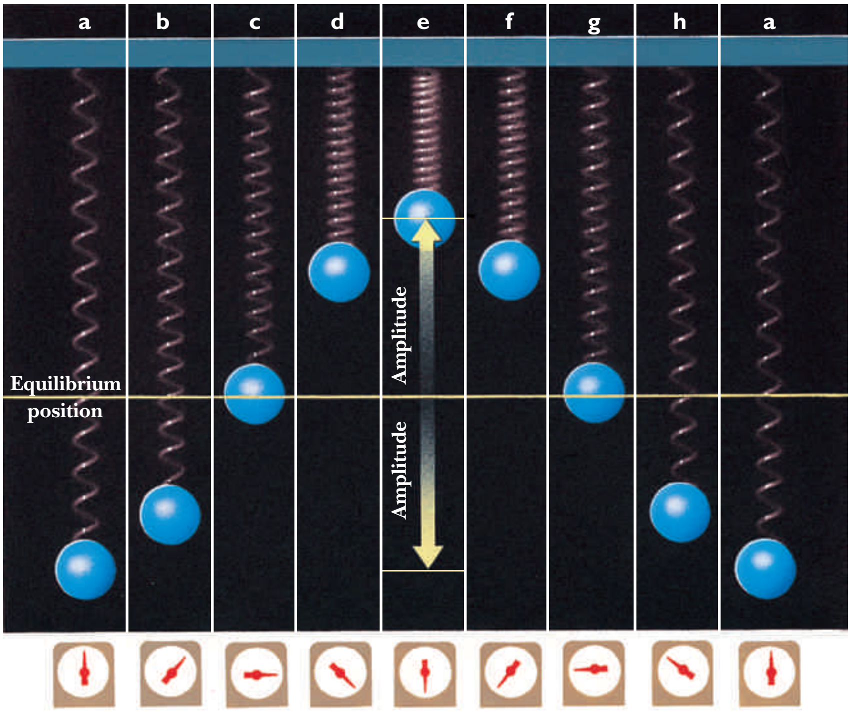


Figure 15-2 A time sequence showing one complete cycle for the vibration of a mass on a spring. The clocks show that equal time intervals separate the images.

This distance is marked in Figure 15-2 and is known as the **amplitude** of the vibration. In real situations the amplitude decreases and eventually the motion dies out because of the frictional effects that convert mechanical energy into thermal energy.

We can describe the time dependence of the vibration equally well by giving its **frequency** f, the number of cycles that occur during a unit of time. Frequency is often measured in cycles per second, or hertz (Hz). For example, concert A (the note that orchestras use for tuning) has a frequency of about 440 hertz, household electricity oscillates at 60 hertz, and your favorite FM station broadcasts radio waves near 100 million hertz.

There is a simple relationship between the frequency f and the period T—one is the reciprocal of the other:

$$f = \frac{1}{T}$$

◀ frequency $= \dfrac{1}{\text{period}}$

$$T = \frac{1}{f}$$

◀ period $= \dfrac{1}{\text{frequency}}$

To illustrate this relationship, let's calculate the period of a spring vibrating at a frequency of 4 hertz:

$$T = \frac{1}{f} = \frac{1}{4\text{ Hz}} = \frac{1}{4\text{ cycles/s}} = \frac{1}{4}\text{ s}$$

This calculation shows that a frequency of 4 cycles per second corresponds to a period of $\frac{1}{4}$ second. This makes sense because a spring vibrating four times per second should take $\frac{1}{4}$ of a second for each cycle. (When we state the period, we know it refers to one cycle and don't write "second per cycle.")

Are You On the Bus?

Q: What is the period of a mass that vibrates with a frequency of 10 times per second?

A: Because the period is the reciprocal of the frequency, we have

$$T = \frac{1}{f} = \frac{1}{10\ \text{Hz}} = 0.1\ \text{s}$$

We may guess that the time it takes to complete one cycle would change as the amplitude changes, but experiments show that the period remains essentially constant. It is fascinating that the amplitude of the motion does *not* affect the period and frequency. (Again, we have to be careful not to stretch the system "too much.") This means that a vibrating guitar string always plays the same frequency regardless of how hard the string is plucked.

Although the period for a mass vibrating on the end of a spring does not depend on the amplitude of the vibration, we may expect the period to change if we switch springs or masses. The stiffness of the spring and the size of the mass do change the rate of vibration.

The stiffness of a spring is characterized by how much force is needed to stretch it by a unit length. For moderate amounts of stretch or compression, this value is a constant known as the **spring constant** k. In SI units this constant is measured in newtons per meter. Larger values correspond to stiffer springs.

In trying to guess the relationship between the spring constant, mass, and period, we would expect the period to decrease as the spring constant increases because a stiffer spring means more force and therefore a quicker return to the equilibrium position. Furthermore, we would expect the period to increase as the mass increases because the inertia of a larger mass will slow the motion.

WORKING IT OUT *Period of a Mass on a Spring*

The mathematical relationship for the period of a mass on a spring can be obtained theoretically and is verified by experiment:

period of a mass on a spring ▶

$$T = 2\pi\sqrt{\frac{m}{k}}$$

where π is approximately 3.14.

As an example, consider a 0.2-kg mass hanging from a spring with a spring constant of 5 N/m:

$$T = 2\pi\sqrt{\frac{m}{k}} = 2\pi\sqrt{\frac{0.2\ \text{kg}}{5\ \text{N/m}}} = 6.28\sqrt{\frac{1}{25}\ \text{s}^2} = 1.26\ \text{s}$$

Therefore, this mass–spring combination vibrates with a period of 1.26 s, or a frequency of 0.793 Hz.

Are You On the Bus?

Q: What is the period of a 0.1-kg mass hanging from a spring with a spring constant of 0.9 N/m?

A: 2.09 s.

The Pendulum

The pendulum is another simple system that oscillates. Students are often surprised to learn (or to discover by experimenting) that the period of oscillation does not depend on the amplitude of the swing. To a very good approximation, large- and small-amplitude oscillations have the same period if we keep their amplitudes less than 30 degrees. This amazing property of pendula was first discovered by Galileo when he was a teenager sitting in church watching a swinging chandelier. (Clearly, he was not paying attention to the service.) Galileo tested his hypothesis by constructing two pendula of the same length and swinging them with different amplitudes. They swung together, verifying his hypothesis.

Let's consider the forces on a pendulum when it has been pulled to the right, as shown in Figure 15-3. The component of gravity acting along the string is balanced by the tension in the string. Therefore, the net force is the component of gravity at right angles to the string and directed toward the lower left. This restoring force causes the pendulum bob to accelerate toward the left. Although the restoring force on the bob is zero at the lowest point of the swing, the bob passes through this point (the equilibrium position) because of its inertia. The restoring force now points toward the right and slows the bob.

We found in free fall that objects with different masses fall with the same acceleration because the gravitational force is proportional to the mass. Therefore, we may expect that the motion of a pendulum would not depend on the mass of the bob. This prediction is true and can be verified easily by making two pendula of the same length with bobs of the same size made out of different materials so that they have different masses. The two pendula will swing side by side.

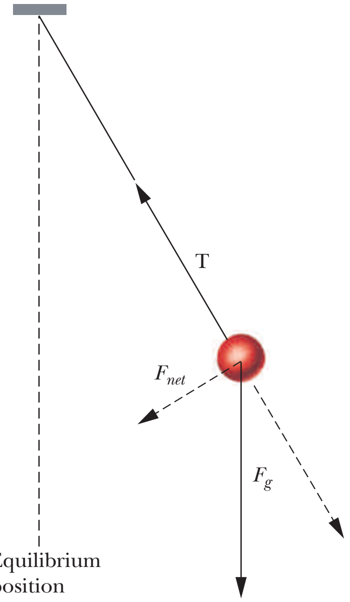


Figure 15-3 The net force on the pendulum bob accelerates it toward the equilibrium position.

Q: Why do we suggest using different materials?

A: If we use the same type of material, the size has to be different to get different masses. Different sizes may also affect the period. When doing an experiment, it is important to keep all but one factor constant.

We also know from our experiences with pendula that the period depends on the length of the pendulum; longer pendula have longer periods. Therefore, the length of the pendulum can be changed to adjust the period.

Because the restoring force for a pendulum is a component of the gravitational force, you may expect that the period depends on the strength of gravity, much as the period of the mass on the spring depends on the spring constant. This hunch is correct and can be verified by taking a pendulum to the Moon, where the acceleration due to gravity is only one-sixth as large as that on Earth.

Richarde Megna/Fundamental Photographs, NYC

A strobe photograph of a pendulum taken at 20 flashes per second. Note that the pendulum bob moves the fastest at the bottom of the swing.

Clocks

Keeping time is a process of counting the number of repetitions of a regular, recurring process, so it is reasonable that periodic motions have been important to timekeepers. Devising accurate methods for keeping time has kept many scientists, engineers, and inventors busy throughout history. The earliest methods for keeping time depended on the motions in the heavens. The day was determined by the length of time it took the Sun to make successive crossings of a north–south line and was monitored with a sundial. The month was

WORKING IT OUT | *Period of a Pendulum*

The period of a pendulum is given by

period of a pendulum ▶

$$T = 2\pi\sqrt{\frac{L}{g}}$$

As an example, consider a pendulum with a length of 10 m:

$$T = 2\pi\sqrt{\frac{L}{g}} = 2\pi\sqrt{\frac{10\ \text{m}}{10\ \text{m/s}^2}} = 6.28\sqrt{1\ \text{s}^2} = 6.28\ \text{s}$$

Therefore, this pendulum would oscillate with a period of 6.28 s.

Q: What would you expect for the period of a 1.7-m pendulum on the Moon?

A: 6.28 s.

© Cengage Learning/David Rogers

A replica of an early mechanical clock.

determined by the length of time it took the Moon to go through its phases. The year was the length of time it took to cycle through the seasons and was monitored with a calendar, a method of counting days.

