundamental Forces

WEBLINK 27.4

Physicists made important progress in understanding the weak force during the 1970s. Steven Weinberg, Sheldon Glashow, and Abdus Salam proposed a theory with three types of intermediate vector bosons, which are exchanged in various manifestations of the weak force. These three particles were actually discovered in experiments in the 1980s, providing strong support for the theory.

A startling prediction of the Weinberg-Glashow-Salam theory is that the weak force and the electromagnetic force should be identical to each other for particles with energies greater than 100 GeV. (One GeV equals 10^9, or 1 billion, electron volts; see Section 5-5.) In other words, if particles are slammed together with a total energy greater than 100 GeV, then electromagnetic interactions become indistinguishable from weak interactions. We say that above 100 GeV the electromagnetic force and the weak force are "unified" into a single **electroweak force.**

This unification occurs because the three types of intermediate vector bosons behave just like photons above 100 GeV. At such high energies, these three particles actually lose their mass, and the weak force becomes a long-range force with the same intrinsic strength as electromagnetism. Physicists describe this similarity by saying that "symmetry is restored" above 100 GeV.

In the world around us, however, the typical energies with which particles interact are very much lower, on the order of 1 eV or less. Below 100 GeV, intermediate vector bosons behave like massive particles, but photons are always massless. Because intermediate vector bosons and photons are not similar at low energies, we say that "symmetry is broken" below 100 GeV, which is why the electromagnetic and the weak forces behave so differently in the world around us. In the language of physics, a **spontaneous symmetry breaking** occurs.

In the 1970s, Sheldon Glashow, Howard Georgi, Jogesh Pati, and Abdus Salam proposed **grand unified theories** (or **GUTs**), which predict that the strong force becomes unified with the weak and electromagnetic forces at energies above 10^{14} GeV. In other words, if particles were to collide at energies greater than 10^{14} GeV, the strong, weak, and electromagnetic interactions would all be long-range forces and would be indistinguishable from each other.

Many physicists suspect that all four forces may be unified at energies greater than 10^{19} GeV (Figure 27-6*a*). That is, if particles were to collide at these colossal energies, there would be no difference between the gravitational, electromagnetic, and nuclear forces. However, no one has yet succeeded in working out the details of such a **supergrand unified theory,** which is sometimes called a **theory of everything** (or **TOE**).

Physicists use particle accelerators to examine the unification of the weak and electromagnetic forces by slamming particles together at energies around 100 GeV (Figure 27-6*b*). There is probably no hope, however, of ever constructing machines capable of making particles collide with energies approaching 10^{19} GeV. It may thus be impossible to test directly the grand and supergrand unified theories in a laboratory. However, the universe immediately after the Big Bang was so hot and its particles were moving so fast that they did indeed collide with energies on the order of 10^{19} GeV. The evolution of the universe during its earliest moments has thus become a laboratory for testing some of the most elegant, sophisticated theories in physics.

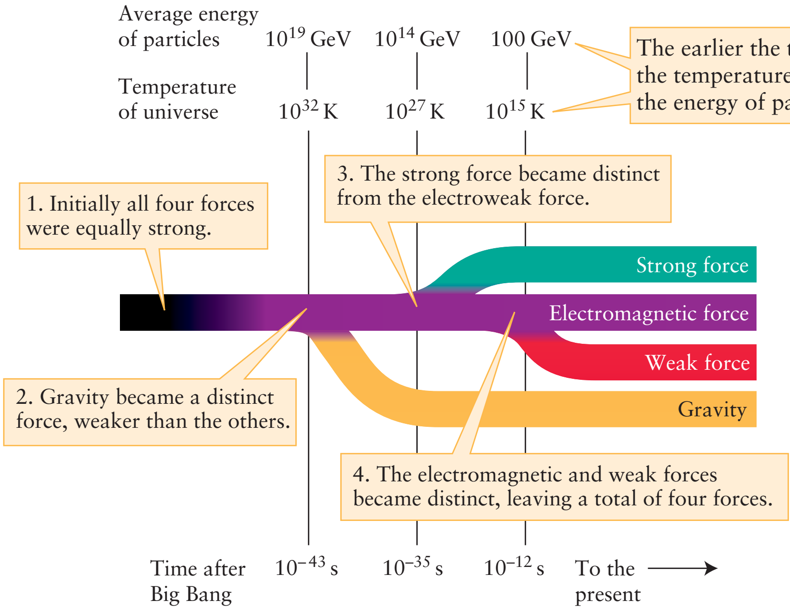
The Fundamental Forces in the Early Universe

Figure 27-6*a* shows how the various forces changed during the first fraction of a second after the Big Bang. Before the Planck time (from $t = 0$ s to $t = 10^{-43}$ s), particles collided with energies greater than 10^{19} GeV, and all four forces were unified. Because we do not yet have a TOE that properly describes the behavior of gravity, we remain ignorant of what was going on during the first 10^{-43} second of the universe's existence. We know, however, that by the end of the Planck time, the expansion and cooling of the universe had caused the energy of particles to fall to 10^{19} GeV. At energies below this, gravity is thought not to be unified with the other three forces.

In the language of physics, at $t = 10^{-43}$ s there was a spontaneous symmetry breaking in which gravity was "frozen out" of the otherwise unified hot soup that filled all space. In such a "soup," the typical energy of a particle (E) is related to temperature (T) by $E = kT$, where k is the Boltzmann constant (roughly 10^{-4} eV/K, or 10^{-13} GeV/K). Thus, the temperature of the universe was 10^{32} K when gravity emerged as a separate force.

As the universe expanded, its temperature decreased and the energy of particles decreased as well. (We discussed this property of gases in Box 21-1.) At $t = 10^{-35}$ s, the energy of particles in the universe had fallen to 10^{14} GeV, equivalent to a temperature of 10^{27} K. At energies and temperatures below these, the strong force is no longer unified with the electromagnetic and weak forces. Thus, at $t = 10^{-35}$ s, there was a second spontaneous symmetry breaking, at which time the strong force "froze out."

The inflationary epoch is thought to have begun at this point. Physicists hypothesize that before the strong force decoupled from the electroweak force, the universe was in an unstable state called a **false vacuum.** In this state, physicists hypothesize that the energy associated with a quantity called the *inflaton field* had a nonzero value (Figure 27-7*a*). (Just as the space around a magnet is permeated by a magnetic field, like that shown in Figure 7-13*a*, the entire universe is thought to be permeated by the inflaton field.) This state was unstable in the same sense as a ball perched atop a cone with a pointed top: The ball will stay there if left undisturbed, but will roll downhill if even slightly disturbed. In an analogous way, it is thought that at the time that the strong and electroweak forces decoupled, the universe "rolled downhill"

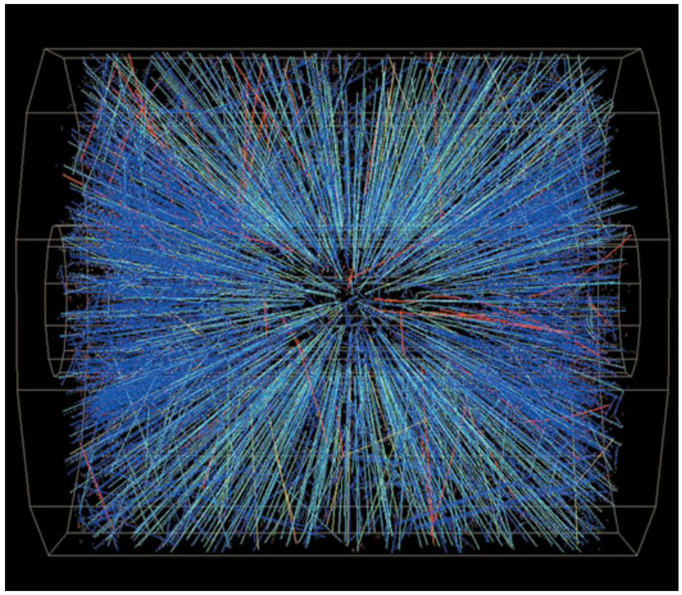


(a) How the four forces behave at different energies and temperatures

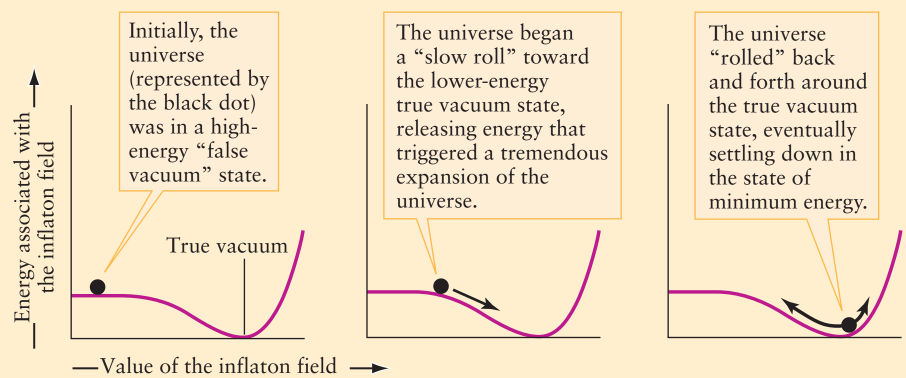
WEB LINK 27.5

Figure 27-6

Unification of the Four Forces **(a)** The strength of the four fundamental forces depends on the energy of the particles that interact. As shown in this schematic diagram, the higher the energy, the more the forces resemble each other. Also included here are the temperature of the universe and the time after the Big Bang when the strengths of the forces are thought to have been equal. **(b)** This shower of subatomic particles was produced by a collision between two gold nuclei, each of which was traveling at 99.995% the speed of light. The energy per proton or neutron was 100 GeV, near the energy at which the electromagnetic and weak forces become unified. (Brookhaven National Laboratory)



(b) The result of a high-energy collision between gold nuclei



(a) The initial false vacuum (b) Inflationary expansion (c) Reheating

Figure 27-7

Inflation: Transitioning from the False Vacuum to the True Vacuum **(a)** The energy of the vacuum is thought to be determined by the value of a quantity called the inflaton field. As shown by the red curve in this diagram, this energy is at a minimum for a certain value of this field. The state of lowest energy is called the true vacuum. **(b)** Inflation is associated with the universe making a transition from its initial false-vacuum state to a true-vacuum state. This expansion took place over a period of about 10^{-32} s, during which the universe cooled to a temperature of about 3 K. **(c)** As the universe "settled in" to the true vacuum state, energy was released that reheated the universe to a temperature of 10^{27} K.

to the *true vacuum,* a state of lower energy. The universe's transition to the true vacuum released energy that caused it to expand tremendously in a brief interval of time (Figure 27-6*b*). By the time the inflationary epoch had ended, about 10^{-32} s after the Big Bang, the universe had increased in scale by a factor of roughly 10^{50}.

The rapid expansion of the universe also gave rise to a rapid cooling. At the end of the inflationary epoch, the temperature of the universe may have dropped to about 3 K, about the same as the temperature that the cosmic background radiation has today. But as the universe finally settled into the vacuum-energy state, an additional amount of energy was released that went into *reheating* the universe to a temperature of 10^{27} K—about the same as it had before inflation began (Figure 27-7*c*). Thus, inflation caused the universe to expand tremendously while having no net effect on its temperature.

After the end of the inflationary epoch at $t = 10^{-32}$ s, the universe continued to expand and cool at a more sedate rate. At $t = 10^{-12}$ s, the temperature of the universe had dropped to 10^{15} K, the energy of the particles had fallen to 100 GeV, and there was a final spontaneous symmetry breaking and "freeze-out" that separated the electromagnetic force from the weak force. From that moment on, all four forces have interacted with particles essentially as they do today.

27-3 During inflation, all the mass and energy in the universe burst forth from the vacuum of space

Inflation was a brief but stupendous expansion of the universe soon after the beginning of time. Physicists now realize that inflation helps explain where all the matter and radiation in the universe came from. To see how a violent expansion of space could create particles, we must first understand what quantum mechanics tells us about space.

Quantum Mechanics and the Heisenberg Uncertainty Principle

Quantum mechanics is the branch of physics that explains the behavior of nature on the atomic scale and smaller. For example, quantum mechanics tells us how to calculate the structure of atoms and the interactions between atomic nuclei. **Elementary particle physics** is the branch of quantum mechanics that deals with individual subatomic particles and their interactions, including the strong, weak, and electromagnetic forces that we discussed in Section 27-2.

The submicroscopic world of quantum mechanics is significantly different from the ordinary world around us. In the ordinary world we have no trouble knowing where things are. You know where your house is; you know where your car is; you know where this book is. In the subatomic world of electrons and nuclei, however, you can no longer speak with this same confidence. A certain amount of fuzziness, or uncertainty, enters into the description of reality at the incredibly small dimensions of the quantum world.

To appreciate the reasons for this uncertainty, imagine trying to measure the position of a single electron. To find out where it is located, you must observe it. And to observe it, you must shine a light on it. However, the electron is so tiny and has such a small mass that the photons in your beam of light possess enough energy to give the electron a mighty kick. As soon as a photon strikes the electron, the electron recoils in some unpredictable direction. Consequently, no matter how carefully you try to measure the precise location of an electron, you necessarily introduce some uncertainty into the speed of that electron.

These ideas are at the heart of the **Heisenberg uncertainty principle,** first formulated in 1927 by the German physicist Werner Heisenberg, one of the founders of quantum mechanics. This principle states that there is a reciprocal uncertainty between position and momentum (equal to the mass of a particle times its velocity). The more precisely you try to measure the position of a particle, the more unsure you become of how the particle is moving. Conversely, the more accurately you determine the momentum of a particle, the less sure you are of its location. These restrictions are not a result of errors in making measurements; they are fundamental limitations imposed by the nature of the universe.

There is an analogous uncertainty involving energy and time. You cannot know the energy of a system with infinite precision at every moment in time. Over short time intervals, there can be great uncertainty about the amounts of energy in the subatomic world. Specifically, let ΔE be the smallest possible uncertainty in energy measured over a short interval of time Δt. (Astronomers and physicists often use the capital Greek letter delta, Δ, as a prefix to denote a small quantity or a small change in a quantity.)

Heisenberg uncertainty principle for energy and time

$$\Delta E \times \Delta t = \frac{h}{2\pi}$$

ΔE = uncertainty in energy

Δt = time interval over which energy is measured

h = Planck's constant = 6.625×10^{-34} J s

This says that the shorter the time interval Δt, the greater the energy uncertainty ΔE must be in order to ensure that the product of ΔE and Δt is equal to $h/2\pi$.

We might look upon the Heisenberg uncertainty principle as merely an unfortunate limitation on our ability to know everything with infinite precision. But, in fact, this principle provides startling insights into the nature of the universe.

Spontaneously Created Matter and Antimatter

We have seen that one of the important conclusions of Einstein's special theory of relativity is the equivalence of mass and energy: $E = mc^2$ (see Section 16-1). There is nothing uncertain about the speed of light (c), which is an absolute constant. Therefore, any uncertainty in the energy of a physical system can be attributed to an uncertainty Δm in the mass. Thus,

$$\Delta E = \Delta m \times c^2$$

Combining this expression with the previous equation, we obtain

Heisenberg uncertainty principle for mass and time

$$\Delta m \times \Delta t = \frac{h}{2\pi c^2}$$

Δm = uncertainty in mass

Δt = time interval over which mass is measured

h = Planck's constant = 6.625×10^{-34} J s

c = speed of light = 3.00×10^8 m/s

This result is astonishing. It means that over a very brief interval Δt of time, we cannot be sure how much matter there is in a particular location, even in "empty space." During this brief moment, matter can spontaneously appear and then disappear. The greater the amount of matter (Δm) that appears spontaneously, the shorter the time interval (Δt) it can exist before disappearing into nothingness. This bizarre state of affairs is a natural consequence of quantum mechanics.

Quantum mechanics reveals that "empty" space is not empty, but is seething with particles and antiparticles that appear and then annihilate

No particle can appear spontaneously by itself, however. For each particle created, so is a second, almost identical **antiparticle.** In other words, equal amounts of matter and **antimatter** come into existence and then disappear.

Despite its exotic name, there is actually nothing terribly mysterious about antimatter. A particle and an antiparticle are identical in almost every respect; their main distinction is that they carry opposite electric charges. For example, an ordinary electron (e^-) carries a negative charge; the corresponding antiparticle has the same mass but a positive charge, which is why it is called a **positron** (e^+). Because particles and antiparticles come and go in pairs, the total electric charge in the universe remains constant. Particles that have no electric charge can also have corresponding antiparticles. An example is the neutrino (ν); we met its antiparticle, the antineutrino ($\overline{\nu}$), in Section 27-2. The antineutrino is also electrically neutral, but differs from the neutrino in having opposite values of other, more subtle physical properties.

A spontaneously created particle-antiparticle pair lasts for only an incredibly brief time. For example, consider an electron and a positron, each with a mass 9.11×10^{-31} kg. If we rewrite the Heisenberg uncertainty principle for mass and time to solve for Δt and then substitute the combined mass of $2 \times 9.11 \times 10^{-31}$ kg, we find that a spontaneously created electron-positron pair can last for a time

$$\Delta t = \frac{1}{\Delta m}\,\frac{h}{2\pi c^2} = \frac{1}{2 \times 9.11 \times 10^{-31}} \times \frac{6.625 \times 10^{-34}}{2\pi(3.00 \times 10^8)^2}$$
$$= 6.43 \times 10^{-22} \text{ s}$$

In other words, an electron and a positron can spontaneously appear and then disappear without violating any laws of physics—but they can remain in existence for no longer than 6.43×10^{-22} s. The more massive the particle, the shorter the time interval it can exist. For example, the proton has about 2000 times more mass than the electron. Pairs of protons and their antiparticles, called **antiprotons,** can appear and disappear spontaneously, but they exist for only 1/2000 as long as pairs of electrons and positrons do.

Virtual Pairs

Spontaneous creation can and does happen absolutely anywhere and at any time, not just under the unusual conditions of the early universe. (It is happening right now in the space between this book and your eyes.) Quantum mechanics tells us that if a process is not strictly forbidden, then it must occur. Pairs of every conceivable particle and antiparticle are constantly being created and destroyed at every location across the universe. However, we have no way of observing these pairs directly without violating the uncertainty principle. For this reason, they are called **virtual pairs.** They do not "really" exist in the same sense as ordinary particles; they "virtually" exist. The particles that are exchanged in the four fundamental forces (see Section 27-2) are also virtual particles.

Although virtual pairs of particles and antiparticles cannot be observed directly, their effects have nonetheless been detected. Imagine, for example, an electron in orbit about the nucleus of an atom, such as a hydrogen atom. Ideally, the electron should follow its orbit in a smooth, unhampered fashion. However, because of the constant brief appearance and disappearance of pairs of particles and antiparticles, minuscule electric fields exist for extremely short intervals of time. These tiny, fleeting electric fields cause the electron to jiggle slightly in its orbit. This jiggling produces slight changes in the energies of different electron orbits in the hydrogen atom, which manifest themselves as a minuscule shift in the wavelengths of the hydrogen spectral lines. (We discussed the connection between the energies of electron orbits in hydrogen and the hydrogen spectrum in Section 5-8.)

This shift was first detected in 1947 by the American physicists Willis Lamb and R. C. Retherford and today is known as the **Lamb shift.** The Lamb shift provides powerful support for the idea that every point in space, all across the universe, is seething with virtual pairs of particles and antiparticles. In this sense, "empty space" is actually not empty at all. Figure 27-8 sketches the constant appearance and disappearance of virtual particles and antiparticles.

In some circumstances, virtual pairs can become *real* pairs of particles and antiparticles, a phenomenon called **pair production.** It has been known for years that highly energetic gamma rays (photons) can convert their energy into pairs of particles and antiparticles. Quite simply, the gamma ray disappears upon colliding with a second photon, and a particle and an antiparticle appear in its place. These particles and antiparticles come from nature's ample supply of virtual pairs. The gamma rays provide a virtual pair with so much energy that the virtual particles can appear as real particles in the real world.

Pair production is routinely observed in high-energy nuclear experiments (see Figure 27-6*b*). Indeed, it is one of the ways in which physicists manufacture exotic species of particles and antiparticles. The only requirement is that nature's balance sheet be

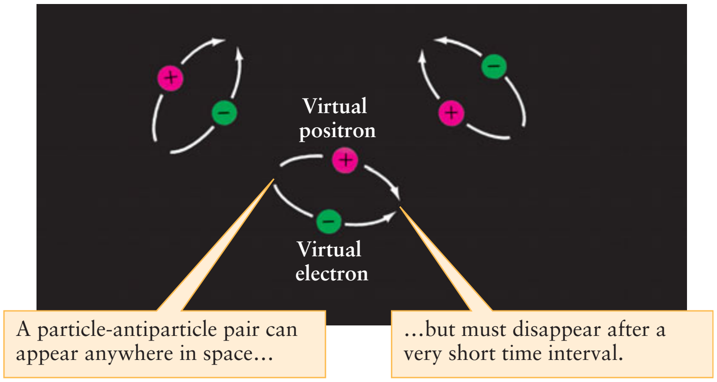


Figure 27-8

Virtual Pairs Pairs of particles and antiparticles can appear and then disappear anywhere in space provided that each pair exists only for a very short time interval, as dictated by the uncertainty principle. In this sketch, electrons are shown in green and positrons are shown in red.

satisfied. To create a particle and an antiparticle having a total mass M, the incoming gamma-ray photons must possess an amount of energy E that is greater than or equal to Mc^2. If the photons carry too little energy (less than Mc^2), pair production will not occur. Likewise, the more energetic the photons, the more massive the particles and antiparticles that can be manufactured. Figure 27-9*a* shows this process of pair production. Figure 27-9*b* shows the inverse process of **annihilation,** in which a particle and antiparticle collide with each other and are converted into high-energy gamma rays.

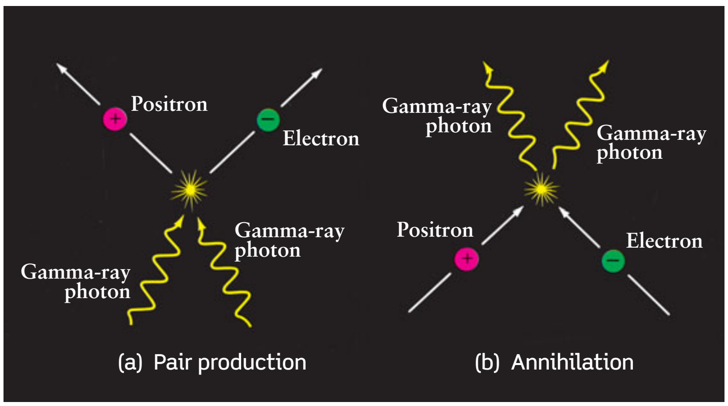


Figure 27-9

Pair Production and Annihilation **(a)** Pairs of virtual particles can be converted into real particles by high-energy gamma-ray photons. In this illustration, an electron (shown in green) and a positron (in red) are produced. This process can take place only if the combined energy of the two photons is no less than Mc^2, where M is the total mass of the electron and positron. **(b)** Conversely, a particle and an antiparticle can annihilate each other and be transformed into energy in the form of gamma rays.

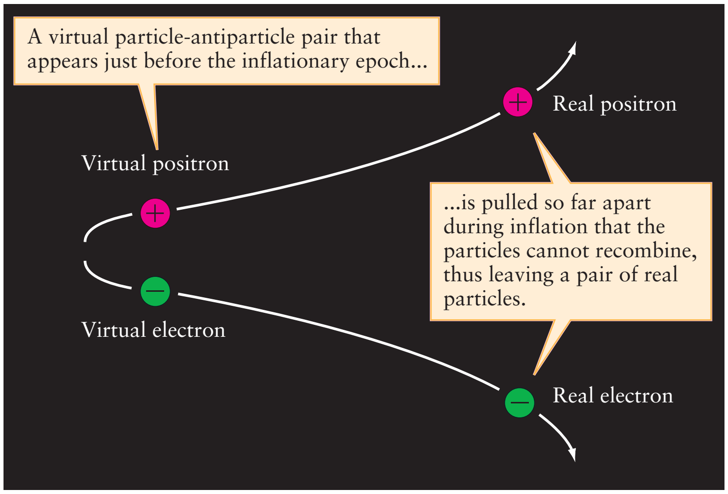


Figure 27-10

Inflation: From Virtual to Real Particles The universe expanded by such a tremendous factor during the inflationary epoch that the members of virtual particle-antiparticle pairs could no longer find each other. As a result, these virtual particles and antiparticles became real particles and antiparticles.

Around 1980 physicists began applying these ideas to their thinking about the creation of the universe. During the inflationary epoch, space was expanding with explosive vigor. As we have seen, however, all space is seething with virtual pairs of particles and antiparticles. Normally, a particle and an antiparticle have no trouble getting back together in a time interval (Δt) short enough to be in compliance with the uncertainty principle. During inflation, however, the universe expanded so fast that particles were rapidly separated from their corresponding antiparticles. Deprived of the opportunity to recombine, these virtual particles had to become *real* particles in the real world. In this way, the universe was flooded with particles and antiparticles created by the violent expansion of space (Figure 27-10).

27-4 As the early universe expanded and cooled, most of the matter and antimatter annihilated each other

As soon as the flood of matter and antimatter appeared in the universe, collisions between particles and antiparticles began to produce numerous high-energy gamma rays. As these gamma rays collided, they promptly turned back into the particles and antiparticles from which they came. As a result, the rate of pair production soon equaled the rate of annihilation. For example, for every electron and positron that annihilated each other to create gamma rays (Figure 27-9*b*), two gamma rays collided elsewhere to produce an electron and a positron (Figure 27-9*a*). In other words, annihilation and pair production reactions proceeded with equal vigor, and as many particles and antiparticles were being created as were being destroyed.

As the universe continued to expand, all the gamma-ray photons became increasingly redshifted. As a result, the temperature of the radiation field fell. Due to their frequent interaction, radiation and particles of all kinds were in **thermal equilibrium:** All particle species, including photons, were at the same temperature. Hence, as the radiation temperature decreased, the temperature of particles of different types decreased as well.

From Quark Confinement to Particle-Antiparticle Annihilation

The first change in the population of particles and antiparticles occurred at $t = 10^{-6}$ s, when the temperature was 10^{13} K and particles were colliding with energies of roughly 1 GeV. Prior to this moment, particles collided so violently that individual protons and neutrons could not exist, being constantly fragmented into quarks. After this time, appropriately called the period of **quark confinement,** quarks were finally able to stick together and became confined within individual protons and neutrons.

The cosmic background radiation we see today was spawned from a vast sea of particles and antiparticles in the early universe

LOOKING DEEPER 27.1

As the universe continued to expand, temperatures eventually became so low that the gamma rays no longer had enough energy to create particular kinds of particles and antiparticles. We say that the temperature dropped below the particular particle's **threshold temperature.** Collisions between these types of particles and antiparticles continued to add photons to the cosmic-radiation background, but collisions between photons could no longer replenish the supply of particles and antiparticles.

At the same time that quark confinement became possible so that protons and neutrons appeared, the universe also became cooler than the 10^{13}-K threshold temperatures of both protons and neutrons. No new protons or neutrons were formed by pair production, but the annihilation of protons by antiprotons and of neutrons by antineutrons continued vigorously everywhere throughout space. This wholesale annihilation dramatically lowered the matter content (particles and antiparticles) of the universe, while simultaneously increasing the radiation (photon) content.

A little later, when the universe was about 1 second old, its temperature fell below 6×10^9 K, the threshold temperature for electrons and positrons. A similar annihilation of pairs of electrons and positrons further decreased the matter content of the universe while raising its radiation content. This radiation field, which fills all space, is the *primordial fireball* discussed in Section 26-5. This fireball, which dominated the universe for the next 380,000 years, therefore derived much of its energy from the annihilation of particles and antiparticles during the first second after the Big Bang.

Now we have a dilemma. If there had been perfect symmetry between particles and antiparticles, then for every proton there should have been an antiproton. For every electron, there should likewise have been a positron. Consequently, by the time the universe was 1 second old, every particle would have been annihilated by an antiparticle, leaving no matter at all in the universe.

Obviously, this did not happen. The planets, stars, and galaxies we see in the sky are made of matter, not antimatter. (If there were substantial amounts of antimatter in the universe, it would eventually collide with ordinary matter. We would then see copious amounts of gamma rays being emitted from the entire sky. While we do observe gamma-ray photons from various locations in the universe, they are neither numerous enough nor of the right energy to indicate the presence of much antimatter.) Thus, there must have been an excess of matter over antimatter immediately after the Big Bang, so that the particles outnumbered the antiparticles.

We can estimate the extent of this asymmetry between matter and antimatter. As noted in Section 26-5, there are roughly 10^9 photons today in the microwave background for each proton and neutron in the universe. Thus, for every 10^9 antiprotons, there must have been 10^9 plus one ordinary protons, leaving one surviving proton after annihilation. Similarly, for every 10^9 positrons, there must have been 10^9 plus one ordinary electrons. Theories of elementary particles and their interactions do indeed predict a slight preference for matter over antimatter of just this sort.

27-5 A background of neutrinos and most of the helium in the universe are relics of the primordial fireball

The early universe must have been populated with vast numbers of neutrinos (ν) and their antiparticles, the antineutrinos ($\bar{\nu}$). These particles have a very small mass, so their threshold temperature is quite low. These particles take part in the nuclear reactions that transform neutrons into protons and vice versa. For example, a neutron can decay into a proton by emitting an electron and an antineutrino:

$$n \rightarrow p + e^- + \bar{\nu}$$

This radioactive decay happens quickly (its half-life is about 10.5 minutes), which is why we do not find free neutrons floating around in the universe today. In the first 2 seconds after the Big Bang, however, neutrons were also created by collisions between protons and electrons:

$$p + e^- \rightarrow n + \nu$$

This reaction kept the number of neutrons approximately equal to the number of protons. This balance was maintained only as long as collisions between protons and electrons were frequent. But the number of electrons decreased precipitously as the temperature fell below 6×10^9 K and electrons and positrons annihilated each other (Section 27-4). By the time the universe was about 2 seconds old, no new neutrons were being formed, the natural tendency for neutrons to decay into protons took over, and the number of neutrons began to decline.

Spawning the First Nuclei

WEB LINK 27.8

Before many of the neutrons could decay into protons, they began to combine with protons to form nuclei. Nuclei of helium, the first element more massive than hydrogen, consist of either two protons and two neutrons (^{4}He) or two protons and a single neutron (^{3}He). It is exceedingly improbable that two protons and one or two neutrons should all simultaneously collide with one another to form a helium nucleus. Instead, helium nuclei are built in a series of steps. The first step is to have a single proton and a single neutron combine to form deuterium (^{2}H), sometimes called "heavy hydrogen." A photon (γ) is emitted in this process, so we write this reaction as

The first atomic nuclei formed within a quarter-hour after the Big Bang

$$p + n \rightarrow {}^2H + \gamma$$

Forming deuterium does not immediately lead to the formation of helium, however. The problem is that deuterium nuclei are easily destroyed, because a proton and a neutron do not stick together very well. Indeed, in the early universe, high-energy gamma rays easily broke deuterium nuclei back down into independent protons and neutrons. As a result, the synthesis of helium could not get beyond the first step. This block to the creation of helium is called the **deuterium bottleneck.**

When the universe was about 3 minutes old, the background radiation had cooled enough that its photons no longer had enough energy to break up the deuterium. By this time, most of the neutrons had decayed into protons, and protons outnumbered neutrons by about 6 to 1. Because deuterium nuclei could now survive, the remaining neutrons combined with protons and rapidly produced helium. (The *Cosmic Connections* figure for Chapter 16 depicts a similar sequence of reactions that take place in the core of the present-day Sun.)

The result was what we find in the universe today—about one helium atom for every ten hydrogen atoms. In addition to helium, nuclei of lithium (Li, which has three protons) and beryllium (Be, which has four) were also produced in small numbers. The process of building up nuclei such as deuterium and helium from protons and neutrons is called **nucleosynthesis** (Figure 27-11).

Because nuclei have positive electric charges, bringing them together to form more massive nuclei requires that they overcome their mutual electric repulsion. They are unable to do so if they are moving too slowly, which will be the case if the temperature is too low. As a result, by about 15 minutes after the Big Bang the universe was no longer hot enough for nucleosynthesis to take place. Only the four lightest elements (hydrogen, helium, lithium, and beryllium) were present in appreciable numbers. The heavier elements would be formed only much later, once stars had formed and nuclear reactions within those stars could manufacture carbon, nitrogen, oxygen, and all the other elements.

CAUTION! Keep in mind that only *nuclei* formed in the first 15 minutes of the history of the universe. It would be another 380,000 years before temperatures became low enough for these nuclei to combine with electrons to form atoms.

The Neutrino-Antineutrino Background

While nuclei were being formed in the early universe, what happened to all those primordial neutrinos and antineutrinos that had interacted so vigorously with the protons and neutrons before the universe was 2 seconds old? The answer is that by $t = 2$ s, matter was sufficiently spread out so that the universe became

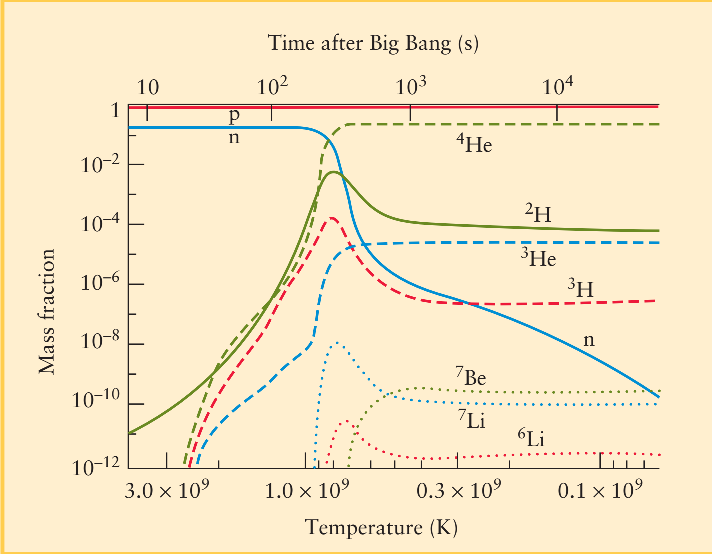


Figure 27-11

Nucleosynthesis in the Early Universe This graph shows how nuclei were produced between 10 seconds and 10 hours after the Big Bang. The vertical axis shows the fraction of the total mass that was in each type of particle or nucleus (p = proton, n = neutron, ^{2}H = deuterium, ^{3}He and ^{4}He = helium, ^{6}Li and ^{7}Li = lithium, ^{7}Be = beryllium). Very few nuclei were formed before the universe was 10 seconds old, thanks to the phenomenon of the deuterium bottleneck, which occurred at times earlier than those shown here. By about 10^3 seconds (roughly 15 minutes) after the Big Bang, the temperature had dropped below 4×10^8 K, no further nucleosynthesis was possible, and the relative amounts of different nuclei stabilized. The number of free neutrons declined rapidly as these particles decayed into protons, electrons, and antineutrinos. (Adapted from R. V. Wagoner)

transparent to neutrinos and antineutrinos. From that time on, neutrinos and antineutrinos could travel across the universe unimpeded. (Recall from Section 16-4 that the Earth itself is virtually transparent to neutrinos from the Sun.)

The neutrinos and antineutrinos that were liberated at $t = 2$ s should now fill the universe much as the cosmic microwave background does. Indeed, these ancient neutrinos and antineutrinos may be about as populous today as the photons in the microwave background (of which there are 4.1×10^8 per cubic meter). The neutrino-antineutrino background should be slightly cooler than the photon background, which received extra energy from electron-positron annihilations. Physicists estimate that the current temperature of the neutrino-antineutrino background is about 2 K, as opposed to 2.725 K for the microwave background. Unfortunately, because neutrinos and antineutrinos are so difficult to detect, we do not yet have direct evidence of the neutrino-antineutrino background.

27-6 Galaxies and the first stars formed from density fluctuations in the early universe

The distribution of matter in the universe today is quite lumpy. Stars are grouped together in galaxies, galaxies into clusters, and clusters into superclusters that stretch across 50 Mpc (150 million ly) or more (see Section 24-6). Furthermore, galaxies seem to be concentrated along enormous sheets, which in turn surround voids measuring 30 to 120 Mpc (100 million to 400 million ly) across. Figures 24-23 and 24-24 show these features, which characterize the large-scale structure of the universe. How did this large-scale structure arise from the chaos of the primordial fireball? When did stars first appear in the universe? And when and how did galaxies first form?

Density Fluctuations and the Jeans Length

At first glance the origin of large-scale structure seems puzzling, because the early universe must have been exceedingly smooth. To see why, think back to the era of recombination that occurred 380,000 years after the Big Bang (see Section 26-5). Before recombination, high-energy photons were constantly and vigorously colliding with charged particles throughout all space. After recombination, the universe became transparent, and these photons stopped interacting with the matter in the universe. Astronomers say that matter "decoupled" from radiation during the era of recombination. Because the cosmic microwave background is extremely isotropic, we can conclude that the matter with which these photons once collided so frequently must also have been spread smoothly across space.

The distribution of matter during the early universe could not have been *perfectly* uniform, however. If it had been, it would still have to be absolutely uniform today; there would now be only a few atoms per cubic meter of space, with no stars and no galaxies. Consequently, there must have been slight lumpiness, or **density fluctuations**, in the distribution of matter in the early universe. These are thought to have originated in the very early universe, even before the inflationary epoch. Infinitesimally small quantum fluctuations in density, which are allowed by the Heisenberg uncertainty principle (see Section 27-3), were stretched during inflation to appreciable size. Through the action of gravity, these fluctuations eventually grew to become the galaxies and clusters of galaxies that we see today throughout the universe. As we saw in Section 26-5 and Section 26-8, the pattern of density fluctuations became imprinted on the cosmic background radiation during the era of recombination. Figure 26-14 and Figure 27-7 show a map of these fluctuations obtained from the WMAP microwave background spacecraft.

Our understanding of how gravity can amplify density fluctuations dates back to 1902, when the British physicist James Jeans solved a problem first proposed by Isaac Newton. Suppose that you have a gas with only very tiny fluctuations in density, as shown in Figure 27-12*a*. These regions of higher density will then gravitationally attract nearby material and thus gain mass. As this happens, however, the pressure of the gas inside these regions will also increase, which can make these regions expand and disperse. The problem that Jeans attacked was this: Under what conditions does gravity overwhelm gas pressure so that a permanent object can form?

Jeans proved that an object will grow from a density fluctuation provided that the fluctuation extends over a distance that exceeds the so-called **Jeans length** (L_J):

Jeans length for density fluctuations

$$L_J = \sqrt{\frac{\pi k T}{m G \rho_m}}$$

L_J = Jeans length

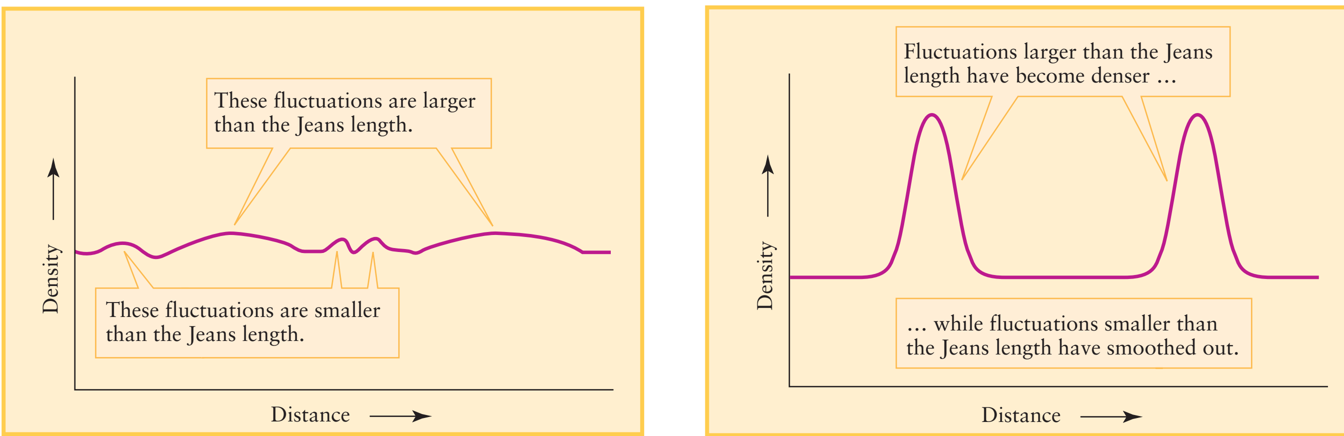
k = Boltzmann constant = 1.38×10^{-23} J/K

T = temperature of the gas (in kelvins)

m = mass of a single particle in the gas (in kilograms)

G = universal constant of gravitation = 6.67×10^{-11} N · m²/kg²

ρ_m = average density of matter in the gas



(a) At an early time

(b) At a later time

Figure 27-12

The Growth of Density Fluctuations **(a)** This conceptual illustration shows small density fluctuations in the distribution of matter shortly after the era of recombination. **(b)** If the size of a fluctuation is greater than the Jeans length (L_J), it becomes gravitationally unstable and can grow in amplitude.