As science and commerce advanced, the need grew for increasingly accurate methods of determining time. An early method for determining medium intervals of time was to monitor the flow of a substance such as sand in an hourglass or water in a water clock. Neither of these, however, was very accurate, and because they were not periodic, they had to be restarted for each time interval. It is interesting to note that Galileo kept time with a homemade water clock in many of his early studies of falling objects.

The next generation of clocks took on a different character, employing oscillations as their basic timekeeping mechanism. Galileo's determination that the period of a pendulum does not depend on the amplitude of its swing led to Christiaan Huygens's development of the pendulum clock in 1656, 14 years after Galileo's death. One of the difficulties Huygens encountered was to develop a mechanism for supplying energy to the pendulum to maintain its swing.

Seafarers spurred the development of clocks that would keep accurate time over long periods. To determine longitude requires measuring the positions of prominent stars and comparing these positions with their positions as seen from Greenwich, England, *at the same time*. Because pendulum clocks did not work on swaying ships, several cash prizes were offered for the design and construction of suitable clocks. Beginning in 1728, John Harrison, an English instrument maker, developed a series of clocks that met the criteria, but he was not able to collect his money until 1765. One of Harrison's clocks was accurate to a few seconds after 5 months at sea.

Any periodic vibration can be used to run clocks. Grandfather clocks use pendulums to regulate the hands and are powered by hanging weights. Mechanical watches have a balance wheel fastened to a spring. Electric clocks use 60-hertz alternating electric current. Digital clocks use the vibrations of quartz crystals or resonating electric circuits.

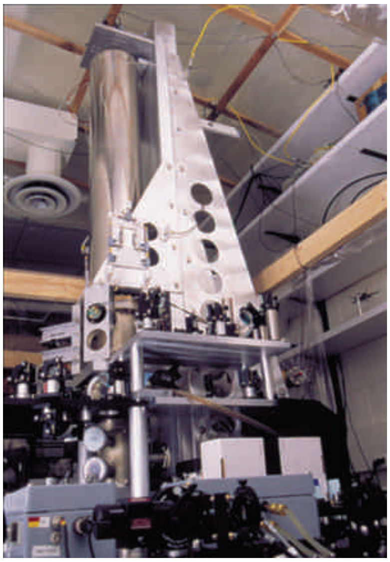
Modern time is kept with atomic clocks, which use the frequencies of atomic transitions (see Chapter 23) and are extremely insensitive to such changes in the clocks' environment as pressure and temperature. Atomic clocks are accurate to better than a second in 60 million years.

Resonance

We discovered with the mass on a spring and the pendulum that each system had a distinctive, natural frequency. The natural frequency of the pendulum is determined by its length and the acceleration due to gravity. Pulling the bob back and releasing it produces an oscillation at this particular frequency.

A child on a swing is an example of a life-size pendulum. If the child does not "pump" her legs and if no one pushes her, the amplitude of the swing continually decreases, and the child comes to rest. As every child knows, however, pumping or pushing greatly increases the amplitude. A less obvious fact is that the size of the effort—be it from pumping or pushing—is not important, but its timing is crucial. The inputs must be given at the natural frequency of the swing. If the child pumps at random times, the swinging dies out. This phenomenon of a large increase in the amplitude when a periodic force is applied to a system at its natural frequency is called **resonance.**

Resonance can also be achieved by using impulses at other special frequencies, but each of these has a definite relationship to the natural frequency. For example, if you push the child on the swing every other time, you are providing inputs at one-half of the natural frequency; every third time gives inputs at one-third of the natural frequency; and so on. Each of these frequencies causes resonance.



© Geoffrey Wheeler

Modern time is kept by extremely accurate atomic clocks such as this F1 operated by the National Institute of Standards and Technology. It is accurate to 1 second in 80 million years.

Q: What happens to the amplitude of the swing if you push at twice the natural frequency?

A: In this case you would be pushing twice for each cycle. One of the pushes would negate the other, and the swing would stop.

More complex systems also have natural frequencies. If someone strikes a spoon on a table, you are not likely to mistake its sound for that of a tuning fork. The vibrations of the spoon produce sounds that are characteristic of the spoon. All objects have natural frequencies. The factors that determine these frequencies are rather complex. In general, the dominant factors are the stiffness of the material, the mass of the material, and the size of the object.

Resonance can have either good or bad effects. Although your radio receives signals from many stations simultaneously, it plays only one station at a time. The radio can be tuned so that its resonant frequency matches the broadcast frequency of your favorite station. Tuning puts the radio in resonance with one particular broadcast frequency and out of resonance with the frequencies of the competing stations. On the other hand, if the radio has an inferior speaker with one or two strong resonant frequencies, it will distort the sounds from the radio station by not giving all frequencies equal amplification.

Suppose you have a collection of pendula of different lengths, as shown in Figure 15-4. Notice that two of these pendula have the same length and thus the same natural frequency. The pendula are not independent because they are all tied to a common string. The motion of one of them is felt by all the others through pulls by the string. If you start the left-hand pendulum swinging, its back-and-forth motion creates a tug on the common string with a frequency equal to its natural, or resonant, frequency. Pendula with different frequencies jiggle a little bit but are not affected much. However, the pendulum with the same frequency resonates with the input frequency, drastically increasing its amplitude. In exactly the same way, objects resonate when input frequencies are the same as any of their natural frequencies.

© Cengage Learning/George Semple

If children pump at the right frequencies, they can increase the amplitudes of their motions.

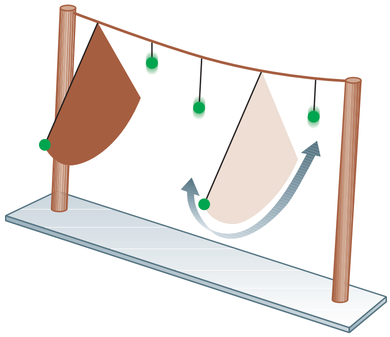


Figure 15-4 Pendula with the same natural frequency resonate with each other.

Everyday Physics | *Tacoma Narrows Bridge*

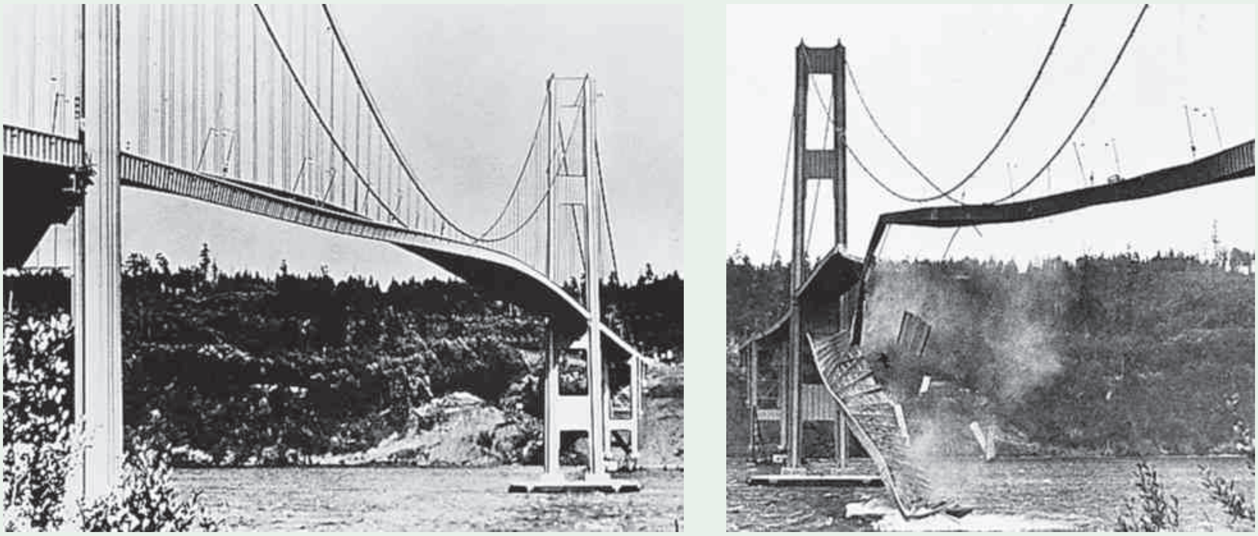
Resonant effects can sometimes have disastrous consequences. In 1940 a new bridge across one of the arms of Puget Sound in the state of Washington was opened to traffic. It was a suspension bridge with a central span of 850 meters (2800 feet). Because the bridge was designed for two lanes, it had a width of only 12 meters (40 feet). Within a few months after it opened, early-morning winds in the Sound caused the bridge to oscillate in standing-wave patterns that were so large in amplitude that the bridge failed structurally and fell into the water below, as shown in the figure.

But why did this bridge fail when other suspension bridges are still standing (including the bridge that now spans the Sound at the location of the original)? The bridge was long, narrow, and particularly flexible. Motorists often complained about the vertical oscillations and nicknamed the bridge "Galloping Gertie." However, the amplitudes of the vertical oscillations were relatively small until that fateful morning. The wind was blowing along the arm of the Sound (perpendicular to the length of the bridge) at moderate to high velocities but was not near gale force. One may speculate that fluctuations in the wind speed matched the natural frequency of the bridge, causing it to resonate. However, the wind was reasonably steady, and wind fluctuations are normally quite random. Furthermore, the forces would be horizontal, and the oscillations were vertical.

The best explanation involves the formation and shedding of vortices in the wind blowing past the bridge. *Vortices* are the eddies that you get near the ends of the oars when you row a boat. Vortices rotate in opposite directions in the wind blowing over and under the bridge. As each vortex is shed, it exerts a vertical impulse on the bridge. Therefore, if the frequency of vortex formation and shedding is near the natural frequency of the bridge for vertical oscillations, a standing wave will form just like those on a guitar string. (The frequency does not have to match exactly; it only needs to be close. How close depends on the details of the bridge construction.)