Figure 27-13 R I V U X G

A Globular Cluster A typical globular cluster contains 10^5 to 10^6 stars, each with an average mass of about 1 $M_\odot$, so the total mass of a typical cluster is 10^5–10^6 $M_\odot$. Cluster diameters range from about 6 to 120 pc (20 to 400 ly). Because these masses and diameters are comparable to the Jeans length (L_J) during the era of recombination, astronomers suspect that globular clusters were among the first objects to form in the universe. (Hubble Heritage Team/AURA/STScI/NASA)

Density fluctuations that extend across a distance larger than the Jeans length tend to grow, while fluctuations smaller than L_J tend to disappear (Figure 27-12*b*).

We can apply the Jeans formula to the conditions that prevailed during the era of recombination, when $T = 3000$ K and $\rho_m = 10^{-18}$ kg/m^3. Taking m to be the mass of the hydrogen atom ($m = 1.67 \times 10^{-27}$ kg), we find that $L_J = 100$ light-years, the diameter of a typical globular cluster (Figure 27-13). Furthermore, the mass contained in a cube whose sides are 1 Jeans length in size (equal to the product of the density, ρ_m, and the volume of the cube, $L_J{}^3$) is about 5×10^5 $M_\odot$, equal to the mass of a typical globular cluster. This suggests that globular clusters were among the first objects to form after recombination.

Population III Stars: The "Zeroth" Generation

We saw in Section 19-4 that globular clusters contain the most ancient stars we can find in the present-day universe. These are Population II stars with a low percentage of metals (elements heaver than hydrogen and helium), and are of an earlier stellar generation than metal-rich Population I stars like the Sun (see Section 19-5). However, the Population II stars in globular clusters *cannot* be the very first stars to have formed after the Big Bang. Those first stars could have contained only hydrogen, helium, and tiny amounts of lithium and beryllium; as Figure 27-11 shows, these were the only elements whose nuclei formed in the early universe. Hence, these original stars would have contained an even smaller percentage of metals than the Population II stars found in globular clusters. Such "zeroth-generation" stars are called **Population III stars.**

The "zeroth generation" of stars were much more massive and luminous than stars today

ANIMATION 27.1

Like stars in the present-day universe, Population III stars would have formed from clouds of gas. These ancient gas clouds were composed almost exclusively of hydrogen and helium atoms, and such clouds have higher internal pressures than do metal-rich clouds of the same temperature. A star can only form when the mutual gravitation of the various parts of a cloud (which tends to make the cloud collapse) overcomes the internal pressure of the cloud (which tends to prevent collapse). Hence, Population III stars could only form if their mass (and hence their mutual gravitation) was rather large. Calculations suggest that these stars had masses from 30 to 1000 $M_\odot$, compared to the range of 0.4 to roughly 120 $M_\odot$ for modern stars. Even the smallest Population III star would rank among the largest stars observed today.

Although no Population III stars have yet been observed directly, we have at least indirect evidence that they existed. Infrared images from the Spitzer Space Telescope like the one that opens this chapter reveal a distant infrared background of starlight that is what we would expect to see from these zeroth-generation stars.

Another bit of evidence follows from the tremendous energy output of these stars: A 1000-$M_\odot$ Population III star would have been millions of times more luminous than the Sun and have a surface temperature in excess of 10^5 K, causing it to emit a flood of short-wavelength, high-energy photons. The photons from even a small number of such stars would have ionized most of the atoms in the universe, leaving electrons and nuclei of hydrogen and helium. This process is called **reionization,** a name that reminds us the universe had previously been ionized prior to recombination at t = 380,000 years. We saw in Section 26-8 that the photons of the cosmic microwave background are scattered by free electrons, and the effects of this scattering can be detected in maps of the background radiation. Data from the WMAP spacecraft suggest that reionization took place around 400 million years after the Big Bang, which in turn suggests that Population III stars formed around that time.

Since Population III stars were all very massive, their lifetimes were short and none could have survived to the present day. But during their short lifetimes, thermonuclear reactions within these stars produced elements heavier than beryllium for the first time in the history of the universe. What is more, calculations suggest that when these stars exploded—and due to their great mass, all of them did—they did not leave a white dwarf, neutron star, or black hole behind. Instead, all of their mass was ejected into space to be incorporated into the next generation of stars. The presence

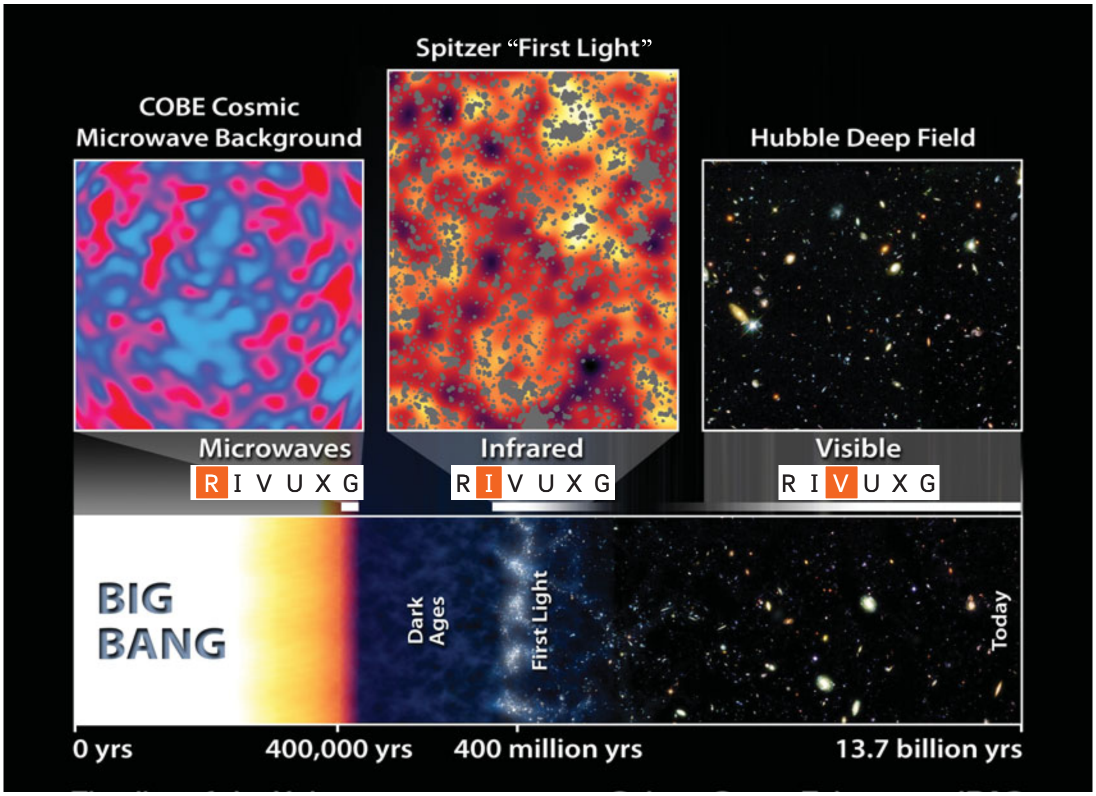


Figure 27-14

A Timeline of Light in the Universe The oldest light that we can see today is the cosmic background radiation, which comes from a time 380,000 years after the Big Bang when the universe first became transparent. This light has a redshift of about $z = 1100$ and appears in the microwave spectrum. Some 400 million years later at a redshift of about $z = 11$ the first stars appeared; their light is now redshifted to infrared wavelengths. Galaxies formed more recently and can be seen at visible wavelengths. (NASA; JPL-Caltech; and A. Kashlinsky, Goddard Space Flight Center)

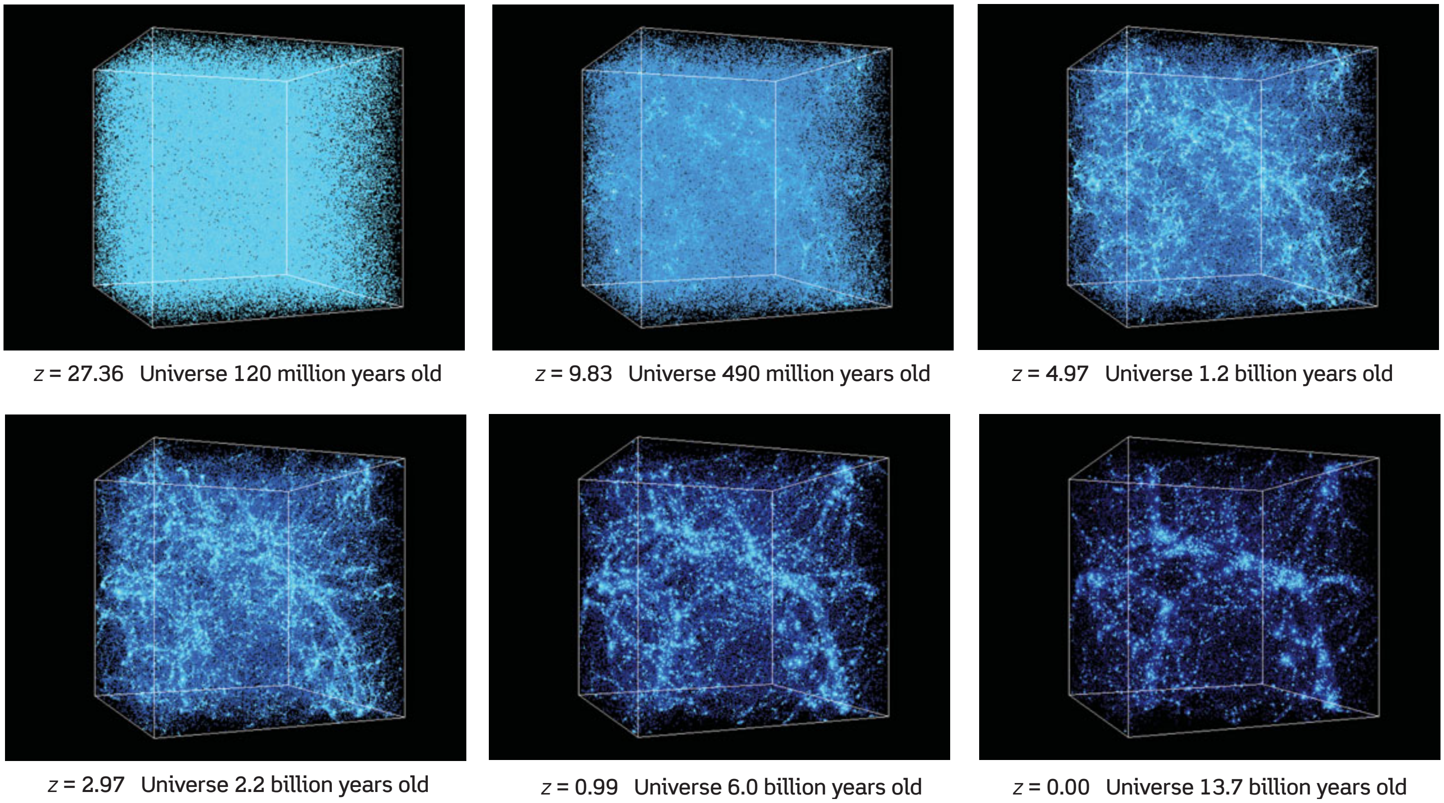
of heavy elements in the ejected material meant that when this material subsequently formed into clouds, the internal pressure of these clouds was lowered substantially and it became possible for low-mass stars to form. This laid the foundation for today's universe, in which dim, low-mass stars are common and luminous, massive stars are the exception (see Figure 17-5).

The era from recombination at $t = 380{,}000$ years to the first stars at $t = 400$ million years is called the **dark ages.** The only photons present at that time were those that make up today's cosmic background radiation. The dark ages ended when the universe was filled for the first time with the light from stars (Figure 27-14).

Forming Large-Scale Structure

Once clumps the size of globular clusters had formed in the universe, how did they form into galaxies, clusters of galaxies, and larger structures? One issue that complicates this matter is the presence of dark energy, which acts to accelerate the expansion of the universe (see Section 26-6). This accelerated expansion pulls clumps of material away from each other and makes it more difficult for them to coalesce into larger structures. Another complication is that about 85% of the mass in the universe is in the form of dark matter, whose nature is not known (see Section 24-8). Researchers have hypothesized different types of dark matter in the hope of explaining the large-scale structure that we see. Neutrinos are an example of **hot dark matter,** so named because it consists of lightweight particles traveling at high speeds. **Cold dark matter,** on the other hand, consists of massive particles traveling at slow speeds. Examples include WIMPs (which we discussed in Section 23-4) as well as other exotic, speculative particles.

WEB LINK 27.9 Scientists use supercomputer simulations to see how different types of dark matter would influence the development of large-scale structure. Figure 27-15 shows the results of such a simulation for a flat universe with dark energy and *cold* dark matter. The simulation follows the motions of 2 million particles of cold dark matter in a box that expands as the universe expands. The box at the lower right of the figure, representing the present time (redshift $z = 0$) is 43 Mpc (160 million ly) a side. At earlier times, the box represents a volume whose side is smaller by a factor $1/(1 + z)$. For example, each side of the box for $z = 0.99$ is actually 1/1.99 as long as the box for $z = 0$.



ANIMATION 27.2 ANIMATION 27.3

Figure 27-15

A Cold Dark Matter Simulation with Dark Energy These six views show the evolution of dark matter particles in a large, box-shaped volume of space. The box actually expands with time to follow the expansion of the universe; in this figure, the boxes have been rescaled so they all appear at the same size. Small fluctuations in density are put into the simulation at the beginning (at upper left); these evolve over time to form structures that resemble those actually observed in our present-day universe (z = 0.00, shown at the lower right). (Simulations performed at the National Center for Supercomputer Applications by Andrey Kravtsov/University of Chicago and Anatoly Klypin/NMSU)

The simulation begins 120 million years after the Big Bang with an almost perfectly uniform distribution of particles, mimicking the tiny density fluctuations that must have been present just after inflation. A supercomputer then calculates how these particles move, based on Newton's laws in an expanding universe. As time goes on, the fluctuations grow into small, bright clumps whose sizes and masses are similar to those of galaxies. A large filament also forms, spanning the entire box from left to right. The simulation shows that no additional structures formed after the universe was about 6 billion years old, corresponding to redshift $z = 1$. The explanation is that after this time, the accelerating expansion of the universe becomes more important than gravitational attraction. The final frame of the simulation strongly resembles actual maps of galaxies in our present-day universe (see Figure 24-24).

Simulations similar to those in Figure 27-15 have also been carried out using *hot* dark matter in the form of neutrinos. A massless neutrino would always travel at the speed of light, just as a photon does. However, experiments show that neutrinos do have a small mass. (This nonzero mass is what allows one type of neutrino to transform into another. We saw in Section 16-4 that such transformations provided the explanation to the long-standing solar neutrino problem.) Hence, neutrinos travel slower than light, and slow down as the universe expands and cools. Slow-moving neutrinos would accumulate over time within density fluctuations, and the gravitational pull of these neutrinos on surrounding matter could eventually lead to the formation of clusters of galaxies.

A primary difference between simulations based on cold and hot dark matter is the way in which galaxies form. In calculations based on cold dark matter, the formation of galaxies takes place from the "bottom up." In these simulations, the densest gas undergoes collapse early in the history of the universe and stars begin to form. The regions of star formation stream along the filaments (Figure 27-16). When they meet at the intersections between filaments, they merge and group together into galaxies, then clusters of galaxies, then superclusters. But in calculations based on hot dark matter, galaxies form from the "top down."

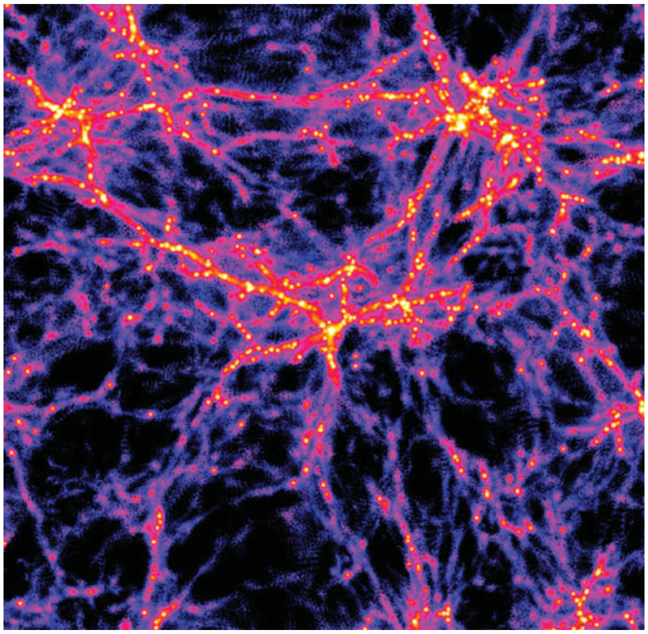


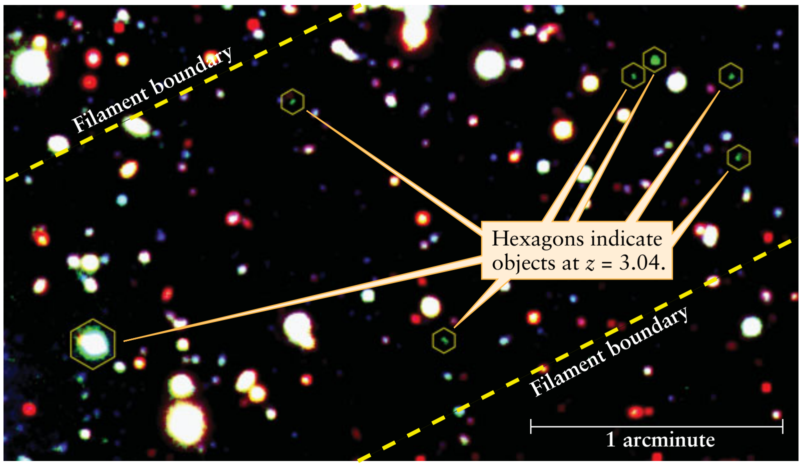
Figure 27-16

"Bottom-Up" Galaxy Formation: Simulation This image is taken from a cold dark matter simulation like that shown in Figure 27-15. A portion of the universe is shown at a time 2.2 billion years after the Big Bang, corresponding to redshift $z = 3.04$. The colors indicate the density of gas: Yellow is highest, red is medium, and blue is the lowest density. Over time, the gas tends to pile up at points where filaments intersect, forming galaxies and clusters of galaxies. (T. Theuns, MPA Garching/ESO)

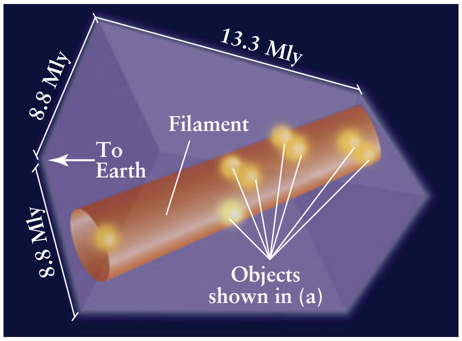
Huge supercluster-sized sheets of matter form first and then fragment into galaxies. Observations of remote galaxies show that galaxies actually formed from the "bottom up" scenario. One piece of evidence for this is the image in Figure 27-17*a*, which shows a handful of "galaxy building blocks" at $z = 3.04$ (when the universe was 2.2 billion years old). These "building blocks," which have not yet coalesced into galaxies, lie within a long filament (see Figure 27-17*b*) that resembles those shown in the simulation of Figure 27-16. Figure 27-18 shows a collection of "building blocks" at a later stage in the process of merging into a galaxy. These observations strongly suggest that the dominant form of dark matter is cold, not hot.

What Large-Scale Structure Reveals

How might the universe have evolved if it had contained different amounts of cold dark matter and dark energy? Figure 27-19 shows some simulations designed to explore these possibilities. If the density of matter in the universe is kept constant, simulations predict approximately the same structure for different values of the dark energy density parameter Ω_Λ defined in Section 26-6 (see Figure 27-19*a* and Figure 27-19*b*). But if too large a matter density is used in the simulation, the voids between galaxies are smaller than what we actually observe in our universe (Figure 27-19*c*). Hence, observations of galaxy clustering coupled with supercomputer simulations of galaxy formation help determine the matter density of our universe. (We made use of this idea in Section 26-7. The brown band in Figure 26-19 shows the constraints on cosmological parameters from these observations and simulations of galaxies.)



(a) High-redshift objects that lie within a filament



(b) Illustration of the filament

Figure 27-17 R I V U X G

"Bottom-Up" Galaxy Formation: Observation **(a)** The hexagons in this image from the Very Large Telescope show the positions of a number of sub-galaxy-sized objects at a redshift $z = 3.04$, the same as in the simulation shown in Figure 27-16. Excited hydrogen atoms in these objects emit ultraviolet photons, which are redshifted to visible wavelengths. This gives these objects a characteristic green color. (The object at lower left actually lies in front of a much brighter quasar.) **(b)** The objects in (a) all lie within an immense filament. The purple box shows the volume of space studied in this observing program. The dimensions are given in millions of light-years (Mly). (European Southern Observatory)

Figure 27-18 R I V U X G

A Galaxy Under Construction This Hubble Space Telescope image shows dozens of small galaxies in the process of merging into a single large galaxy. We see this galaxy at a redshift $z = 2.2$, corresponding to a time 3.1 billion years after the Big Bang. (NASA; ESA; G. Miley and R. Overzier, Leiden Observatory; and the ACS Science Team)

The best match to the observed distribution of galaxy clusters and to the cosmic background radiation data (see Figure 26-21) is a model like that shown in Figure 27-15, with dark energy and cold dark matter in the proportions listed in Table 26-2.

The *Cosmic Connections* figure summarizes the past history of our universe down to the present day. As we discussed in Section 26-7, the *future* of our universe is less certain, and depends on the detailed character of dark energy. More detailed data about galaxy clusters and the cosmic background radiation will be needed to pin down the future evolution of our universe.

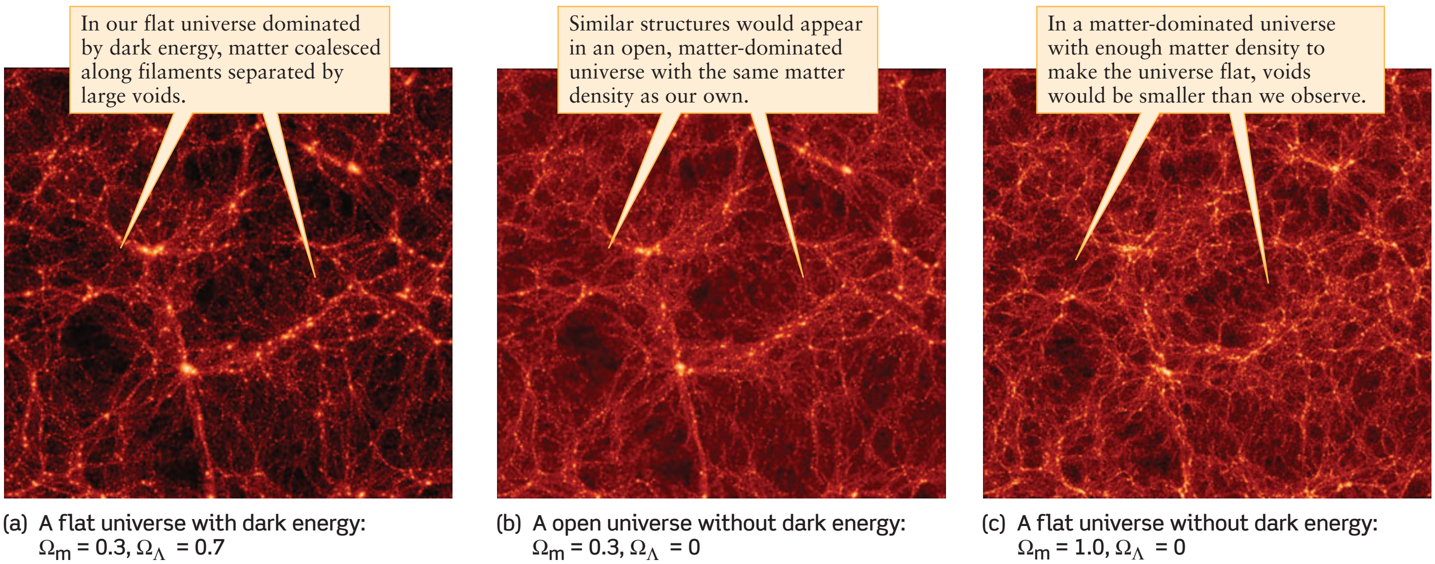


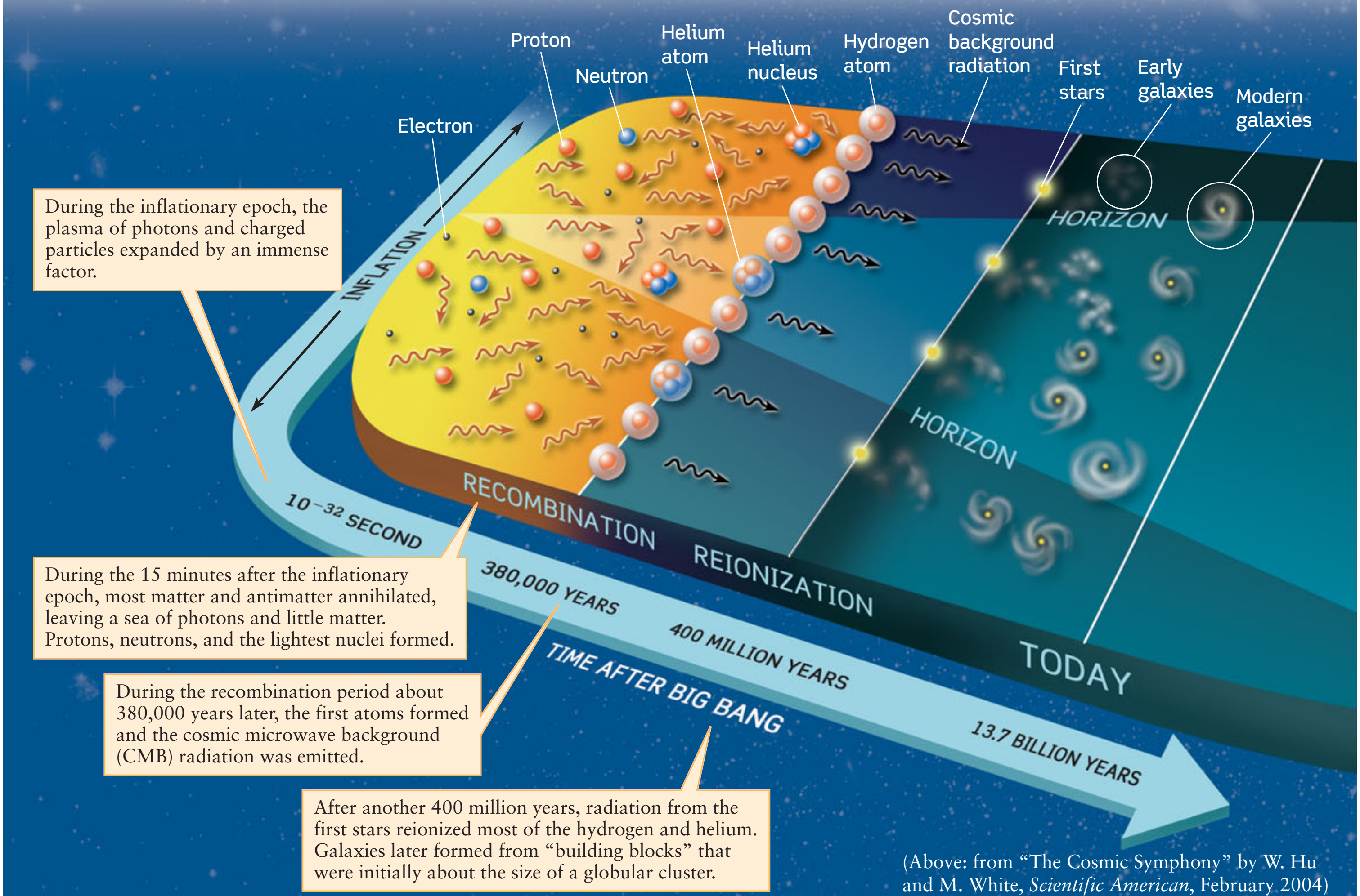
Figure 27-19

Using Simulations to Constrain the Matter Density of the Universe Cold dark matter simulations like those in Figures 27-15 and 27-16 help astronomers determine the value of the matter density parameter Ω_m. These three simulations show a portion of the universe at $z = 0$. **(a)** A simulation with $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$, close to the values for our universe, gives a good match to the observed distribution of filaments and voids. **(b)** Nearly as good a match is obtained if we keep $\Omega_m = 0.3$, but eliminate dark energy so that $\Omega_\Lambda = 0$. **(c)** If we use a larger value of Ω_m, the distribution of matter in the simulation is a poor match to our universe. (Simulation by the Virgo Supercomputing Center using computers based at the Computer Center of the Max-Planck-Institute in Garching and at the Edinburgh Parallel Computer Centre)

COSMIC CONNECTIONS

The History of the Universe

Research in astronomy, elementary particle physics, and nuclear physics has allowed scientists to piece together the grand sweep of events over the first billion years of cosmic history. The greatest drama occurred in the first few minutes, and these graphs have been drawn to emphasize those earliest moments.



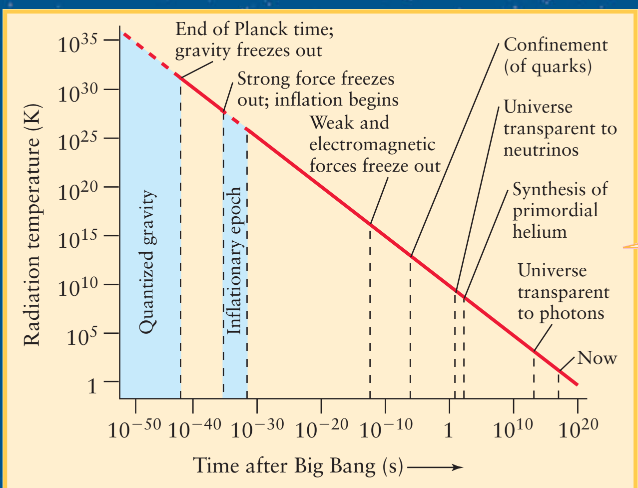
During the inflationary epoch, the plasma of photons and charged particles expanded by an immense factor.

During the 15 minutes after the inflationary epoch, most matter and antimatter annihilated, leaving a sea of photons and little matter. Protons, neutrons, and the lightest nuclei formed.

During the recombination period about 380,000 years later, the first atoms formed and the cosmic microwave background (CMB) radiation was emitted.

After another 400 million years, radiation from the first stars reionized most of the hydrogen and helium. Galaxies later formed from "building blocks" that were initially about the size of a globular cluster.

(Above: from "The Cosmic Symphony" by W. Hu and M. White, *Scientific American*, February 2004)



This graph shows the temperature history of the universe.

- As the universe cooled, the four forces "froze out" of their unified state as a result of spontaneous symmetry breaking.
- Neutrons and protons froze out of the hot "quark soup" during the quark confinement stage, which occured 10^{-6} second after the Big Bang.

27-7 Theories that attempt to unify the fundamental forces predict that the universe may have 11 dimensions

While we have a growing understanding of the early universe, there remains a veil obscuring the first 10^{-43} s after the Big Bang. These very *first* moments in the history of the universe, whose duration was the Planck time, determined everything that would come after. To understand this brief interval, we need to construct a quantum-mechanical theory that unifies gravity with the other fundamental forces of nature and that reconciles quantum mechanics with gravity. While this remains an unfinished task, remarkable progress has been made in recent years. One major breakthrough is that physicists have had to abandon the idea that there are only three dimensions of space.

Beyond Four Dimensions: Kaluza-Klein Theories

In his special and general theories of relativity, Einstein combined time with the three dimensions of ordinary space, resulting in a four-dimensional combination called *spacetime* (see Section 22-1). In 1919, the Polish physicist Theodor Kaluza proposed the existence of a *fifth* dimension. Kaluza hoped to describe both gravity and electromagnetism in terms of the curvature of five-dimensional spacetime, just as Einstein had explained gravity by itself in terms of the curvature of four-dimensional spacetime (see Section 22-2).

A particle always follows the straightest possible path in the four space dimensions of Kaluza's theory. But in the three dimensions of ordinary space, the path appears curved. Hence, it appears to us that the particle has been deflected by gravitational and electromagnetic forces. Kaluza's hypothetical fifth dimension exists at every point in ordinary space but is curled up so tightly, like a very tiny loop, that it is not directly observable.

In 1926, the Swedish physicist Oskar Klein attempted to make Kaluza's five-dimensional theory compatible with quantum mechanics. While he was not successful, Klein discovered that particles of different masses could be identified with different vibrations of the tiny loop of Kaluza's fifth dimension. Today, any quantum-mechanical theory that uses more than four dimensions to provide a unified description of the forces of nature is called a **Kaluza-Klein theory.**

When Kaluza and Klein developed their theories, gravity and electromagnetism were the only known forces of nature. Today we know of four fundamental forces, which suggests that modern Kaluza-Klein theories should have even more than five dimensions. Edward Witten at Princeton University has argued that a geometric theory for describing all four forces would work best with 11 dimensions, ten of space and one of time. At every point in ordinary space and at every moment of time, the seven "extra" dimensions must be rolled up, like the loops of Kaluza and Klein, into compact structures far too tiny for us to detect (Figure 27-20).

Subatomic particles may actually be multidimensional membranes

While we do not yet have a definite 11-dimensional theory, physicists know something of how it would work. Particles would travel along the straightest possible paths in spacetime, but their paths in ordinary three-dimensional space would appear curved. From our perspective, we interpret these curved paths as the result of the four forces acting on the particles.

M-Theory and Speculative Models of the Universe

WEB LINK 27.10

Theoretical physicists have shown that if there are indeed 11 dimensions, there must exist particles so massive that they have not yet been discovered. These speculative particles may be the dark matter that pervades the

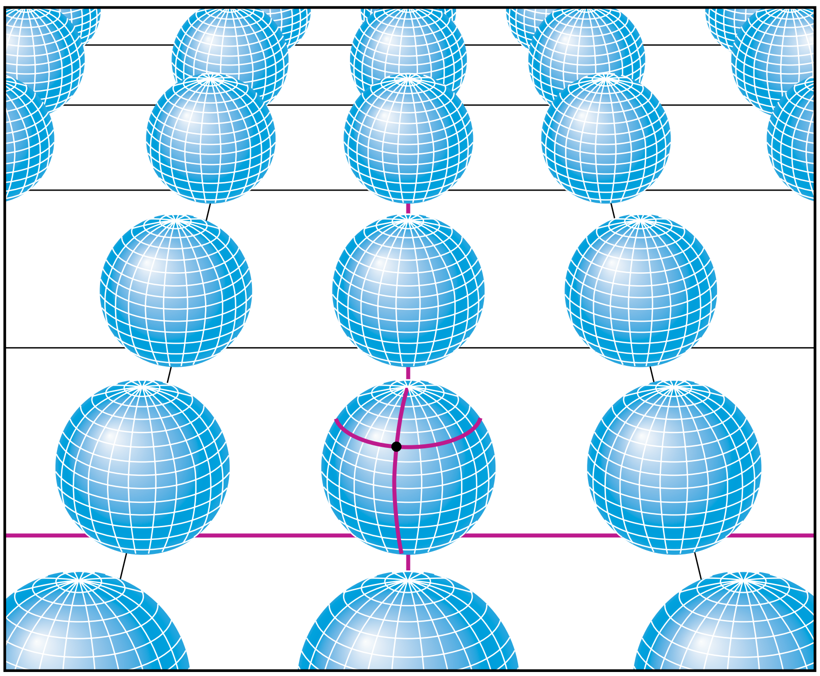


Figure 27-20

Hidden Dimensions of Space Hidden dimensions of space might exist provided they are curled up so tightly that we cannot observe them. This drawing shows how an ordinary two-dimensional plane might contain two additional dimensions. At every point on the plane, there is a very tiny sphere so small that it cannot be seen. To pinpoint a particular location, you need to give not only a position on the plane but also a position on the sphere that is tangent to the plane at that point, as indicated by the red lines. (Adapted from D. Freedman and P. Van Nieuwenhuizen) universe and holds together clusters of galaxies. Even more bizarre, the new theories no longer regard fundamental particles, such as electrons and quarks, as tiny points of mass. Instead, these particles may actually be multidimensional membranes, wrapped so tightly around the extra dimensions of space that they appear to us as points.

Andrew Strominger at Harvard University has shown that some of these membranes may fold on themselves in such a way that not even light can escape from them. In other words, these membranes may be black holes! This theory of membranes in 11 dimensions, or **M-theory,** may well explain the most exotic aspects of quantum mechanics, cosmology, and gravity. (M-theory encompasses a number of related theories known as *superstring theories*.)

One of the many speculative ideas inspired by M-theory is the *cyclic model* (known in an earlier version as the *ekpyrotic model,* from the Greek word for "conflagration"). In this model, our universe is one of two four-dimensional membranes (three space dimensions and one time dimension) that move with respect to each other along a fifth, "hidden" dimension. When the two membranes collide, a Big Bang results. The physics of the collision is such that each membrane automatically has the critical density, so that our universe is automatically flat. This provides a solution to the flatness problem that does not require inflation, so there is no inflationary epoch in the cyclic model. The cyclic model offers a natural explanation for dark matter: It is matter in the other membrane, the particles of which we cannot see but can detect through their gravitation. Dark energy in the cyclic model is a manifestation of the field that controls the interaction between the two membranes. The model is called *cyclic* because the membranes move apart and then back together in a rhythmic way. We would see a universe that expands for a time after the Big Bang, then eventually collapses back on itself in a Big Crunch. This is followed by another Big Bang, and the cycle repeats.

The cyclic model and the inflationary model are both *scientific* theories, because they both make predictions about things that can be observed. In particular, the two models make different predictions about the redshifts of very distant galaxies and about the polarization of the cosmic background radiation. Improved measurements of both of these will help us understand which of these models—each exotic in its own right—is a better description of our universe.

At present, M-theory is only the outline of a "theory of everything." We do not yet know how to describe a quark or an electron in terms of the higher dimensions of spacetime, nor can we yet explain the interactions among such particles. A full description of the first 10^{-43} s after the Big Bang is still far off. However, theoretical physicists around the world are devoting their efforts to these challenging problems. In league with astronomers and experimental physicists, they continue the age-old quest to understand the fundamental character of the universe.

We shall not cease from exploration
And the end of all our exploring
Will be to arrive where we started
And know the place for the first time.
T. S. Eliot, *Four Quartets*

Key Words

Key Ideas

Cosmic Inflation: A brief period of rapid expansion, called inflation, is thought to have occurred immediately after the Big Bang. During a tiny fraction of a second, the universe expanded to a size many times larger than it would have reached through its normal expansion rate.

• Inflation explains why the universe is nearly flat and the 2.725-K microwave background is almost perfectly isotropic.

The Four Forces and Their Unification: Four basic forces—gravity, electromagnetism, the strong force, and the weak force—explain all the interactions observed in the universe.

• Grand unified theories (GUTs) are attempts to explain three of the forces in terms of a single consistent set of physical laws. A supergrand unified theory would explain all four forces.

• GUTs suggest that all four fundamental forces were equivalent just after the Big Bang. However, because we have no satisfactory supergrand unified theory, we can as yet say nothing about the nature of the universe during this period before the Planck time ($t = 10^{-43}$ s after the Big Bang).

• At the Planck time, gravity froze out to become a distinctive force in a spontaneous symmetry breaking. During a second spontaneous symmetry breaking, the strong nuclear force became a distinct force; this transition triggered the rapid inflation of the universe. A final spontaneous symmetry breaking separated the electromagnetic force from the weak nuclear force; from that moment on, the universe behaved as it does today.

Particles and Antiparticles: Heisenberg's uncertainty principle states that the amount of uncertainty in the mass of a subatomic particle increases as it is observed for shorter and shorter time periods.

• Because of the uncertainty principle, particle-antiparticle pairs can spontaneously form and disappear within a fraction of a second. These pairs, whose presence can be detected only indirectly, are called virtual pairs.

• A virtual pair can become a real particle-antiparticle pair when high-energy photons collide. In this process, called pair production, the photons disappear, and their energy is replaced by the mass of the particle-antiparticle pair. In the process of annihilation, a colliding particle-antiparticle pair disappears and high-energy photons appear.

The Origin of Matter: Just after the inflationary epoch, the universe was filled with particles and antiparticles formed by pair production and with numerous high-energy photons formed by annihilation. A state of thermal equilibrium existed in this hot plasma.

• As the universe expanded, its temperature decreased. When the temperature fell below the threshold temperature required to produce each kind of particle, annihilation of that kind of particle began to dominate over production.

• Matter is much more prevalent than antimatter in the present-day universe. This is because particles and antiparticles were not created in exactly equal numbers just after the Planck time.

Nucleosynthesis: Helium could not have been produced until the cosmological redshift eliminated most of the high-energy photons. These photons created a deuterium bottleneck by breaking down deuterons before they could combine to form helium.

Density Fluctuations and the Origin of Stars and Galaxies: The large-scale structure of the universe arose from primordial density fluctuations.

• The first stars were much more massive and luminous than stars in the present-day universe. The material that they ejected into space seeded the cosmos for all later generations of stars.

• Galaxies are generally located on the surfaces of roughly spherical voids. Models based on dark energy and cold dark matter give good agreement with details of this large-scale structure.

The Frontier of Knowledge: The search for a theory that unifies gravity with the other fundamental forces suggests that the universe actually has 11 dimensions (ten of space and one of time), seven of which are folded on themselves so that we cannot see them. The idea of higher dimensions has motivated alternative cosmological models.

Questions

Review Questions

1. What is the horizon problem? What is the flatness problem? How can these problems be resolved by the idea of inflation?
2. The inflationary epoch lasted a mere 10^{-32} second. Why, then, is it worthy of so much attention by scientists?
3. In what ways is inflation similar to the present-day expansion of the universe? In what ways is it different?
4. Explain why the inflationary model does not violate the principle that the speed of light in a vacuum represents the ultimate speed limit.
5. Describe an example of each of the four basic types of interactions in the physical universe. Do you think it possible that a fifth force might be discovered someday? Explain your answer.
6. If gravity is intrinsically so weak compared to the strong force, why do we say that gravity rather than the strong force keeps the planets in orbit around the Sun?
7. Explain how changes in the energy of the vacuum can account for the rapid expansion during the inflationary epoch.
8. What is the Heisenberg uncertainty principle? How does it lead to the idea that all space is filled with virtual particle-antiparticle pairs?
9. What is the difference between an electron and a positron?
10. Is it possible for a single hydrogen atom, with a positively charged proton and a negatively charged electron, to be created as a virtual pair? Why or why not?
11. Which can exist for a longer time, a virtual electron-positron pair or a virtual proton-antiproton pair? Explain your reasoning.
12. Explain why antimatter was present in copious amounts in the early universe but is very rare today.
13. What is meant by the threshold temperature of a particle?
14. Explain the connection between the fact that humans exist and the imbalance between matter and antimatter in the early universe.
15. Explain the connection between particles and antiparticles in the early universe and the cosmic microwave background that we observe today.
16. What is the deuterium bottleneck? Why was it important during the formation of nuclei in the early universe?
17. Why were only the four lightest chemical elements produced in the early universe?
18. The first stars in the universe are thought to have appeared some 400 million (4×10^8) years after the Big Bang. Once these stars formed, thermonuclear fusion reactions began in their interiors. Explain why these were the first fusion reactions to occur since the universe was 15 minutes old.
19. Why is it reasonable to suppose that all space is filled with a neutrino background analogous to the cosmic microwave background?
20. What is the Jeans length? Why is it significant for the formation of structure in the universe?
21. Why do astronomers suspect that globular clusters were among the first objects to form in the history of the universe? Why not something larger and more massive?
22. What are Population III stars? How do they differ from stars found in the present-day universe? Why are they so difficult to detect directly?
23. Describe the large-scale structure of the universe as revealed by the distribution of clusters and superclusters of galaxies.
24. What is the difference between hot and cold dark matter? How do astronomers decide which was more important in the formation of large-scale structures such as clusters of galaxies?
25. How did the presence of dark energy help to "turn off" the process of structure formation in the universe?

26. Describe the observational evidence for (**a**) the Big Bang, (**b**) the inflationary epoch, (**c**) the confinement of quarks, and (**d**) the era of recombination.

Advanced Questions

Problem-solving tips and tools

We described Wien's law for blackbody radiation in Section 5-4. If light with wavelength λ_0 is emitted by an object at redshift z, the wavelength that we measure is $\lambda = \lambda_0(1 + z)$. As explained in Section 26-6, the critical density ρ_c is equal to $3H_0{}^2/8\pi G$; that is, ρ_c is proportional to the square of the Hubble constant H_0. If H_0 = 73 km/s/Mpc, the critical density is equal to about 1.0×10^{-26} kg/m^3. It is also useful to know that the mass of a proton is 1.67×10^{-27} kg, the mass of an electron is 9.11×10^{-31} kg, the mass of the Sun is 1.99×10^{30} kg, that 1 m^3 = 10^6 cm^3, that 1 light-year = 9.46×10^{15} m, and that 1 GeV = 10^3 MeV = 10^9 eV.

27. An electron has a lifetime of 1.0×10^{-8} s in a given energy state before it makes a transition to a lower state. What is the uncertainty in the energy of the photon emitted in this process?
28. How many times stronger than the weak force is the electromagnetic force? How many times stronger than the electromagnetic force is the strong force? Use this information to suggest one reason why the electromagnetic and weak forces can become unified at a lower energy than do the electroweak and strong forces.
29. How long can a proton-antiproton pair exist without violating the principle of the conservation of mass?
30. The mass of the intermediate vector boson W^+ (and of its antiparticle, the W^-) is 85.6 times the mass of the proton. The weak nuclear force involves the exchange of the W^+ and the W^-. (**a**) Find the rest energy of the W^+. Give your answer in GeV. (**b**) Find the threshold temperature for the W^+ and W^-. (**c**) From Figure 27-6, how long after the Big Bang did W^+ and W^- particles begin to disappear from the universe? Explain.
31. Using the physical conditions present in the universe during the era of recombination (T = 3000 K and $\rho_m = 10^{-18}$ kg/m^3), show by calculation that the Jeans length for the universe at that time was about 100 ly and that the total mass contained in a sphere with this diameter was about 4×10^5 $M_\odot$.
32. (**a**) If a Population III star had a surface temperature of 10^5 K, what was its wavelength of maximum emission? In what part of the electromagnetic spectrum does this wavelength lie? (**b**) To ionize a hydrogen atom requires a photon of wavelength 91.2 nm or shorter. Explain how Population III stars caused reionization. (**c**) If reionization occurred at z = 11, what do we measure the wavelength of maximum emission of a Population III star to be? In what part of the electromagnetic spectrum does this wavelength lie? (**d**) The image that opens this chapter was made using infrared wavelengths. Suggest why these wavelengths were chosen.
33. (**a**) If the Hubble constant is 73 km/s/Mpc, the critical density ρ_c is 1.0×10^{-26} kg/m^3. The average density of dark matter is known to be about 0.20 times the critical density. Suppose that massive neutrinos constitute this dark matter, and the average density of neutrinos throughout space is 100 neutrinos per cubic centimeter. (In fact, the density of neutrinos is far less than this.) Under these assumptions, what must be the mass of the neutrino? Give your answers in kilograms and as a fraction of the mass of the electron. (**b**) Why do astronomers think that massive neutrinos are *not* the dominant type of dark matter in the universe?
34. A typical dark nebula (see Figure 18-4) has a temperature of 30 K and a density of about 10^{-12} kg/m^3. (**a**) Calculate the Jeans length for such a dark nebula, assuming that the nebula is mostly composed of hydrogen. Express your answer in meters and in light-years. (**b**) A typical dark nebula is several light-years across. Is it likely that density fluctuations within such nebulae will grow with time? (**c**) Explain how your answer to (b) relates to the idea that protostars form within dark nebulae (see Section 18-3).
35. At the time labeled z = 4.97 in Figure 27-15, how large was the length of each side of the box used in the simulation compared to its size in the present day (z = 0)? How much greater was the density at z = 4.97 than the present-day density?

Discussion Questions

36. If you hold an iron rod next to a strong magnet, the rod will become magnetized; one end will be a north pole and the other a south pole. But if you heat the iron rod to 1043 K (770°C = 1418°F) or higher, it will lose its magnetization and there will be no preferred magnetic direction in the rod. This demagnetization is an example of *restoring* a spontaneously broken symmetry. Explain why.
37. Some GUTs predict that the proton is unstable, although with a half-life far longer than the present age of the universe. What would it be like to live at a time when protons were decaying in large numbers?

Web/eBook Questions

38. Search the World Wide Web for information about the top quark. What kind of particle is it? How does it compare with the up and down quarks found in protons and neutrons? Why did physicists work so hard to try to find it?
39. Search the World Wide Web for information about primordial deuterium (that is, deuterium that was formed in the very early universe). Why are astronomers interested in knowing how abundant primordial deuterium is in the universe? What techniques do they use to detect it?
40. Search the World Wide Web for information about the South Pole Telescope. What is the purpose of this telescope? Why is it to be sited at the South Pole? How will it help us understand the early universe?

Activities

Observing Projects

STARRY NIGHT

41. Use the *Starry Night Enthusiast*™ program to observe globular clusters. First display the entire celestial sphere by selecting **Favourites > Guides > Atlas.** Select **View > Deep Space > Messier Objects** to display this set of diffuse objects in the sky. (**a**) Open the **Find** pane and locate and examine the following globular clusters. In each case, find the approximate angular diameter

of the cluster: (i) M3; (ii) M12; (iii) M13. (**b**) Speculate on how these clusters would appear if you could see them at the same distance at the time of recombination, before the first stars formed.

42. Use the *Starry Night Enthusiast*™ program to examine the distribution of galaxies in our local universe. Select **Favourites > Deep Space > Tully Database** to display the 3-dimensional distribution of the 28,000 galaxies nearest to the Milky Way. Stop **Time** and remove the image of the astronaut's feet by clicking on **View > Feet.** The Milky Way is at the center of the box. You can rotate the box by putting the mouse cursor over the image, holding down the mouse button and Shift key, and moving the mouse. (On a two-button mouse, hold down the left mouse button.) As you rotate this cube of galaxies, you will note the apparent lack of galaxies in one plane, the so-called Zone of Avoidance, caused by the obscuration of the light from distant galaxies in these directions by our own Milky Way Galaxy. You can zoom in or out using the buttons at the upper right of the toolbar. Note particularly the appearance of walls of galaxies, which surround voids in which few galaxies are found, and the clustering of galaxies at the interstices of these walls. Compare the box to the simulated present-day universe shown at the lower right of Figure 27-19. What are the similarities? What are the differences?

Collaborative Exercises

43. The four fundamental forces of nature are the strong force, the weak force, the gravitational force, and the electromagnetic force. List four things at your school that rely on one of these fundamental forces, and explain how each thing is dependent on one of the fundamental forces.

44. Consider the following hypothetical scenario adapted from a daytime, cable television talk show. Chris states that Pat borrowed Chris's telescope without permission. Tyler purchased balloons and a new telescope eyepiece without telling Chris. Sean borrowed star maps from the library, with the library's permission, but without telling Pat. Eventually, when the four met on Sunday evening, Chris was crying and speechless. Can you create a "grand unified theory" that explains this entire situation?

45. The *Cosmic Connection* figure shows the history of the universe in the form of a graph of the temperature versus the time after the Big Bang. Create a similar history of your class, starting with estimated outside temperature on the vertical axis and number of days since the beginning of the academic term on the horizontal axis. Include dates for major exams and assignments up through today. In different color ink, show your predictions for temperatures, days, and events from today until the end of the course.

New Horizons in the Cosmic Microwave Background

by John Ruhl

The WMAP image of the cosmic microwave background radiation (CMB) in Figure 26-14 gives us a beautiful vision of the universe when it was only a few hundred thousand years old. This map was not—in any sense of the phrase—made overnight. We knew for more than a decade that measurements like these would be a wonderful tool for cosmology. Many independent groups toiled, and many years passed, before cutting-edge technology and hard-earned experience made these measurements possible. The result, using data from WMAP and many other ground-based and balloon-borne instruments with higher angular resolution, has helped pin down parameters of the standard cosmological model such as the density parameter Ω_Λ, the dark energy density parameter Ω_0, and the Hubble constant H_0 to a few percent. This is truly precision cosmology—which not long ago would have been an oxymoron!

While the temperature of the CMB (2.725 K) has been well measured, a new frontier awaits. We expect, and in fact two groups of researchers have now found, that the CMB photons are slightly polarized (see Figure 27-4). The level of polarization is extremely small (a few percent of the temperature fluctuations) and very challenging to measure, but it provides us with a new way of probing the early universe. More precise measurements of that polarization will be able to tell us a few very interesting things about the early universe.

First, the prediction of polarization assumes that our current model of the very young universe is correct—specifically, that we understand the nature of the lumps that eventually collapsed into galaxies and clusters of galaxies, and that we understand the process of the decoupling of photons and matter that occurred when the universe was a few hundred thousand years old. We are always looking for ways to test this model, and the polarization provides a very strong one! Second, polarization information can be combined with the measure temperature anisotropies to improve our measurements of cosmological parameters. We recently flew an instrument around Antarctica to make more precise measurements of the CMB polarizations—we are analyzing the data even now—and other groups are ramping up their efforts as well.

There is another, very long-shot—and long-term—goal of polarization measurements that, if successful, would lead to an incredible discovery. The simplest models of inflation predict that gravitational waves (see Section 22-2) were produced as inflation ends. These waves can impress their unique "fingerprints" in the pattern of CMB polarization. If inflation occurred at a sufficiently high energy scale (near 10^{16} GeV), this gravity wave "fingerprint" may be strong enough for us to detect. But if inflation happened at a much lower energy scale, there would be too few gravitational waves for us to see the signature.

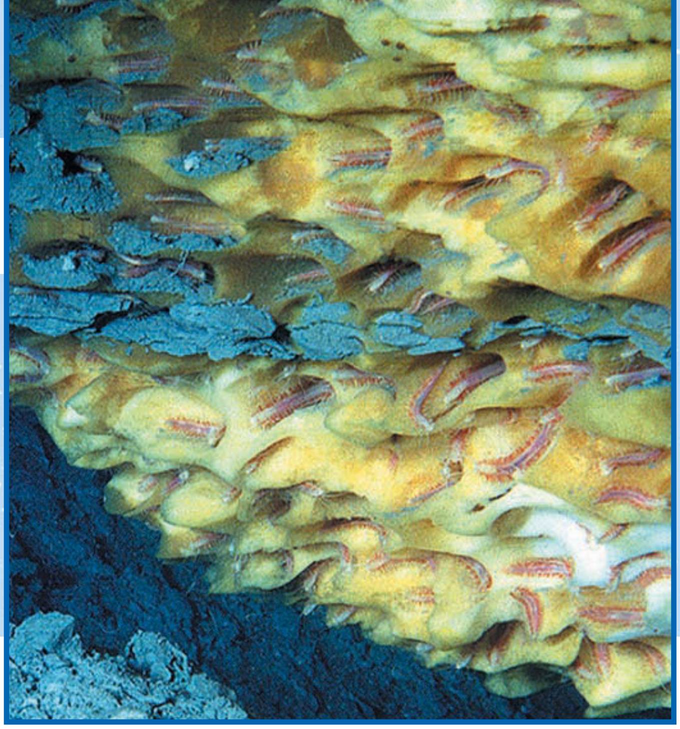
This measurement is very challenging, and not for the faint of heart. But if that signature can be detected, it would give us a fantastic probe of inflation!

Professor John Ruhl has been studying the cosmic microwave background for over a decade, using a variety of ground-based and balloon-borne instruments. His work has taken him to Antarctica eight times, including many seasons at "high and dry" South Pole, one of the world's best sites for millimeter-wave astronomy. Professor Ruhl picked up the astrophysics bug as an undergraduate physics major at the University of Michigan, and did his graduate work in cosmology at Princeton University. Since 2002 he has been a professor of physics and astronomy at Case Western Reserve University.

This page intentionally left blank

28

The Search for Extraterrestrial Life



WEB LINK 28.1

R I V U X G

An "alien" life-form on Earth: pink, eyeless worms in an underwater mound of yellow solid methane. (Dr. Charles Fisher, Eberly College of Science, Pennsylvania State University)

One of the most compelling questions in science is also one of the simplest: Are we alone? That is, does life exist beyond the Earth? As yet, we have no definitive answer to this question. None of our spacecraft has found life elsewhere in the solar system, and radio telescopes have yet to detect signals of intelligent origin coming from space. Reports of aliens visiting our planet and abducting humans make compelling science fiction, but none of these reports has ever been verified.

Yet there are reasons to suspect that life might indeed exist beyond the Earth. One is that biologists find living organisms in some of the most "unearthly" environments on our planet. An example (shown at the top of this page) is at the bottom of the Gulf of Mexico, where the crushing pressure and low temperature cause methane—normally a gas—to form solid, yellowish mounds. Amazingly, these mounds teem with colonies of pink, eyeless, alien-looking worms the size of your thumb. If life can flourish here, might it not also flourish in the equally hostile conditions found on other worlds?

In this chapter we will look for places in our solar system where life may once have originated, and where it may exist today. We will see how scientists estimate the chances of finding life beyond our solar system, and how they search for signals from other intelligent species. And we will learn how a new generation of telescopes may make it possible to detect the presence of even single-celled organisms on worlds many light-years away.

28-1 The chemical building blocks of life are found in space

WEB LINK 28.2

Suppose you were the first visitor to a new and alien planet. How would you recognize which of the strange objects around you were living, and which were inanimate? Questions such as these are central to **astrobiology,** the study of life in the universe. Most astrobiologists suspect that if we find living organisms on other worlds, they will be "life as we know it"—that is, their biochemistry will be based on the unique properties of the carbon atom, as is the case for all terrestrial life.

Organic Molecules in the Universe

Why carbon? The reason is that carbon has the most versatile chemistry of any element. Carbon atoms can form chemical bonds to create especially long and complex molecules (Figure 28-1). These carbon-based compounds, called **organic molecules,** include all the molecules of which living organisms are made. (Silicon has some chemical similarities to carbon, and it can also

Learning Goals

By reading the sections of this chapter, you will learn

28-1 How comets and meteorites could have aided the origin of life on Earth

28-2 On which other worlds of our solar system life might have evolved

28-3 About the controversy over proposed fossil life from Mars

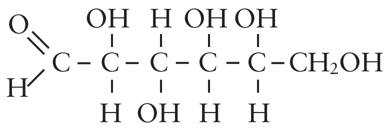
28-4 How astronomers estimate the number of other civilizations in the Galaxy

28-5 Why radio telescopes may be a useful tool for contacting alien civilizations

28-6 How astronomers hope to detect Earthlike planets orbiting other stars

Z – X – X – X – X – Y

(a) Linear molecule



(b) Glucose

Figure 28-1

Complex Molecules and Carbon **(a)** Atoms that can bond to only two other atoms, like the atoms denoted X shown here, can form a chain of atoms called a linear molecule. The chain stops where we introduce an atom, such as those labeled Y and Z, that can bond to only one other atom. **(b)** A carbon atom (denoted C) can bond with up to *four* other atoms. Hence, carbon atoms can form more complex, nonlinear molecules like glucose. All organic molecules that are found in living organisms have backbones of carbon atoms.

form complex molecules. But as Figure 28-2 shows, complex silicon molecules do not have the right properties to make up complex systems such as living organisms.)

Organic molecules can be linked together to form elaborate structures, such as chains, lattices, and fibers. Some of these structures are capable of complex, self-regulating chemical reactions. Furthermore, the primary constituents of organic molecules—carbon, hydrogen, nitrogen, oxygen, sulfur, and phosphorus—are among the most abundant elements in the universe. The versatility and abundance of carbon suggest that extraterrestrial life is also likely to be based on organic chemistry.

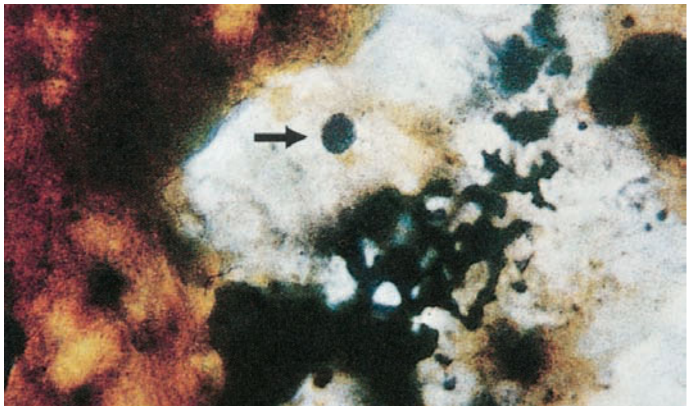
If life is based on organic molecules, then these molecules must initially be present on a planet in order for life to arise from nonliving matter. We now understand that many organic molecules originate from nonbiological processes in interstellar space. One such molecule is carbon monoxide (CO), which is made when a carbon atom and an oxygen atom collide and bond together. Carbon monoxide is found in abundance within giant interstellar clouds that lie along the spiral arms of our Milky Way Galaxy (see Figure 1-7) as well as in other galaxies (see Figure 1-9). Carbon atoms have also combined with other elements to produce an impressive array of interstellar organic molecules, including ethyl alcohol (CH_3CH_2OH), formaldehyde (H_2CO), methyl cyanoacetylene (CH_3C_3N), and acetaldehyde (CH_3CHO). Radio astronomers have detected these by looking for the telltale

(a) (b)

Figure 28-2 R I V U X G

Why Silicon is Unsuitable for Making Living Organisms Like carbon, silicon atoms can bond with up to four other atoms. However, the resulting compounds are either too soft or too hard, or too reactive or too inert, to be suitable for use in living organisms. **(a)** Silicone has a backbone of silicon and oxygen atoms. The molecules form a gel or liquid rather than a solid, and are too inert to undergo the rapid chemical changes required of molecules in organisms. **(b)** Molecules can also be made with a silicon-carbon-oxygen backbone, but the results (like this quartz crystal) are too rigid for use in organisms. (a: Richard Megna/Fundamental Photographs, New York; b: © Mark A. Schneider/Visuals Unlimited)

(a)



(b)

Figure 28-3 R I V U X G

A Carbonaceous Chondrite (a) Carbonaceous chondrites are primitive meteorites that date back to the very beginning of the solar system. This sample is a piece of the Allende meteorite, a large carbonaceous chondrite that fell in Mexico in 1969. **(b)** Chemical analyses of newly fallen specimens disclose that they are rich in organic molecules (arrow), many of which are the chemical building blocks of life. (a: From the collection of Ronald A. Oriti; b: Harvard-Smithsonian Center for Astrophysics)

microwave emission lines of carbon-based chemicals in interstellar clouds.

The planets of our solar system formed out of interstellar material (see Section 8-5), and some of the organic molecules in that material must have ended up on the planets' surfaces. Evidence for this comes from meteorites called **carbonaceous chondrites,** like the one shown in Figure 28-3*a*. These are ancient meteorites that date from the formation of the solar system and that are often found to contain a variety of carbon-based molecules (Figure 28-3*b*). The spectra of comets (see Section 7-5)—which are also among the oldest objects in the solar system—show that they, too, contain an assortment of organic compounds.

Comets and meteoroids were much more numerous in the early solar system than they are today, and they were correspondingly more likely to collide with a planet. These collisions would have seeded the planets with organic compounds from the very beginning of our solar system's history. Similar processes are thought to take place in other planetary systems, which are thought to form in basically the same way as did our own (see Figure 8-13 and the image that opens Chapter 8).

The Miller-Urey Experiment

WEB LINK 28.3

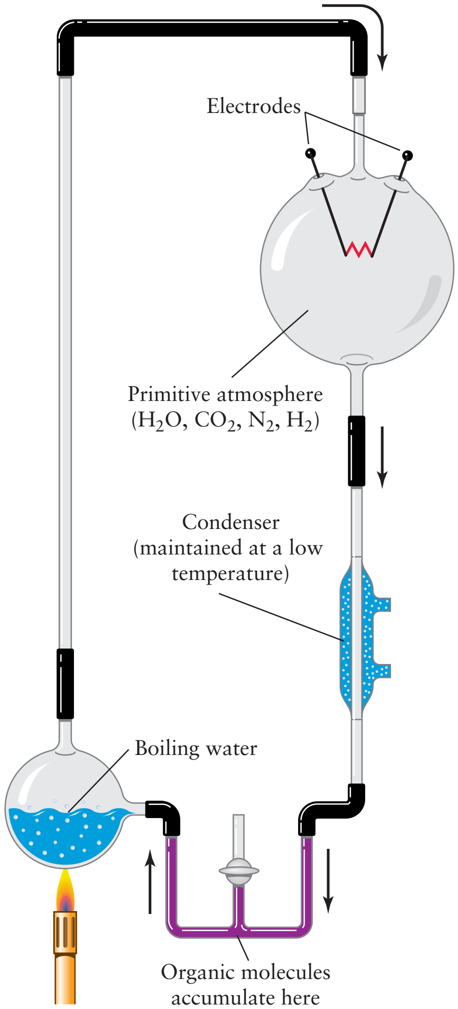
Comets and meteorites would not have been the only sources of organic material on the young planets of our solar system. In 1952, the American chemists Stanley Miller and Harold Urey demonstrated that under conditions that are thought to have prevailed on the primitive Earth, simple chemicals can combine to form the chemical building blocks of life. In a closed container, they prepared a sample of "atmosphere": a mixture of hydrogen (H_2), ammonia (NH_3), methane (CH_4), and water vapor (H_2O), the most common molecules in the solar system. Miller and Urey then exposed this mixture of gases to an electric arc (to simulate atmospheric lightning) for a week. At the end of this period, the inside of the container had become coated with a reddish-brown substance rich in amino acids and other compounds essential to life.

Since Miller and Urey's original experiment, most scientists have come to the conclusion that Earth's primordial atmosphere was composed of carbon dioxide (CO_2), nitrogen (N_2), and water vapor outgassed from volcanoes, along with some hydrogen. Modern versions of the Miller-Urey experiment (Figure 28-4) using these common gases have also succeeded in synthesizing a wide variety of organic compounds. The combination of comets and meteorites falling from space and chemical synthesis in the atmosphere could have made the chemical building blocks of life available in substantial quantities on the young Earth.

Organic molecules do not have to come from living organisms—they can also be synthesized in nature from simple chemicals

CAUTION! It is important to emphasize that scientists have *not* created life in a test tube. While organic molecules may have been available on the ancient Earth, biologists have yet to figure out how these molecules gathered themselves into cells and developed systems for self-replication. Nevertheless, because so many chemical components of life are so easily synthesized under conditions that simulate the primordial Earth, it seems reasonable to suppose that life could have originated as the result of chemical processes. Furthermore, because the molecules that combine to form these compounds are rather common, it seems equally reasonable that life could have originated in the same way on other planets.

Organic building blocks are commonplace throughout the universe, but this does not guarantee that life is equally commonplace. If a planet's environment is hostile, life may never get started or may quickly be extinguished. But we now have evidence



INTERACTIVE EXERCISE 28.1

Figure 28-4

An Updated Miller-Urey Experiment Modern versions of this classic experiment prove that numerous organic compounds important to life can be synthesized from gases that were present in the Earth's primordial atmosphere. This experiment supports the hypothesis that life on the Earth arose as a result of ordinary chemical reactions.

that Jupiter-sized planets orbit other stars (see Section 8-6 and Geoff Marcy's essay "Alien Planets" following Chapter 8) and that additional planetary systems are forming around young stars (see Section 8-4, especially Figure 8-8). It seems probable that there are Earthlike planets orbiting other stars, and that conditions on some of these worlds may be suitable for life as we know it.

28-2 Europa and Mars have the potential for life to have evolved

If life evolved on Earth from nonliving organic molecules, might the same process have taken place elsewhere in our solar system? Scientists have carefully scrutinized the planets and satellites in an attempt to answer this question, and most of the answers have been disappointing.

The Importance of Liquid Water

One major problem is that liquid water is essential for the survival of life as we know it. The water need not be pleasant by human standards—terrestrial organisms have been found in water that is boiling hot, fiercely acidic, or ice cold (see the image that opens this chapter)—but it must be liquid. In order for water on a planet's surface to remain liquid, the temperature cannot be too hot or too cold. Furthermore, there must be a relatively thick atmosphere to provide enough pressure to keep liquid water from evaporating. Of all the worlds of the present-day solar system, only Earth has the right conditions for water to remain liquid on its surface.

WEB LINK 28.4

However, there is now compelling evidence that Europa, one of the large satellites of Jupiter (see Table 7-2 and Figure 7-4), has an ocean of water *beneath* its icy surface. As it orbits Jupiter, Europa is caught in a tug-of-war between Jupiter's gravitational influence and those of the other large satellites. This flexes the interior of Europa, and this flexing generates enough heat to keep subsurface water from freezing. Chunks of ice on the surface can float around on this underground ocean, rearranging themselves into a pattern that reveals the liquid water beneath.

No one knows whether life exists in Europa's ocean. But interest in this exotic little world is great, and scientists have proposed several missions to explore Europa in more detail.

Searching for Life on Mars

The next best possibility for the existence of life is Mars. The present-day Martian atmosphere is so thin that water can exist only as ice or as a vapor. However, images made from Martian orbit show dried-up streambeds, flash flood channels, and sediment deposits. These features are evidence that the Martian atmosphere was once thicker and that water once coursed over the planet's surface. Could life have evolved on Mars during its "wet" period? If so, could life—even in the form of microorganisms—have survived as the Martian atmosphere thinned and the surface water either froze or evaporated?

Some areas on Mars were covered with liquid water for extended periods

In 1976, two spacecraft landed on different parts of Mars in search of answers to these questions. *Viking Lander 1* and *Viking Lander 2* each carried a scoop at the end of a mechanical arm to retrieve surface samples (Figure 28-5). These samples were deposited into a compact on-board biological laboratory that carried out three different tests for Martian microorganisms.

1. The *gas-exchange experiment* was designed to detect any processes that might be broadly considered as respiration. A surface sample was placed in a sealed container along with a controlled amount of gas and nutrients. The gases in the container were then monitored to see if their chemical composition changed.

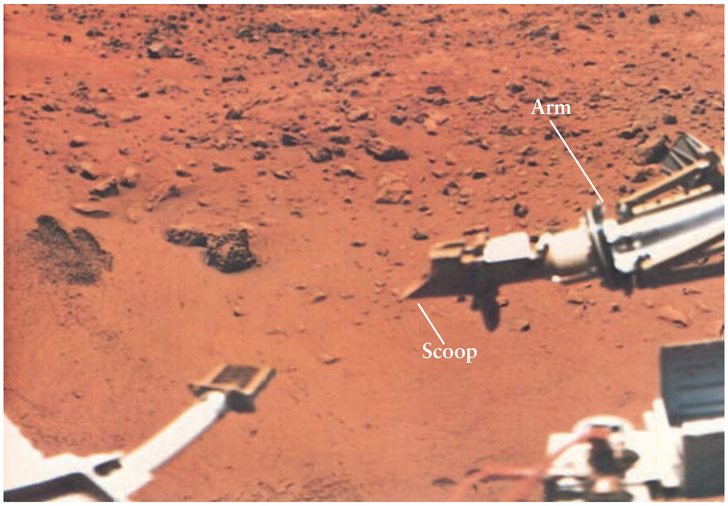


Figure 28-5 RI**V**UXG

Digging in the Martian Surface This view from the *Viking Lander 1* spacecraft shows the mechanical arm with its small scoop against the backdrop of the Martian terrain. The scoop was able to dig about 30 cm (12 in.) beneath the surface. (NASA)

2. The *labeled-release experiment* was designed to detect metabolic processes. A sample was moistened with nutrients containing radioactive carbon atoms. If any organisms in the sample consumed the nutrients, their waste products should include gases containing the telltale radioactive carbon.

3. The *pyrolytic-release experiment* was designed to detect photosynthesis, the biological process by which terrestrial plants use solar energy to help synthesize organic compounds from carbon dioxide. In the *Viking* experiments, a surface sample was placed in a container along with radioactive carbon dioxide and exposed to artificial sunlight. If plantlike photosynthesis occurred, microorganisms in the sample would take in some of the radioactive carbon from the gas.

The first data returned from these experiments caused great excitement, for in almost every case, rapid and extensive changes were detected inside the sealed containers. Further analysis of the data, however, led to the conclusion that these changes were due solely to nonbiological chemical processes. It appears that the Martian surface is rich in unstable chemicals that react with water to release oxygen gas. Because the present-day surface of Mars is bone-dry, these chemicals had nothing to react with until they were placed inside the moist interior of the *Viking Lander* laboratory.

At best, the results from the *Viking Lander* biological experiments were inconclusive. Perhaps life never existed on Mars at all. Or perhaps it did originate there, but failed to survive the thinning of the Martian atmosphere, the unstable chemistry of the planet's surface, and exposure to ultraviolet radiation from the Sun. (Unlike Earth, Mars has no ozone layer to block ultraviolet rays.) Another possibility is that Martian microorganisms have survived only in certain locations that the *Viking Landers* did not sample, such as isolated spots on the surface or deep beneath the ground. And yet another option is that there is life on Mars, but the experimental apparatus on board the *Viking Lander* spacecraft was not sophisticated enough to detect it.

An entirely different set of biological experiments were designed for the British spacecraft *Beagle 2*, which landed on Mars in December 2003. (The spacecraft's name commemorated HMS *Beagle*, the survey ship from which Charles Darwin made many of the observations that led to the theory of evolution.)

Unlike the *Viking Lander* experiments, the apparatus on board *Beagle 2* was designed to test for the presence of either living or dead microorganisms. To do this, the spacecraft was to bore into the interiors of rocks to gather pristine, undisturbed samples, then heat these samples in the presence of oxygen gas. All carbon compounds decompose and form carbon dioxide (CO_2) when treated in this way, but biologically important molecules signal their presence by decomposing at a lower temperature. As a further test for the chemicals of life, *Beagle 2* was to check to see how many of the CO_2 molecules contain the isotope ^{12}C (which appears preferentially in biological molecules) and how many contain ^{13}C (which does not). (See Box 5-5 for a description of isotopes.) A positive result to these experiments would indicate that Martian rocks contain microorganisms that either survive to the present day or that died out at some point in the past.

Another *Beagle 2* experiment was to search for traces of methane in the Martian atmosphere. Microorganisms on Earth can gain energy by converting carbon dioxide to methane, and presumably Martian microorganisms could do the same. Left to itself, methane rapidly decomposes in the Martian atmosphere. Hence, if any methane is found on Mars, it must necessarily have been freshly formed—which would strongly suggest that life exists on Mars today.

Unfortunately, scientists on Earth were unable to establish contact with *Beagle 2* after landing, and so no results were returned from these experiments. But scientists hope to send a replacement lander to Mars using a set of experiments like those on board *Beagle 2*.

New Evidence for Martian Water

While the promise of *Beagle 2* will have to be fulfilled by a future mission, scientists searching for evidence of Martian life have been encouraged by other spacecraft that have provided new evidence of water on Mars. The European Space Agency's *Mars Express* spacecraft, which went into orbit around Mars in December 2003, used its infrared cameras to examine the ice cap at the Martian south pole (Figure 28-6). These cameras allowed scientists to see through the ice cap's surface layer of frozen carbon dioxide and reveal an underlying layer with the characteristic spectrum of water ice. (Figure 7-4 shows a similar spectrum obtained from Europa.) In January 2004, NASA successfully landed two robotic rovers named *Spirit* and *Opportunity* at two very different sites on opposite sides of Mars. While the terrain around the *Spirit* landing site appears to have been dry for billions of years, *Opportunity* landed in an area that appears to have been under water for extended periods (Figure 28-7*a*). Measurements made by

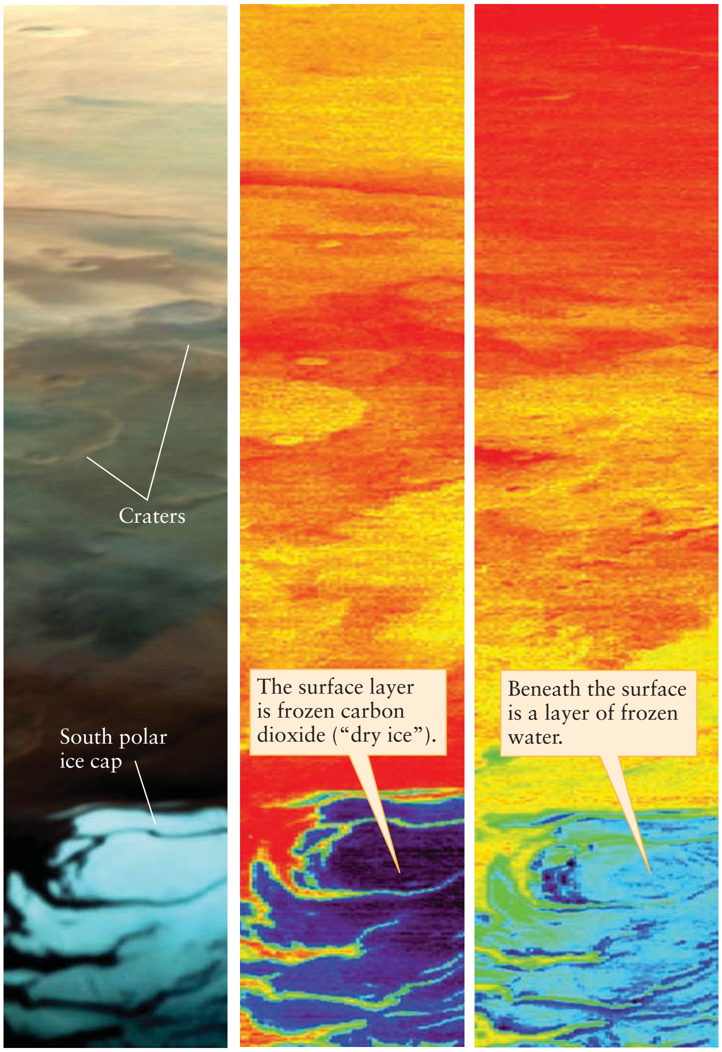


Figure 28-6

Water at the Martian South Pole **(a)** This visible-light image shows the south polar cap of Mars, but does not indicate its chemical composition. But by using a camera tuned to different wavelengths of infrared light, the *Mars Express* spacecraft was able to identify the distinctive reflections of **(b)** an upper layer of carbon dioxide ice and **(c)** a deeper layer of water ice. Other observations have shown that there is also water ice at the Martian north pole. (ESA-OMEGA)

Opportunity confirm that some of the very dark surface material at its landing site contains an iron-rich mineral called gray hematite (Figure 28-7*b*). On Earth, deposits of gray hematite are commonly found at the bottoms of lakes or mineral hot springs. The presence of gray hematite at the *Opportunity* site reinforces the argument that Mars once had liquid water on its surface, and helps hold open the possibility that living organisms could have evolved on Mars.