The bridge would have been fine except for another unfortunate circumstance. Besides the vertical standing wave, there were also torsional, or twisting, standing waves on the bridge. Normally, the frequencies of the two standing waves are quite different. But for the Tacoma Narrows Bridge, the two frequencies were fairly close (eight per minute for the vertical motion compared with ten per minute for the twisting motion). This allowed some of the energy from the vertical motion to be transferred to the twisting motion that eventually led to the mechanical failure of the bridge.

1. What was the periodic driving force that caused the standing wave to be formed on the bridge?
2. What uncommon characteristic of the Tacoma Narrows Bridge ultimately led to its mechanical failure?



Special Collections Div. Univ. of Washington Libraries, photo by Farquharson (both)

The Tacoma Narrows Bridge collapsed when winds set up resonant vibrations.

Waves: Vibrations That Move

Most **waves** begin with a disturbance of some material. Some disturbances, such as the clapping of hands, are onetime, abrupt events, whereas others,

Figure 15-5 A wave pulse travels along a line of dominoes.

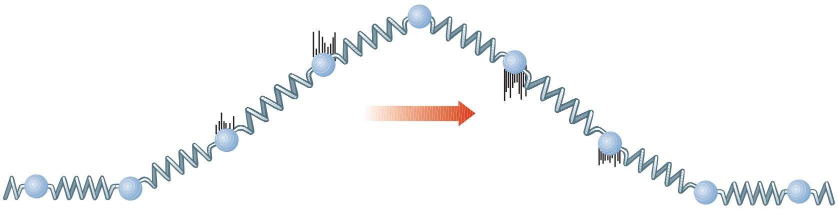


Figure 15-6 A wave disturbance can move along a chain of balls and springs.

such as the back-and-forth vibration of a guitar string, are periodic events. The simplest wave is a single pulse that moves outward as a result of a single disturbance.

Imagine a long row of dominoes lined up as shown in Figure 15-5. Once a domino is pushed over, it hits its neighbor, and its neighbor hits its neighbor, and so on, sending the disturbance along the line of dominoes. The key point is that "something" moves along the line of dominoes—from the beginning to the end—but it is not any individual domino.

Actually, the domino example is not completely analogous to what happens in most situations involving waves because there is no mechanism for restoring the dominoes in preparation for the next pulse. We can correct this omission by imagining a long chain of balls connected by identical springs, as shown in Figure 15-6. As the disturbance moves from left to right, individual balls are lifted up from their equilibrium positions and then returned to these equilibrium positions. The springs allow each ball to pull its neighbor away from equilibrium, just as the dominoes passed the disturbance from neighbor to neighbor by striking each other. After the pulse passes, the springs provide the restoring force that returns each ball to equilibrium. Notice that the pulse travels along the chain of balls without any of the balls moving in the direction of the pulse. Figure 15-6 shows the shape of the chain of balls and springs at the time the center ball reaches its maximum displacement.

In a similar manner, a pebble dropped into a pond depresses a small portion of the surface. Each vibrating portion of the surface generates disturbances in the surrounding water. As the process continues, the disturbance moves outward in circular patterns, such as those shown in Figure 15-7.

This type of disturbance, or pulse, occurs in a number of common, everyday events. A crowd transmits single pulses when a small group begins pushing. This push spreads outward through the crowd much as a ripple moves over a pond's surface. Similarly, the disturbance produced by a clap sends a single sound pulse through the air. Other examples of nonrecurrent waves include tidal bores, tidal waves, explosions, and light pulses emitted by supernovas (exploding stars).

Ian O'Leary/Stone/Getty

Figure 15-7 Water drops produce disturbances that move outward in circular patterns.

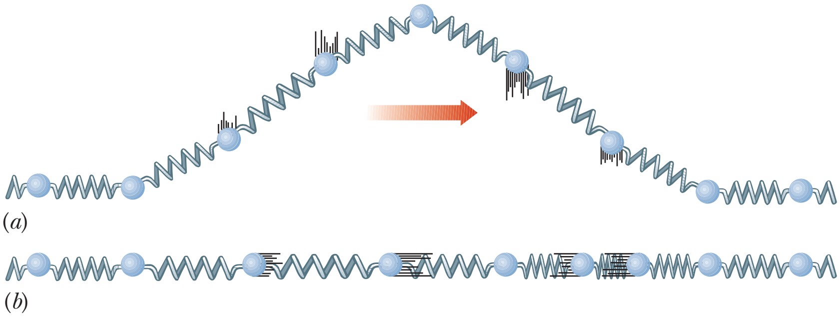


Figure 15-8 In a transverse wave (a), the medium moves perpendicular to the direction of propagation of the wave, whereas in a longitudinal wave (b), the medium's motion is parallel to the direction of propagation.

Although a wave moves outward from the original disturbance, there is no overall motion of the material. As the wave travels through the medium, the particles of the material vibrate about their equilibrium positions. Although the wave travels down the chain, the individual balls of the chain return to their original positions. The wave transports energy rather than matter from one place to another. The energy of an undisturbed particle in front of the wave is increased as the wave passes by and then returns to its original value. In a real medium, however, some of the energy of the wave is left behind as thermal energy in the medium.

There are two basic wave types. A wave in which the vibration of the medium is perpendicular to the motion of the wave is called a **transverse wave.** Waves on a rope are transverse waves. A wave in which the vibration of the medium is along the same direction as the motion of the wave is called a **longitudinal wave.** Both types can exist in the chain of balls. If a ball is moved vertically, a transverse wave is generated [Figure 15-8(a)]. If the ball is moved horizontally, the wave is longitudinal [Figure 15-8(b)].

Transverse waves can move only through a material that has some rigidity; transverse waves cannot exist within a fluid because the molecules simply slip by each other. Longitudinal waves, on the other hand, can move through most materials because the materials can be compressed and have restoring forces.

Q: Is it possible to have transverse waves on the surface of water?

A: Transverse surface waves are possible because the force of gravity tends to restore the surface to its flat equilibrium shape. Actually, the motion of the individual water molecules is a combination of transverse motion and longitudinal motion; the water molecules follow elliptical paths.

One-Dimensional Waves

Because all waves have similar properties, we can look at waves that are easy to study and then make generalizations about other waves. Imagine a clothesline tied to a post, as in Figure 15-9. A flick of the wrist generates a single wave pulse that travels away from you. On an idealized rope, the wave pulse would maintain its shape and size. On a real rope, the wave pulse slowly spreads out. We will ignore this spreading in our discussion. The wave's speed can be calculated by dividing the distance the pulse travels by the time it takes.

The speed of the wave can be changed. If you pull harder on the rope, the pulse moves faster; the speed increases as the tension in the rope increases. The speed also depends on the mass of the rope; a rope with more mass per

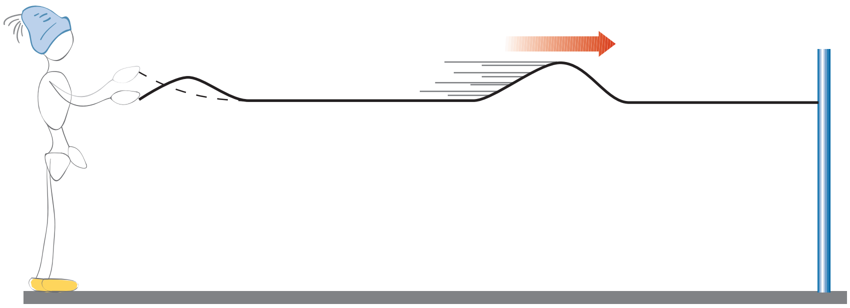


Figure 15-9 While pieces of the rope vibrate up and down, the wave moves along the rope.

unit length has a slower wave speed. Surprisingly, the amplitude of the pulse does not have much effect on the speed.

These observations make sense if we consider the vibrations of a small portion of the rope. The piece of rope is initially at rest and moves as the leading edge of the pulse arrives. How fast the rope returns to its equilibrium position determines how the pulse passes through the region (and hence, the speed of the pulse). The more massive the rope, the more sluggishly it moves. Also, if the rope is under a larger tension, the restoring forces on the piece of rope are larger and cause it to return to its equilibrium position more quickly.

FLAWED REASONING

A physics teacher has offered his class a prize if they can send a transverse pulse down a long spring and then send a second pulse down the same spring in such a manner as to catch up with the first pulse. Three students have taken up the challenge.

Trever: "I will make the first pulse with a slow movement of my hand and then make a second pulse with a very quick jerk on the spring. That should send the second pulse down the spring at a quicker speed."

Lindzee: "I think the amplitude of the pulse is what matters, not how fast you move your hand. Send the first pulse down with a big amplitude and then send the second pulse down with a small amplitude."

Courtnee: "The textbook claims that pulse speeds don't depend on how the pulse was created but only on the tension and the mass density of the spring. We can't change the mass density after we send the first pulse, but we could tighten the spring. Send the first pulse and then pull the spring tighter before we send the second pulse."

All three students are wrong. **Find the flaws in their claims.**

ANSWER Courtnee correctly points out the flaws in her classmates' reasoning. The speed of a pulse down a stretched spring does not depend on the size or shape of the pulse or the manner in which it was created. Courtnee's suggestion of stretching the spring will indeed increase the speed of transverse pulses on the spring, but this will speed up *both* pulses. The physics teacher tricked them with an impossible challenge.

When a pulse hits the end that is attached to the post, it "bounces" off and heads back. This reflected pulse has the same shape as the incident pulse but is inverted, as shown in Figure 15-10. If the incident pulse is an "up" pulse (a **crest**), the reflected pulse is a "down" pulse (a **trough**). If the end of the rope is free to move up and down, the pulse still reflects, but no inversion takes place.

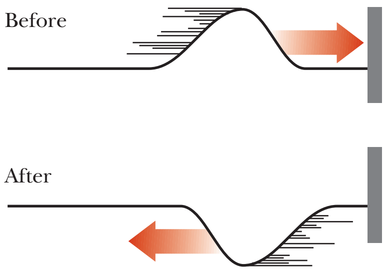


Figure 15-10 A wave pulse is inverted when it reflects from a fixed end. Note that the steep edge leads on the way in and on the way out.