The Martian Civilization That Never Was

In 1976, while the *Viking Lander*s were carrying out their biological experiments on the Martian surface, the companion *Viking Orbiter* spacecraft photographed some surface features that at first glance seemed to have been crafted by *intelligent* life on Mars. The *Viking Orbiter 1* image in Figure 28-8*a* shows what appears to be a humanlike face, perhaps the product of an advanced and artistic civilization. However, when the more

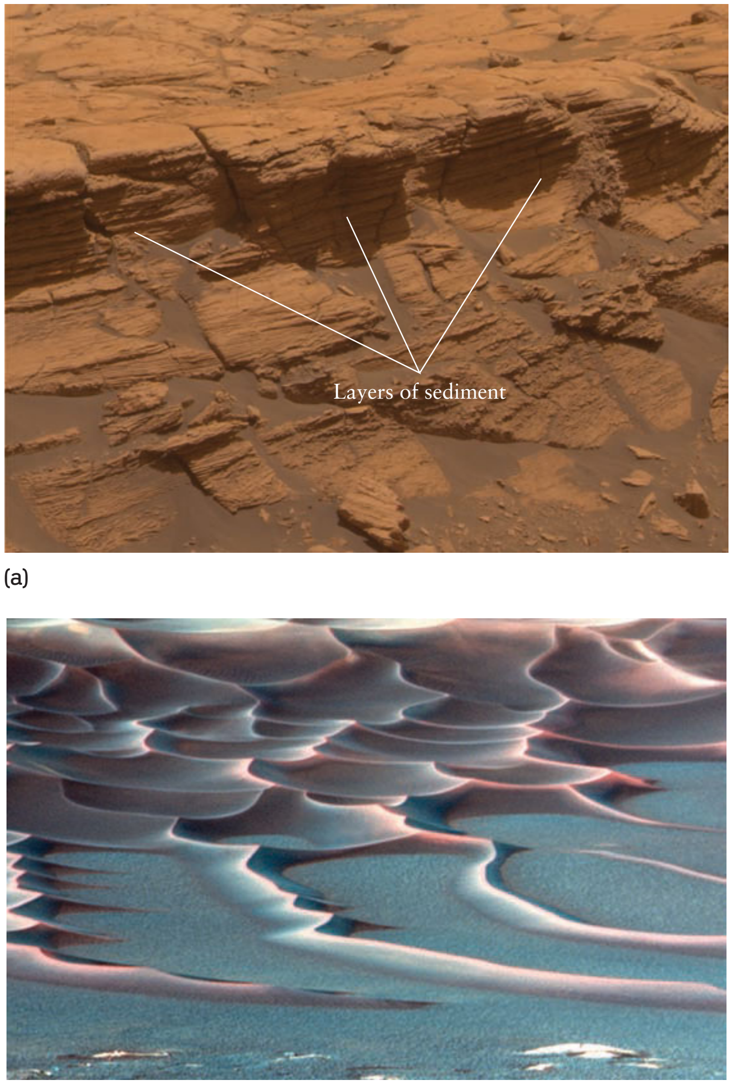
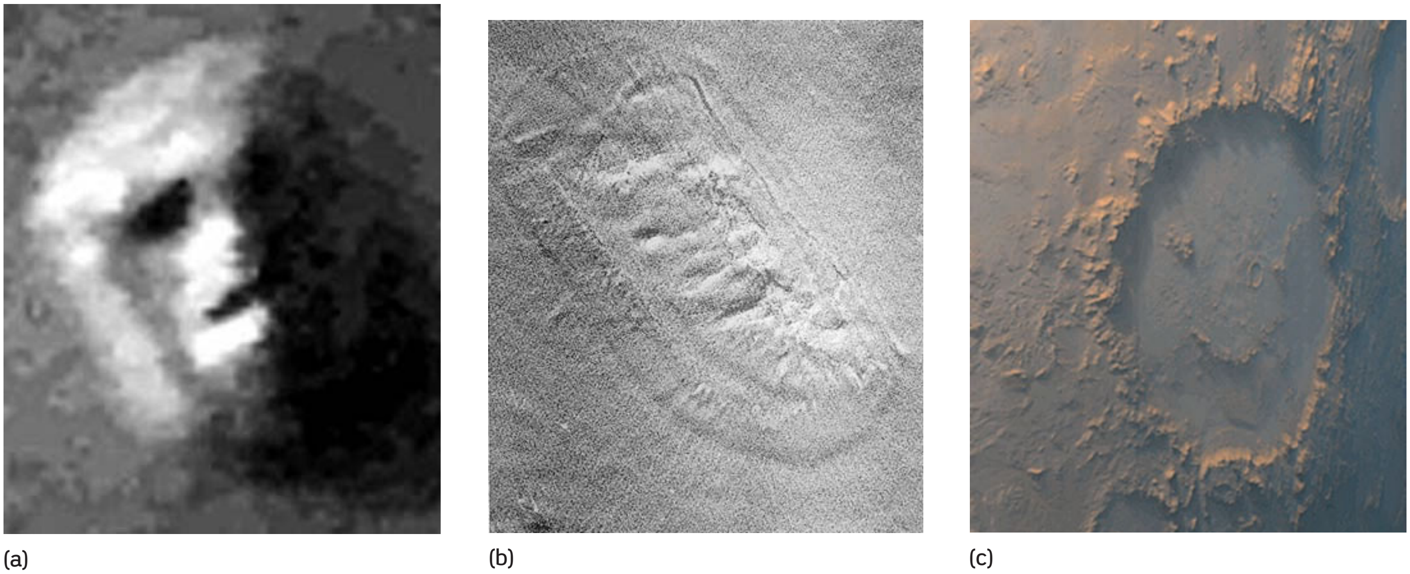


Figure 28-7 R I V U X G

Evidence of Ancient Martian Water **(a)** The Mars rover *Opportunity* photographed these sedimentary layers in a region called Meridiani Planum. Some of the layers are made of dust deposited by the Martian winds, but others were laid down by minerals that precipitated out of standing water. **(b)** In this false-color image from *Opportunity* of Martian sand dunes, the bluish color shows the presence of millimeter-sized spheres of gray hematite. Such spheres naturally form in water-soaked deposits. (NASA/JPL/Cornell)



WEB LINK 28.9

Figure 28-8 R I V U X G

A "Face" on Mars? (a) This 1976 image from *Viking Orbiter 1* shows a Martian surface feature that resembles a human face. Some suggested that this feature might have been made by intelligent beings. **(b)** This 1998 *Mars Global Surveyor* (MGS) image, made under different lighting conditions with a far superior camera, reveals the "face" to be just an eroded hill. **(c)** This MGS image shows features of natural origin within a 215-km (134-mi) wide crater on Mars. Can you see this "face"? (a: NSSDC/NASA and Dr. Michael H. Carr; b, c: Malin Space Science Systems/NASA)

advanced *Mars Global Surveyor* spacecraft viewed the surface in 1998 using a superior camera (Figure 28-8*b*), it found no evidence for facial features.

Scientists are universally convinced that the "face" and other apparent patterns in the *Viking Orbiter* images were created by shadows on wind-blown hills. Microscopic life may once have existed on Mars, and may yet exist today, but there is no evidence that the red planet has ever been the home of intelligent beings.

28-3 Meteorites from Mars have been scrutinized for life-forms

While spacecraft can carry biological experiments to other worlds such as Mars, many astrobiologists look forward to the day when a spacecraft will return Martian samples to laboratories on Earth. Until that day arrives, we have the next best thing: A dozen meteorites that appear to have formed on Mars have been found at a variety of locations on the Earth.

These meteorites are called **SNC meteorites** after the names given to the first three examples found (Shergotty, Nakhla, and Chassigny). What identifies SNC meteorites as having come from Mars is the chemical composition of trace amounts of gas trapped within them. This composition is very different from that of the Earth's atmosphere, but is a nearly perfect match to the composition of the Martian atmosphere found by the *Viking Landers*.

How could a rock have traveled from Mars to Earth? When a large piece of space debris collides with a planet's surface and forms an impact crater, most of the material thrown upward by the impact falls back onto the planet's surface. But some extraordinarily powerful impacts produce large craters—on Mars, roughly 100 km in diameter or larger. These tremendous impacts eject some rocks with such speed that they escape the planet's gravitational attraction and fly off into space.

There are numerous large craters on Mars, so a good number of Martian rocks have probably been blasted into space over the planet's history. These ejected rocks then go into elliptical orbits around the Sun. A few such rocks will have orbits that put them on a collision course with the Earth, and these are the ones that scientists find as SNC meteorites.

Using the radioactive age-dating technique (see Section 8-3), scientists find that most SNC meteorites are between 200 million and 1.3 billion years old, much younger than the 4.56-billion-year age of the solar system. But one SNC meteorite, denoted by the serial number ALH 84001 and found in Antarctica in 1984, was discovered in 1993 to be 4.5 billion years old (Figure 28-9*a*). Thus, ALH 84001 is a truly ancient piece of Mars. Analysis of ALH 84001 suggests that it was fractured by an impact between 3.8 and 4.0 billion years ago, was ejected from Mars by another impact 16 million years ago, and landed in Antarctica a mere 13,000 years ago.

ALH 84001 is the only known specimen of a rock that was on Mars during the era when liquid water existed on the planet's surface. Scientists have therefore investigated its chemical composition carefully, in the hope that this rock may contain clues to the amount of water that once flowed on the Martian surface. One such clue is the presence of rounded grains of minerals called carbonates, which can form only in the presence of water.

(a)

(b)

WEB LINK 28.10

Figure 28-9 R I V U X G

A Meteorite from Mars **(a)** This 1.9-kg meteorite, known as ALH 84001, formed on Mars some 4.5 billion years ago. About 16 million years ago a massive impact blasted it into space, where it drifted in orbit around the Sun until landing in Antarctica 13,000 years ago. The small cube at lower right is 1 cm (0.4 in.) across. **(b)** This electron microscope image, magnified some 100,000 times, shows tubular structures about 100 nanometers (10^{-7} m) in length found within the Martian meteorite ALH 84001. One controversial interpretation is that these are the fossils of microorganisms that lived on Mars billions of years ago. (a: NASA Johnson Space Center; b: *Science,* NASA)

In 1996, David McKay and Everett Gibson of the NASA Johnson Space Center, along with several collaborators, reported the results of a two-year study of the carbonate grains in ALH 84001. They made three remarkable findings. First, in and around the carbonate grains were large numbers of elongated, tubelike structures resembling fossilized microorganisms (Figure 28-9*b*). Second, the carbonate grains contain very pure crystals of iron sulfide and magnetite. These two compounds are rarely found together (especially in the presence of carbonates) but can be produced by certain types of bacteria. Indeed, about one-fourth of the magnetite crystals found in ALH 84001 are of a type that on Earth are formed only by bacteria. Third, the carbonates contain organic molecules—just the sort, in fact, that result from the decay of microorganisms.

McKay and Gibson concluded that the structures seen in Figure 28-9*b* are fossilized remains of microorganisms. If so, these organisms lived and died on Mars billions of years ago, during the era when liquid water was abundant.

Are McKay and Gibson's conclusions correct? Their claims of ancient life on Mars are extraordinary, and they require extraordinary proof. With only one rock like ALH 84001 known to science, however, such proof is hard to come by, and many scientists are skeptical. They argue that the structures found in ALH 84001 could have been formed in other ways that do not require the existence of Martian microorganisms. Future spacecraft may help resolve the controversy by examining rocks on the Martian surface. For now, the existence of microscopic life on Mars in the distant past remains an open question.

Claims that scientists have found Martian microorganisms are intriguing but very controversial

28-4 The Drake equation helps scientists estimate how many civilizations may inhabit our galaxy

We have seen that only a few locations in our solar system may have been suitable for the origin of life. But what about other planetary systems? The development of life on the Earth seems to suggest that extraterrestrial life, including intelligent species, might evolve on terrestrial planets around other stars, given sufficient time and hospitable conditions. How can we learn whether such worlds exist, given the tremendous distances that separate us from them? This is the great challenge facing the **search for extraterrestrial intelligence,** or **SETI.**

Are we alone, or does the Galaxy teem with intelligent life? Or is the truth somewhere in between?

Close Encounters Versus Remote Communication

WEB LINK 28.11

A tenet of modern folklore is the belief that alien civilizations do exist, and that their spacecraft have visited Earth. Indeed, surveys show that between one-third and one-half of all Americans believe in unidentified flying objects (UFOs). A somewhat smaller percentage believes that aliens have landed on Earth. But, in fact, there is *no* scientifically verifiable evidence of alien visitations. As an example, many UFO proponents believe that the U.S. government is hiding evidence of an alien spacecraft that crashed near Roswell, New Mexico, in 1947. However, the bits of "spacecraft wreckage" found near Roswell turn out to be nothing more than remnants of an unmanned research balloon. To find real evidence of the presence or absence of intelligent civilizations on worlds orbiting other stars, we must look elsewhere.

With our present technology, sending even a small unmanned spacecraft to another star requires a flight time of tens of thousands of years. Speculative design studies have been made for unmanned probes that could reach other stars within a century or less, but these are prohibitively expensive. Instead, many astronomers hope to learn about extraterrestrial civilizations by detecting radio transmissions from them. Radio waves are a logical choice for interstellar communication because they can travel immense distances without being significantly degraded by the interstellar medium, the thin gas and dust found between the stars (see Section 8-1).

Over the past several decades, astronomers have proposed various ways to search for alien radio transmissions, and several searches have been undertaken. In 1960, Frank Drake first used a radio telescope at the National Radio Astronomy Observatory in West Virginia to listen to two Sunlike stars, Tau Ceti and Epsilon Eridani, without success. More than 60 more extensive SETI searches have taken place since then, using radio telescopes around the world. Occasionally, a search has detected an unusual or powerful signal. But none has ever repeated, as a signal of intelligent origin might be expected to do. To date, we have no confirmed evidence of radio transmissions from another world.

Are They Out There?

INTERACTIVE EXERCISE 28.2

Should we be discouraged by this failure to make contact? What are the chances that a radio astronomer might someday detect radio signals from an extraterrestrial civilization? The first person to tackle this issue was Frank Drake, who proposed that the number of technologically advanced civilizations in the Galaxy could be estimated by a simple equation. This is now called the **Drake equation:**

Drake equation

$$N = R_* \, f_p \, n_e \, f_l \, f_i \, f_c \, L$$

N = number of technologically advanced civilizations in the Galaxy whose messages we might be able to detect

R_* = the rate at which solar-type stars form in the Galaxy

f_p = the fraction of stars that have planets

n_e = the number of planets per solar system that are Earthlike (that is, suitable for life)

f_l = the fraction of those Earthlike planets on which life actually arises

f_i = the fraction of those life-forms that evolve into intelligent species

f_c = the fraction of those species that develop adequate technology and then choose to send messages out into space

L = the lifetime of a technologically advanced civilization

The Drake equation is enlightening because it expresses the number of extraterrestrial civilizations in a simple series of terms. We can estimate some of these terms from what we know about stars, stellar evolution, and planetary orbits. The *Cosmic Connections* figure depicts some of these considerations.

For example, the first two factors, R_* and f_p, can be determined by observation. In estimating R_*, we should probably exclude stars with masses greater than about 1.5 times that of the Sun. These more massive stars use up the hydrogen in their cores in 3 billion (3×10^9) years or less. On Earth, by contrast, human intelligence developed only within the last million years or so, some 4.56 billion years after the formation of the solar system. If that is typical of the time needed to evolve higher life-forms, then a star of 1.5 solar masses or more probably fades away or explodes into a supernova before creatures as intelligent as we can evolve on any of that star's planets.

Although stars less massive than the Sun have much longer lifetimes, they, too, seem unsuited for life because they are so dim. Only planets very near a low-mass star would be sufficiently warm for life as we know it, and a planet that close is subject to strong tidal forces from its star. We saw in Section 4-8 how the Earth's tidal forces keep the Moon locked in synchronous rotation, with one face continually facing the Earth. In the same way, a planet that orbits too close to its star would have one hemisphere that always faced the star, while the other hemisphere would be in perpetual, frigid darkness.

This leaves us with stars not too different from the Sun. (Like Goldilocks sampling the three bears' porridge, we must have a star that is not too hot and not too cold, but just right.) Based on statistical studies of star formation in the Milky Way, some astronomers estimate that roughly one of these Sunlike stars forms in the Galaxy each year in the galactic **habitable zone** (see the *Cosmic Connections* figure). This sets R_* at 1 per year.

As we saw in Sections 8-4 and 8-5, the planets in our solar system formed as a natural consequence of the birth of the Sun. We have also seen evidence suggesting that planetary formation may be commonplace around single stars (see Figure 8-8). Many astronomers suspect that most Sunlike stars probably have planets, and so they give f_p a value of 1.

Unfortunately, the rest of the terms in the Drake equation are very uncertain. Let's play with some hypothetical values. The chances that a planetary system has an Earthlike world suitable for life are not known. Were we to consider our own solar system as representative, we could put n_e at 1. Let's be more conservative, however, and suppose that 1 in 10 solar-type stars is orbited by a habitable planet, making $n_e = 0.1$. From what we know about the evolution of life on the Earth, we might assume that, given appropriate conditions, the development of life is a certainty, which would make $f_l = 1$. This is an area of intense interest to astrobiologists.

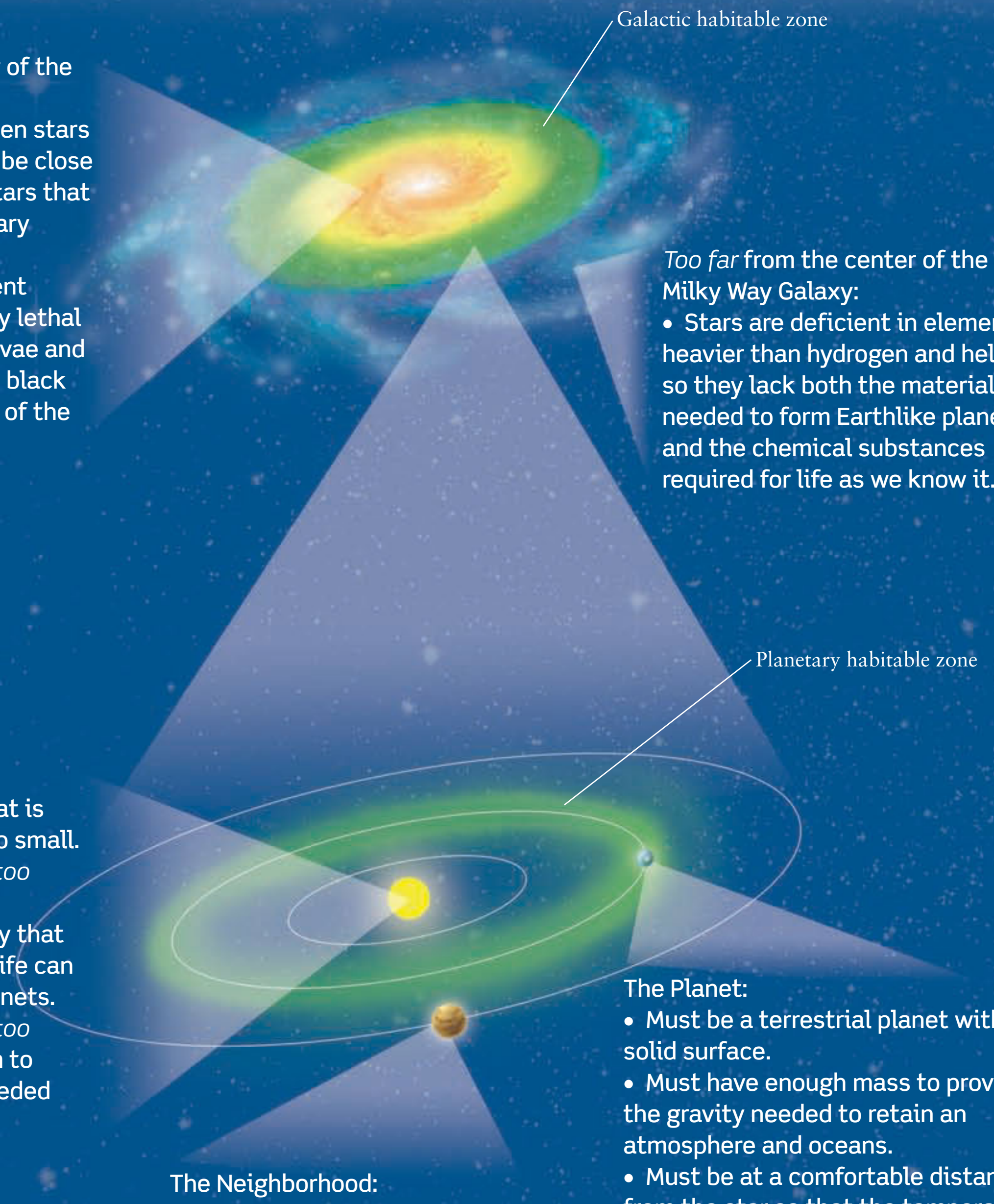
For the sake of argument, we might also assume that evolution might naturally lead to the development of intelligence (a conjecture that is hotly debated) and also make $f_i = 1$. It's anyone's guess as to whether these intelligent extraterrestrial beings would attempt communication with other civilizations in the Galaxy, but were we to assume that they would, f_c would be put at 1 also.

The last variable, L, involving the longevity of a civilization, is the most uncertain of all. Looking at our own example, we see

COSMIC CONNECTIONS

Habitable Zones for Life

Intelligent civilizations in our Milky Way Galaxy can evolve only in a certain region called the galactic habitable zone. In that zone, a suitable planet must lie within the planetary habitable zone of its parent star. (After C. H. Lineweaver, Y. Fenner, and B. K. Gibson)



Too close to the center of the Milky Way Galaxy:
- The distances between stars are small, so there can be close encounters between stars that would disrupt a planetary system.
- There are also frequent outbursts of potentially lethal radiation from supernovae and from the supermassive black hole at the very center of the Galaxy.

Too far from the center of the Milky Way Galaxy:
- Stars are deficient in elements heavier than hydrogen and helium, so they lack both the materials needed to form Earthlike planets and the chemical substances required for life as we know it.

The Star:
- Must have a mass that is neither too large, or too small.
- If the star's mass is *too large*, it will use up its hydrogen fuel so rapidly that it will burn out before life can evolve on any of its planets.
- If the star's mass is *too small*, it will be too dim to provide the warmth needed for life.

The Neighborhood:
- There needs to be one or more large Jovian planets whose gravitational forces will clear away comets and meteors.

The Planet:
- Must be a terrestrial planet with a solid surface.
- Must have enough mass to provide the gravity needed to retain an atmosphere and oceans.
- Must be at a comfortable distance from the star so that the temperatures are neither too high nor too low.
- Must be in a stable, nearly circular orbit. (A highly elliptical orbit would cause excessively large temperature swings as the planet moved toward and away from the star.)

a planet whose atmosphere and oceans are increasingly polluted by creatures that possess nuclear weapons. If we are typical, perhaps L is as short as 100 years. Putting all these numbers together, we arrive at

$$N = 1/\text{year} \times 1 \times 0.1 \times 1 \times 1 \times 1 \times 100 \text{ years} = 10$$

In other words, out of the hundreds of billions of stars in the Galaxy, we would estimate that there are only 10 technologically advanced civilizations from which we might receive communications.

A wide range of values has been proposed for the terms in the Drake equation, and these various guesses produce vastly different estimates of N. Some scientists argue that there is exactly one advanced civilization in the Galaxy and that we are it. Others speculate that there may be hundreds or thousands of planets inhabited by intelligent creatures. If we wish to know whether our Galaxy is devoid of other intelligence, teeming with civilizations, or something in between, we must keep searching the skies.

28-5 Radio searches for alien civilizations are under way

Even if only a few alien civilizations are scattered across the Galaxy, we have the technology to detect radio transmissions from them. But if other civilizations are trying to communicate with us using radio waves, what frequency are they using? This is an important question, because if we fail to tune our radio telescopes to the right frequency, we might never know whether the aliens are out there.

A reasonable choice would be a frequency that is fairly free of interference from extraneous sources. SETI pioneer Bernard Oliver was the first to draw attention to a range of relatively noise-free frequencies in the neighborhood of the microwave emission lines of hydrogen (H) and hydroxide (OH) (Figure 28-10). This region of the microwave spectrum is called the **water hole,** because H and OH together make H_2O, or water.

In 1989, NASA began work on the High Resolution Microwave Survey (HRMS), an ambitious project to scan the entire sky at frequencies spanning the water hole from 10^3 to 10^4 MHz. HRMS would have observed more than 800 nearby solar-type stars over a narrower frequency range in the hope of detecting signals that were either pulsed (like Morse code) or continuous (like the carrier wave for a TV or radio broadcast). The sophisticated signal-processing technology of HRMS would have been able to sift through tens of millions of individual frequency channels simultaneously. It would even have been able to detect the minute Doppler shifts in a signal coming from an alien planet as that planet spun on its axis and moved around its star.

Sadly, just one year after HRMS began operation in 1992, the U.S. Congress imposed a mandate requiring that NASA no longer support HRMS or any other radio searches for extraterrestrial intelligence. This decision, which was made on budgetary grounds, saved a few million dollars—an entirely negligible amount compared to the total NASA budget. Ironically, the senator who spearheaded this move was from the state of Nevada, where tax dollars have been spent to signpost a remote desert road as "The Extraterrestrial Highway."

Even though NASA funding is no longer available, several teams of scientists remain actively involved in SETI programs. Funding for these projects has come from nongovernmental organizations such as the Planetary Society and from private individuals. Since 1995 the SETI Institute in California has been carrying out Project Phoenix, the

Millions of personal computers have helped scan for alien radio transmissions

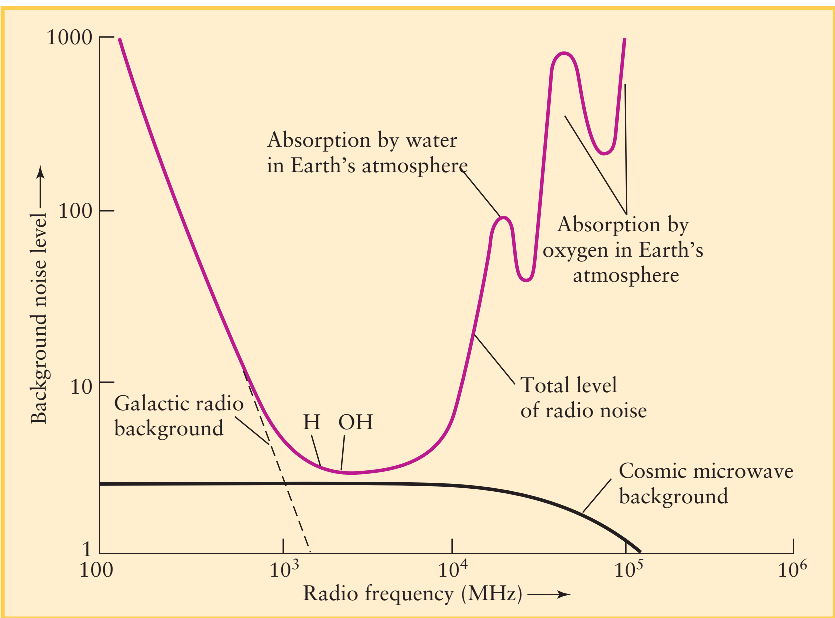


Figure 28-10

The Water Hole This graph shows the background noise level from the sky at various radio and microwave frequencies. The so-called water hole is a range of radio frequencies from about 10^3 to 10^4 megahertz (MHz) in which there is little noise and little absorption by the Earth's atmosphere. Some scientists suggest that this noise-free region would be well suited for interstellar communication. Within the water hole itself, the principal source of noise is the afterglow of the Big Bang, called the cosmic microwave background. To put this graph in perspective, a frequency of 100 MHz corresponds to "100" on a FM radio, and 10^3 MHz is a frequency used for various types of radar. (Adapted from C. Sagan and F. Drake) direct successor to HRMS. When complete, this project will have surveyed a thousand Sunlike stars within 200 light-years at millions of radio frequencies. At Harvard University, BETA (the *B*illion-channel *E*xtra*T*errestrial *A*ssay) is scanning the sky at even more individual frequencies within the water hole. Other multifrequency searches are being carried out under the auspices of the University of Western Sydney in Australia and the University of California.

WEB LINK 28.14 A major challenge facing SETI is the tremendous amount of computer time needed to analyze the mountains of data returned by radio searches. To this end, scientists at the University of California, Berkeley, have recruited nearly 5 million personal computer users to participate in a project called SETI@home. Each user receives actual data from a detector called SERENDIP IV (*S*earch for *E*xtraterrestrial *R*adio *E*missions from *N*earby, *D*eveloped, *I*ntelligent *P*opulations) and a data analysis program that also acts as a screensaver. When the computer's screensaver is on, the program runs, the data are analyzed, and the results are reported via the Internet to the researchers at Berkeley. The program then downloads new data to be analyzed. As of 2006, SETI@home users had provided as much computer time as a single computer working full-time for 2,000,000 years!

All current SETI projects make use of existing radio telescopes and must share telescope time with astronomical researchers. But by 2008, the SETI Institute plans to have completed and put into operation a radio telescope that will be dedicated solely to the search for intelligent signals. This telescope, called the Allen Telescope Array, will actually be hundreds of relatively small and inexpensive radio dishes working together. Perhaps this new array will be the first to detect a signal from a distant civilization.

28-6 Infrared telescopes in space will soon begin searching for Earthlike planets

Although no longer involved in SETI, NASA is planning a major effort to search for Earthlike planets suitable for the evolution of an advanced civilization. Such a search poses a major challenge. Astronomers have discovered many Jupiter-sized planets by detecting the "wobble" that these planets produce in their parent star (see Section 8-6). But a planet the size of the Earth would exert only a weak gravitational force on its parent star, and the resulting "wobble" is too small for us to detect with present technology. Earth-sized planets are also too dim to be seen in visible light against the glare of their parent star.

Searching for Planetary Transits

WEB LINK 28.15 An alternative technique will be used by an orbiting telescope called Kepler, which is scheduled for launch in 2008. If a star is orbited by a planet whose orbital plane is oriented edge-on to our line of sight, once per orbit the planet will pass in front of the star in an event called a *transit*. This causes a temporary dimming of the light we see from that star. As we discussed in Section 8-6, astronomers have used Earth-based telescopes to observe dimming of this kind from Jupiter-sized planets transiting their parent stars. To detect the much slighter dimming caused by the transit of a small, Earth-sized planet, Kepler will use specialized detectors. Furthermore, by observing from orbit using a telescope with a wide field of view, Kepler will be able to continuously monitor thousands of stars at once.

If Kepler detects stars that dim slightly as expected, astronomers will have to make sure that the dimming really is due to a transiting planet and not some other cause. (One key test will be whether the dimming repeats with a definite period, as would be expected for a planet in a periodic orbit.) Once this is done, astronomers will be able to determine the transiting planets' sizes (which determines how much dimming takes place), as well as the sizes of their orbits (which can be calculated from Kepler's third law by using the orbital period of the planet, which is the same as the time interval between successive transits). Given the distance from the planet to its parent star and the star's luminosity, astronomers will even be able to estimate the planet's average temperature.

Searching Using Images and Spectra

WEB LINK 28.16 The results from Kepler may help astronomers select stars to study in more detail using Darwin, a more advanced orbiting telescope under development by the European Space Agency. Planned for a 2015 launch, Darwin will search for Earthlike planets by detecting their infrared radiation. The rationale is that stars like the Sun emit much less infrared radiation than visible light, while planets are relatively strong emitters of infrared. Hence, observing in the infrared makes it less difficult (although still technically challenging) to detect planets orbiting a star.

Darwin will also analyze the infrared spectra of any planets that it finds, in the hope of seeing the characteristic absorption of atmospheric gases such as ozone, carbon dioxide, and water vapor (Figure 28-11). The relative amounts of these gases, as determined from a planet's spectrum, can reveal whether life is present on that planet.

Darwin will need to achieve enough resolution to detect individual planets. One proposed mission design makes use of interferometry. We discussed this technique for improving the resolution of telescopes in Section 6-6. By combining the light from three widely spaced dishes, each at least 3 m in diameter, Darwin will make the sharpest infrared images of any telescope in history. NASA has proposed a similar planet-finding interferometry mission called Terrestrial Planet Finder, but the funding for this mission is uncertain.

Future telescope technology may make it possible to resolve details on planets orbiting other stars

A more speculative project is an infrared telescope with sufficient resolution that some detail would be visible in the image of an extrasolar planet. One concept for such a mission would consist of five Darwin-type telescopes flying in a geometrical formation some 6000 km across (equal to the radius of the Earth). All five telescopes would collect light from the same extrasolar planet, then reflect it onto a single mirror. The combined light would go to detectors on board a sixth spacecraft. The

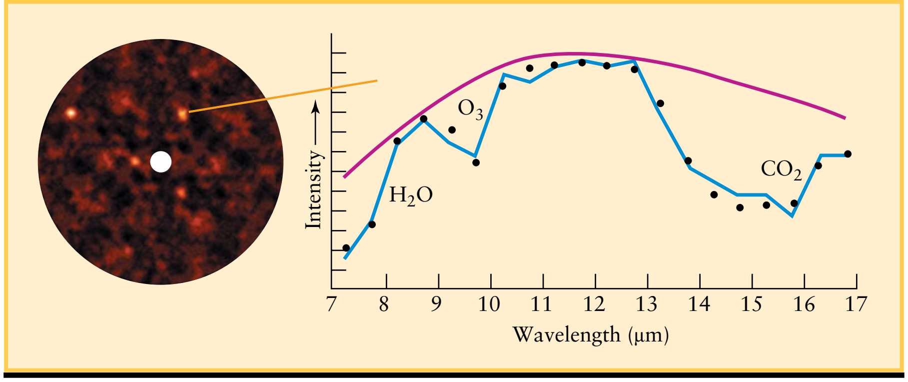


Figure 28-11

The Spectrum of a Simulated Planet The image on the left is a simulation of what the Terrestrial Planet Finder infrared telescope might see once it is launched. The white dot at the center is a nearby Sunlike star, and the smaller dots around it are planets orbiting the star. On the right is the simulated infrared spectrum of one of the planets, showing broad absorption lines of water vapor (H_2O), ozone (O_3), and carbon dioxide (CO_2). While all these molecules can be created by nonbiological processes, the presence of life will change the relative amounts of each molecule in the planet's atmosphere. Thus, the infrared spectrum of such planets will make it possible to identify worlds on which life may have evolved. (Jet Propulsion Laboratory)

technology needed for such an ambitious mission does not yet exist, but may become available within a few decades.

Sometime during the twenty-first century, missions such as Darwin may answer the question "Are there worlds like Earth orbiting other stars?" If the answer is yes, radio searches for intelligent signals will gain even more impetus.

The potential rewards from such searches are great. Detecting a message from an alien civilization could dramatically change the course of our own civilization, through the sharing of scientific information with another species or an awakening of social or humanistic enlightenment. In only a few years our technology, industry, and social structure might advance the equivalent of centuries into the future. Such changes would touch every person on the Earth. Mindful of these profound implications, scientists push ahead with the search for extraterrestrial intelligence.

Key Words

astrobiology, p. 28-1
carbonaceous chondrite, p. 28-3
Drake equation, p. 28-9
habitable zone, p. 28-9
organic molecules, p. 28-1
search for extraterrestrial intelligence (SETI), p. 28-8
SNC meteorite, p. 28-7
water hole, p. 28-11

Key Ideas

Organic Molecules in the Universe: All life on Earth, and presumably on other worlds, depends on organic (carbon-based) molecules. These molecules occur naturally throughout interstellar space.

• The organic molecules needed for life to originate were probably brought to the young Earth by comets or meteorites. Another likely source for organic molecules is chemical reactions in the Earth's primitive atmosphere. Similar processes may occur on other worlds.

Life in the Solar System: Besides Earth, only two worlds in our solar system—the planet Mars and Jupiter's satellite Europa—may have had the right conditions for the origin of life.

• Mars once had liquid water on its surface, though it has none today. Life may have originated on Mars during the liquid water era.

• The *Viking Lander* spacecraft searched for microorganisms on the Martian surface, but found no conclusive sign of their presence. The unsuccessful *Beagle 2* mission to Mars was to carry out a different set of biological experiments on samples taken from the interiors of rocks; these experiments may be attempted again on a future mission.

• An ancient Martian rock that came to Earth as a meteorite shows circumstantial evidence that microorganisms once existed on Mars. Additional rock samples are needed to provide corroboration.

• Europa appears to have extensive liquid water beneath its icy surface. Future missions may search for the presence of life there.

Radio Searches for Extraterrestrial Intelligence: Astronomers have carried out a number of searches for radio signals from other stars. No signs of intelligent life have yet been detected, but searches are continuing and using increasingly sophisticated techniques.

• The Drake equation is a tool for estimating the number of intelligent, communicative civilizations in our Galaxy.

Telescope Searches for Earthlike Planets: A new generation of orbiting telescopes may be able to detect terrestrial planets around nearby stars. If such planets are found, their infrared spectra may reveal the presence or absence of life.

Questions

Review Questions

1. Why are extreme life-forms on Earth, such as those shown in the photograph that opens this chapter, of interest to astrobiologists?
2. What is meant by "life as we know it"? Why do astrobiologists suspect that extraterrestrial life is likely to be of this form?
3. How have astronomers discovered organic molecules in interstellar space? Does this discovery mean that life of some sort exists in the space between the stars?
4. Mercury, Venus, and the Moon are all considered unlikely places to find life. Suggest why this should be.
5. Many science-fiction stories and movies—including *The War of the Worlds, Invaders from Mars, Mars Attacks!*, and *Martians, Go Home*—involve invasions of Earth by intelligent beings from Mars. Why Mars rather than any of the other planets?
6. Summarize the differences in philosophy between the biological experiments on board the *Viking Landers* and those on board *Beagle 2*.
7. What arguments can you give against the idea that the "face" on Mars (Figure 28-8) is of intelligent origin? What arguments can you give in favor of this idea?
8. Suppose someone brought you a rock that he claimed was a Martian meteorite. What scientific tests would you recommend be done to test this claim?
9. Why are most searches for extraterrestrial intelligence made using radio telescopes? Why are most of these carried out at frequencies between 10^3 MHz and 10^4 MHz?
10. Explain why infrared telescopes like those proposed for Darwin and Terrestrial Planet Finder need to be placed in space.

Advanced Questions

Problem-solving tips and tools

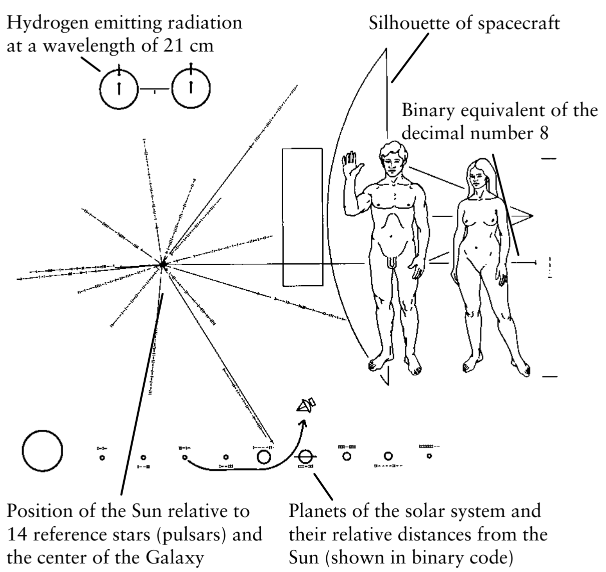
The small-angle formula, discussed in Box 1-1, will be useful. Section 5-2 gives the relationship between wavelength and frequency, while Section 5-9 and Box 5-6 discuss the Doppler effect. Section 6-3 gives the relationship between the angular resolution of a telescope, the telescope diameter, and the wavelength used. You will find useful data about the planets in Appendix 1.