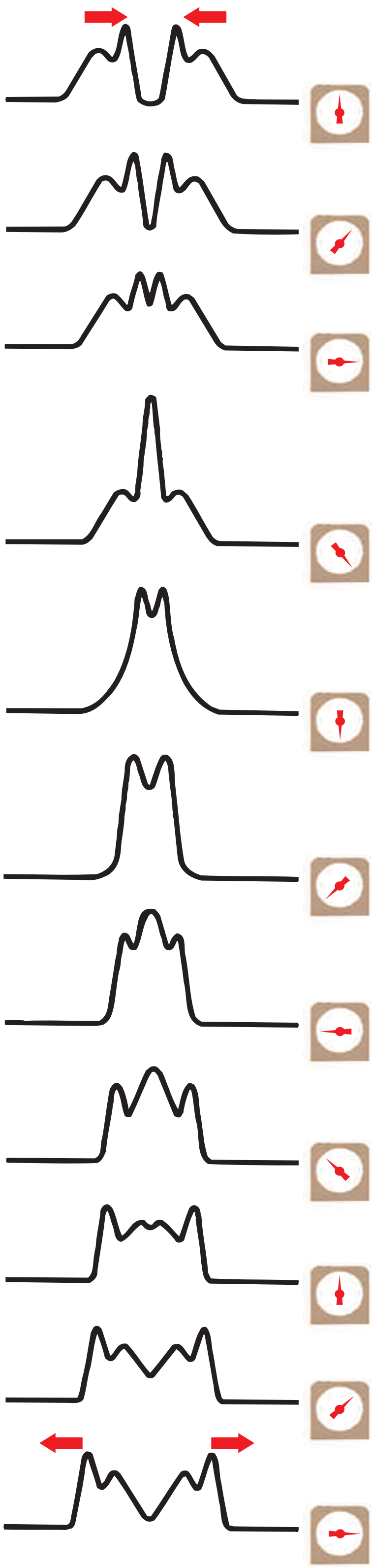


Figure 15-11 In contrast to a wave, the blue half of the ball leads on the way in and trails after reflection from the wall.

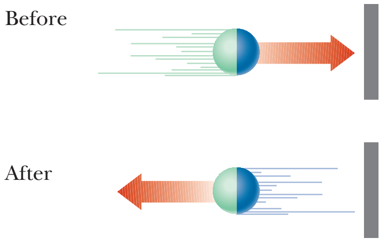


Figure 15-12 The two wave pulses on the rope pass through each other as if the other were not present.

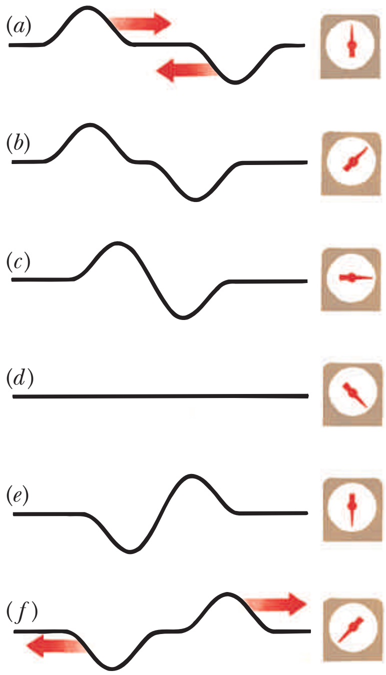


Figure 15-13 This time sequence shows that the superposition of two wave pulses yields shapes that are the sum of the individual shapes.

What about the front and back of the wave pulse? To observe this you would generate a pulse that is not symmetric. The pulses shown in Figure 15-10 are steeper in front than in back. The steeper edge is away from you when the pulse moves down the rope and toward you when the reflected pulse returns. The leading edge continues to lead.

These inversions contrast with the behavior of a ball when it "reflects" from a wall. If the ball is not spinning, the top of the ball remains on top, but the leading edge is interchanged. Figure 15-11 shows that the blue half leads before the collision, whereas the green half leads afterward.

Superposition

Suppose you send a crest down the rope and, when it reflects as a trough, you send a second crest to meet it. An amazing thing happens when they meet. The waves pass through each other as if the other were not there. This is shown in Figure 15-12. Each pulse retains its own shape, clearly demonstrating that the pulses are not affected by the "collision." A similar thing happens when you throw two pebbles into a pond. Even though the wave patterns overlap, you can still see a set of circular patterns move outward from *each* splash.

In contrast, imagine what would happen if two particles—say, two Volkswagens—were to meet. Particles don't exhibit this special property of waves. It would certainly be a strange world if waves did not pass through each other. Two singers singing at the same time would garble each other's music, or the sounds from one might bounce off those from the other.

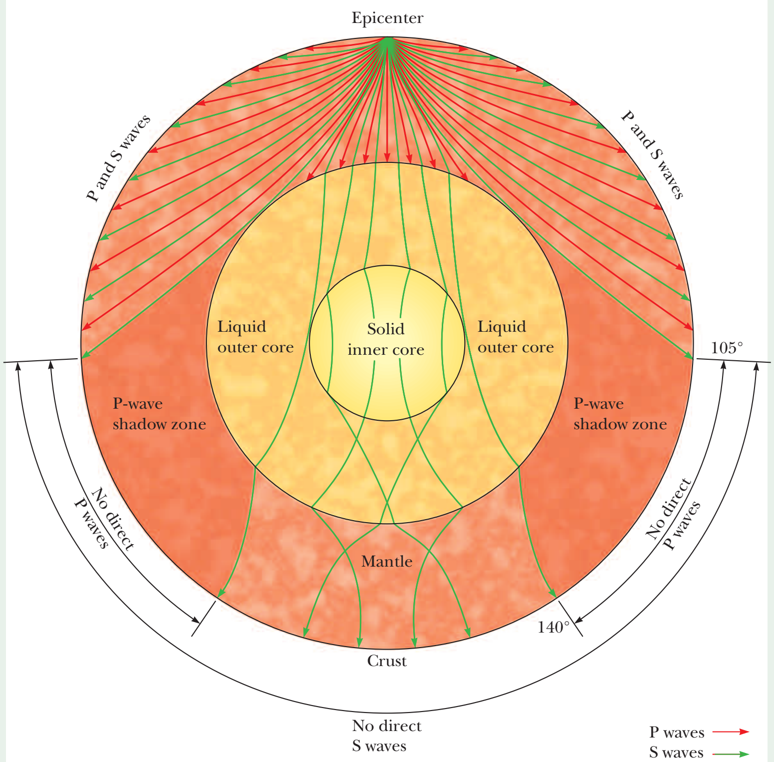
During the time the wave pulses pass through each other, the resulting disturbance is a combination of the individual ones; it is a **superposition** of the pulses. As shown in Figure 15-13, the distance of the medium from the equilibrium position, the **displacement,** is the algebraic sum of the displacements of the individual wave pulses. If we consider displacements above the equilibrium position as positive and those below as negative, we can obtain the shape of the resultant disturbance by adding these numbers at each location along the rope.

Everyday Physics | *Probing Earth*

Imagine drawing a circle to represent Earth. Further, imagine drawing a dot to show how far Earth's interior has been explored by direct drilling and sampling techniques. Where would you place the dot? The dot should be placed on the original circle. Earth is about 6400 kilometers (4000 miles) in radius, and we have drilled into its interior only about 12.2 kilometers, less than 0.2% of the distance to the center. Therefore, we must learn about Earth's interior using indirect means such as looking at signals from explosions and earthquakes.

Three kinds of wave are produced in an earthquake. One type travels along the surface, and the other two travel through Earth's interior; one of the interior waves is a longitudinal wave, and the other is a transverse wave. These waves move outward in all directions from the earthquake site and are received at numerous earthquake-monitoring sites around the world. The detection of these waves and their arrival times provides clues about Earth's interior.

Two major things happen to the waves: First, partial reflections occur at boundaries between distinctly different regions. Second, the waves change speed as the physical conditions—such as the elasticity and density—change. Changing a wave's speed usually results in the wave changing direction, a phenomenon known as refraction, which we will discuss in Chapter 18. As the waves go deeper into the interior, they speed up, causing them to change direction.



Cross section of Earth showing the paths of some waves produced by an earthquake.

The longitudinal waves are called the primary waves, or P waves, and are created by the alternating expansion and compression of the rocks near the source of the shock. This push–pull vibration can be transmitted through solids, liquids, and gases. P waves move with the highest speeds and therefore are the first to arrive at a seismograph station. P waves move at about 5 kilometers per second (11,000 miles per hour!) near the surface and speed up to about 7 kilometers per second toward the base of the upper crust.

The secondary waves, or S waves, are transverse waves. In this case the rock movement is perpendicular to the direction the wave is traveling. S waves travel through solids but cannot propagate through liquids and gases because fluids lack rigidity.

We wouldn't be able to infer much about Earth's interior if only one signal arrived at each site. There would be a number of paths that could account for the characteristics and timing of the signal. Fortunately, many sites receive multiple signals that allow large computers to piece the information together to form a model of Earth's interior. Information that is not received is also important. After an earthquake, many sites do not receive any transverse signals. This tells us that they are located in a shadow region behind a liquid core, as shown in the figure.

1. Would *transverse* or *longitudinal* waves be used to communicate with a submarine that was submerged in the ocean? Explain your choice.
2. What characteristic of all waves, transverse or longitudinal, explains the P wave shadow zone that is observed?

Source: H. Levin, *The Earth through Time* (Philadelphia: Saunders, 1992).

If two crests overlap, the disturbance is bigger than either one alone. A crest and a trough produce a smaller disturbance. If the crest and the trough are the same size and have symmetric shapes, they completely cancel at the instant of total overlap. A high-speed photograph taken at this instant yields a picture of a straight rope. This phenomenon is illustrated in Figure 15-12(d). This is not as strange as it may seem. If we take a high-speed photograph of a pendulum just as it swings through the equilibrium position, it would appear that the pendulum was not moving but simply hanging straight down. In either case, longer exposures would blur, showing the motion.