11. In 1802, when it seemed likely to many scholars that there was life on Mars, the German mathematician Karl Friedrich Gauss proposed that we signal the Martian inhabitants by drawing huge geometric patterns in the snows of Siberia. His plan was never carried out. (**a**) Suppose patterns had been drawn that were 1000 km across. What minimum diameter would the objective of a Martian telescope need to have to be able to resolve these patterns? Assume that the observations are made at a wavelength of 550 nm, and assume that Earth and Mars are at their minimum separation. (**b**) Ideally, the patterns used would be ones that could not be mistaken for natural formations. They should also indicate that they were created by an advanced civilization. What sort of patterns would you have chosen?
12. Assume that all the terms in the Drake equation have the values given in the text, except for N and L. (**a**) If there are 1000 civilizations in the Galaxy today, what must be the average lifetime of a technological civilization? (**b**) What if there are a million such civilizations?
13. (**a**) Of the visually brightest stars in the sky listed in Appendix 5, which might be candidates for having Earthlike planets on which intelligent civilizations have evolved? Explain your selection criteria. (**b**) Repeat part (a) for the nearest stars, listed in Appendix 4.
14. It has been suggested that extraterrestrial civilizations would choose to communicate at a wavelength of 21 cm. Hydrogen atoms in interstellar space naturally emit at this wavelength, so astronomers studying the distribution of hydrogen around the Galaxy would already have their radio telescopes tuned to receive extraterrestrial signals. (**a**) Calculate the frequency of this radiation in megahertz. Is this inside or outside the water hole? (**b**) Discuss the merits of this suggestion.
15. Imagine that a civilization in another planetary system is sending a radio signal toward Earth. As our planet moves in its orbit around the Sun, the wavelength of the signal we receive will change due to the Doppler effect. This gives SETI scientists a way to distinguish stray signals of terrestrial origin (which will not show this kind of wavelength change) from interstellar signals. (**a**) Use the data in Appendix 1 to calculate the speed of the Earth in its orbit. For simplicity, assume the orbit is circular. (**b**) If the alien civilization is transmitting at a frequency of 3000 MHz, what wavelength (in meters) would we receive if the Earth were moving neither toward nor away from their planet? (**c**) The maximum Doppler shift occurs if the Earth's orbital motion takes it directly toward or directly away from the alien planet. How large is that maximum wavelength shift? Express your answer both in meters and as a percentage of the unshifted wavelength you found in (b). (**d**) Discuss why it is important that SETI radio receivers be able to measure frequency and wavelength to very high precision.
16. Astronomers have proposed using interferometry to make an extremely high-resolution telescope. This proposal involves placing a number of infrared telescopes in space, separating them by thousands of kilometers, and combining the light from the individual telescopes. One design of this kind has an effective diameter of 6000 km and uses infrared radiation with a wavelength of 10 mm. If it is used to observe an Earthlike planet orbiting the star Epsilon Eridani, 3.22 parsecs (10.5 light-years) from Earth, what is the size of the smallest detail that this system will be able to resolve on the face of that planet? Give your answer in kilometers.

Discussion Questions

17. Suppose someone told you that the *Viking Landers* failed to detect life on Mars simply because the tests were designed to detect terrestrial life-forms, not Martian life-forms. How would you respond?

18. Science-fiction television shows and movies often depict aliens as looking very much like humans. Discuss the likelihood that intelligent creatures from another world would have (**a**) a biochemistry similar to our own, (**b**) two legs and two arms, and (**c**) about the same dimensions as a human.
19. The late, great science-fiction editor John W. Campbell exhorted his authors to write stories about organisms that think as well as humans, but not *like* humans. Discuss the possibility that an intelligent being from another world might be so alien in its thought processes that we could not communicate with it.
20. If a planet always kept the same face toward its star, just as the Moon always keeps the same face toward Earth, most of the planet's surface would be uninhabitable. Discuss why.
21. How do you think our society would respond to the discovery of intelligent messages coming from a civilization on a planet orbiting another star? Explain your reasoning.
22. What do you think will set the limit on the lifetime of our technological civilization? Explain your reasoning.
23. The first of all Earth spacecraft to venture into interstellar space were *Pioneer 10* and *Pioneer 11*, which were launched in 1972 and 1973, respectively. Their missions took them past Jupiter and Saturn and eventually beyond the solar system. Both spacecraft carry a metal plaque with artwork (reproduced below) that shows where the spacecraft is from and what sort of creatures designed it. If an alien civilization were someday to find one of these spacecraft, which of the features on the plaque do you think would be easily understandable to them? Explain.



Web/eBook Questions

24. Any living creatures in the subsurface ocean of Europa would have to survive without sunlight. Instead, they might obtain energy from Europa's inner heat. Search the World Wide Web for information about "black smokers," which are associated with high-temperature vents at the bottom of Earth's oceans. What kind of life is found around black smokers? How do these life-forms differ from the more familiar organisms found in the upper levels of the ocean?
25. Search the World Wide Web for information about the *Mars Express* orbiter and the *Spirit* and *Opportunity* rovers. What discoveries have these missions made about water on Mars? Have they found any evidence that liquid water has existed on Mars in the recent past? Describe the evidence, if any.
26. Like other popular media, the World Wide Web is full of claims of the existence of "extraterrestrial intelligence"—namely, UFO sightings and alien abductions. (**a**) Choose a Web site of this kind and analyze its content using the idea of *Occam's razor,* the principle that if there is more than one viable explanation for a phenomenon, one should choose the simplest explanation that fits all the observed facts. (**b**) Read what a skeptical Web site has to say about UFO sightings. A good example is the Web site of the Committee for the Scientific Investigation of Claims of the Paranormal, or CSICOP. After considering what you have read on both sides of the UFO debate, discuss your opinions about whether aliens really have landed on Earth.
27. **The Drake Equation.** Access the Active Integrated Media Module "The Drake Equation" in Chapter 28 of the *Universe* Web site or eBook. (**a**) For each of the terms in the Drake equation, choose a value that seems reasonable to you. How did you choose these values? Using the module, what do you find for the number of civilizations in our Galaxy? From your calculation, are civilizations common or uncommon in our Galaxy? (**b**) Using the module, choose a set of values that give $N = 10^6$ (a million civilizations). What values did you use? Which of these seem reasonable to you, and why?

Activities

Observing Projects

28. Use the *Starry Night Enthusiast*™ program to view the Earth as it might be seen by a visiting spacecraft. First, select **Viewing Location . . .** in the **Options** menu and set the viewing location to your city or town from the list of cities provided or click on the **Map** tab in the **Viewing Location** pane and use the mouse to click on your approximate position on the world map. Then click the **Set Location** button. Set the local time to 12:00:00 P.M. (noon). To see the Earth from space, use the up and down elevation buttons on the toolbar to raise yourself above the surface until you can see the entire Earth. You can use the scrollbars (select **View > Show Scrollbars**) on the right side and bottom of the window to center the Earth in your view. The Earth can be rotated to allow you to see different locations by clicking and moving the mouse while its icon, a four-way arrow, is over the Earth's image. (**a**) Describe any features you see that suggest life could exist on Earth. Explain your reasoning. (**b**) Using the controls at the right-hand end of the toolbar, zoom in to show more detail around your city or town. The amount of detail is comparable to the view from a spacecraft a few million kilometers away. Can you see any

evidence that life does exist on Earth? (**c**) From a distance of a few million kilometers, are there any measurements that a spacecraft could carry out to prove that life exists on Earth? Explain your reasoning.

29. Use the *Starry Night Enthusiast™* program to examine the planet Mars. Open the **Favourites** pane and double-click on **Solar System > Mars** to view this planet from about 6800 km above its surface. (Click on **View > Feet** to remove the astronaut's spacesuit from the view.) You can zoom in or out on Mars using the buttons in the Zoom section at the right of the toolbar. You can rotate Mars by placing the mouse cursor over the image and moving the mouse while holding down the mouse button. (On a two-button mouse, hold down the left mouse button.) Rotate Mars and zoom in and out to familiarize yourself with the different surface features. Based on what you observe, where on the Martian surface would you choose to land a spacecraft to search for the presence of life? Explain how you made your choice.

Collaborative Exercise

30. Imagine that astronomers have discovered intelligent life in a nearby star system. Your group is submitting a proposal for who on Earth should speak for the planet and what 50-word message should be conveyed. Prepare a maximum one-page proposal that states (**a**) who should speak for Earth and why; (**b**) what this person should say in 50 words; and (**c**) why this message is the most important compared to other things that could be said. Only serious responses receive full credit.

A Biologist's View of Astrobiology

by Kevin W. Plaxco

In the early 1960s, flush with the excitement of Sputnik and the first manned space missions, the Nobel Prize-winning biologist Joshua Lederberg coined the word "exobiology" to describe the scientific study of extraterrestrial life. But after a brief flurry of popularity in the 1960s and 1970s (due partly to the pioneering research and public outreach work of Carl Sagan), exobiology fell out of fashion among scientists; it was, after all, the only field of scientific study without any actual subject material to research.

In the last decade, however, the field of astrobiology has filled the void left by exobiology's demise by being simultaneously more encompassing and more practical. Astrobiology removes exobiology's distinction between life on earth and life "out there" and focuses on broader, but more tractable, questions regarding the relationship between life (any life, anywhere) and the universe. In a nutshell, astrobiology uses our significant (if incomplete) knowledge of life on Earth to address three broad questions about life in the universe:

- Which of the physical attributes of our universe were necessary in order for it to support the origins and further evolution of life?
- How did life arise and evolve on Earth, and how might it have arisen and evolved elsewhere?
- Where else might life have arisen in our universe, and how do we best search for it?

Much of the worth and appeal of astrobiology lie in the fact that these questions, which are among the most fundamental posed by science, address the most profound issues regarding our origins and our nature.

Perhaps not surprisingly, some aspects of astrobiology are far better understood than others. We understand, for example, the broad details of how the almost metal-free hydrogen and helium left over after the Big Bang was enriched in the metallic elements by fusion reactions occurring in the more massive stars and how these elements—absolutely critical for biology because only they support the complex chemistry required for life—were then returned to the cosmos in supernova explosions. Similarly, we understand in detail how rocky planets—which we see as potentially habitable—form out of the dust and gas of prestellar nebulae. In contrast, however, much of the *biology* in astrobiology remains speculative at best. There is, for example, nothing even remotely resembling consensus regarding how life first arose on Earth much less how—or if—it might have arisen elsewhere in the cosmos.

Perhaps equally unsurprisingly, some branches of the astrobiology community are more "optimistic" than others. Astronomers generally view astrobiology in terms of the processes by which potentially habitable planets form. As these processes are reasonably well understood, astronomers tend to be rather upbeat regarding the events surrounding life's origins. The radio astronomer Frank Drake, who pioneered the search for extraterrestrial intelligence (SETI) in the 1960s, is a relatively extreme example of this hopefulness: When trying to estimate the number of intelligent, "communicating" civilizations in the Galaxy, he assumed that rocky, water-soaked planets almost invariably give rise to life. Biochemists, in contrast, tend to view the key issues in astrobiology as relating to the first steps in the origins of life—that is, the detailed processes by which inanimate chemicals first combine to form a living, evolving, self-replicating system. To date, about a half dozen theories have been put forth to explain how this might have occurred on Earth, but all are fraught with enormous scientific difficulties. Specifically, every single theory of the origins of life postulated to date either utterly lacks experimental evidence to back it up or requires events that are so improbable they make even the "astronomical" number of planets in the universe look small by comparison. When faced with what is, in effect, a complete unknown, biochemists tend to be much more cautious about estimating the probability of life arising and the chances of our being alone in the cosmos. To quote Francis Crick, co-discoverer of the structure of DNA, "We cannot decide whether the origin of life on Earth was an extremely unlikely event or almost a certainty, or any possibility in between these two extremes."

Irrespective of our degree of personal optimism or pessimism, though, no one would be more thrilled than us biologists by the detection of life that had arisen independently of life on Earth. Such a find would revolutionize our field by providing the first, tangible evidence of whether there are viable routes to the evolution of life other than the DNA/proteins/cells path that it took on Earth. But even barring such a momentous discovery as the detection of extraterrestrial life, astrobiology still provides a significant service to biology: By placing life into a broader context, astrobiology provides the perspective that we need in order to truly understand who we are and where we come from.

Prior to becoming a professor of biochemistry at the University of California, Santa Barbara, Kevin Plaxco earned his B.Sc. in chemistry from the University of California at Riverside and his Ph.D. in molecular biology from the California Institute of Technology. While his research focuses predominantly on the study of proteins, the "macromolecular machines" that perform the lion's share of the action in our cells, he has also performed NASA-funded research into methods for the detection of extraterrestrial life. In addition, he teaches a course in astrobiology and has published a popular science book on the subject.

This page intentionally left blank

Appendices

Appendix 1 The Planets: Orbital Data

Planet	Semimajor axis (10^6 km)	Semimajor axis (AU)	Sidereal period (years)	Sidereal period (days)	Synodic period (days)	Average orbital speed (km/s)	Orbital eccentricity	Inclination of orbit to ecliptic (°)
Mercury	57.9	0.387	0.241	87.969	115.88	47.9	0.206	7.00
Venus	108.2	0.723	0.615	224.70	583.92	35.0	0.007	3.39
Earth	149.6	1.000	1.000	365.256	—	29.79	0.017	0.00
Mars	227.9	1.524	1.88	686.98	779.94	24.1	0.093	1.85
Jupiter	778.3	5.203	11.86		398.9	13.1	0.048	1.30
Saturn	1429	9.554	29.46		378.1	9.64	0.053	2.48
Uranus	2871	19.194	84.10		369.7	6.83	0.043	0.77
Neptune	4498	30.066	164.86		367.5	5.5	0.010	1.77

Appendix 2 The Planets: Physical Data

Planet	Equatorial diameter (km)	Equatorial diameter (Earth = 1)	Mass (kg)	Mass (Earth = 1)	Average density (kg/m^3)	Rotation period* (solar days)	Inclination of equator to orbit (°)	Surface gravity (Earth = 1)	Albedo	Escape speed (km/s)
Mercury	4880	0.383	3.302×10^{23}	0.0553	5430	58.646	0.5	0.38	0.12	4.3
Venus	12,104	0.949	4.868×10^{24}	0.815	5243	243.01^R	177.4	0.91	0.59	10.4
Earth	12,756	1.000	5.974×10^{24}	1.000	5515	0.997	23.45	1.000	0.39	11.2
Mars	6794	0.533	6.418×10^{23}	0.107	3934	1.026	25.19	0.38	0.15	5.0
Jupiter	142,984	11.209	1.899×10^{27}	317.8	1326	0.414	3.12	2.36	0.44	60.2
Saturn	120,536	9.449	5.685×10^{26}	95.16	687	0.444	26.73	1.1	0.47	35.5
Uranus	51,118	4.007	8.682×10^{25}	14.53	1318	0.718^R	97.86	0.92	0.56	21.3
Neptune	49,528	3.883	1.024×10^{26}	17.15	1638	0.671	29.56	1.1	0.51	23.5

**For Jupiter, Saturn, Uranus, and Neptune, the internal rotation period is given. A superscript* R *means that the rotation is retrograde (opposite the planet's orbital motion).*

Appendix 3 Satellites of the Planets

Planet	Satellite	Date of discovery	Average distance from center of planet (km)	Orbital (sidereal) period (days)*	Orbital eccentricity	Size of satellite (km)**	Mass (kg)
Earth	Moon	—	384,400	27.322	0.0549	3476	7.349×10^{22}
Mars	Phobos	1877	9378	0.319	0.01	$28 \times 23 \times 20$	1.1×10^{16}
	Deimos	1877	23,460	1.263	0.00	$16 \times 12 \times 10$	1.8×10^{15}
Jupiter	Metis	1979	128,000	0.295	0.0012	44	1.2×10^{17}
	Adrastea	1979	129,000	0.298	0.0018	$24 \times 20 \times 16$	7.5×10^{15}
	Amalthea	1892	181,400	0.498	0.0031	$270 \times 200 \times 155$	2.1×10^{18}
	Thebe	1979	221,900	0.675	0.0177	98	1.5×10^{18}
	Io	1610	421,600	1.769	0.0041	3642	8.932×10^{22}
	Europa	1610	670,900	3.551	0.0094	3120	4.791×10^{22}
	Ganymede	1610	1,070,000	7.155	0.0011	5268	1.482×10^{23}
	Callisto	1610	1,883,000	16.689	0.0074	4800	1.077×10^{23}
	Themisto	2000	7,284,000	130.02	0.2426	9	6.9×10^{14}
	Leda	1974	11,165,000	240.92	0.1636	18	1.1×10^{16}
	Himalia	1904	11,461,000	250.56	0.1623	184	6.7×10^{18}
	Lysithea	1938	11,717,000	259.20	0.1124	38	6.3×10^{16}
	Elara	1905	11,741,000	259.64	0.2174	78	8.7×10^{17}
	Kallichore	2000	12,555,000	456.10	0.2480	4	1.5×10^{13}
	S/2003 J12	2003	15,912,000	489.52^{R}	0.6056	1	1.5×10^{12}
	Carpo	2003	16,989,000	456.10	0.4297	3	4.5×10^{13}
	Euporie	2001	19,304,000	550.74^{R}	0.1432	2	1.5×10^{13}
	S/2003 J3	2003	20,221,000	583.88^{R}	0.1970	2	1.5×10^{13}
	S/2003 J18	2003	20,514,000	596.59^{R}	0.0148	2	1.5×10^{13}
	Orthosie	2001	20,720,000	622.56^{R}	0.2808	2	1.5×10^{13}
	Euanthe	2001	20,797,000	620.49^{R}	0.2321	3	4.5×10^{13}
	Harpalyke	2000	20,858,000	623.31^{R}	0.2268	4	1.2×10^{14}
	Praxidike	2000	20,907,000	625.38^{R}	0.2308	7	4.3×10^{14}
	Thyone	2001	20,939,000	627.21^{R}	0.2286	4	9.0×10^{13}
	S/2003 J16	2003	20,963,000	616.36^{R}	0.2245	2	1.5×10^{13}
	Iocaste	2000	21,061,000	631.60^{R}	0.2160	5	1.9×10^{14}
	Mneme	2003	21,069,000	620.04^{R}	0.2273	2	1.5×10^{13}
	Hermippe	2001	21,131,000	633.90^{R}	0.2096	4	9.0×10^{13}
	Thelxinoe	2003	21,162,000	628.09^{R}	0.2206	2	1.5×10^{13}
	Helike	2003	21,263,000	634.77^{R}	0.1558	4	9.0×10^{13}
	Ananke	1951	21,276,000	629.77^{R}	0.2435	28	3.0×10^{16}
	S/2003 J15	2003	22,627,000	689.77^{R}	0.1916	2	1.5×10^{13}
	Eurydome	2001	22,865,000	717.33^{R}	0.2759	3	4.5×10^{13}
	Arche	2002	22,931,000	723.90^{R}	0.2588	3	4.5×10^{13}
	S/2003 J17	2003	23,001,000	714.47^{R}	0.2379	2	1.5×10^{13}
	Pasithee	2001	23,004,000	719.44^{R}	0.2675	2	1.5×10^{13}
	S/2003 J10	2003	23,042,000	716.25^{R}	0.4295	2	1.5×10^{13}
	Chaldene	2000	23,100,000	723.70^{R}	0.2519	4	7.5×10^{13}
	Isonoe	2000	23,155,000	726.25^{R}	0.2471	4	7.5×10^{13}

Appendix 3 Satellites of the Planets

Planet	Satellite	Date of discovery	Average distance from center of planet (km)	Orbital (sidereal) period (days)*	Orbital eccentricity	Size of satellite (km)**	Mass (kg)
JUPITER (continued)	Erinome	2000	23,196,000	728.51^R	0.2665	3	4.5×10^{13}
	Kale	2001	23,217,000	729.47^R	0.2599	2	1.5×10^{13}
	Aitne	2001	23,229,000	730.18^R	0.2643	3	4.5×10^{13}
	Taygete	2000	23,280,000	732.41^R	0.2525	5	1.6×10^{14}
	S/2003 J9	2003	23,384,000	733.29^R	0.2632	1	1.5×10^{12}
	Carme	1938	23,404,000	734.17^R	0.2533	46	1.3×10^{17}
	Sponde	2001	23,487,000	748.34^R	0.3121	2	1.5×10^{13}
	Megaclite	2000	23,493,000	752.88^R	0.4197	6	2.1×10^{14}
	S/2003 J5	2003	23,495,000	738.73^R	0.2478	4	9.0×10^{13}
	S/2003 J19	2003	23,533,000	740.42^R	0.2557	2	1.5×10^{13}
	S/2003 J23	2003	23,563,000	732.44^R	0.2714	2	1.5×10^{13}
	Kalyke	2000	23,566,000	742.03^R	0.2465	5	1.9×10^{14}
	S/2003 J14	2003	23,614,000	779.23^R	0.3439	2	1.5×10^{13}
	Pasiphaë	1908	23,624,000	743.63^R	0.4090	58	3.0×10^{17}
	Eukelade	2003	23,661,000	746.39^R	0.2721	4	9.0×10^{13}
	S/2003 J4	2003	23,930,000	755.24^R	0.3618	2	1.5×10^{13}
	Sinope	1914	23,939,000	758.90^R	0.2495	38	7.5×10^{16}
	Hegemone	2003	23,947,000	739.60^R	0.3276	3	4.5×10^{13}
	Aoede	2003	23,981,000	761.50^R	0.4322	4	9.0×10^{13}
	Kallichore	2003	24,043,000	764.74^R	0.2640	2	1.5×10^{13}
	Autonoe	2000	24,046,000	760.95^R	0.3168	4	9.0×10^{13}
	Callirrhoe	1999	24,103,000	758.77^R	0.2828	7	8.7×10^{14}
	Cyllene	2003	24,349,000	751.91^R	0.3189	2	1.5×10^{13}
	S/2003 J2	2003	29,541,000	979.99^R	0.2255	2	1.5×10^{13}
SATURN	Pan	1981	133,580	0.575	0	$35 \times 35 \times 23$	4.9×10^{15}
	Daphnis	2005	136,500	0.594	0	7	3×10^{14}
	Atlas	1980	137,670	0.602	0.0012	$46 \times 38 \times 19$	6.6×10^{15}
	Prometheus	1980	139,380	0.613	0.0022	$119 \times 87 \times 61$	1.6×10^{17}
	Pandora	1980	141,720	0.629	0.0042	$103 \times 80 \times 64$	1.4×10^{17}
	Epimetheus	1980	151,410	0.694	0.0098	$135 \times 108 \times 105$	5.3×10^{17}
	Janus	1980	151,460	0.695	0.0068	$193 \times 173 \times 137$	1.9×10^{18}
	Mimas	1789	185,540	0.942	0.0196	397	3.8×10^{19}
	Methone	2004	194,440	1.01	0.0001	3	2×10^{13}
	Pallene	2004	212,280	1.154	0.004	4	4×10^{13}
	Enceladus	1789	238,040	1.37	0.0047	504	1.1×10^{20}
	Tethys	1684	294,670	1.888	0.0001	1066	6.2×10^{20}
	Telesto	1980	294,710	1.888	0.0002	$29 \times 22 \times 20$	8×10^{15}
	Calypso	1980	294,710	1.888	0.0005	$30 \times 23 \times 14$	5×10^{15}
	Polydeuces	2004	377,200	2.737	0.0192	3.5	3×10^{13}
	Dione	1684	377,420	2.737	0.0022	1123	1.1×10^{21}
	Helene	1980	377,420	2.737	0.0071	$36 \times 32 \times 30$	2×10^{16}

Appendix 3 Satellites of the Planets

Planet	Satellite	Date of discovery	Average distance from center of planet (km)	Orbital (sidereal) period (days)*	Orbital eccentricity	Size of satellite (km)**	Mass (kg)
SATURN (continued)	Rhea	1672	527,070	4.518	0.001	1528	2.3×10^{21}
	Titan	1655	1,221,870	15.95	0.0288	5150	1.34×10^{23}
	Hyperion	1848	1,500,880	21.28	0.0274	$360 \times 280 \times 225$	5.7×10^{18}
	Iapetus	1671	3,560,840	79.33	0.0283	1472	2.0×10^{21}
	Kiviuq	2000	11,111,000	449.22	0.3288	16	3×10^{15}
	Ijiraq	2000	11,124,000	451.43	0.3163	12	1×10^{15}
	Phoebe	1898	12,947,780	550.31^{R}	0.1635	$230 \times 220 \times 210$	8.3×10^{18}
	Paaliaq	2000	15,200,000	686.93	0.3631	19	4×10^{15}
	Skathi	2000	15,541,000	728.21^{R}	0.2701	6	2×10^{14}
	Albiorix	2000	16,182,000	783.46	0.477	32	3×10^{16}
	S/2004 S11	2004	17,119,000	834.84	0.4691	6	2×10^{14}
	Erriapo	2000	17,343,000	871.18	0.4724	10	9×10^{14}
	Siarnaq	2000	17,531,000	895.55	0.2961	40	4×10^{16}
	S/2006 S8	2006	17,610,000	869^{R}	0.418	6	2×10^{14}
	Tarvos	2000	17,983,000	926.23	0.5305	15	3×10^{15}
	S/2006 S4	2006	18,105,000	905^{R}	0.374	6	2×10^{14}
	S/2004 S19	2004	18,217,125	912^{R}	0.36	8	8×10^{14}
	S/2004 S13	2004	18,403,000	933.45^{R}	0.2586	6	2×10^{14}
	S/2006 S6	2006	18,600,000	942^{R}	0.192	6	2×10^{14}
	Mundilfari	2000	18,685,000	952.67^{R}	0.21	7	3×10^{14}
	S/2006 S1	2006	18,981,135	970^{R}	0.13	6	2×10^{14}
	Narvi	2003	19,007,000	1003.93^{R}	0.4309	7	3×10^{14}
	S/2004 S15	2004	19,338,000	1005.93^{R}	0.1428	6	2×10^{14}
	S/2004 S17	2004	19,447,000	1014.7^{R}	0.1793	4	4×10^{13}
	Suttungr	2000	19,459,000	1016.67^{R}	0.114	7	3×10^{14}
	S/2004 S14	2004	19,856,000	1038.67^{R}	0.3715	6	2×10^{14}
	S/2004 S12	2004	19,878,000	1046.16^{R}	0.3261	5	2×10^{14}
	S/2004 S18	2004	20,129,000	1083.57^{R}	0.5214	7	3×10^{14}
	S/2004 S9	2004	20,390,000	1086.1^{R}	0.2397	5	2×10^{14}
	Thrymr	2000	20,474,000	1094.23^{R}	0.4652	7	3×10^{14}
	S/2004 S10	2004	20,735,000	1116.47^{R}	0.252	6	2×10^{14}
	S/2004 S7	2004	20,999,000	1140.28^{R}	0.5299	6	2×10^{14}
	S/2006 S3	2006	21,132,000	1142^{R}	0.471	6	2×10^{14}
	S/2006 S7	2006	22,290,000	1237^{R}	0.368	6	2×10^{14}
	S/2006 S2	2006	22,350,000	1245^{R}	0.341	7	3×10^{14}
	S/2004 S16	2004	22,453,000	1260.28^{R}	0.1364	4	4×10^{13}
	Ymir	2000	23,040,000	1315.21^{R}	0.335	18	3×10^{15}
	S/2006 S5	2006	23,190,000	1314^{R}	0.139	6	2×10^{14}
	S/2004 S8	2004	25,108,000	1490.87^{R}	0.2064	6	2×10^{14}
URANUS	Cordelia	1986	49,800	0.335	0.0003	50×36	4.4×10^{16}
	Ophelia	1986	53,800	0.376	0.0099	54×38	5.3×10^{16}

Appendix 3 Satellites of the Planets

Planet	Satellite	Date of discovery	Average distance from center of planet (km)	Orbital (sidereal) period (days)*	Orbital eccentricity	Size of satellite (km)**	Mass (kg)
URANUS (continued)	Bianca	1986	59,200	0.435	0.0009	64×46	9.2×10^{16}
	Cressida	1986	61,800	0.464	0.0004	92×74	3.4×10^{17}
	Desdemona	1986	62,700	0.474	0.0001	90×54	1.8×10^{17}
	Juliet	1986	64,400	0.493	0.0007	150×74	5.6×10^{17}
	Portia	1986	66,100	0.513	0.0001	156×126	1.7×10^{18}
	Rosalind	1986	69,900	0.558	0.0001	72	2.5×10^{17}
	Cupid	2003	74,800	0.618	0.0013	18	3.8×10^{15}
	Belinda	1986	75,300	0.624	0.0001	128×64	3.6×10^{17}
	Perdita	1986	76,420	0.638	0.003	26	1.3×10^{16}
	Puck	1985	86,000	0.762	0.0001	162	2.9×10^{18}
	Mab	2003	97,734	0.923	0.0025	24	1.0×10^{16}
	Miranda	1948	129,900	1.413	0.0013	471	6.59×10^{19}
	Ariel	1851	190,900	2.52	0.0012	1158	1.35×10^{21}
	Umbriel	1851	266,000	4.144	0.0039	1169	1.2×10^{21}
	Titania	1787	436,300	8.706	0.0011	1578	3.53×10^{21}
	Oberon	1787	583,500	13.46	0.0014	1522	3.01×10^{21}
	Francisco	2001	4,276,000	266.56	0.1459	22	1.4×10^{15}
	Caliban	1997	7,231,000	579.73^{R}	0.1587	72	7.4×10^{17}
	Stephano	1999	8,004,000	677.36^{R}	0.2292	32	6.0×10^{15}
	Trinculo	2001	8,504,000	749.24^{R}	0.22	18	7.5×10^{14}
	Sycorax	1997	12,179,000	1288.3^{R}	0.5224	150	5.4×10^{18}
	Margaret	2003	14,345,000	1687.01	0.6608	20	1.0×10^{15}
	Prospero	1999	16,256,000	1978.29^{R}	0.4448	25	2.1×10^{16}
	Setebos	1999	17,418,000	2225.21^{R}	0.5914	24	2.1×10^{16}
	Ferdinand	2003	20,901,000	2887.21^{R}	0.3682	21	4.4×10^{16}
NEPTUNE	Naiad	1989	48,227	0.294	0.0004	$96 \times 60 \times 52$	1.9×10^{17}
	Thalassa	1989	50,075	0.311	0.0002	$108 \times 100 \times 52$	3.5×10^{17}
	Despina	1989	52,526	0.335	0.0002	$180 \times 150 \times 130$	2.1×10^{18}
	Galatea	1989	61,953	0.429	0	$204 \times 184 \times 144$	2.1×10^{18}
	Larissa	1989	73,548	0.555	0.0014	$216 \times 204 \times 164$	4.2×10^{18}
	Proteus	1989	117,647	1.122	0.0005	$440 \times 416 \times 404$	4.4×10^{19}
	Triton	1846	354,800	5.877^{R}	0	2706	2.15×10^{22}
	Nereid	1949	5,513,400	360.14	0.7512	340	3.1×10^{19}
	S/2002 N1	2002	15,728,000	1879.71^{R}	0.5711	62	1.8×10^{17}
	S/2002 N2	2002	22,422,000	2914.07	0.2931	44	6.3×10^{16}
	S/2002 N3	2002	23,571,000	3167.85	0.4237	42	5.5×10^{16}
	Psamathe	2003	46,695,000	9115.91^{R}	0.4499	24	1.0×10^{16}
	S/2002 N4	2002	48,387,000	9373.99^{R}	0.4945	60	1.6×10^{17}

This table was compiled from data provided by the Jet Propulsion Laboratory.

**A superscript* R *means that the satellite orbits in a retrograde direction (opposite to the planet's rotation).*

***The size of a spherical satellite is equal to its diameter.*

Appendix 4 The Nearest Stars

Name	Parallax (arcsec)	Distance (parsecs)	Distance (light-years)	Spectral Type	Proper motion (arcsec/yr)	Apparent visual magnitude	Absolute visual magnitude	Mass (Sun = 1)
Proxima Centauri	0.772	1.30	4.22	M5.5 V	3.853	+11.09	+15.53	0.107
Alpha Centauri A	0.747	1.34	4.36	G2 V	3.710	−0.01	+4.36	1.144
Alpha Centauri B	0.747	1.34	4.36	K0 V	3.724	+1.34	+5.71	0.916
Barnard's Star	0.547	1.83	5.96	M4.0 V	10.358	+9.53	+13.22	0.166
Wolf 359	0.419	2.39	7.78	M6.0 V	4.696	+13.44	+16.55	0.092
Lalande 21185	0.393	2.54	8.29	M2.0 V	4.802	+7.47	+10.44	0.464
Sirius A	0.380	2.63	8.58	A1 V	1.339	−1.43	+1.47	1.991
Sirius B	0.380	2.63	8.58	white dwarf	1.339	+8.44	+11.34	0.500
UV Ceti	0.374	2.68	8.73	M5.5 V	3.368	+12.54	+15.40	0.109
BL Ceti	0.374	2.68	8.73	M6.0 V	3.368	+12.99	+15.85	0.102
Ross 154	0.337	2.97	9.68	M3.5 V	0.666	+10.43	+13.07	0.171
Ross 248	0.316	3.16	10.32	M5.5 V	1.617	+12.29	+14.79	0.121
Epsilon Eridani	0.310	3.23	10.52	K2 V	0.977	+3.73	+6.19	0.850
Lacailee 9352	0.304	3.29	10.74	M1.5 V	6.896	+7.34	+9.75	0.529
Ross 128	0.299	3.35	10.92	M4.0 V	1.361	+11.13	+13.51	0.156
EZ Aquarii A	0.290	3.45	11.27	M5.0 V	3.254	+13.33	+15.64	0.105
EZ Aquarii B	0.290	3.45	11.27	—	3.254	+13.27	+15.58	0.106
EZ Aquarii C	0.290	3.45	11.27	—	3.254	+14.03	+16.34	0.095
Procyon A	0.286	3.50	11.40	F5 IV-V	1.259	+0.38	+2.66	1.569
Procyon B	0.286	3.50	11.40	white dwarf	1.259	+10.70	+12.98	0.500
61 Cygni A	0.286	3.50	11.40	K5.0 V	5.281	+5.21	+7.49	0.703
61 Cygni B	0.286	3.50	11.40	K7.0 V	5.172	+6.03	+8.31	0.630
GJ725 A	0.283	3.53	11.53	M3.0 V	2.238	+8.90	+11.16	0.351
GJ725 B	0.283	3.53	11.53	M3.5 V	2.313	+9.69	+11.95	0.259
GX Andromedae	0.281	3.56	11.62	M1.5 V	2.918	+8.08	+10.32	0.486
GQ Andromedae	0.281	3.56	11.62	M3.5 V	2.918	+11.06	+13.30	0.163
Epsilon Indi A	0.276	3.63	11.82	K5 V	4.704	+4.69	+6.89	0.766
Epsilon Indi B	0.276	3.63	11.82	T1.0	4.823			0.044
Epsilon Indi C	0.276	3.63	11.82	T6.0	4.823			0.028
DX Cancri	0.276	3.63	11.83	M6.5 V	1.290	+14.78	+16.98	0.087
Tau Ceti	0.274	3.64	11.89	G8 V	1.922	+3.49	+5.68	0.921
RECONS 1	0.272	3.68	11.99	M5.5 V	0.814	+13.03	+15.21	0.113
YZ Ceti	0.269	3.72	12.13	M4.5 V	1.372	+12.02	+14.17	0.136
Luyten's Star	0.264	3.79	12.37	M3.5 V	3.738	+9.86	+11.97	0.257
Kapteyn's Star	0.255	3.92	12.78	M1.5 V	8.670	+8.84	+10.87	0.393
AX Microscopium	0.253	3.95	12.87	M0.0 V	3.455	+6.67	+8.69	0.600

This table, compiled from data reported by the Research Consortium on Nearby Stars, lists all known stars within 4.00 parsecs (13.05 light-years).

**Stars that are components of multiple star systems are labeled A, B, and C.*

Appendix 5 The Visually Brightest Stars

Name	Distance: Designation	Distance: (parsecs)	Distance: (light-years)	Spectral type	Radial velocity (km/s)*	Proper motion (arcsec/year)	Apparent visual magnitude	Apparent visual brightness (Sirius = 1)**	Absolute visual magnitude
Sirius A	α CMa A	2.63	8.58	A1 V	−7.6	1.34	−1.43	1.000	+1.46
Canopus	α Car	95.9	313	F0 II	+20.5	0.03	−0.72	0.520	−5.63
Arcturus	α Boo	11.3	36.7	K1.5 III	−5.2	2.28	−0.04	0.278	−0.30
Alpha Centauri A	α Cen A	1.34	4.36	G2 V	−25	3.71	−0.01	0.270	+4.36
Vega	α Lyr	7.76	25.3	A0 V	−13.9	0.35	+0.03	0.261	+0.58
Capella	α Aur	12.9	42.2	G5 III	+30.2	0.43	+0.08	0.249	−0.48
Rigel	α Ori A	237	773	B8 Ia	+20.7	0.002	+0.12	0.240	−6.75
Procyon	α CMi A	3.50	11.4	F5 IV-V	−3.2	1.26	+0.34	0.196	+2.62
Achernar	α Eri	44.1	144	B3 V	+16	0.10	+0.50	0.169	−2.72
Betelgeuse	α Ori	131	427	M1 Iab	+21	0.03	+0.58	0.157	−5.01
Hadar	β Cen	161	525	B1 III	+5.9	0.04	+0.60	0.154	−5.43
Altair	α Aql	51.4	168	A7 V	−26.1	0.66	+0.77	0.132	−2.79
Aldebaran	α Tau A	20.0	65.1	K5 III	+54.3	0.20	+0.85	0.122	−0.65
Spica	α Vir	80.4	262	B1 III-IV	+1	0.05	+1.04	0.103	−3.49
Antares	α Sco A	185	604	M1.5 Iab	−3.4	0.03	+1.09	0.098	−5.25
Pollux	β Gem	10.3	33.7	K0 IIIb	+3.3	0.63	+1.15	0.093	+1.08
Fomalhaut	α PsA	7.69	25.1	A3 V	+6.5	0.37	+1.16	0.092	+1.73
Deneb	α Cyg	990	3230	A2 Ia	−4.5	0.002	+1.25	0.085	−8.73
Mimosa	β Cru	108	353	B0.5 IV	+15.6	0.05	+1.297	0.081	−3.87
Regulus	α Leo A	23.8	77.5	B7 V	+5.9	0.25	+1.35	0.077	−0.53

Data in this table were compiled from SIMBAD database operated at the Centre de Données Astronomiques de Strasbourg, France.

**A positive radial velocity means that the star is receding; a negative radial velocity means that the star is approaching.*

***This is a ratio of the star's apparent brightness to that of Sirius, the brightest star in the night sky.*

Note: Acrux, or α Cru (the brightest star in Crux, the Southern Cross) appears to the naked eye as a star of apparent magnitude +0.87, the same as Aldebaran. However, it does not appear in this table because Acrux is actually a binary star system. The blue-white component stars of this binary system have apparent magnitudes of +1.4 and +1.9, and so they are dimmer than any of the stars listed here.

Appendix 6 Some Important Astronomical Quantities

Astronomical Unit	1 AU = 1.4960×10^{11} m
	= 1.4960×10^{8} km
Light-year	1 ly = 9.4605×10^{15} m
	= 63,240 AU
Parsec	1 pc = 3.2616 ly
	= 3.0857×10^{16} m
	= 206,265 AU
Year	1 y = 365.2564 days
	= 3.156×10^{7} s
Solar mass	1 $M_{\odot} = 1.989 \times 10^{30}$ kg
Solar Radius	1 $R_{\odot} = 6.9599 \times 10^{8}$ m
Solar Luminosity	1 $L_{\odot} = 3.90 \times 10^{26}$ W

Appendix 7 Some Important Physical Constants

Speed of light	$c = 2.9979 \times 10^{8}$ m/s
Gravitational constant	$G = 6.673 \times 10^{-11}$ N m^2/kg^2
Planck's constant	$h = 6.6261 \times 10^{-34}$ J s
	$= 4.1357 \times 10^{-15}$ eV s
Boltzmann constant	$k = 1.3807 \times 10^{-23}$ J/K
	$= 8.6173 \times 10^{-5}$ eV/K
Stefan-Boltzmann constant	$\sigma = 5.6704 \times 10^{-8}$ W m^{-2} K^{-4}
Mass of electron	$m_e = 9.1094 \times 10^{-31}$ kg
Mass of proton	$m_p = 1.6726 \times 10^{-27}$ kg
Mass of neutron	$m_n = 1.6749 \times 10^{-27}$ kg
Mass of hydrogen atom	$m_H = 1.6735 \times 10^{-27}$ kg
Rydberg constant	$R = 1.0973 \times 10^{7}$ m^{-1}
Electron volt	1 eV $= 1.6022 \times 10^{-19}$ J

Appendix 8 Some Useful Mathematics

Area of a rectangle of sides a and b	$A = ab$
Volume of a rectangular solid of sides a, b, and c	$V = abc$
Hypotenuse of a right triangle whose other sides are a and b	$c = \sqrt{a^2 + b^2}$
Circumference of a circle of radius r	$C = 2\pi r$
Area of a circle of radius r	$A = \pi r^2$
Surface ares of a sphere of radius r	$A = 4\pi r^2$
Volume of a sphere of radius r	$V = 4\pi r^3/3$
Value of π	$\pi = 3.1415926536$

Glossary

You can find more information about each term in the indicated chapter or chapters. See the Key Words section in each chapter for the specific page on which the meaning of a given term is described.