Periodic Waves

A rope moved up and down with a steady frequency and amplitude generates a train of wave pulses. All the pulses have the same size and shape as they travel down the rope. The drawing in Figure 15-14 shows a **periodic wave** moving to the right. New effects emerge when we examine periodic waves. For one thing, unlike the single pulse, periodic waves have a frequency. The frequency of the wave is the oscillation frequency of any piece of the medium.

An important property of a periodic wave is the distance between identical positions on adjacent wave pulses, called the **wavelength** of the periodic wave. This is the smallest distance for which the wave pattern repeats. It may be measured between two adjacent crests, or two adjacent troughs, or any two identical spots on adjacent pulses, as shown in Figure 15-15. The symbol used for wavelength is the Greek letter lambda λ.

The speed of the wave can be determined by measuring how far a particular crest travels in a certain time. In many situations, however, the speed is too fast, the wavelength too short, or the amplitude too small to allow us to follow the motion of a single crest. We then use an alternative procedure.

Suppose you take a number of photographs at the same frequency as the vertical vibration of any portion of the rope. You would find that all the pictures look the same. During the time the shutter of the camera was closed, each portion of the rope went through a complete cycle, ending in the position it had during the previous photograph. But this means that each crest moved from its original position to the position of the crest in front of it. That

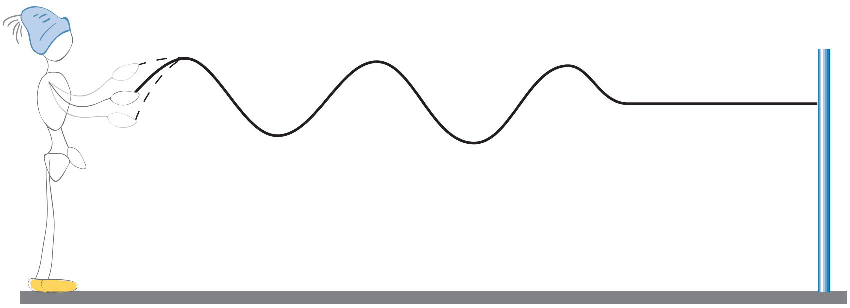


Figure 15-14 A periodic wave on a rope can be generated by moving the end up and down with a constant frequency.

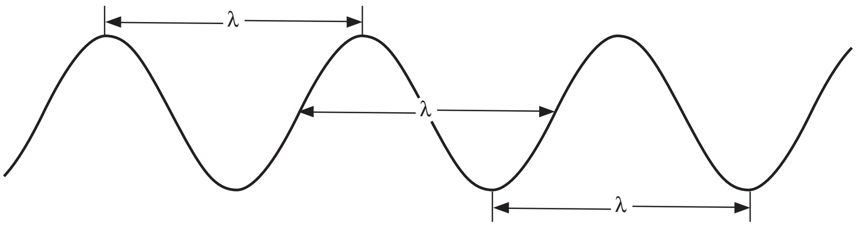


Figure 15-15 The wavelength of a periodic wave is the distance between any two identical spots on the wave.

is, the crest moved a distance equal to the wavelength λ. Because the time between exposures is equal to the period T, the wave's speed v is

$$v = \frac{\lambda}{T}$$

Because the frequency is just the reciprocal of the period, we can change the equation to read

$$v = \lambda f$$

◀ **speed = wavelength × frequency**

Although we developed this relationship for waves on a rope, there is nothing special about these waves. This relationship holds for all periodic waves, such as radio waves, sound waves, and water waves.

WORKING IT OUT *Speed of a Wave*

If you know any two of the three quantities in the wave equation, you can use this relationship to calculate the third. As an example, let's calculate the speed of a wave that has a frequency of 40 Hz and a wavelength of $\frac{3}{4}$ m. Multiplying the wavelength and the frequency gives us the speed:

$$v = \lambda f = (\tfrac{3}{4}\ \text{m})(40\ \text{Hz}) = 30\ \text{m/s}$$

Q: If water waves have a frequency of 5 Hz and a wavelength of 8 cm, what is the wave speed?

A: $v = \lambda f =$ (8 cm)(5 Hz) = 40 cm/s.

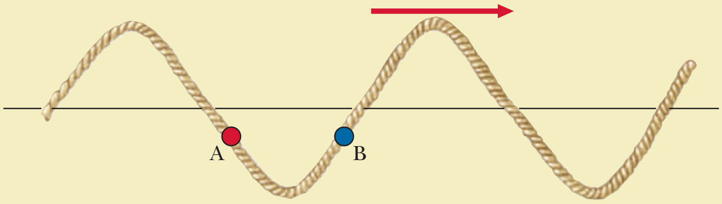
Standing Waves

When a periodic wave is confined, new effects emerge because of the superposition of the reflected waves with the original ones. Let's return to the example of a periodic wave moving down a rope toward a rigid post. When the periodic wave reflects from the post, it superimposes with the wave heading toward the post. The complete pattern results from the superposition of the original wave and reflections from both ends. In general, we get a complicated pattern with a small amplitude, but certain frequencies cause the rope to vibrate with a large amplitude. Figure 15-16 shows multiple images of a resonating rope. Although the superimposing waves move along the rope, they produce a resonant pattern that does not move along the rope. Because the pattern appears to stand still (in the horizontal direction), it is known as a **standing wave.**

It may seem strange that two identical waves traveling in opposite directions combine to produce a vibrational pattern that doesn't travel along the rope. We can see how this happens by using the superposition principle to find the results of combining the two traveling waves. Let's start at a time when the crests of the traveling wave moving to the right (the blue line in Figure 15-17) line up with the crests of the wave moving to the left (yellow line). (The blue and yellow lines lie on top of each other and are shown as a single green line.) Adding the displacements of the two traveling waves yields a wave that has the same basic shape, but twice the amplitude. This is shown by the black line in Figure 15-17(a).

FLAWED REASONING

The following question appears on the midterm exam: "A periodic wave is traveling to the right on a long, stretched rope. Two small pieces of yarn are tied to the rope, one at point A and the other at point B, as shown in the figure:



Draw an arrow for each piece of yarn, indicating the direction of its velocity when the picture was taken."

Brielle gives the following answer to this question: "Because the wave is moving to the right, the pieces of yarn must also be moving to the right. The wave is carried by the rope."

What is wrong with Brielle's reasoning, and what is the correct answer to the exam question?

ANSWER The wave is a transverse wave, meaning that the medium moves perpendicular to the direction of the wave. Therefore, the pieces of yarn can only move up or down. If we look at the wave at a slightly later time (when it has moved a little to the right), we see that the yarn at point A has moved upward and the yarn at point B has moved downward:

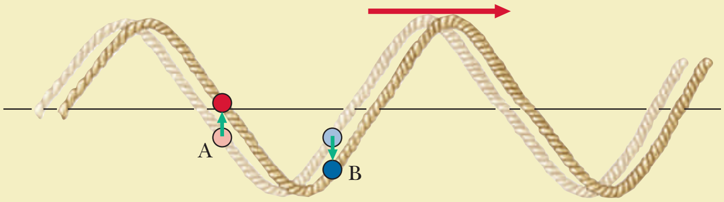


Figure 15-17(b) shows the situation a short time later. The blue wave has moved to the right, and the yellow wave has moved the same distance to the left. The superposition at this time produces a shape that still looks like one of the traveling waves but does not have as large an amplitude as before. A short time later, the crests of one wave line up with the troughs of the other. At this time the two waves cancel each other, and the rope is straight. Although the rope is straight at this instant, some parts of the rope are moving up while others are moving down. The remaining drawings in Figure 15-17 show how this pattern changes through the rest of the cycle as time progresses.

Notice that some portions of the rope do not move. Even though each traveling wave by itself would cause all pieces of the rope to move, the waves interfere to produce no motion at these points. Such locations are known as **nodes** and are located on the vertical lines indicated by N in Figure 15-17. The positions on the rope that have the largest amplitude are known as **antinodes** and are on the lines marked by A. Notice that the nodes and antinodes alternate and are equally spaced.

There is a relationship between the shape of the resonant pattern, or standing wave, and the moving periodic waves that superimpose to create it. The

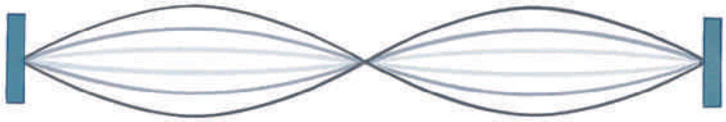


Figure 15-16 A strobe drawing of a standing wave on a rope shows how the shape of the rope changes with time. The shape does not move to the left or right.

"wavelength" of the standing wave is equal to the wavelength of the underlying periodic wave. This can be seen in Figure 15-17.

Unlike the pendulum or the mass on a spring where there was only one resonant frequency, periodic waves that are confined have many different resonant frequencies. Strobe photographs of the standing wave with the lowest frequency, or **fundamental frequency,** show that the rope has shapes like those drawn in Figure 15-18. The images show how the shape of the rope changes during one-half cycle. At position 1 the rope has the largest possible crest. The displacement continually decreases until it becomes zero and the rope is straight (between positions 3 and 4). The rope's inertia causes it to overshoot and form a trough that grows in size. At the end of the half cycle, the rope is in position 6 and beginning its upward journey. The process then repeats.

This pattern has the lowest frequency and thus the longest wavelength of the resonant modes. Notice that one-half wavelength is equal to the length of the rope, or the wavelength is twice the length of the rope. Because this is also the wavelength of the traveling waves, the longest resonant wavelength on a rope with nodes at each end is twice the length of the rope.

If we slowly increase the frequency of the traveling wave, the amplitude quickly decreases. The vibrational patterns are rather indistinct until the next resonant frequency is reached. This new resonant frequency has twice the frequency of the fundamental and is known as the second **harmonic.** Six shapes of the rope for this standing-wave pattern are shown in Figure 15-19. When the rope is low on the left side, it is high on the right, and vice versa. Because the rope has the shape of a full wavelength, the wavelength of the traveling waves is equal to the length of the rope. Note that this pattern has one more node and one more antinode than the fundamental standing wave.

We can continue this line of reasoning. A third resonant frequency can be reached by again raising the frequency of the traveling wave. This third harmonic frequency is equal to three times the fundamental frequency. This standing wave has three antinodes and four nodes, including the two at the ends of the rope. Other resonant frequencies occur at whole-number multiples of the fundamental frequency.

Remember that the product of the frequency and the wavelength is a constant. This is a consequence of the fact that the speeds of all waves on this rope are the same. Therefore, if we increase the frequency by some multiple while keeping the speed the same, the wavelength must decrease by the same multiple. The fundamental wavelength is the largest, and its associated frequency is the smallest. As we march through higher and higher frequencies, we get shorter and shorter wavelengths. The wavelengths of the higher harmonics are obtained by dividing the fundamental wavelength by successive whole

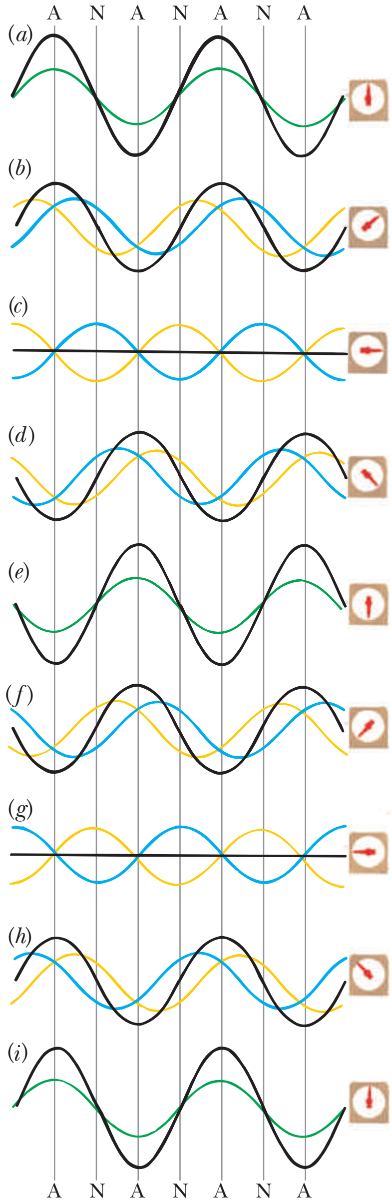


Figure 15-17 This set of strobe drawings shows how two traveling waves (the blue and yellow lines) combine to form a standing wave (the black line). Only the black line would be visible in a photograph.

Q: How does the distance between adjacent nodes or antinodes compare with the wavelength?

A: Because one antinode is up when the adjacent ones are down, each antinodal region corresponds to a crest or a trough. Therefore, the distance between adjacent antinodes or adjacent nodes is one-half wavelength. The distance between adjacent nodes and antinodes is one-quarter wavelength.

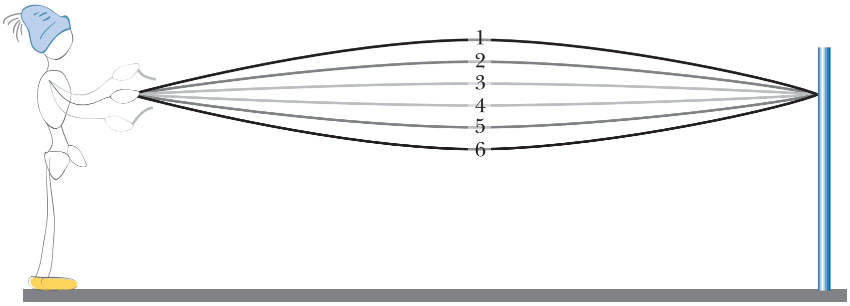


Figure 15-18 The shapes of a rope oscillating as a standing wave of the lowest frequency.

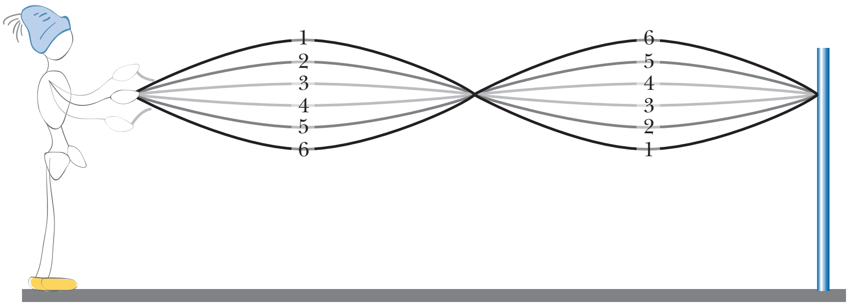


Figure 15-19 The shapes of a rope for a standing wave with the second resonant frequency.

numbers; the wavelength of the second harmonic is one-half the wavelength of the fundamental wavelength.

Q: How does the wavelength of the third harmonic compare with the length of the rope?

A: The wavelength of the third harmonic is one-third the length of the fundamental wavelength. Because the fundamental wavelength is twice the length of the rope, the wavelength of the third harmonic would be two-thirds the length of the rope.

Interference

Standing waves on a rope are an example of the superposition, or **interference,** of waves in one dimension. If we use a two-dimensional medium—say, the surface of water—we can generate some new effects.

Suppose we use two wave generators to create periodic waves on the surface of water in a ripple tank like the one in Figure 15-20(a). Because the two waves travel in the same medium, they have the same speed. We also assume that the two sources have the same frequency and that they are **in phase;** that is, both sources produce crests at the same time, troughs at the same time, and so on. The superposition of these waves creates the interference pattern shown in the photograph and drawing of Figure 15-20. The bright regions are produced by the crests, whereas the dark regions are produced by the troughs.

In some places, crest meets crest to form a supercrest, and one-half period later, trough meets trough to form a supertrough. This meeting point is a region of large amplitude; the two waves form antinodal regions. In other places, crest and trough meet. Here, if the two waves have about the same amplitude, they cancel each other, resulting in little or no amplitude; the two waves form nodal regions.

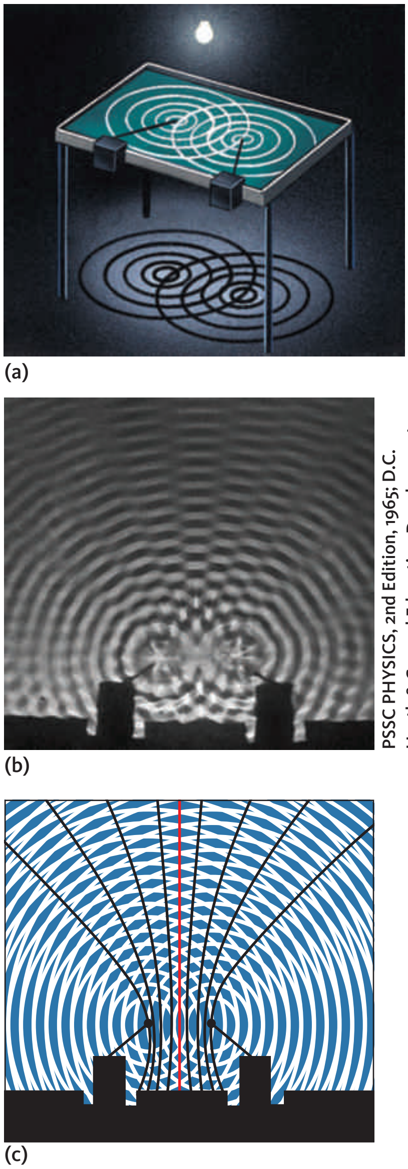
Because of the periodic nature of the waves, the nodal and antinodal regions have fixed locations. These stationary interference patterns can be observed only if the two sources emit waves of the same frequency; otherwise, one wave continually falls behind the other, and the relationships between the two waves change. The two wave sources do not have to be in phase; there can be a time delay between the generation of crests by one source and the other as long as the time delay is constant. For simplicity we usually assume that this time delay is zero; that is, the two sources are in phase.

The regions of crests and troughs lie along lines. One such antinodal line lies along the perpendicular to the midpoint of the line joining the two sources. This is the vertical red line in Figure 15-20(c). This central antinodal line is the same distance from the two sources. Therefore, crests generated at the same time at the two sources arrive at the same time at this midpoint to form supercrests. Similarly, two troughs arrive together, creating a supertrough.

Consider a point *P* off to the right side of the central line, as shown in Figure 15-21. Although nothing changes at the sources, we get a different result. Crests from the two sources no longer arrive at the same time. Crests from the left-hand source must travel a greater distance and therefore take longer to get to *P*. The amount of delay depends on the difference in the two path lengths.

If the point *P* is chosen such that the distances to the sources differ by an amount equal to one-half wavelength, crests overlap with troughs and troughs overlap with crests at point *P*. The waves cancel. There are many points that have this path difference. They form nodal lines that lie along each side of the central line, as shown by the black lines in Figure 15-20(c). An antinodal line occurs when the path lengths differ by one wavelength; the next nodal line when the paths differ by $1\frac{1}{2}$ wavelengths, and so on.

The photograph in Figure 15-22 shows the interference pattern for water waves with a longer wavelength than those in Figure 15-20. The nodal lines are now more widely spaced for the same source separation. Therefore, longer wavelengths produce wider patterns. Actually, the width of the pattern depends on the relative size of the wavelength and the source separation. As the ratio of the wavelength to the separation gets bigger, the nodal lines spread out. If the wavelength is much larger than the separation, the pattern is essentially that of a single source, whereas if it is much smaller, the nodal lines are so close together they cannot be seen.



PSSC PHYSICS, 2nd Edition, 1965; D.C. Heath & Co. and Education Development Center, Inc., Newton, MA.