A ring One of three prominent rings encircling Saturn. (Chapter 12)

absolute zero A temperature of −273°C (or 0 K), at which all molecular motion stops; the lowest possible temperature. (Chapter 5)

absorption line spectrum Dark lines superimposed on a continuous spectrum. (Chapter 5)

acceleration The rate at which an object's velocity changes due to a change in speed, a change in direction, or both. (Chapter 4)

accretion The gradual accumulation of matter in one location, typically due to the action of gravity. (Chapter 8)

active optics A technique for improving a telescopic image by altering the telescope's optics to compensate for variations in air temperature or flexing of the telescope mount. (Chapter 6)

adaptive optics A technique for improving a telescopic image by altering the telescope's optics in a way that compensates for distortion caused by the Earth's atmosphere. (Chapter 6)

aerosol Tiny droplets of liquid dispersed in a gas. (Chapter 13)

albedo The fraction of sunlight that a planet, asteroid, or satellite reflects. (Chapter 9)

amino acids The chemical building blocks of proteins. (Chapter 15)

angle The opening between two lines that meet at a point. (Chapter 1)

angular diameter The angle subtended by the diameter of an object. (Chapter 1)

angular distance The angle between two points in the sky. (Chapter 1)

angular measure The size of an angle, usually expressed in degrees, arcminutes, and arcseconds. (Chapter 1)

angular momentum See *conservation of angular momentum*.

angular resolution The angular size of the smallest feature that can be distinguished with a telescope. (Chapter 6)

angular size See *angular diameter*.

annular eclipse An eclipse of the Sun in which the Moon is too distant to cover the Sun completely, so that a ring of sunlight is seen around the Moon at mid-eclipse. (Chapter 3)

anorthosite Rock commonly found in ancient, cratered highlands on the Moon. (Chapter 10)

Antarctic Circle A circle of latitude $23\frac{1}{2}°$ north of the Earth's south pole. (Chapter 2)

aphelion The point in its orbit where a planet is farthest from the Sun. (Chapter 4)

apogee The point in its orbit where a satellite or the Moon is farthest from the Earth. (Chapter 3)

apparent solar day The interval between two successive transits of the Sun's center across the local meridian. (Chapter 2)

apparent solar time Time reckoned by the position of the Sun in the sky. (Chapter 2)

arcminute One-sixtieth (1/60) of a degree, designated by the symbol ′. (Chapter 1)

arcsecond One-sixtieth (1/60) of an arcminute or 1/3600 of a degree, designated by the symbol ″. (Chapter 1)

Arctic Circle A circle of latitude $23\frac{1}{2}°$ south of the Earth's north pole. (Chapter 2)

asteroid One of tens of thousands of small, rocky, planetlike objects in orbit about the Sun. Also called minor planets. (Chapter 7, Chapter 15)

asteroid belt A region between the orbits of Mars and Jupiter that encompasses the orbits of many asteroids. (Chapter 7, Chapter 15)

asthenosphere A warm, plastic layer of the mantle beneath the lithosphere of the Earth. (Chapter 9)

astrobiology The study of life in the universe. (Chapter 28)

astrometric method A technique for detecting extrasolar planets by looking for stars that "wobble" periodically. (Chapter 8)

astronomical unit (AU) The semimajor axis of the Earth's orbit; the average distance between the Earth and the Sun. (Chapter 1)

atmosphere (atm) A unit of atmospheric pressure. (Chapter 9)

atmospheric pressure The force per unit area exerted by a planet's atmosphere. (Chapter 9)

atom The smallest particle of an element that has the properties characterizing that element. (Chapter 5)

atomic number The number of protons in the nucleus of an atom of a particular element. (Chapter 5, Chapter 8)

AU See *astronomical unit*.

aurora (plural **aurorae**) Light radiated by atoms and ions in the Earth's upper atmosphere, mostly in the polar regions. (Chapter 9)

aurora australis Aurorae seen from southern latitudes; the southern lights. (Chapter 9)

aurora borealis Aurorae seen from northern latitudes; the northern lights. (Chapter 9)

autumnal equinox The intersection of the ecliptic and the celestial equator where the Sun crosses the equator from north to south. Also used to refer to the date on which the Sun passes through this intersection. (Chapter 2)

average density The mass of an object divided by its volume. (Chapter 7)

B ring One of three prominent rings encircling Saturn. (Chapter 12)

Balmer line An emission or absorption line in the spectrum of hydrogen caused by an electron transition between the second and higher energy levels. (Chapter 5)

Balmer series The entire pattern of Balmer lines. (Chapter 5)

baseline In interferometry, the distance between two telescopes whose signals are combined to give a higher-resolution image. (Chapter 6)

belt A dark band in Jupiter's atmosphere. (Chapter 12)

Big Bang An explosion of all space that took place roughly 13.7 billion years ago and that marks the beginning of the universe. (Chapter 1)

biosphere The layer of soil, water, and air surrounding the Earth in which living organisms thrive. (Chapter 9)

black hole An object whose gravity is so strong that the escape speed exceeds the speed of light. (Chapter 1)

blackbody A hypothetical perfect radiator that absorbs and re-emits all radiation falling upon it. (Chapter 5)

blackbody curve The intensity of radiation emitted by a blackbody plotted as a function of wavelength or frequency. (Chapter 5)

blackbody radiation The radiation emitted by a perfect blackbody. (Chapter 5)

blueshift A decrease in the wavelength of photons emitted by an approaching source of light. (Chapter 5)

Bohr orbits In the model of the atom described by Niels Bohr, the only orbits in which electrons are allowed to move about the nucleus. (Chapter 5)

Bok globule A small, roundish, dark nebula. (Chapter 20)

bright terrain (Ganymede) Young, reflective, relatively crater-free terrain on the surface of Ganymede. (Chapter 13)

brown ovals Elongated, brownish features usually seen in Jupiter's northern hemisphere. (Chapter 12)

C ring One of three prominent rings encircling Saturn. (Chapter 12)

capture theory The hypothesis that the Moon was gravitationally captured by the Earth. (Chapter 10)

carbonaceous chondrite A type of meteorite that has a high abundance of carbon and volatile compounds. (Chapter 15, Chapter 28)

Cassegrain focus An optical arrangement in a reflecting telescope in which light rays are reflected by a secondary mirror to a focus behind the primary mirror. (Chapter 6)

Cassini division An apparent gap between Saturn's A and B rings. (Chapter 12)

CCD See *charge-coupled device*.

celestial equator A great circle on the celestial sphere 90° from the celestial poles. (Chapter 2)

celestial poles The points about which the celestial sphere appears to rotate. See also *north celestial pole* and *south celestial pole*. (Chapter 2)

celestial sphere An imaginary sphere of very large radius centered on an observer; the apparent sphere of the sky. (Chapter 2)

center of mass The point between a star and a planet, or between two stars, around which both objects orbit. (Chapter 8)

charge-coupled device (CCD) A type of solid-state device designed to detect photons. (Chapter 6)

chemical composition A description of which chemical substances make up a given object. (Chapter 7)

chemical differentiation The process by which the heavier elements in a planet sink toward its center while lighter elements rise toward its surface. (Chapter 8)

chemical element See *element*.

chondrule A glassy, roughly spherical blob found within meteorites. (Chapter 8)

chromatic aberration An optical defect whereby different colors of light passing through a lens are focused at different locations. (Chapter 6)

chromosphere A layer in the atmosphere of the Sun between the photosphere and the corona. (Chapter 16)

circumpolar A term describing a star that neither rises nor sets but appears to rotate around one of the celestial poles. (Chapter 2)

CNO cycle A series of nuclear reactions in which carbon is used as a catalyst to transform hydrogen into helium. (Chapter 16)

co-creation theory The hypothesis that the Earth and the Moon formed at the same time from the same material. (Chapter 10)

collisional ejection theory The hypothesis that the Moon formed from material ejected from the Earth by the impact of a large asteroid. (Chapter 10)

coma (of a comet) The diffuse gaseous component of the head of a comet. (Chapter 15)

coma (optical) The distortion of off-axis images formed by a parabolic mirror. (Chapter 6)

comet A small body of ice and dust in orbit about the Sun. While passing near the Sun, a comet's vaporized ices give rise to a coma and tail. (Chapter 7, Chapter 15)

compound A substance consisting of two or more chemical elements in a definite proportion. (Chapter 5)

condensation temperature The temperature at which a particular substance in a low-pressure gas condenses into a solid. (Chapter 8)

conduction The transfer of heat by directly passing energy from atom to atom. (Chapter 16)

conic section The curve of intersection between a circular cone and a plane; this curve can be a circle, ellipse, parabola, or hyperbola. (Chapter 4)

conjunction The geometric arrangement of a planet in the same part of the sky as the Sun, so that the planet is at an elongation of 0°. (Chapter 4)

conservation of angular momentum A law of physics stating that in an isolated system, the total amount of angular momentum—a measure of the amount of rotation—remains constant. (Chapter 8)

constellation A configuration of stars in the same region of the sky. (Chapter 2)

continuous spectrum A spectrum of light over a range of wavelengths without any spectral lines. (Chapter 5)

convection The transfer of energy by moving currents of fluid or gas containing that energy. (Chapter 9, Chapter 16)

convection cell A circulating loop of gas or liquid that transports heat from a warm region to a cool region. (Chapter 9)

convection current The pattern of motion in a gas or liquid in which convection is taking place. (Chapter 9)

convective zone The region in a star where convection is the dominant means of energy transport. (Chapter 16)

core (of the Earth) The iron-rich inner region of the Earth's interior. (Chapter 9)

core accretion model The hypothesis that each of the Jovian planets formed by accretion of gas onto a rocky core. (Chapter 8)

corona (of the Sun) The Sun's outer atmosphere, which has a high temperature and a low density. (Chapter 16)

coronal hole A region in the Sun's corona that is deficient in hot gases. (Chapter 16)

coronal mass ejection An event in which billions of tons of gas from the Sun's corona is suddenly blasted into space at high speed. (Chapter 16)

cosmic rays Fast-moving subatomic particles that enter our solar system from interstellar space. (Chapter 11)

coudé focus An optical arrangement with a reflecting telescope. A series of mirrors is used to direct light to a remote focus away from the moving parts of the telescope. (Chapter 6)

crater See *impact crater*.

crust (of a planet) The surface layer of a terrestrial planet. (Chapter 9)

crustal dichotomy (Mars) The contrast between the young northern lowlands and older southern highlands on Mars. (Chapter 11)

crystal A material in which atoms are arranged in orderly rows. (Chapter 9)

current sheet A broad, flat region in Jupiter's magnetosphere that contains an abundance of charged particles. (Chapter 12)

D ring One of several faint rings encircling Saturn. (Chapter 12)

dark terrain (Ganymede) Older, heavily cratered, dark-colored terrain on the surface of Ganymede. (Chapter 13)

decametric radiation Radiation from Jupiter whose wavelength is about 10 meters. (Chapter 12)

decimetric radiation Radiation from Jupiter whose wavelength is about a tenth of a meter. (Chapter 12)

declination Angular distance of a celestial object north or south of the celestial equator. (Chapter 2)

deferent A stationary circle in the Ptolemaic system along which another circle (an epicycle) moves, carrying a planet, the Sun, or the Moon. (Chapter 4)

degree A basic unit of angular measure, designated by the symbol °. (Chapter 1)

degree Celsius A basic unit of temperature, designated by the symbol °C and used on a scale where water freezes at 0° and boils at 100°. (Chapter 5)

degree Fahrenheit A basic unit of temperature, designated by the symbol °F and used on a scale where water freezes at 32° and boils at 212°. (Chapter 5)

density See *average density*.

differential rotation The rotation of a nonrigid object in which parts adjacent to each other at a given time do not always stay close together. (Chapter 12, Chapter 16)

differentiated asteroid An asteroid in which chemical differentiation has taken place, so that denser material is toward the asteroid's center. (Chapter 15)

differentiation See *chemical differentiation*.

diffraction The spreading out of light passing through an aperture or opening in an opaque object. (Chapter 6)

diffraction grating An optical device, consisting of thousands of closely spaced lines etched in glass or metal, that disperses light into a spectrum. (Chapter 6)

direct motion The apparent eastward movement of a planet seen against the background stars. (Chapter 4)

disk instability model The hypothesis that gases in the solar nebula coalesced rapidly to form the Jovian planets. (Chapter 8)

diurnal motion Any apparent motion in the sky that repeats on a daily basis, such as the rising and setting of stars. (Chapter 2)

Doppler effect The apparent change in wavelength of radiation due to relative motion between the source and the observer along the line of sight. (Chapter 5)

Drake equation An equation used to estimate the number of intelligent civilizations in the Galaxy with whom we might communicate. (Chapter 28)

dust devil Whirlwinds found in dry or desert areas on both Earth and Mars. (Chapter 11)

dust tail The tail of a comet that is composed primarily of dust grains. (Chapter 15)

dwarf planet A solar system body that is large enough to be spherical in shape and have a circular orbit around the Sun, but not large enough to clear its own path of other bodies. The term is used for Ceres, Pluto, and Eris. (Chapter 15)

dynamo The mechanism whereby electric currents within an astronomical body generate a magnetic field. (Chapter 7)

E ring A very broad, faint ring encircling Saturn. (Chapter 12)

earthquake A sudden vibratory motion of the Earth's surface. (Chapter 9)

eccentricity A number between 0 and 1 that describes the shape of an ellipse. (Chapter 4)

eclipse The cutting off of part or all the light from one celestial object by another. (Chapter 3)

eclipse path The track of the tip of the Moon's shadow along the Earth's surface during a total or annular solar eclipse. (Chapter 3)

eclipse year The interval between successive passages of the Sun through the same node of the Moon's orbit. (Chapter 3)

ecliptic The apparent annual path of the Sun on the celestial sphere. (Chapter 2)

ecliptic plane The plane of Earth's orbit around the Sun. (Chapter 2)

electromagnetic radiation Radiation consisting of oscillating electric and magnetic fields. Examples include gamma rays, X rays, visible light, ultraviolet and infrared radiation, radio waves, and microwaves. (Chapter 5)

electromagnetic spectrum The entire array of electromagnetic radiation. (Chapter 5)

electromagnetism Electric and magnetic phenomena, including electromagnetic radiation. (Chapter 5)

electron A subatomic particle with a negative charge and a small mass, usually found in orbits about the nuclei of atoms. (Chapter 5)

electron volt (eV) The energy acquired by an electron accelerated through an electric potential of one volt. (Chapter 5)

element A chemical that cannot be broken down into more basic chemicals. (Chapter 5)

ellipse A conic section obtained by cutting completely through a circular cone with a plane. (Chapter 4)

elongation The angular distance between a planet and the Sun as viewed from Earth. (Chapter 4)

emission line spectrum A spectrum that contains bright emission lines. (Chapter 5)

Encke gap A narrow gap in Saturn's A ring. (Chapter 12)

energy flux The rate of energy flow, usually measured in joules per square meter per second. (Chapter 5)

energy level In an atom, a particular amount of energy possessed by an atom above the atom's least energetic state. (Chapter 5)

energy-level diagram A diagram showing the arrangement of an atom's energy levels. (Chapter 5)

epicenter The location on the Earth's surface directly over the focus of an earthquake. (Chapter 9)

epicycle A moving circle in the Ptolemaic system about which a planet revolves. (Chapter 4)

epoch The date used to define the coordinate system for objects on the sky. (Chapter 2)

equal areas, law of See *Kepler's second law*.

equinox One of the intersections of the ecliptic and the celestial equator. Also used to refer to the date on which the Sun passes through such an intersection. (Chapter 2) See also *autumnal equinox* and *vernal equinox*.

escape speed The speed needed by an object (such as a spaceship) to leave a second object (such as a planet or star) permanently and to escape into interplanetary space. (Chapter 7)

eV See *electron volt*.

excited state A state of an atom, ion, or molecule with a higher energy than the ground state. (Chapter 5)

exponent A number placed above and after another number to denote the power to which the latter is to be raised, as n in 10^n. (Chapter 1)

extinction (interstellar) See *interstellar extinction*.

extrasolar planet A planet orbiting a star other the Sun. (Chapter 8)

eyepiece lens A magnifying lens used to view the image produced at the focus of a telescope. (Chapter 6)

F ring A thin, faint ring encircling Saturn just beyond the A ring. (Chapter 12)

false color In astronomical images, color used to denote different values of intensity, temperature, or other quantities. (Chapter 6)

far side (of the Moon) The side of the Moon that faces perpetually away from the Earth. (Chapter 10)

favorable opposition An opposition of Mars that affords good Earth-based views of the planet. (Chapter 11)

filament A portion of the Sun's chromosphere that arches to high altitudes. (Chapter 16)

first quarter moon The phase of the Moon that occurs when the Moon is 90° east of the Sun. (Chapter 3)

fission theory The hypothesis that the Moon was pulled out of a rapidly rotating proto-Earth. (Chapter 10)

flake tectonics A model of a planetary interior, particularly Venus, in which a thin crust remains stationary but wrinkles and flakes in response to interior convection currents. (Chapter 11)

flare See *solar flare*.

focal length The distance from a lens or mirror to the point where converging light rays meet. (Chapter 6)

focal plane The plane in which a lens or mirror forms an image of a distant object. (Chapter 6)

focal point The point at which a lens or mirror forms an image of a distant point of light. (Chapter 6)

focus (of an ellipse) (plural **foci**) One of two points inside an ellipse such that the combined distance from the two foci to any point on the ellipse is a constant. (Chapter 4)

focus (of a lens or mirror) The point to which light rays converge after passing through a lens or being reflected from a mirror. (Chapter 6)

force A push or pull that acts on an object. (Chapter 4)

frequency The number of crests or troughs of a wave that cross a given point per unit time. Also, the number of vibrations per unit time. (Chapter 5)

full moon A phase of the Moon during which its full daylight hemisphere can be seen from Earth. (Chapter 3)

fusion crust The coating on a stony meteorite caused by the heating of the meteorite as it descended through the Earth's atmosphere. (Chapter 15)

G ring A thin, faint ring encircling Saturn. (Chapter 12)

galaxy A large assemblage of stars, nebulae, and interstellar gas and dust. (Chapter 1)

Galilean satellites The four large moons of Jupiter. (Chapter 13)

gamma rays The most energetic form of electromagnetic radiation. (Chapter 5)

geocentric model An Earth-centered theory of the universe. (Chapter 4)

global warming The upward trend of the Earth's average temperature caused by increased amounts of greenhouse gases in the atmosphere. (Chapter 9)

granulation The rice grain-like structure found in the solar photosphere. (Chapter 16)

granule A convective cell in the solar photosphere. (Chapter 16)

grating See *diffraction grating*.

gravitational force See *gravity*.

gravity The force with which all matter attracts all other matter. (Chapter 4)

Great Red Spot A prominent high-pressure system in Jupiter's southern hemisphere. (Chapter 12)

greatest eastern elongation The configuration of an inferior planet at its greatest angular distance east of the Sun. (Chapter 4, Chapter 11)

greatest western elongation The configuration of an inferior planet at its greatest angular distance west of the Sun. (Chapter 4, Chapter 11)

greenhouse effect The trapping of infrared radiation near a planet's surface by the planet's atmosphere. (Chapter 9)

greenhouse gas A substance whose presence in a planet's atmosphere enhances the greenhouse effect. (Chapter 9)

ground state The state of an atom, ion, or molecule with the least possible energy. (Chapter 5)

habitable zone Regions of a galaxy or around a star in which conditions may be suitable for life to have developed. (Chapter 28)

half-life The time required for one-half of a quantity of a radioactive substance to decay. (Chapter 8)

heliocentric model A Sun-centered theory of the universe. (Chapter 4)

helioseismology The study of the vibrations of the Sun as a whole. (Chapter 16)

highlands (on Mars) See *southern highlands*.

highlands (on the Moon) See *lunar highlands*.

Hirayama family A group of asteroids that have nearly identical orbits about the Sun. (Chapter 15)

hot spot (on Jupiter) An unusually warm and cloud-free part of Jupiter's atmosphere. (Chapter 12)

hot-spot volcanism Volcanic activity that occurs over a hot region buried deep within a planet. (Chapter 11)

hydrocarbon Any one of a variety of chemical compounds composed of hydrogen and carbon. (Chapter 13)

hydrogen envelope A huge, tenuous sphere of gas surrounding the head of a comet. (Chapter 15)

hydrogen fusion The thermonuclear conversion of hydrogen into helium. (Chapter 16)

hydrostatic equilibrium A balance between the weight of a layer in a star and the pressure that supports it. (Chapter 16)

hyperbola A conic section formed by cutting a circular cone with a plane at an angle steeper than the sides of the cone. (Chapter 4)

hypothesis An idea or collection of ideas that seems to explain a specified phenomenon; a conjecture. (Chapter 1)

ice rafts (Europa) Segments of Europa's icy crust that have been moved by tectonic disturbances. (Chapter 13)

ices Solid materials with low condensation temperatures, including ices of water, methane, and ammonia. (Chapter 7)

igneous rock A rock that formed from the solidification of molten lava or magma. (Chapter 9)

imaging The process of recording the image made by a telescope of a distant object. (Chapter 6)

impact breccia A type of rock formed from other rocks that were broken apart, mixed, and fused together by a series of meteoritic impacts. (Chapter 10)

impact crater A circular depression on a planet or satellite caused by the impact of a meteoroid. (Chapter 7, Chapter 10)

inertia, law of See *Newton's first law*.

inferior conjunction The configuration when an inferior planet is between the Sun and Earth. (Chapter 4)

inferior planet A planet that is closer to the Sun than the Earth is. (Chapter 4)

infrared radiation Electromagnetic radiation of wavelength longer than visible light but shorter than radio waves. (Chapter 5)

inner core (of the Earth) The solid innermost portion of the Earth's iron-rich core. (Chapter 9)

interferometry A technique of combining the observations of two or more telescopes to produce images better than one telescope alone could make. (Chapter 6)

intermediate-period comet A comet with an orbital period between 20 and 200 years. (Chapter 15)

internal rotation period The period with which the core of a Jovian planet rotates. (Chapter 12)

interstellar medium Gas and dust in interstellar space. (Chapter 8)

Io torus A doughnut-shaped ring of gas circling Jupiter at the distance of Io's orbit. (Chapter 13)

ion tail The relatively straight tail of a comet produced by the solar wind acting on ions. (Chapter 15)

ionization The process by which a neutral atom becomes an electrically charged ion through the loss or gain of electrons. (Chapter 5)

iron meteorite A meteorite composed primarily of iron. (Chapter 15)

isotope Any of several forms for the same chemical element whose nuclei all have the same number of protons but different numbers of neutrons. (Chapter 5)

jet An extended line of fast-moving gas ejected from the vicinity of a star or a black hole. (Chapter 8)

joule (J) A unit of energy. (Chapter 5)

Jovian planet Low-density planets composed primarily of hydrogen and helium, including Jupiter, Saturn, Uranus, and Neptune. (Chapter 7)

Jupiter-family comet A comet with an orbital period of less than 20 years. (Chapter 15)

kelvin (K) A unit of temperature on the Kelvin temperature scale, equivalent to a degree Celsius. (Chapter 5)

Kelvin-Helmholtz contraction The contraction of a gaseous body, such as a star or nebula, during which gravitational energy is transformed into thermal energy. (Chapter 8)

Kepler's first law The statement that each planet moves around the Sun in an elliptical orbit with the Sun at one focus of the ellipse. (Chapter 4)

Kepler's second law The statement that a planet sweeps out equal areas in equal times as it orbits the Sun; also called the law of equal areas. (Chapter 4)

Kepler's third law A relationship between the period of an orbiting object and the semimajor axis of its elliptical orbit. (Chapter 4)

kiloparsec (kpc) One thousand parsecs; about 3260 light-years. (Chapter 1)

kinetic energy The energy possessed by an object because of its motion. (Chapter 7)

Kirchhoff's laws Three statements about circumstances that produce absorption lines, emission lines, and continuous spectra. (Chapter 5)

Kirkwood gaps Gaps in the spacing of asteroid orbits, discovered by Daniel Kirkwood. (Chapter 15)

Kuiper belt A region that extends from around the orbit of Pluto to about 500 AU from the Sun where many icy objects orbit the Sun. (Chapter 7, Chapter 14, Chapter 15)

lava Molten rock flowing on the surface of a planet. (Chapter 9)

law of equal areas See *Kepler's second law*.

law of inertia See *Newton's first law*.

law of universal gravitation A formula deduced by Isaac Newton that expresses the strength of the force of gravity that two masses exert on each other. (Chapter 4)

laws of physics A set of physical principles with which we can understand natural phenomena and the nature of the universe. (Chapter 1)

lens A piece of transparent material (usually glass) that can bend light and bring it to a focus. (Chapter 6)

libration An apparent rocking of the Moon whereby an Earth-based observer can, over time, see slightly more than one-half the Moon's surface. (Chapter 10)

light-gathering power A measure of the amount of radiation brought to a focus by a telescope. (Chapter 6)

light pollution Light from cities and towns that degrades telescope images. (Chapter 6)

light scattering The process by which light bounces off particles in its path. (Chapter 5, Chapter 12)

light-year (ly) The distance light travels in a vacuum in one year. (Chapter 1)

limb darkening The phenomenon whereby the Sun looks darker near its apparent edge, or limb, than near the center of its disk. (Chapter 16)

line of nodes The line where the plane of the Earth's orbit intersects the plane of the Moon's orbit. (Chapter 3)

liquid metallic hydrogen Hydrogen compressed to such a density that it behaves like a liquid metal. (Chapter 7, Chapter 12)

lithosphere The solid, upper layer of the Earth; essentially the Earth's crust. (Chapter 9)

local meridian See *meridian*.

long-period comet A comet that takes hundreds of thousands of years or more to complete one orbit of the Sun. (Chapter 15)

lower meridian The half of the meridian that lies below the horizon. (Chapter 2)

lowlands (on Mars) See *northern lowlands*.

luminosity The rate at which electromagnetic radiation is emitted from a star or other object. (Chapter 5, Chapter 16)

lunar eclipse An eclipse of the Moon by the Earth; a passage of the Moon through the Earth's shadow. (Chapter 3)

lunar highlands Ancient, high-elevation, heavily cratered terrain on the Moon. (Chapter 10)

lunar month See *synodic month*.

lunar phase The appearance of the illuminated area of the Moon as seen from Earth. (Chapter 3)

Lyman series A series of spectral lines of hydrogen produced by electron transitions to and from the lowest energy state of the hydrogen atom. (Chapter 5)

magma Molten rock beneath a planet's surface. (Chapter 9)

magnetic axis A line connecting the north and south magnetic poles of a planet or star possessing a magnetic field. (Chapter 14)

magnetic-dynamo model A theory that explains the solar cycle as a result of the Sun's differential rotation acting on the Sun's magnetic field. (Chapter 16)

magnetic reconnection An event where two oppositely directed magnetic fields approach and cancel, thus releasing energy. (Chapter 16)

magnetogram An image of the Sun that shows regions of different magnetic polarity. (Chapter 16)

magnetometer A device for measuring magnetic fields. (Chapter 7)

magnetopause That region of a planet's magnetosphere where the magnetic field counterbalances the pressure from the solar wind. (Chapter 9)

magnetosphere The region around a planet occupied by its magnetic field. (Chapter 9)

magnification The factor by which the apparent angular size of an object is increased when viewed through a telescope. (Chapter 6)

magnifying power See *magnification*.

major axis (of an ellipse) The longest diameter of an ellipse. (Chapter 4)

mantle (of a planet) That portion of a terrestrial planet located between its crust and core. (Chapter 9)

mare (plural **maria**) Latin for "sea"; a large, relatively crater-free plain on the Moon. (Chapter 10)

mare basalt A type of lunar rock commonly found in the mare basins. (Chapter 10)

mass A measure of the total amount of material in an object. (Chapter 4)

mean solar day The interval between successive meridian passages of the mean Sun; the average length of a solar day. (Chapter 2)

mean sun A fictitious object that moves eastward at a constant speed along the celestial equator, completing one circuit of the sky with respect to the vernal equinox in one tropical year. (Chapter 2)

medium (plural **media**) A material through which light travels. (Chapter 6)

medium, interstellar See *interstellar medium*.

megaparsec (Mpc) One million parsecs. (Chapter 1)

melting point The temperature at which a substance changes from solid to liquid. (Chapter 9)

meridian (or **local meridian**) The great circle on the celestial sphere that passes through an observer's zenith and the north and south celestial poles. (Chapter 2)

meridian transit The crossing of the meridian by any astronomical object. (Chapter 2)

mesosphere A layer in the Earth's atmosphere above the stratosphere. (Chapter 9)

metamorphic rock A rock whose properties and appearance have been transformed by the action of pressure and heat beneath the Earth's surface. (Chapter 9)

meteor The luminous phenomenon seen when a meteoroid enters the Earth's atmosphere; a "shooting star." (Chapter 15)

meteor shower Many meteors that seem to radiate from a common point in the sky. (Chapter 15)

meteorite A fragment of a meteoroid that has survived passage through the Earth's atmosphere. (Chapter 1, Chapter 8, Chapter 15)

meteoritic swarm A collection of meteoroids moving together along an orbit about the Sun. (Chapter 15)

meteoroid A small rock in interplanetary space. (Chapter 7, Chapter 15)

microwaves Short-wavelength radio waves. (Chapter 5)

mineral A naturally occurring solid composed of a single element or chemical combination of elements, often in the form of crystals. (Chapter 9)

minor planet See *asteroid*.

minute of arc See *arcminute*.

model A hypothesis that has withstood experimental or observational tests; or, the results of a theoretical calculation that gives the values of temperature, pressure, density, and so forth throughout the interior of an object such as a planet or star. (Chapter 1)

molecule A combination of two or more atoms. (Chapter 5)

moonquake Sudden, vibratory motion of the Moon's surface. (Chapter 10)

nanometer (nm) One billionth of a meter: 1 nm = 10^{-9} meter = 10^{-6} millimeter = 10^{-3} μm. (Chapter 5)

neap tide An ocean tide that occurs when the Moon is near first-quarter or third-quarter phase. (Chapter 4)

near-Earth object (NEO) An asteroid whose orbit lies wholly or partly within the orbit of Mars. (Chapter 15)

nebula A cloud of interstellar gas and dust. (Chapter 1)

nebular hypothesis The idea that the Sun and the rest of the solar system formed from a cloud of interstellar material. (Chapter 8)

nebulosity See *nebula*.

negative hydrogen ion A hydrogen atom that has acquired a second electron. (Chapter 16)

NEO See *near-Earth object*.

neutrino A subatomic particle with no electric charge and very little mass, yet one that is important in many nuclear reactions. (Chapter 16)

neutron A subatomic particle with no electric charge and with a mass nearly equal to that of the proton. (Chapter 5)

new moon The phase of the Moon when the dark hemisphere of the Moon faces the Earth. (Chapter 3)

Newtonian mechanics The branch of physics based on Newton's laws of motion. (Chapter 1, Chapter 4)

Newtonian reflector A reflecting telescope that uses a small mirror to deflect the image to one side of the telescope tube. (Chapter 6)

Newton's first law of motion The statement that a body remains at rest, or moves in a straight line at a constant speed, unless acted upon by a net outside force; the law of inertia. (Chapter 4)

Newton's form of Kepler's third law A relationship between the period of two objects orbiting each other, the semimajor axis of their orbit, and the masses of the objects. (Chapter 4)

Newton's second law of motion A relationship between the acceleration of an object, the object's mass, and the net outside force acting on the mass. (Chapter 4)

Newton's third law of motion The statement that whenever one body exerts a force on a second body, the second body exerts an equal and opposite force on the first body. (Chapter 4)

noble gas An element whose atoms do not combine into molecules. (Chapter 12)

node See *line of nodes*.

nonthermal radiation Radiation other than that emitted by a heated body. (Chapter 12)

north celestial pole The point directly above the Earth's north pole where the Earth's axis of rotation, if extended, would intersect the celestial sphere. (Chapter 2)

northern lights See *aurora borealis*.

northern lowlands (on Mars) Relatively young and crater-free terrain in the Martian northern hemisphere. (Chapter 11)

nucleus (of an atom) The massive part of an atom, composed of protons and neutrons, about which electrons revolve. (Chapter 5)

nucleus (of a comet) A collection of ices and dust that constitute the solid part of a comet. (Chapter 15)

objective lens The principal lens of a refracting telescope. (Chapter 6)

objective mirror The principal mirror of a reflecting telescope. (Chapter 6)

oblate Flattened at the poles. (Chapter 12)

oblateness A measure of how much a flattened sphere (or spheroid) differs from a perfect sphere. (Chapter 12)

Occam's razor The notion that a straightforward explanation of a phenomenon is more likely to be correct than a convoluted one. (Chapter 4)

occultation The eclipsing of an astronomical object by the Moon or a planet. (Chapter 13, Chapter 14)

oceanic rift A crack in the ocean floor that exudes lava. (Chapter 9)

1-to-1 spin-orbit coupling See *synchronous rotation*.

Oort cloud A presumed accumulation of comets and cometary material surrounding the Sun at distances of roughly 50,000 AU. (Chapter 7, Chapter 15)

opposition The configuration of a planet when it is at an elongation of 180° and thus appears opposite the Sun in the sky. (Chapter 4)

optical telescope A telescope designed to detect visible light. (Chapter 6)

optical window The range of visible wavelengths to which the Earth's atmosphere is transparent. (Chapter 6)

organic molecules Molecules containing carbon, some of which are the molecules of which living organisms are made. (Chapter 28)

outer core (of the Earth) The outer, molten portion of the Earth's iron-rich core. (Chapter 9)

outgassing The release of gases into a planet's atmosphere by volcanic activity. (Chapter 9)

ozone A type of oxygen whose molecules contain three oxygen atoms. (Chapter 9)

ozone hole A region of the Earth's atmosphere over Antarctica where the concentration of ozone is abnormally low. (Chapter 9)

ozone layer A layer in the Earth's upper atmosphere where the concentration of ozone is high enough to prevent much ultraviolet light from reaching the surface. (Chapter 9)

P wave One of three kinds of seismic waves produced by an earthquake; a primary wave. (Chapter 9)

parabola A conic section formed by cutting a circular cone at an angle parallel to one of the sides of the cone. (Chapter 4)

parallax The apparent displacement of an object due to the motion of the observer. (Chapter 4)

parsec (pc) A unit of distance; 3.26 light-years. (Chapter 1)

partial lunar eclipse A lunar eclipse in which the Moon does not appear completely covered. (Chapter 3)

partial solar eclipse A solar eclipse in which the Sun does not appear completely covered. (Chapter 3)

Paschen series A series of spectral lines of hydrogen produced by electron transitions between the third and higher energy levels. (Chapter 5)

penumbra (of a shadow) (plural **penumbrae**) The portion of a shadow in which only part of the light source is covered by an opaque body. (Chapter 3)

penumbral eclipse A lunar eclipse in which the Moon passes only through the Earth's penumbra. (Chapter 3)

perigee The point in its orbit where a satellite or the Moon is nearest the Earth. (Chapter 3)

perihelion The point in its orbit where a planet or comet is nearest the Sun. (Chapter 4)

period (of a planet) The interval of time between successive geometric arrangements of a planet and an astronomical object, such as the Sun. (Chapter 4)

periodic table A listing of the chemical elements according to their properties, invented by Dmitri Mendeleev. (Chapter 5)

photoelectric effect The phenomenon whereby certain metals emit electrons when exposed to short-wavelength light. (Chapter 5)

photometry The measurement of light intensities. (Chapter 6)

photon A discrete unit of electromagnetic energy. (Chapter 5)

photosphere The region in the solar atmosphere from which most of the visible light escapes into space. (Chapter 16)

photosynthesis A biochemical process in which solar energy is converted into chemical energy, carbon dioxide and water are absorbed, and oxygen is released. (Chapter 9)

physics, laws of See *laws of physics*.

pixel A picture element. (Chapter 6)

plage A bright region in the solar atmosphere as observed in the monochromatic light of a spectral line. (Chapter 16)

Planck's law The relationship between the energy of a photon and its wavelength or frequency; $E = hc/\lambda = h\nu$. (Chapter 5)

plane of the ecliptic The plane in which the Earth orbits the Sun. (Chapter 3)

planetesimal One of many small bodies of primordial dust and ice that combined to form the planets. (Chapter 8)

plasma A hot ionized gas. (Chapter 12, Chapter 16)

plastic The attribute of being nearly solid yet able to flow. (Chapter 9)

plate A large section of the Earth's lithosphere that moves as a single unit. (Chapter 9)

plate tectonics The motions of large segments (plates) of the Earth's surface over the underlying mantle. (Chapter 9)

plutino One of about 100 objects in the Kuiper belt that orbit the Sun with nearly the same semimajor axis as Pluto. (Chapter 14)

polymer A long molecule consisting of many smaller molecules joined together. (Chapter 13)

positional astronomy The study of the apparent positions of the planets and stars and how those positions change. (Chapter 2)

positron An electron with a positive rather than negative electric charge; the antiparticle of the electron. (Chapter 16)

power of ten The exponent n in 10^n. (Chapter 1)

powers-of-ten notation A shorthand method of writing numbers, involving 10 followed by an exponent. (Chapter 1)

precession (of the Earth) A slow, conical motion of the Earth's axis of rotation caused by the gravitational pull of the Moon and Sun on the Earth's equatorial bulge. (Chapter 2)

precession of the equinoxes The slow westward motion of the equinoxes along the ecliptic due to precession of the Earth. (Chapter 2)

primary mirror See *objective mirror*.

prime focus The point in a telescope where the objective focuses light. (Chapter 6)

primitive asteroid See *undifferentiated asteroid*.

prograde orbit An orbit of a satellite around a planet that is in the same direction as the rotation of the planet. (Chapter 13)

prograde rotation A situation in which an object (such as a planet) rotates in the same direction that it orbits around another object (such as the Sun). (Chapter 11)

prominence Flamelike protrusions seen near the limb of the Sun and extending into the solar corona. (Chapter 16)

proplyd See *protoplanetary disk*.

proton A heavy, positively charged subatomic particle that is one of two principal constituents of atomic nuclei. (Chapter 5)

proton-proton chain A sequence of thermonuclear reactions by which hydrogen nuclei are built up into helium nuclei. (Chapter 16)

protoplanet A Moon-sized object formed by the coalescence of planetesimals. (Chapter 8)

protoplanetary disk (proplyd) A disk of material encircling a protostar or a newborn star. (Chapter 8)

protosun The part of the solar nebula that eventually developed into the Sun. (Chapter 8)

Ptolemaic system The definitive version of the geocentric cosmogony of ancient Greece. (Chapter 4)

pulsar A pulsating radio source thought to be associated with a rapidly rotating neutron star. (Chapter 1)

quantum mechanics The branch of physics dealing with the structure and behavior of atoms and their constituents as well as their interaction with light. (Chapter 5)

quasar A very luminous object with a very large redshift and a starlike appearance. (Chapter 1)

radial velocity That portion of an object's velocity parallel to the line of sight. (Chapter 5)

radial velocity method A technique used to detect extrasolar planets by observing Doppler shifts in the spectrum of the planet's star. (Chapter 8)

radiant (of a meteor shower) The point in the sky from which meteors of a particular shower seem to originate. (Chapter 15)

radiation darkening The darkening of methane ice by electron impacts. (Chapter 14)

radiation pressure Pressure exerted on an object by radiation falling on the object. (Chapter 15)

radiative diffusion The random migration of photons from a star's center toward its surface. (Chapter 16)

radiative zone A region within a star where radiative diffusion is the dominant mode of energy transport. (Chapter 16)

radio telescope A telescope designed to detect radio waves. (Chapter 6)

radio waves The longest-wavelength electromagnetic radiation. (Chapter 5)

radio window The range of radio wavelengths to which the Earth's atmosphere is transparent. (Chapter 6)

radioactive dating A technique for determining the age of a rock sample by measuring the radioactive elements and their decay products in the sample. (Chapter 8)

radioactive decay The process whereby certain atomic nuclei spontaneously transform into other nuclei. (Chapter 8)

redshift The shifting to longer wavelengths of the light from remote galaxies and quasars; the Doppler shift of light from a receding source. (Chapter 5)

reflecting telescope A telescope in which the principal optical component is a concave mirror. (Chapter 6)

reflection The return of light rays by a surface. (Chapter 6)

reflector A reflecting telescope. (Chapter 6)

refracting telescope A telescope in which the principal optical component is a lens. (Chapter 6)

refraction The bending of light rays when they pass from one transparent medium to another. (Chapter 6)

refractor A refracting telescope. (Chapter 6)

refractory element An element with high melting and boiling points. (Chapter 10)

regolith The layer of rock fragments covering the surface of the Moon. (Chapter 10)

residual polar cap An ice-covered polar region on Mars that does not completely evaporate during the Martian summer. (Chapter 11)

respiration A biological process that produces energy by consuming oxygen and releasing carbon dioxide. (Chapter 9)

retrograde motion The apparent westward motion of a planet with respect to background stars. (Chapter 4)

retrograde orbit An orbit of a satellite around a planet that is in the direction opposite to which the planet rotates. (Chapter 13)

retrograde rotation A situation in which an object (such as a planet) rotates in the direction opposite to which it orbits around another object (such as the Sun). (Chapter 11)

rift valley A feature created when a planet's crust breaks apart along a line. (Chapter 11)

right ascension A coordinate for measuring the east-west positions of objects on the celestial sphere. (Chapter 2)

ring particles Small particles that constitute a planetary ring. (Chapter 12)

ringlet One of many narrow bands of particles of which Saturn's ring system is composed. (Chapter 12)