Figure 15-20 (a) A light bulb above a ripple tank produces light and dark lines on the floor due to the water waves. (b) The interference pattern produced by two point sources of the same wavelength and phase. (c) The locations of the nodal lines in this pattern are shown in black; the central antinodal line is shown in red.

Diffraction

In the photographs in Figure 15-23, periodic water waves move toward a barrier. We see that the waves do not go straight through the opening in the barrier but spread out behind the barrier. This bending of a wave is called **diffraction** and is definitely not a property of particles. If a BB gun is fired repeatedly through an opening in a barrier, the pattern it produces is a precise "shadow" of the opening if we assume that no BBs bounce off the opening's edges.

Q: Would the central line still be an antinodal line if the two sources were completely out of phase—that is, if one source generates a crest at the same time as the other generates a trough?

A: The central line would now be a nodal line because crests and troughs would arrive at the same time.

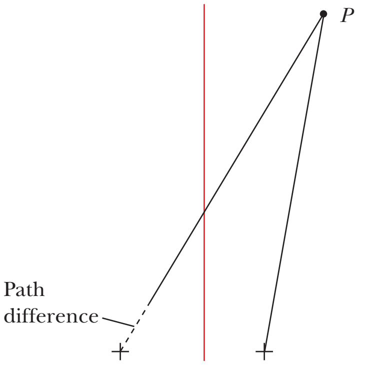
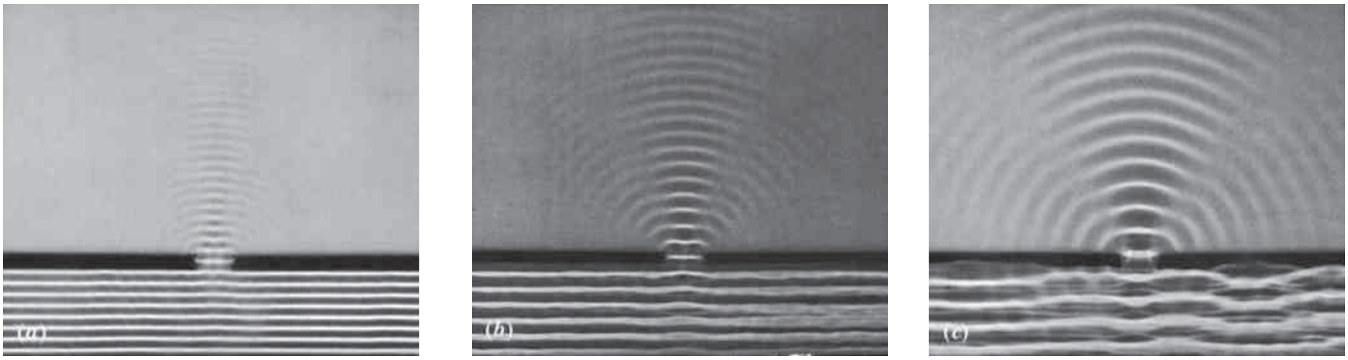


Figure 15-21 Whether the region at *P* is a nodal or antinodal region depends on the difference in the path lengths from the two sources.

Images not available due to copyright restrictions



PSSC PHYSICS, 2nd Edition, 1965; D.C. Heath & Co. and Education Development Center, Inc., Newton, MA. (all)

Figure 15-23 Ripple tank patterns of water moving upward and passing through a narrow barrier. Note that the amount of diffraction increases as the wavelength gets longer.

Courtesy of Sabina Zigman/Benjamin Cardozo High School, and by perm of PHYSICS TEACHER, vol.37, 1/99, p.55

The waves are diffracted as they pass through the openings in the seawall, producing an interesting shape at the beach.

The amount of diffraction depends on the relative sizes of the wavelength and the opening. If the wavelength is much smaller than the opening, very little diffraction is evident. As the wavelength gets closer to the size of the opening, the amount of diffraction gets bigger. In Figure 15-23(c) the opening and the wavelength are approximately the same size, and the diffraction is evident.

Notice that diffraction produces nodal and antinodal lines similar to those observed in the interference patterns from two point sources. In this case there is a broad central antinodal region with nodal lines on each side, and the diffraction pattern is created by different portions of the wave interfering with themselves. The spacing of these lines is determined by the ratio of the wavelength and the width of the opening.

Summary

Vibrations and oscillations are described by the length of time required for one cycle, the period T (or its reciprocal, the frequency f), and the amplitude of the vibration, the maximum distance the object travels from the equilibrium point. When vibrations are small, the period is independent of the amplitude. The pendulum and a mass hanging on a spring are examples of systems that vibrate.

All systems have a distinctive set of natural frequencies. A simple system such as a pendulum has only one natural frequency, whereas more complex systems have many natural frequencies. When a system is excited at a natural frequency, it resonates with a large amplitude.

Waves are vibrations moving through a medium; it is the wave (energy) that moves through the medium, not the medium itself. Transverse waves vibrate perpendicular to the direction of the wave, whereas longitudinal waves vibrate parallel to the direction of the wave. The speed of a periodic wave is equal to the product of its frequency and its wavelength, $v = \lambda f$.

Waves pass through each other as if the other were not there. When they overlap, the shape is the algebraic sum of the displacements of the individual waves. When a periodic wave is confined, resonant patterns known as standing waves can be produced. Portions of the medium that do not move are called nodes, whereas portions with the largest amplitudes are known as antinodes. The fundamental standing wave has the lowest frequency and the longest wavelength.

Two identical periodic-wave sources with a constant phase difference produce an interference pattern consisting of large-amplitude antinodal regions and zero-amplitude nodal regions. The spacing of the interference pattern depends on the relative size of the wavelength and the source separation.

Waves do not go straight through openings or around barriers but spread out. This diffraction pattern contains nodal and antinodal regions and depends on the relative sizes of the wavelength and the opening.

CHAPTER 15 *Revisited*

When waves move in a medium, the medium oscillates in place. No material is transported from one location to another; it is the disturbance that moves. Unlike with particles, when two waves pass through the same region at the same time, the individual disturbances are added together. Afterward, each wave retains its own identity.

Key Terms

amplitude The maximum distance from the equilibrium position that occurs in periodic motion.

antinode One of the positions in a standing-wave or interference pattern where there is maximum movement; that is, the amplitude is a maximum.

crest The peak of a wave disturbance.

cycle One complete repetition of a periodic motion. It may start anyplace in the motion.

diffraction The spreading of waves passing through an opening or around a barrier.

displacement In wave (or oscillatory) motion, the distance of the disturbance (or object) from its equilibrium position.

equilibrium position A position where the net force is zero.

frequency The number of times a periodic motion repeats in a unit of time. It is equal to the inverse of the period.

fundamental frequency The lowest resonant frequency for an oscillating system.

harmonic A frequency that is a whole-number multiple of the fundamental frequency.

in phase Describes two or more waves with the same wavelength and frequency that have their crests lined up.

interference The superposition of waves.

longitudinal wave A wave in which the vibrations of the medium are parallel to the direction the wave is moving.

node One of the positions in a standing-wave or interference pattern where there is no movement; that is, the amplitude is zero.

oscillation A vibration about an equilibrium position or shape.

period The shortest length of time it takes a periodic motion to repeat. It is equal to the inverse of the frequency.

periodic wave A wave in which all the pulses have the same size and shape. The wave pattern repeats itself over a distance of one wavelength and over a time of one period.

resonance A large increase in the amplitude of a vibration when a force is applied at a natural frequency of the medium or object.

spring constant The amount of force required to stretch a spring by one unit of length. The spring constant is measured in newtons per meter.

standing wave The interference pattern produced by two waves of equal amplitude and frequency traveling in opposite directions. The pattern is characterized by alternating nodal and antinodal regions.

superposition The combining of two or more waves at a location in space.

transverse wave A wave in which the vibrations of the medium are perpendicular to the direction the wave is moving.

trough A valley of a wave disturbance.

vibration An oscillation about an equilibrium position or shape.

wave The movement of energy from one place to another without any accompanying matter.

wavelength The shortest repetition length for a periodic wave. For example, it is the distance from crest to crest or trough to trough.

Questions and exercises are paired so that most odd-numbered are followed by a similar even-numbered.
Blue-numbered questions and exercises are answered in Appendix B.
indicates more challenging questions and exercises.
WebAssign **Many Conceptual Questions and Exercises for this chapter may be assigned online at WebAssign.**

Conceptual Questions

1. If the net force on a mass oscillating at the end of a vertical spring is zero at the equilibrium point, why doesn't the mass stop there?
2. If the restoring force on a pendulum is zero when it is vertical, why doesn't it quit swinging at this point?
3. A mass is oscillating up and down on a vertical spring. When the mass is above the equilibrium point and moving downward, in what direction does the net force on the mass act? When the mass is above the equilibrium point and moving upward, what is the direction of the net force on the mass? Explain.
4. A mass is oscillating up and down on a vertical spring. When the mass is below the equilibrium point and moving downward, what is the direction of its acceleration? Is the mass speeding up or slowing down? Explain.
5. A mass is oscillating up and down on a vertical spring. If the mass is increased, will the period of oscillation increase, decrease, or stay the same? Will the frequency increase, decrease, or stay the same? Explain.
6. A grandfather clock (with a pendulum) keeps perfect time on Earth. If you were to transport this clock to the Moon, would its period of oscillation increase, decrease, or stay the same? Would its frequency increase, decrease, or stay the same? Explain.
7. You hang a 1-kilogram block from a spring and find that the spring stretches 15 centimeters. What mass would you need to stretch the spring 45 centimeters?
8. Which spring would you expect to have the greater spring constant, the one in the suspension of your Chevy or the one in your watch? Why?
9. Assume that you pull the mass on a spring 1 centimeter from the equilibrium position, let go, and measure the period of the oscillation. Would you expect the period to be larger, the same, or smaller if you pulled the mass 2 centimeters from the equilibrium position? Why?
10. The amplitude of a real pendulum decreases because of frictional forces. How does the period of this real pendulum change?
11. What is the period of the hand on a clock that measures the seconds? What is its frequency?
12. What is the period of the hand on a clock that measures the minutes? What is its frequency?
13. Suppose your grandfather clock runs too fast. If the mass on the pendulum can be moved up or down, which way would you move it to adjust the clock? Explain your reasoning.
14. How does the natural frequency of a swing change when you move from sitting down to standing up?