Roche limit The smallest distance from a planet or other object at which a second object can be held together by purely gravitational forces. (Chapter 12)

rock A mineral or combination of minerals. (Chapter 9)

runaway greenhouse effect A greenhouse effect in which the temperature continues to increase. (Chapter 11)

runaway icehouse effect A situation in which a decrease in atmospheric temperature causes a further decrease in temperature. (Chapter 11)

S wave One of three kinds of seismic waves produced by an earthquake; a secondary wave. (Chapter 9)

saros A particular cycle of similar eclipses that recur about every 18 years. (Chapter 3)

scarp A line of cliffs formed by the faulting or fracturing of a planet's surface. (Chapter 11)

scattering of light See *light scattering*.

scientific method The basic procedure used by scientists to investigate phenomena. (Chapter 1)

seafloor spreading The separation of plates under the ocean due to lava emerging in an oceanic rift. (Chapter 9)

search for extraterrestrial intelligence (SETI) The scientific search for evidence of intelligent life on other planets. (Chapter 28)

second of arc See *arcsecond*.

sedimentary rock A rock that is formed from material deposited on land by rain or winds, or on the ocean floor. (Chapter 9)

seeing disk The angular diameter of a star's image. (Chapter 6)

seismic wave A vibration traveling through a terrestrial planet, usually associated with earthquake-like phenomena. (Chapter 9)

seismograph A device used to record and measure seismic waves, such as those produced by earthquakes. (Chapter 9)

semimajor axis One-half of the major axis of an ellipse. (Chapter 4)

SETI See *search for extraterrestrial intelligence*.

shepherd satellite A satellite whose gravity restricts the motions of particles in a planetary ring, preventing them from dispersing. (Chapter 12)

shield volcano A volcano with long, gently sloping sides. (Chapter 11)

shock wave An abrupt, localized region of compressed gas caused by an object traveling through the gas at a speed greater than the speed of sound. (Chapter 9)

SI units The International System of Units, based on the meter (m), the second (s), and the kilogram (kg). (Chapter 1)

sidereal clock A clock that measures sidereal time. (Chapter 2)

sidereal day The interval between successive meridian passages of the vernal equinox. (Chapter 2)

sidereal month The period of the Moon's revolution about the Earth with respect to the stars. (Chapter 3)

sidereal period The orbital period of one object about another as measured with respect to the stars. (Chapter 4)

sidereal time Time reckoned by the location of the vernal equinox. (Chapter 2)

sidereal year The orbital period of the Earth about the Sun with respect to the stars. (Chapter 2)

small-angle formula A relationship between the angular and linear sizes of a distant object. (Chapter 1)

SNC meteorite A meteorite that came to Earth from Mars. (Chapter 28)

solar constant The average amount of energy received from the Sun per square meter per second, measured just above the Earth's atmosphere. (Chapter 5)

solar corona Hot, faintly glowing gases seen around the Sun during a total solar eclipse; the uppermost regions of the solar atmosphere. (Chapter 3)

solar cycle See *22-year solar cycle*.

solar eclipse An eclipse of the Sun by the Moon; a passage of the Earth through the Moon's shadow. (Chapter 3)

solar flare A sudden, temporary outburst of light from an extended region of the solar surface. (Chapter 16)

solar nebula The cloud of gas and dust from which the Sun and solar system formed. (Chapter 8)

solar neutrino A neutrino emitted from the core of the Sun. (Chapter 16)

solar neutrino problem The discrepancy between the predicted and observed numbers of solar neutrinos. (Chapter 16)

solar system The Sun, planets and their satellites, asteroids, comets, and related objects that orbit the Sun. (Chapter 1)

solar wind An outward flow of particles (mostly electrons and protons) from the Sun. (Chapter 8, Chapter 16)

south celestial pole The point directly above the Earth's south pole where the Earth's axis of rotation, if extended, would intersect the celestial sphere. (Chapter 2)

southern highlands (on Mars) Older, cratered terrain in the Martian southern hemisphere. (Chapter 11)

southern lights See *aurora australis*.

spectral analysis The identification of chemical substances from the patterns of lines in their spectra. (Chapter 5)

spectral line In a spectrum, an absorption or emission feature that is at a particular wavelength. (Chapter 5)

spectrograph An instrument for photographing a spectrum. (Chapter 6)

spectroscopy The study of spectra and spectral lines. (Chapter 5, Chapter 6, Chapter 7)

spectrum (plural **spectra**) The result of dispersing a beam of electromagnetic radiation so that components with different wavelengths are separated in space. (Chapter 5)

speed Distance traveled divided by the time elapsed to cover that distance. (Chapter 4)

spherical aberration The distortion of an image formed by a telescope due to differing focal lengths of the optical system. (Chapter 6)

spicule A narrow jet of rising gas in the solar chromosphere. (Chapter 16)

spin-orbit coupling See *1-to-1 spin-orbit coupling* and *3-to-2 spin-orbit coupling*.

spring tide An ocean tide that occurs at new moon and full moon phases. (Chapter 4)

stable Lagrange points Locations along Jupiter's orbit where the combined gravitational effects of the Sun and Jupiter cause asteroids to collect. (Chapter 15)

Stefan-Boltzmann law A relationship between the temperature of a blackbody and the rate at which it radiates energy. (Chapter 5)

stony iron meteorite A meteorite composed of both stone and iron. (Chapter 15)

stony meteorite A meteorite composed of stone. (Chapter 15)

stratosphere A layer in the Earth's atmosphere directly above the troposphere. (Chapter 9)

subduction zone A location where colliding tectonic plates cause the Earth's crust to be pulled down into the mantle. (Chapter 9)

subtend To extend over an angle. (Chapter 1)

summer solstice The point on the ecliptic where the Sun is farthest north of the celestial equator. Also used to refer to the date on which the Sun passes through this point. (Chapter 2)

sunspot A temporary cool region in the solar photosphere. (Chapter 16)

sunspot cycle The semiregular 11-year period with which the number of sunspots fluctuates. (Chapter 16)

sunspot maximum/minimum That time during the sunspot cycle when the number of sunspots is highest/lowest. (Chapter 16)

supergranule A large convective feature in the solar atmosphere, usually outlined by spicules. (Chapter 16)

superior conjunction The configuration of a planet being behind the Sun as viewed from the Earth. (Chapter 4)

superior planet A planet that is more distant from the Sun than the Earth is. (Chapter 4)

supernova (plural **supernovae**) A stellar outburst during which a star suddenly increases its brightness roughly a millionfold. (Chapter 1, Chapter 15)

surface wave A type of seismic wave that travels only over the Earth's surface. (Chapter 9)

synchronous rotation The rotation of a body with a period equal to its orbital period; also called 1-to-1 spin-orbit coupling. (Chapter 3, Chapter 10)

synchrotron radiation A type of nonthermal radiation emitted by charged particles moving through a magnetic field. (Chapter 12)

synodic month The period of revolution of the Moon with respect to the Sun; the length of one cycle of lunar phases. Also called the lunar month. (Chapter 3)

synodic period The interval between successive occurrences of the same configuration of a planet. (Chapter 4)

T Tauri wind A flow of particles away from a T Tauri star. (Chapter 8)

tail (of a comet) Gas and dust particles from a comet's nucleus that have been swept away from the comet's head by the radiation pressure of sunlight and the solar wind. (Chapter 15)

tangential velocity That portion of an object's velocity perpendicular to the line of sight. (Chapter 19)

temperature See *degree Celsius, degree Fahrenheit,* and *kelvin.*

terminator The line dividing day and night on the surface of the Moon or a planet; the line of sunset or sunrise. (Chapter 10)

terrae Cratered lunar highlands. (Chapter 10)

terrestrial planet High-density worlds with solid surfaces, including Mercury, Venus, Earth, and Mars. (Chapter 7)

theory A hypothesis that has withstood experimental or observational tests. (Chapter 1)

thermal equilibrium A balance between the input and outflow of heat in a system. (Chapter 16)

thermal radiation The radiation naturally emitted by any object that is not at absolute zero. Blackbody radiation is an idealized case of thermal radiation. (Chapter 12)

thermonuclear fusion The combining of nuclei under conditions of high temperature in a process that releases substantial energy. (Chapter 16)

thermosphere A region in the Earth's atmosphere between the mesosphere and the exosphere. (Chapter 9)

third quarter moon The phase of the Moon that occurs when the Moon is 90° west of the Sun. (Chapter 3)

3-to-2 spin-orbit coupling The rotation of Mercury, which makes three complete rotations on its axis for every two complete orbits around the Sun. (Chapter 11)

tidal force A gravitational force whose strength and/or direction varies over a body and thus tends to deform the body. (Chapter 4, Chapter 12)

tidal heating The heating of the interior of a satellite by continually varying tidal stresses. (Chapter 13)

time zone A region on the Earth where, by agreement, all clocks have the same time. (Chapter 2)

total lunar eclipse A lunar eclipse during which the Moon is completely immersed in the Earth's umbra. (Chapter 3)

total solar eclipse A solar eclipse during which the Sun is completely hidden by the Moon. (Chapter 3)

totality (lunar eclipse) The period during a total lunar eclipse when the Moon is entirely within the Earth's umbra. (Chapter 3)

totality (solar eclipse) The period during a total solar eclipse when the disk of the Sun is completely hidden. (Chapter 3)

transit An event in which an astronomical body moves in front of another. See also *meridian transit* and *solar transit.* (Chapter 8)

transit method A method for detecting extrasolar planets that come between us and their parent star, dimming the star's light. (Chapter 8)

Trans-Neptunian object Any small body of rock and ice that orbits the Sun within the solar system, but beyond the orbit of Neptune. (Chapter 7, Chapter 14)

Trojan asteroid One of several asteroids that share Jupiter's orbit about the Sun. (Chapter 15)

Tropic of Cancer A circle of latitude $23\frac{1}{2}°$ north of the Earth's equator. (Chapter 2)

Tropic of Capricorn A circle of latitude $23\frac{1}{2}°$ south of the Earth's equator. (Chapter 2)

tropical year The period of revolution of the Earth about the Sun with respect to the vernal equinox. (Chapter 2)

troposphere The lowest level in the Earth's atmosphere. (Chapter 9)

22-year solar cycle The semiregular 22-year interval between successive appearances of sunspots at the same latitude and with the same magnetic polarity. (Chapter 16)

ultramafic lava A type of lava enriched in magnesium and iron. These give the lava a higher melting temperature. (Chapter 13)

ultraviolet radiation Electromagnetic radiation of wavelengths shorter than those of visible light but longer than those of X rays. (Chapter 5)

umbra (of a shadow) (plural **umbrae**) The central, completely dark portion of a shadow. (Chapter 3)

undifferentiated asteroid An asteroid within which chemical differentiation did not occur. (Chapter 15)

universal constant of gravitation (*G*) The constant of proportionality in Newton's law of gravitation. (Chapter 4)

upper meridian The half of the meridian that lies above the horizon. (Chapter 2)

Van Allen belts Two doughnut-shaped regions around the Earth where many charged particles (protons and electrons) are trapped by the Earth's magnetic field. (Chapter 9)

velocity The speed and direction of an object's motion. (Chapter 4)

vernal equinox The point on the ecliptic where the Sun crosses the celestial equator from south to north. Also used to refer to the date on which the Sun passes through this intersection. (Chapter 2)

very-long-baseline interferometry (VLBI) A method of connecting widely separated radio telescopes to make very high-resolution observations. (Chapter 6)

visible light Electromagnetic radiation detectable by the human eye. (Chapter 5)

VLBI See *very-long-baseline interferometry*.

volatile element An element with low melting and boiling points. (Chapter 10)

waning crescent moon The phase of the Moon that occurs between third quarter and new moon. (Chapter 3)

waning gibbous moon The phase of the Moon that occurs between full moon and third quarter. (Chapter 3)

water hole A range of frequencies in the microwave spectrum suitable for interstellar radio communication. (Chapter 28)

watt A unit of power, equal to one joule of energy per second. (Chapter 5)

wavelength The distance between two successive wave crests. (Chapter 5)

wavelength of maximum emission The wavelength at which a heated object emits the greatest intensity of radiation. (Chapter 5)

waxing crescent moon The phase of the Moon that occurs between new moon and first quarter. (Chapter 3)

waxing gibbous moon The phase of the Moon that occurs between first quarter and full moon. (Chapter 3)

weight The force with which gravity acts on a body. (Chapter 4)

white ovals Round, whitish feature usually seen in Jupiter's southern hemisphere. (Chapter 12)

Widmanstätten patterns Crystalline structure seen in certain types of meteorites. (Chapter 15)

Wien's law A relationship between the temperature of a blackbody and the wavelength at which it emits the greatest intensity of radiation. (Chapter 5)

winter solstice The point on the ecliptic where the Sun reaches its greatest distance south of the celestial equator. Also used to refer to the date on which the Sun passes through this point. (Chapter 2)

X rays Electromagnetic radiation whose wavelength is between that of ultraviolet light and gamma rays. (Chapter 5)

Zeeman effect A splitting or broadening of spectral lines due to a magnetic field. (Chapter 16)

zenith The point on the celestial sphere directly overhead an observer. (Chapter 2)

zodiac A band of 12 constellations around the sky centered on the ecliptic. (Chapter 12)

zonal winds The pattern of alternating eastward and westward winds found in the atmospheres of Jupiter and Saturn. (Chapter 12)

zone A light-colored band in Jupiter's atmosphere. (Chapter 12)

This page intentionally left blank

Answers to Selected Questions

How to Get the Most from UNIVERSE

1. Paragraphs labeled **Caution! 2.** Paragraphs labeled **Analogy 3. Tools of the Astronomer's Trade** Boxes and **The Heavens on the Earth** Boxes **4.** On the *Universe* Web site **5.** It shows whether the image was made with **R**adio waves, **I**nfrared radiation, **V**isible light, **U**ltraviolet light, **X** rays, or **G**amma rays **6.** On the *Universe* Web site **7.** (**a**) page A-8 (**b**) page A-1 (**c**) page A-4 (**d**) page A-6 **8.** (**a**) page 417 (**b**) page 130 (**c**) page 89 (**d**) page 74 **9.** 0.5 year **10.** At the back of the book, following the index

CAUTION! Only mathematical answers are given in the following, not answers that require interpretation or discussion. Your instructor will expect you to show the steps required to reach each mathematical answer.

Chapter 1

24. 8.5×10^3 km **25.** 2.8×10^7 Suns **26.** About 3×10^{36} times larger **27.** 8.94×10^{56} hydrogen atoms **28.** (**a**) 1.581×10^{-5} ly (**b**) 4.848×10^{-6} pc **29.** 4.99×10^2 s **30.** 4.3×10^9 km **31.** (**a**) 1.59×10^{14} km (**b**) 16.8 years **32.** 4.32×10^{17} s **34.** (**a**) 1.5 m (**b**) 89 m (**c**) 5.4×10^3 m **35.** 6.9 m **36.** 3.4×10^3 km **37.** 3.7 km **38.** 0.320 arcmin

Chapter 2

28. Around 8:02 p.m. **33.** (**a**) About 9 hours **46.** October 25, 1917 **49.** 50° **50.** 3:50 a.m. local time **52.** (**a**) 6:00 p.m. (**b**) September 21

Chapter 3

31. (**a**) 0.91 hour (**b**) 11° **32.** 49 arcsec

Chapter 4

18. Semimajor axis = 0.25 AU, period = 0.125 year **19.** Average = 25 AU, farthest = 50 AU **23.** 6 newtons, 4 m/s^2 **26.** 1/9 as strong **33.** 87.97 days **38.** (**a**) 4 AU (**b**) 8 years **39.** (**a**) 16.0 AU (**b**) 0.5 AU **40.** (**a**) 1.26 AU (**b**) 258 days **41.** Earth exerts a force of 1.98×10^{20} newtons on the Moon **42.** Forces are approximately the same, Earth's acceleration is 100 times greater **43.** $\frac{1}{4}$ as much as on Earth **44.** 62 newtons, 0.13 **45.** 0.5 year **46.** (**a**) 24 hours (**b**) 43,200 km **47.** 119 minutes **49.** (**a**) 5 years (**b**) 2.92 AU **50.** (**a**) 3.43×10^{-5} newton (**b**) 3.21×10^{-5} newton (**c**) 2.2×10^{-6} newton

Chapter 5

2. 7.5 times **3.** 500 s **7.** 0.340 m **8.** 2.61×10^{14} Hz **14.** 135 nm **15.** 3400 K **24.** 9400 nm **25.** 9980°F **26.** About 10 μm **27.** 2.9 nm **28.** 3.9×10^{26} W **29.** 190 times more **30.** (**a**) 4890 nm = 4.89 μm (**b**) 540 times more **31.** (**a**) 5.75×10^8 W/m (**b**) 10,000 K **33.** 2.43×10^{-3} nm **34.** (**a**) 1005 nm **38.** Possible transitions: energy = 1 eV, λ = 1240 nm; energy = 2 eV, λ = 620 nm; energy = 3 eV, λ = 414 nm **39.** 656.6 nm **40.** Coming toward us at 13.0 km/s **41.** 8.6×10^4 km/s

Chapter 6

31. $\frac{1}{25}$ = 0.04 **33.** (**a**) 222× (**b**) 100× (**c**) 36× (**d**) 0.75 arcsec **36.** 300 km (Hubble Space Telescope, Jupiter's moons); 110 km (human eye, our Moon) **37.** (**a**) 34 ly (**b**) 37 km **40.** (**a**) 5.39×10^{-4} m (**c**) 0.56 m

Chapter 7

22. Mass = 6.4×10^{23} kg, average density = 3900 kg/m^3 **24.** (**a**) 1.0×10^{13} kg (**b**) 1.2 m/s **25.** (**a**) 3.3×10^{21} J (**b**) Equivalent to 3.9×10^7 Hiroshima-type weapons **26.** (**a**) 2.02 km/s (**b**) 19.4 km/s **27.** 1.2 km/s **28.** (**a**) 618 km/s **30.** 1.76 years **32.** (**a**) 1000 years (**b**) 17 days **34.** In 100 years, probability is 8.3×10^{-8} (one chance in 12 million); in 10^6 years, probability is 8.3×10^{-4} (one chance in 1200)

Chapter 8

30. 0.40 kg after 1.3 billion years; 0.20 kg after 2.6 billion years; 0.10 kg after 3.9 billion years **31.** 2.6 billion years **33.** (**a**) About 180 AU **34.** 2.9×10^7 AU = 140 parsecs = 460 light-years **35.** (**a**) About 600 AU (**b**) About 5×10^{40} cubic meters (**c**) About 10^{55} atoms (**d**) About 3×10^{14} atoms per cubic meter **36.** 860 years **37.** 2.2×10^{30} kg = 1.1 times the mass of the Sun **38.** (**a**) 12 m/s (**b**) 1.3×10^{-3} arcsec (**c**) 9.0×10^{-5} arcsec

Chapter 9

29. (**a**) 6.8×10^{16} W (**b**) 1.07×10^{17} W (**c**) 209 watts per square meter (**d**) 246 K = −27°C **30.** 4 km **31.** Core: 17%; mantle: 82%; crust: 1% **32.** 0.020 (2% of the total mass) **33.** (**b**) About 15,000 kg/m^3

Chapter 10

26. (**a**) 4671 km (**b**) 1707 km below the surface (**c**) 449 km **30.** 130 newtons on the Moon, 780 newtons on Earth **32.** (**a**) 5.78×10^{-5} newton (**b**) 4.15×10^{-5} newton (**c**) 1.39 **36.** 2.56 seconds **38.** (**b**) 10^6 times greater

Chapter 11

46. 13.0 arcsec **48.** (**a**) 0.16 AU **51.** (**a**) 3.03 m/s (**b**) 1.26 nm **52.** 0.615 AU **55.** For T = 460°C, λ_{max} = 4.0 μm **63.** (**a**) 3770 km (**b**) 370 km **64.** Difference between the round-trip times is 2.0×10^{-4} s **66.** 920 m **73.** Phobos: about 16 arcminutes; Deimos: about 2.7 arcminutes **74.** (**a**) Radius = 20,400 km, altitude = 17,000 km

Chapter 12

37. 12.7 km/s **40.** 8.5×10^{53} hydrogen atoms, 7.1×10^{52} helium atoms **41.** Roughly 600 km/h **43.** 127 K **45.** 59.5 km/s **46.** 8300 newtons **47.** 7.1×10^4 kg/m^3 **51.** (**a**) 14.4 hours for inner edge of A ring, 7.9 hours for inner edge of B ring

Chapter 13

40. 760 km **44.** 2.8×10^{11} (280 billion) years **45.** About 4×10^{-11} (one part in 3×10^{10}) **46.** About 8.1 minutes **49.** Escape speed = 2.6 km/s; mass of molecule = 2.0×10^{-26} kg, corresponding to a molecular weight of 12 **52.** 0.09 arcsec **53.** (**a**) 65.7 hours (**b**) 27.6 arcsec

Chapter 14

29. Sun-Uranus force = 1.39×10^{21} newtons, Neptune-Uranus force = 2.24×10^{17} newtons; Neptune reduces the sunward gravitational pull on Uranus by 1.61×10^{-4}, or 0.0161% **31.** (**a**) 2900 kg/m^3 **32.** (**a**) 8.4 hours **35.** (**a**) 1.13×10^{-3} as bright as on Earth (**b**) 4.10×10^{-4} as bright as on Earth (**c**) 2.77 times as bright **36.** About 7200 km **38.** 0.95 arcsec **43.** Without the boost, a one-way trip would take 30.5 years

Chapter 15

33. 2.82 AU **35.** 30 km **37.** (**a**) 3.6×10^{12} kg (**b**) 0.83 m/s **39.** About 4% **40.** 3.6×10^7 km = 0.24 AU **41.** (**a**) Period = 350 years, lifetime = 3.5×10^4 years (**b**) Period = 1.1×10^4 years, lifetime = 1.1×10^6 years (**c**) Period = 3.5×10^5 years, lifetime = 3.5×10^7 years (**d**) Period = 1.1×10^7 years, lifetime = 1.1×10^9 years **44.** (**a**) 10^{15} kg (**b**) 10^{-16} kg/m^3

Chapter 16

29. (**a**) 1.8×10^{-9} J (**b**) 9.0×10^{16} J (**c**) 5.4×10^{41} J **30.** (**a**) 4.6×10^{-36} s (**b**) 2.3×10^{10} s (**c**) 1.4×10^{15} s = 4.4×10^{7} years **31.** 0.048 (4.8%) of the Sun's mass will be converted from hydrogen to helium; chemical composition of the Sun (by mass) will be 69% hydrogen, 30% helium **32.** (**a**) 8.8×10^{29} kg of hydrogen consumed, 6.2×10^{26} kg lost **34.** (**a**) 1.64×10^{-13} J (**b**) 2.43×10^{-3} nm **35.** 1.4×10^{13} kg/s **39.** 98,600 kg/m^3 **40.** 1.9×10^{-7} nm **42.** (**a**) 1700 nm **44.** For the photosphere, 500 nm; for the chromosphere, 58 nm; for the corona, 1.9 nm **47.** For the umbra, 670 nm; for the penumbra, 580 nm **48.** (**a**) (Flux from patch of penumbra)/(flux from patch of photosphere) = 0.55 (**b**) (Flux from patch of penumbra)/(flux from patch of umbra) = 1.8

Chapter 17

34. (**a**) 9.7 pc (**b**) 0.10 arcsec **35.** 6.54 pc **36.** (**a**) 161 km/s (**b**) 294 km/s **37.** Distance = 105 pc; tangential velocity would have to be about 5200 km/s **38.** 110 km/s **39.** (**a**) +59.4 km/s (**c**) 486.23 nm **45.** 37.0 AU **46.** 6.1 $L_\odot$ **47.** 0.38 pc = 7.9×10^{4} AU **48.** (**a**) +13.8 (**b**) (Luminosity of HIP 72509)/(luminosity of Sun) = 2.4×10^{-4} **49.** +17 **50.** 6300 pc **51.** (**a**) Brightest star has $M = -2.37$; dimmest star has $M = +4.32$ (**b**) $M = +0.79$ **54.** (**b**) $m_B - m_V = -0.23$ (Bellatrix), +0.68 (Sun), +1.86 (Betelgeuse) **55.** 99 $R_\odot$ **59.** The radius of star X is 17 times larger than the radius of star Y **60.** Radius increases by a factor of 2, luminosity increases by a factor of 64 **62.** T = 10,000 K, $R = 3.8\ R_\odot$ **63.** $L = 35\ L_\odot$, so distance = 14 pc **64.** (**b**) 0.26 $R_\odot$ = 1.8×10^{5} km **65.** 9700 years **66.** (**a**) 5.0 pc (**b**) 22.5 AU (**c**) 1.5 $M_\odot$ **67.** (**a**) 40 $M_\odot$ **68.** About 2500 pc, about 125 times greater volume

Chapter 18

30. 0.34% **32.** 3.4×10^{4} atoms per cm^3 **35.** 100 $R_\odot$ = 7.4×10^{8} km = 4.9 AU **37.** (**a**) 250 AU = 3.7×10^{10} km **39.** 1.3×10^{31} m^3 **41.** 2.2×10^{3} km/s, or 7.5×10^{-3} (0.75%) of the speed of light

Chapter 19

29. 657,000 km **30.** (**a**) 618 km/s (**b**) 61.8 km/s **31.** (**a**) 12.0 km/s (**b**) 9.3 km/s 32. 2.3×10^{29} kg, or 0.15 (15%) of the original mass of hydrogen **36.** (**a**) 4.9×10^{7} years (**b**) 3.8×10^{11} years **38.** About 1900 K (1600°C, or 2900°F) **39.** 3.4×10^{7} years **44.** About 650 pc **45.** About 370 pc **49.** (**a**) 1.65×10^{6} km

Chapter 20

36. 0.11 $R_\odot$ **39.** We see the nebula about 5900 to 8200 years after the central star shed its outer layers **41.** (**a**) 97 nm **43.** (**a**) 1.8×10^{9} kg/m^3 (**b**) 6.5×10^{3} km/s **44.** About 4×10^{6} kg **46.** 0.53 AU **48.** (**a**) 1.7×10^{30} kg = 0.84 $M_\odot$ (**b**) 1.1×10^{12} newtons (**c**) 1.5×10^{8} m/s, or 0.5 (50%) of the speed of light **49.** (**a**) 6.4×10^{5} m/s **50.** 3.2×10^{9} km = 22 AU **51.** (Maximum luminosity of SN 1993J)/(maximum luminosity of SN 1987A) = 4.5 **53.** (**a**) 7×10^{-7} $b_\odot$ (**b**) It would be about 700 times brighter than Venus **54.** 1.3×10^{8} pc = 130 Mpc

Chapter 21

33. About 5500 B.C. **35.** (**a**) About 700 A.D. **36.** (**a**) Radius = 4.6 ly, diameter = 9.2 ly **37.** (**b**) about 40 pc **38.** (**b**) Maxiumum correction = 10^{-4} of the pulsar period **39.** (**a**) 1250 years **40.** (**a**) 5.07×10^{-11} seconds per second (b) 2330 years **41.** (**a**) Density of matter in a neutron = 4×10^{17} kg/m^3 **46.** (**a**) 0.097 nm (**b**) 5.8×10^{31} W = 1.5×10^{5} $L_\odot$ **47.** (**b**) 48 nm (**c**) 14 km

Chapter 22

33. 0.6 of the speed of light **34.** 8.3×10^{-8} s **35.** 0.8 of the speed of light **36.** (**a**) 25 years (**b**) 20 light-years as measured by an Earth observer, 12 light-years as measured by the astronaut **38.** 2.8 $M_\odot$ **39.** 5.7×10^{8} years **40.** (**a**) 2.01×10^{6} km **44.** 0.32 year **45.** 0.84 m **46.** (**a**) R_{Sch} = 8.9 mm, density = 2.0×10^{30} kg/m^3 (**b**) R_{Sch} = 3.0 km, density = 1.8×10^{19} kg/m^3 (**c**) $R_{Sch} = 3.5 \times 10^{9}$ km = 24 AU, density = 13 kg/m^3 **47.** 7.4×10^{30} kg **48.** 2.9×10^{17} kg/m^3 **49.** 2.7×10^{38} kg = 1.4×10^{8} $M_\odot$

Chapter 23

30. (**a**) 1.2×10^{12} cubic parsecs (**b**) 1.1×10^{8} cubic parsecs (**c**) Probability = 9.6×10^{-5}; we can expect to see a supernova with 300 pc once every 350,000 years **31.** (**a**) 8.3×10^{9} AU (**b**) 4.0×10^{4} pc **33.** 9.5×10^{-25} J; it takes 3.2×10^{5} such photons to equal the energy of one H_α photon **34.** (**b**) 5700 pc **35.** 21 times **37.** A 10% error in radius results in a 10% error in mass; a 10% error in velocity results in a 20% error in mass **38.** (**a**) 3.1×10^{8} years (**b**) 7.4×10^{11} $M_\odot$ **39.** 2.7×10^{11} $M_\odot$ **44.** (**a**) 1.1×10^{7} km = 0.073 AU (**b**) 9.1×10^{-6} arcsec (**c**) 251 arcsec **47.** (**a**) 6.3 years (**b**) 3.7×10^{6} $M_\odot$

Chapter 24

35. (**a**) 6.9 Mpc (**b**) 8.3 Mpc (**c**) 1.4 Mpc **36.** 9.5 Mpc **37.** (**a**) 1.1×10^{10} km = 70 AU (**b**) 7.0 Mpc **38.** (**a**) 152 Mpc = 4.96×10^{8} ly **39.** 54 km/s/Mpc **40.** (**a**) $z = 0.0252$ (**b**) 106 Mpc **41.** (**a**) 2.85×10^{5} km/s, or 0.951 (95.1%) of the speed of light (**b**) 1.60×10^{6} km/s (**c**) 4020 Mpc = 1.31×10^{10} ly **42.** (**a**) 1.2×10^{70} atoms (**b**) 3.6×10^{-6} atom per cm^3 **43.** 1.2×10^{12} $M_\odot$ **44.** Period = 4.4×10^{8} years; mass = 3.6×10^{11} $M_\odot$

Chapter 25

27. 2.87×10^{5} km/s, or 0.958 (95.8%) of the speed of light **28.** 2.93×10^{5} km/s, or 0.976 (97.6%) of the speed of light **33.** (**a**) 30 years (**b**) 2014 (**c**) 5/3 of the speed of light **34.** (**a**) 1.80×10^{5} km/s, or 0.600 (60.0%) of the speed of light (**b**) 134 hours (**c**) 970 AU **35.** (**a**) 1.2×10^{11} $L_\odot$ **36.** 1.5×10^{8} $M_\odot$ 37. 2.9×10^{9} km = 20 AU

Chapter 26

33. $z = 4$ (**b**) Distances were 1/9 as great and the density of matter was 729 times greater **35.** (**a**) 20 billion years (**b**) 13 billion years (**c**) 9.7 billion years **36.** 1.6×10^{8} km/s/Mpc **37.** 65.0 times denser **38.** 1.06×10^{-3} m = 1.06 mm **39.** 13.6 K **40.** (**a**) 4.6×10^{-23} kg/m^3 **41.** (**a**) 9.5×10^{-18} kg/m^3 (**b**) 4.8×10^{-18} kg/m^3 (**c**) 1.3×10^{-7} kg/m^3 **43.** (**a**) 4.7×10^{-27} kg/m^3 (**b**) 1.9×10^{-26} kg/m^3 **45.** (**b**) $q_0 = -0.595$ (**c**) $\Omega_\Lambda = 0.135$ **46.** (**a**) −0.17 (**b**) +0.12 **47.** 1.3×10^{-53} m^{-2}

Chapter 27

27. 1.1×10^{-26} J **29.** 3.5×10^{-25} s **30.** (**a**) 80.5 GeV (**b**) 9.34×10^{14} K **32.** (**a**) 29 nm (**c**) 348 nm **33.** (**a**) 2.0×10^{-35} kg = 2.2×10^{-5} electron mass **34.** (**a**) 1.1×10^{14} m = 0.011 ly **35.** Length was 0.168 of its present-day value; density was 213 times the present-day value

Chapter 28

11. (**a**) 3.7 cm **12.** (**a**) 10,000 years (b) 10 million years **14.** (**a**) 1430 MHz **15.** (**a**) 29.8 km/s (b) 0.10 m (c) 9.9×10^{-6} m, or 9.9×10^{-3}% of the unshifted wavelength **16.** 200 km

Index

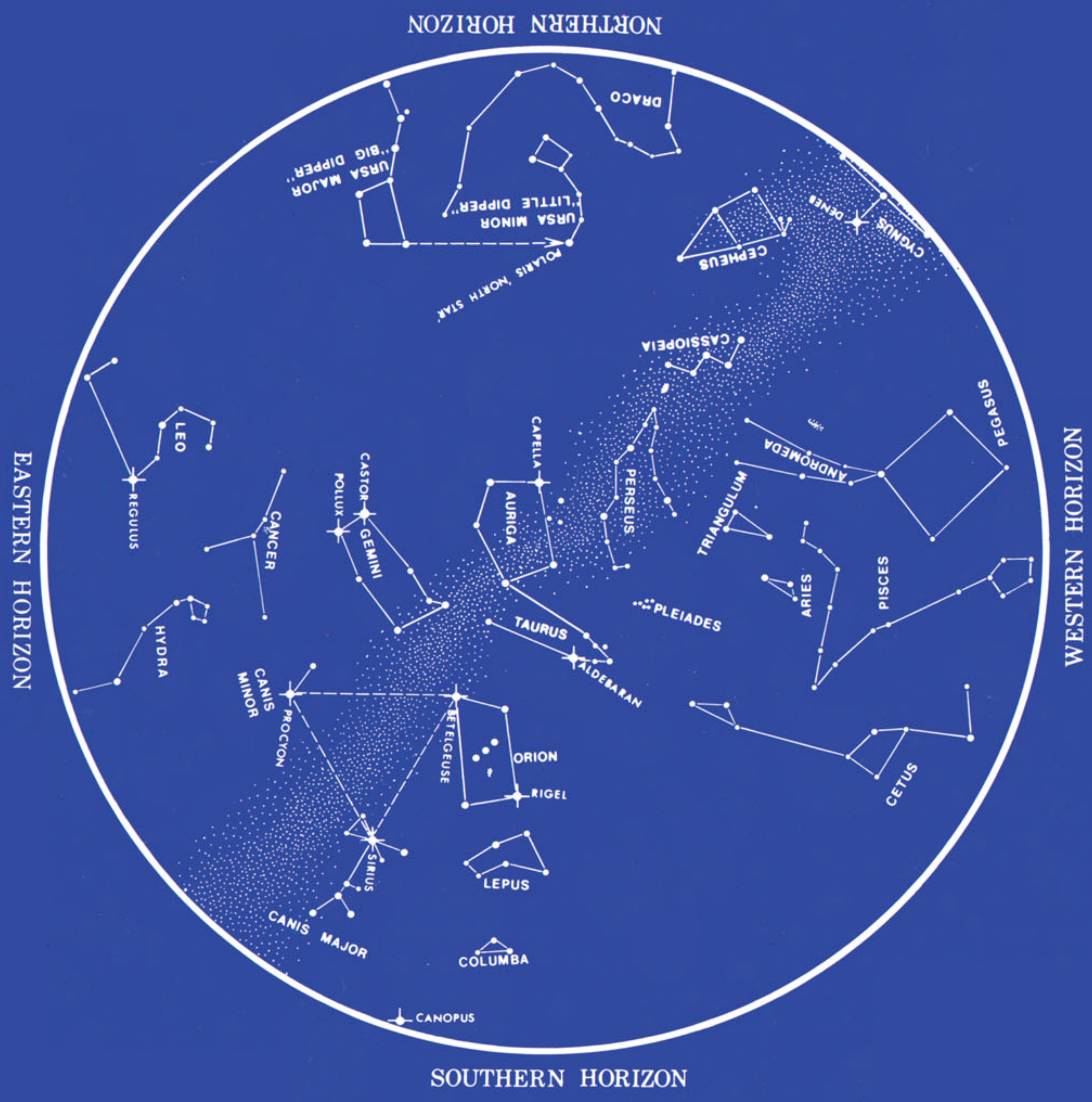
Page numbers in *italics* indicate illustrations.

This page intentionally left blank

Star Charts

The following set of star charts, one for each month of the year, are from *Griffith Observer* magazine. They are useful in the northern hemisphere only. For a set of star charts suitable for use in the southern hemisphere, see the *Universe* Web site (**www.whfreeman.com/universe8e**).

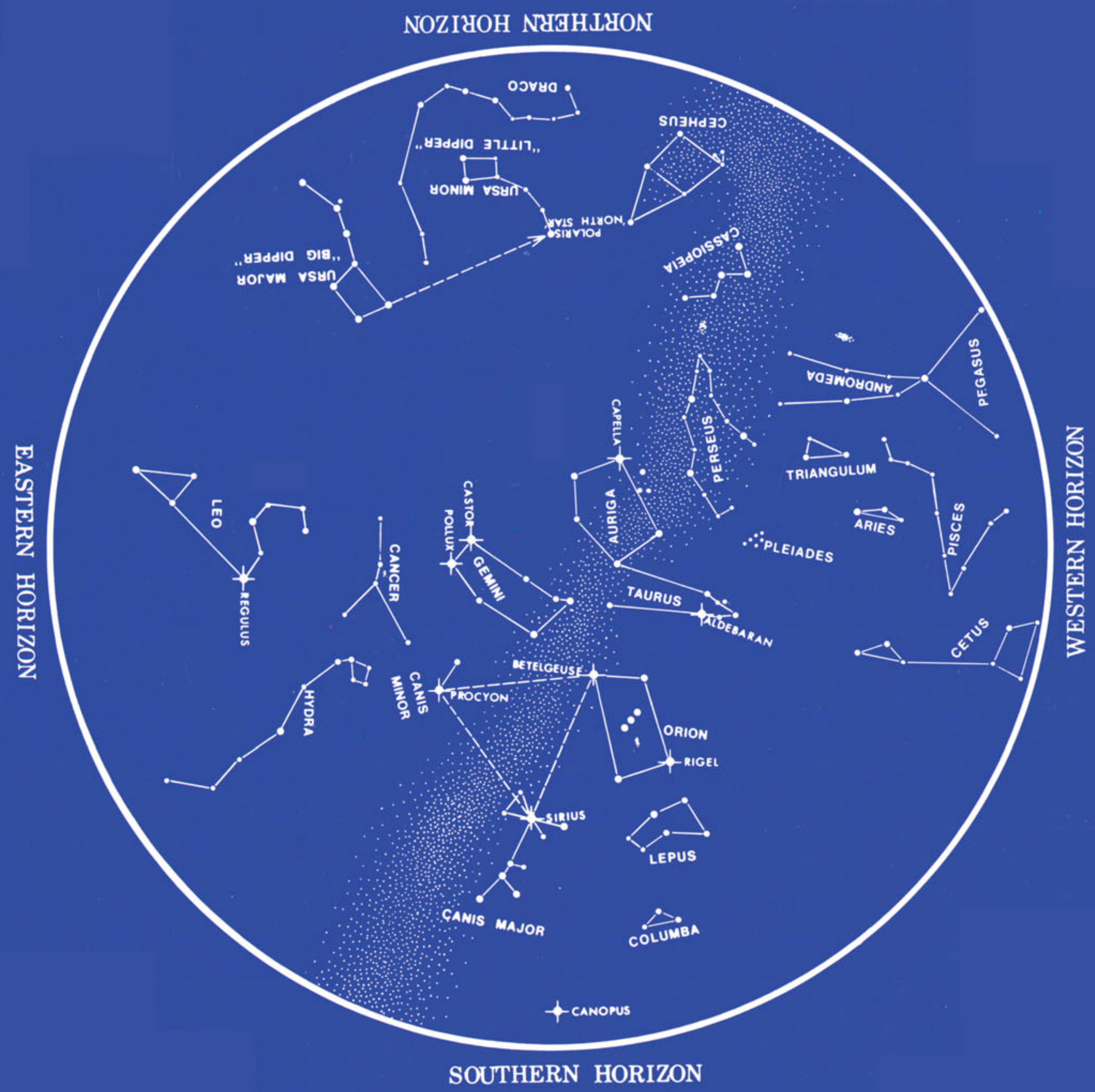
To use these charts, first select the chart that best corresponds to the date and time of your observations. Hold the chart vertically as shown in the above illustration and turn it so that the direction you are facing is shown at the bottom.



THE NIGHT SKY IN JANUARY

Chart time (Local Standard Time):

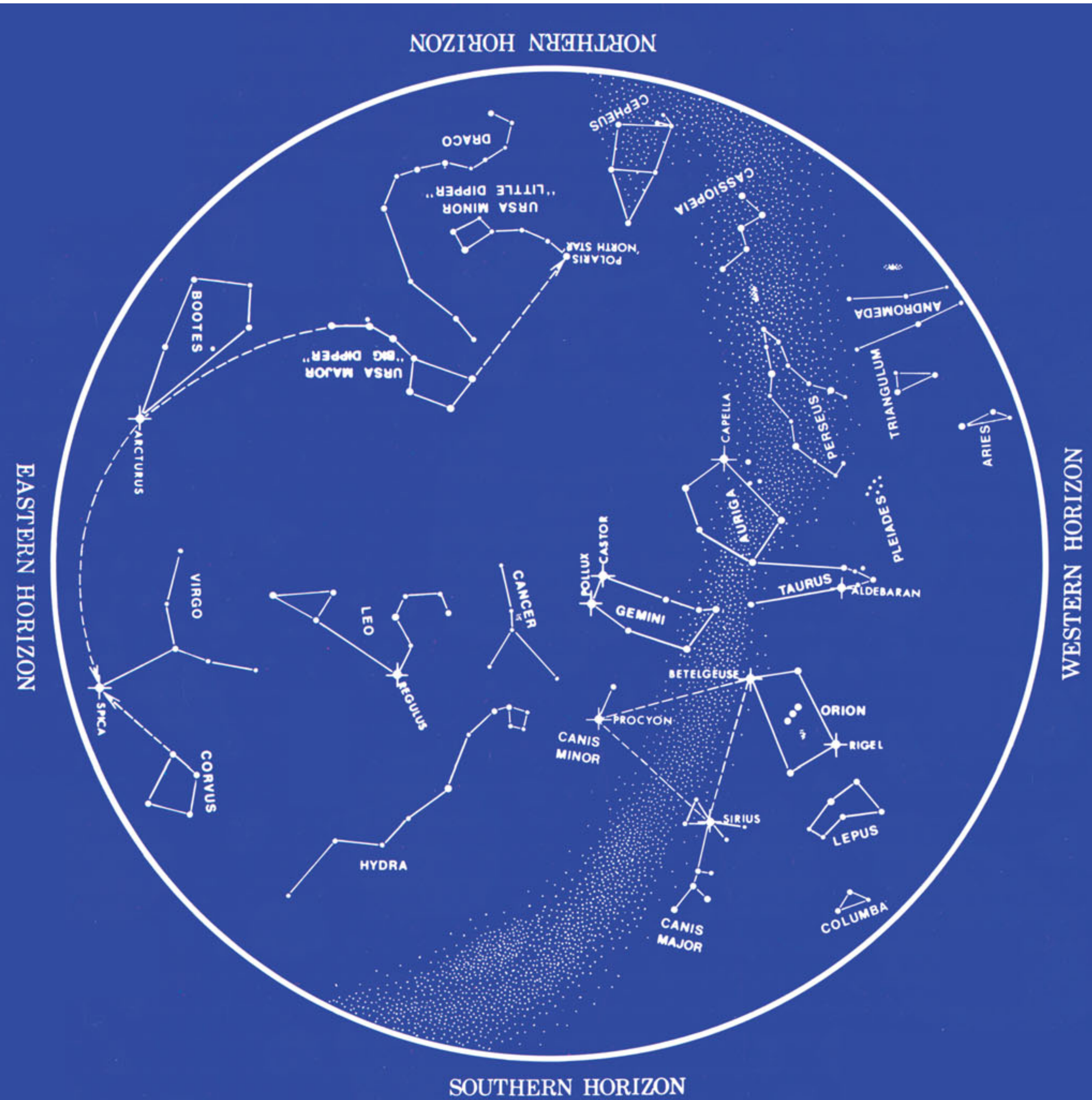
10 pm . . . First of January
9 pm . . . Middle of January
8 pm . . . Last of January



THE NIGHT SKY IN FEBRUARY

Chart time (Local Standard Time):

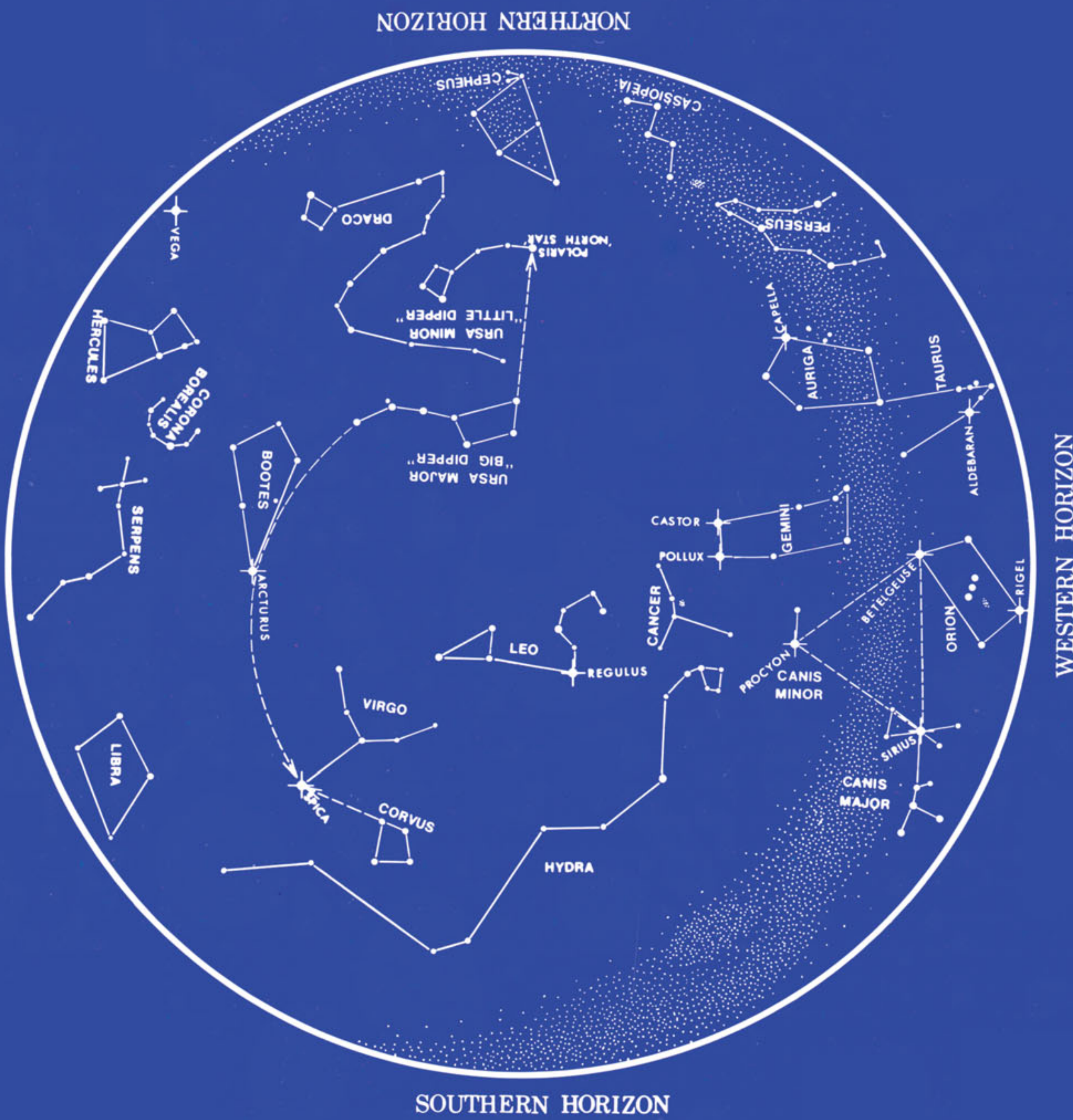
10 pm . . . First of February
9 pm . . . Middle of February
8 pm . . . Last of February



THE NIGHT SKY IN MARCH

Chart time (Local Standard Time):

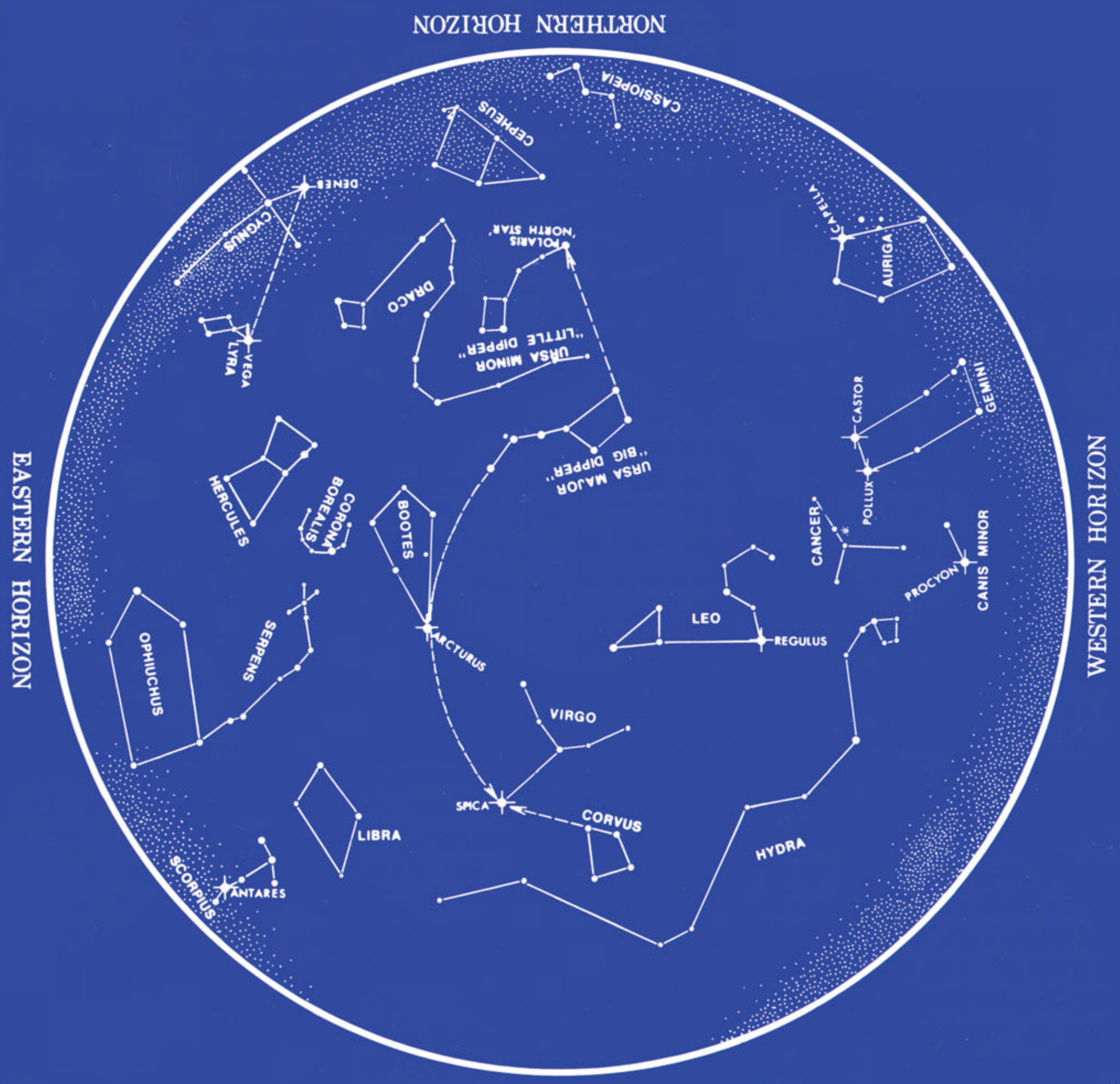
10 pm . . . First of March
9 pm . . . Middle of March
8 pm . . . Last of March



THE NIGHT SKY IN APRIL

Chart time (Daylight Saving Time):

11 pm . . . First of April
10 pm . . . Middle of April
9 pm . . . Last of April

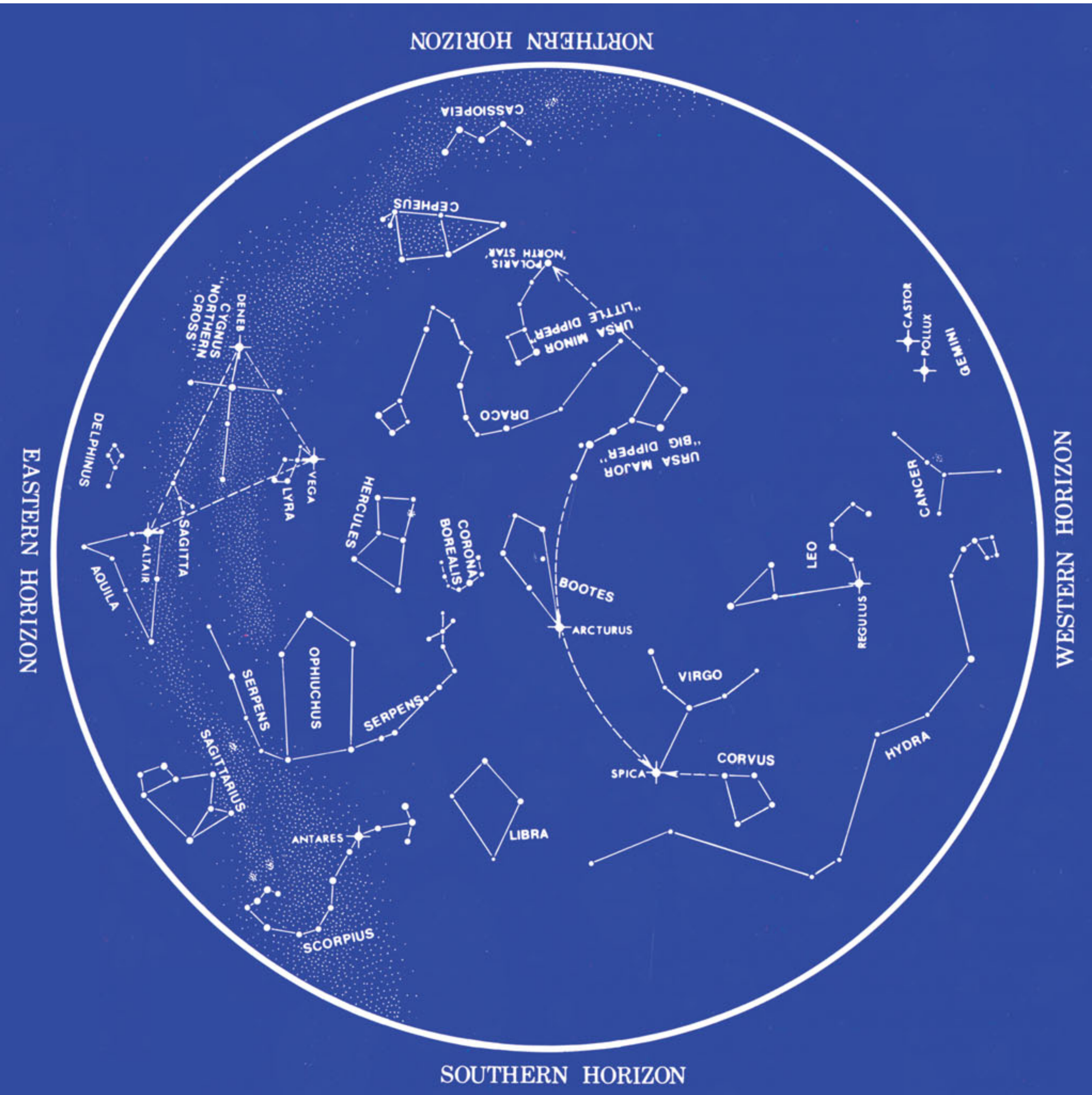


SOUTHERN HORIZON

THE NIGHT SKY IN MAY

Chart time (Daylight Saving Time):

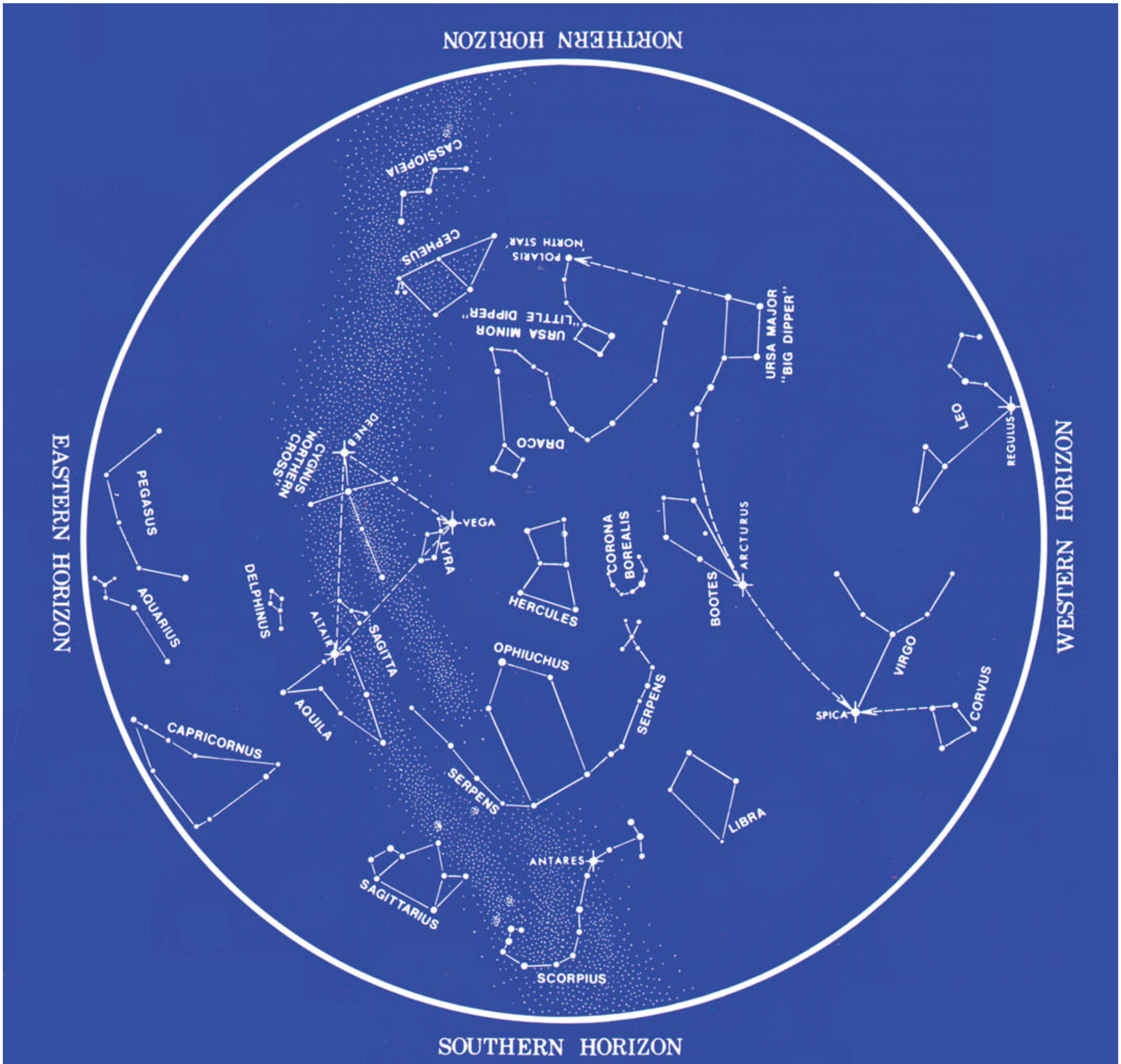
11 pm . . . First of May
10 pm . . . Middle of May
9 pm . . . Last of May



THE NIGHT SKY IN JUNE

Chart time (Daylight Saving Time):

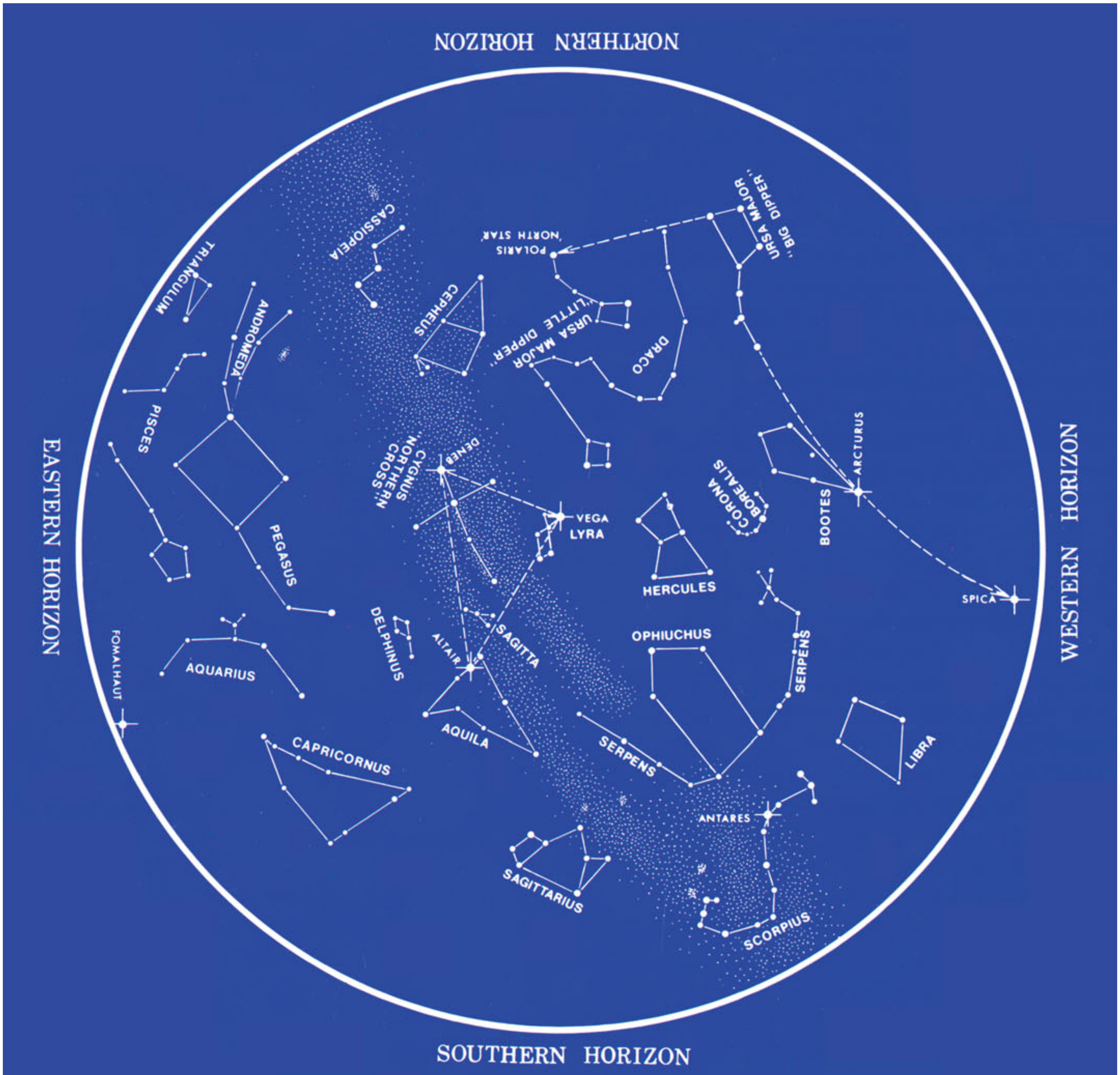
11 pm . . . First of June
10 pm . . . Middle of June
9 pm . . . Last of June



THE NIGHT SKY IN JULY

Chart time (Daylight Saving Time):

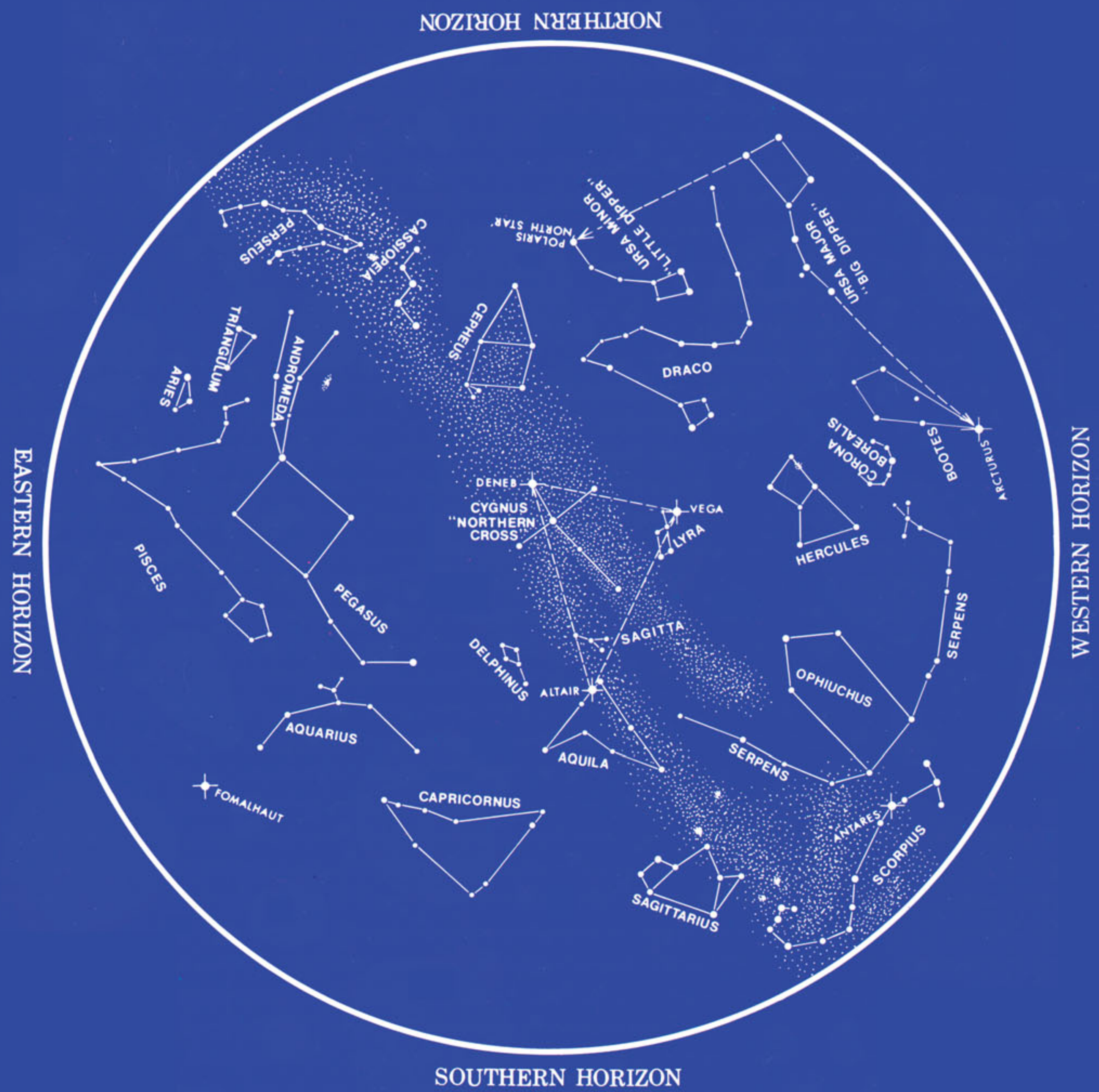
11 pm . . . First of July
10 pm . . . Middle of July
9 pm . . . Last of July



THE NIGHT SKY IN AUGUST

Chart time (Daylight Saving Time):

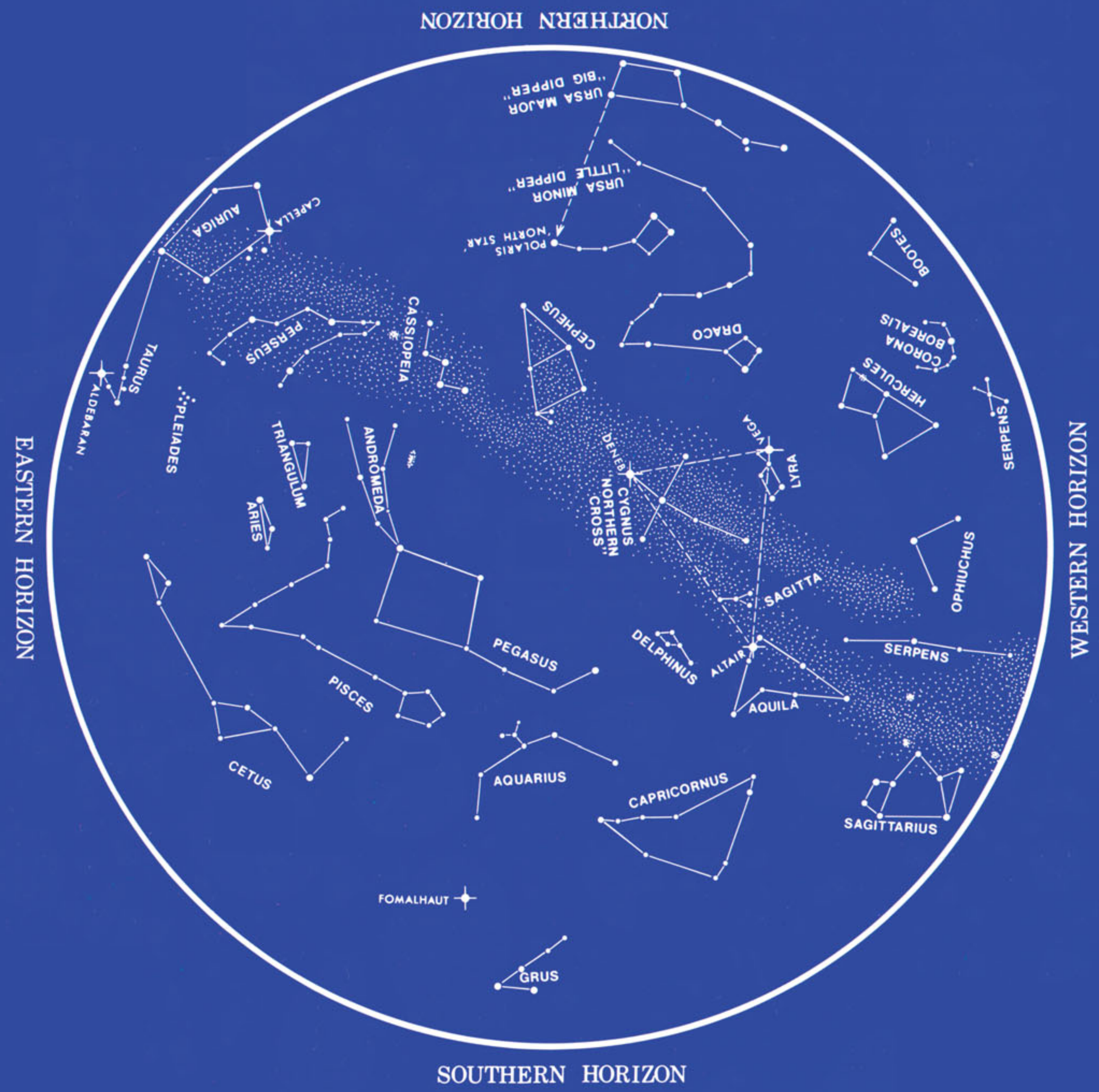
11 pm . . . First of August
10 pm . . . Middle of August
9 pm . . . Last of August



THE NIGHT SKY IN SEPTEMBER

Chart time (Daylight Saving Time):

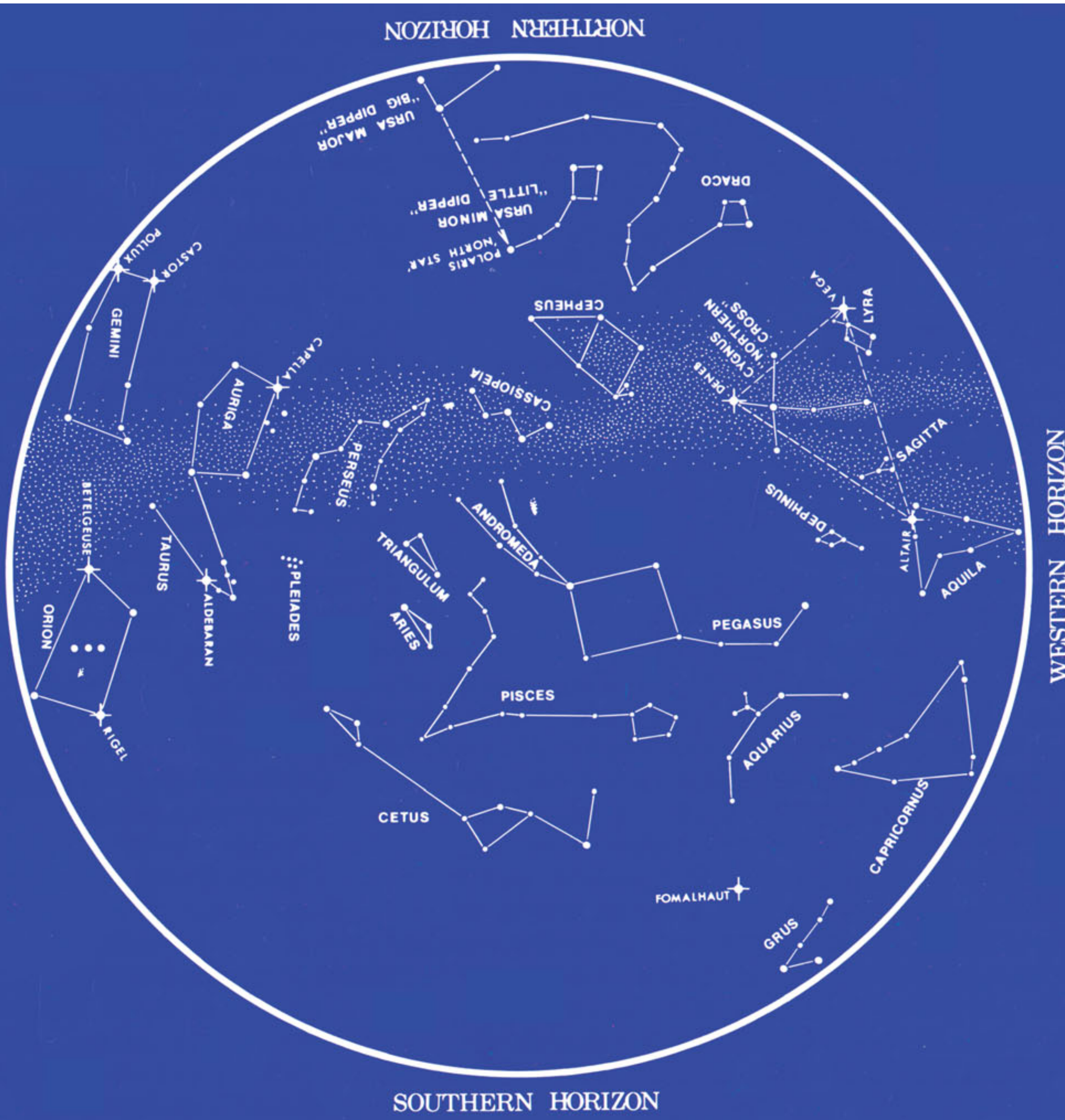
11 pm... First of September
10 pm... Middle of September
9 pm... Last of September



THE NIGHT SKY IN OCTOBER

Chart time (Daylight Saving Time):

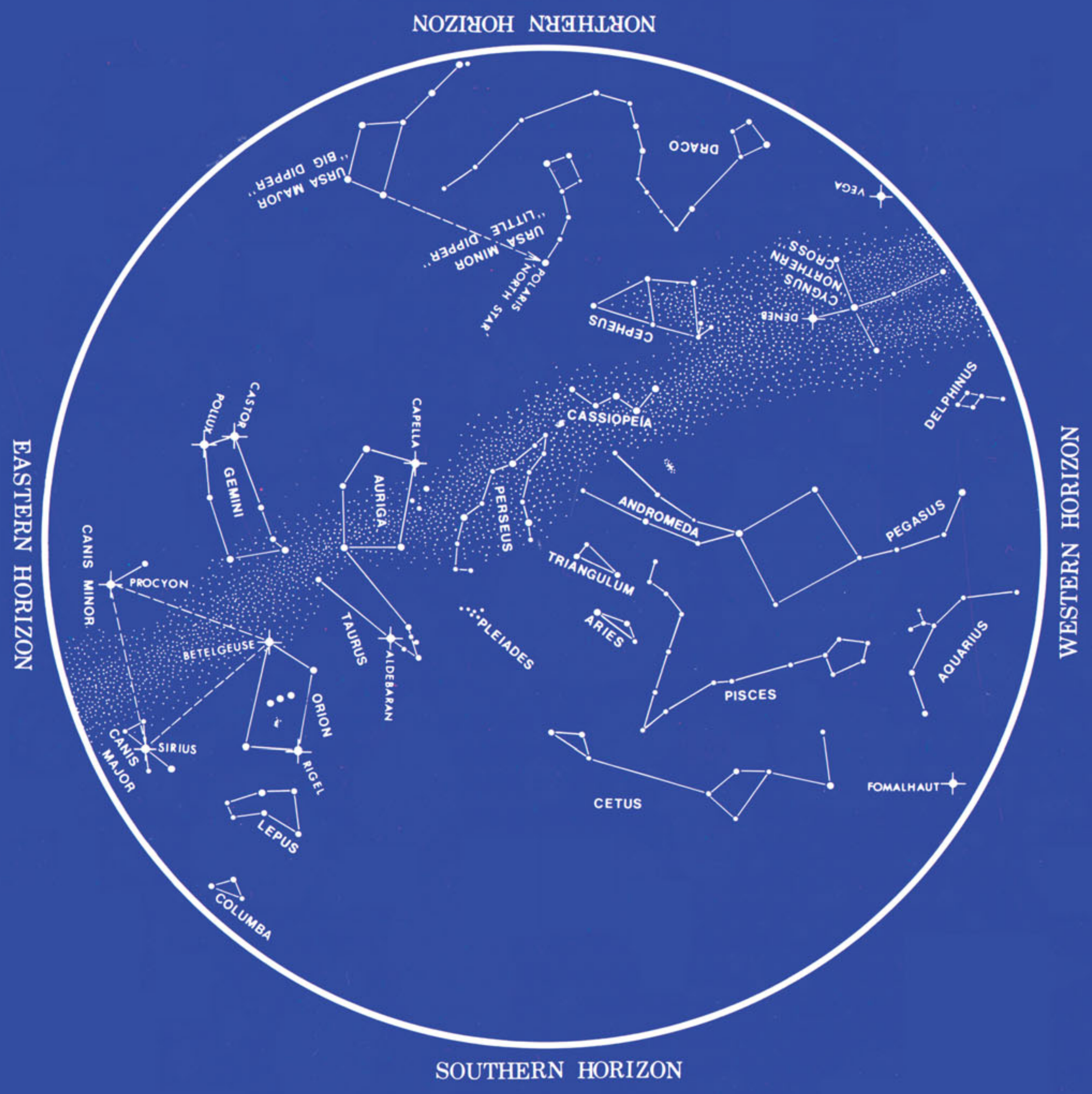
11 pm . . . First of October
10 pm . . . Middle of October
9 pm . . . Last of October



THE NIGHT SKY IN NOVEMBER

Chart time (Local Standard Time):

10 pm . . . First of November
9 pm . . . Middle of November
8 pm . . . Last of November



THE NIGHT SKY IN DECEMBER

Chart time (Local Standard Time):

10 pm . . . First of December
9 pm . . . Middle of December
8 pm . . . Last of December