15. You find that the exhaust system on your 1979 Chrysler Cordoba tends to rattle loudly when the tachometer, which measures the engine's frequency, reads 2000 rpm. It is relatively quiet at frequencies above or below 2000 rpm. Use the concept of resonance to explain this.

16. Why do soldiers "break step" before crossing a suspension bridge?

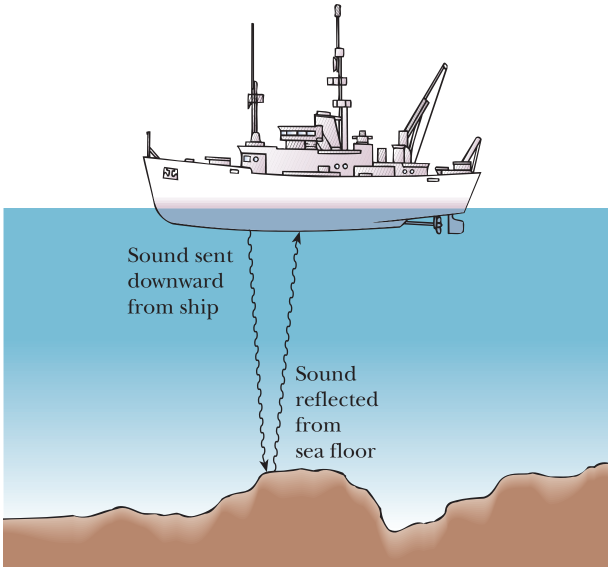
17. You hold one end of a spring in your hand and hang a block from the other end. After lifting the block up slightly and releasing it, you find that it oscillates up and down at a frequency of 2 hertz. At which of the following frequencies could you jiggle your hand up and down and produce resonance: 5 hertz, 4 hertz, 1.5 hertz, 1 hertz, or 0.5 hertz?

18. You stand to the side of the low point of a child's swing and always push the child in the same direction. Which of the following multiples of the fundamental frequency will not produce resonance: $\frac{1}{3}$, $\frac{1}{2}$, 1, or 2?

19. When you yell at your friend, are the air molecules that strike his ear the same ones that were in your lungs? Explain.

20. What is being transported along a clothesline when a wave moves from one end to the other?

21. Sonar devices use underwater sound to explore the ocean floor. Would you expect sonar to be a longitudinal or a transverse wave? Explain.



22. You fasten one end of a long spring to the base of a wall and stretch it out along the floor, holding the other end in your hand. Describe how you would generate a transverse pulse on the spring. Describe how you would generate a longitudinal pulse on the spring.

23. Is it possible for a shout to overtake a whisper? Explain.

24. You generate a small transverse pulse on a long spring stretched between a doorknob and your hand. How could you generate a second pulse that would overtake the first pulse?

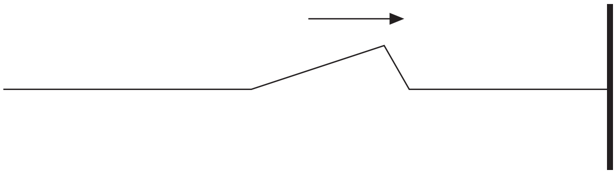
25. Which one or more of the following properties affect the speed of waves along a rope: amplitude of the pulse, shape of the pulse, tension in the rope, or the mass per unit length of the rope? Why?

26. You move your hand up and down to send a pulse along a long spring stretched between a doorknob and your hand. Which of the following would generate a slower-traveling pulse: Moving your hand the same distance as before, but more slowly; moving your hand a smaller distance at the same speed as before; or moving closer to the doorknob to decrease the tension in the spring?

27. You send a pulse of amplitude 5 centimeters down the right side of a spring. A moment later you send an identical pulse on the same side. The first pulse reflects from the fixed end and returns along the spring. When the reflected pulse meets the second pulse, will the resulting amplitude be less than, equal to, or greater than 5 centimeters? Explain your reasoning.

28. Imagine that the string in Figure 15-14 is tied to the pole with a loose loop such that the end is free to move up and down. A pulse of amplitude 10 centimeters is sent down the top of the string, and a moment later a second identical pulse is sent, also on the top. The first pulse reflects from the free boundary and returns along the string. When the reflected pulse meets the second pulse, will the resulting amplitude be less than, equal to, or greater than 10 centimeters? Explain your reasoning.

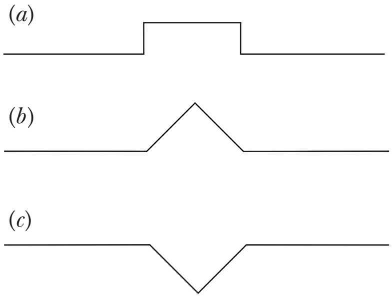
29. The pulse in the following figure is traveling on a string



to the right toward a fixed end. Draw the shape of the pulse after it reflects from the boundary.

30. A pulse in the shape of a crest is sent from left to right along a stretched rope. A trough travels in the opposite direction so that the pulses meet in the middle of the rope. Would you expect to observe a crest or a trough arrive at the right-hand end of the rope? Explain.

31. If shapes (a) and (b) in the following figure correspond to idealized wave pulses on a rope, what shape is produced when they completely overlap?



32. Repeat Question 31 for shapes (a) and (c).

33. Which of the following properties are meaningful for periodic waves but not for single pulses: frequency, wavelength, speed, amplitude?

34. In the following list of properties of periodic waves, which one is independent of the others: frequency, wavelength, speed, amplitude?

35. Two waves have the same speed but one has twice the frequency. Which wave has the longer wavelength? Explain.

36. If the frequency of a periodic wave is cut in half while the speed remains the same, what happens to the wavelength?

37. If the speed of a periodic wave doubles while the period remains the same, what happens to the wavelength?

38. What happens to the wavelength of a periodic wave if both the speed of the wave and the frequency are cut in half?

39. Travelers spaced 10 feet apart are all walking at 3 mph relative to a moving sidewalk. When the moving sidewalk ends, they continue to walk at 3 mph. An observer standing next to the moving sidewalk notes that the travelers are passing by at a frequency of 1 hertz. A second observer stands just beyond the end of the moving sidewalk and notes the frequency at which the travelers pass.


© Rob Melnychuk/Digital Vision/Getty Images

Would this frequency be greater than, equal to, or less than 1 hertz? Is the spacing between the travelers after leaving the moving sidewalk greater than, equal to, or less than 10 feet? Explain.

40. A waterproof electric buzzer has a membrane that vibrates at a constant frequency of 440 hertz. The buzzer is placed in a bucket of water. Knowing that the speed of sound is much greater in water than in air, will the frequency of the sound heard in the air be greater than, equal to, or less than 440 hertz? Will the wavelength of the sound in air be greater than, equal to, or less than what it was in the water? Explain. (*Hint:* Review Question 39 and think of the travelers as the wave crests.)

41. Draw a diagram to represent the standing-wave pattern for the third harmonic of a rope fixed at both ends. How many antinodes are there?

42. Draw a diagram to represent the standing-wave pattern for the fourth harmonic of a rope fixed at both ends. How many nodes are there?

43. How much higher is the frequency of the fifth harmonic on a rope than the fundamental frequency?

44. How much higher is the frequency of the sixth harmonic on a rope than that of the second?

45. How many antinodes are there when a rope fixed at both ends vibrates in its third harmonic?

46. How many nodes are there when a rope fixed at both ends vibrates in its fourth harmonic?

47. Standing waves can be established on a rope that is fixed on one end but free to slide up and down a pole on the other. The fixed end remains a node, while the free end must be an antinode. Draw diagrams to represent the standing-wave patterns for the two lowest frequencies.

48. How does the fundamental wavelength of standing waves on a string with one end fixed and the other free compare to the fundamental wavelength if the same string is held with both ends fixed?

49. How does the wavelength of the fourth harmonic on a rope with both ends fixed compare with the length of the rope?

50. How does the wavelength of the fourth harmonic on a rope with both ends fixed compare with that of the second harmonic?

51. A longitudinal standing wave can be established in a long aluminum rod by stroking it with rosin on your fingers. If the rod is held tightly at its midpoint, what is the wavelength of the fundamental standing wave? Assume that there are antinodes at each end of the rod and a node where the rod is held.

52. What is the wavelength of the fundamental standing wave for the rod in Question 51 if it is held midway between the center and one end? Will the resulting pitch be higher or lower than when the rod was held at its midpoint? Explain.

53. Two point sources produce waves of the same wavelength and are in phase. At a point midway between the sources, would you expect to find a node or an antinode? Explain.

54. Two point sources produce waves of the same wavelength and are completely out of phase (that is, one produces a crest at the same time as the other produces a trough). At a point midway between the sources, would you expect to find a node or an antinode? Why?

55. What happens to the spacing of the antinodal lines in an interference pattern when the two sources are moved farther apart? Explain.

56. As you increase the frequency, what happens to the spacing of the nodal lines in an interference pattern produced by two sources? Explain.

Courtesy of VWR Corporation

57. An interference pattern is produced in a ripple tank. As the two sources are brought closer together, does the separation of the locations of maximum amplitude along the far edge of the tank decrease, increase, or remain the same? Why?

58. As the frequency of the two sources forming an interference pattern in a ripple tank increases, does the separation of the locations of minimum amplitude along the far edge of the tank increase, decrease, or remain the same? Why?

59. What happens to the spacing of the antinodal lines in an interference pattern when the two slits are moved farther apart? Explain.

60. As you increase the frequency, what happens to the spacing of the nodal lines in a diffraction pattern? Explain.

Exercises

61. If a mass on a spring takes 6 s to complete two cycles, what is its period?

62. If a mass on a spring has a frequency of 4 Hz, what is its period?

63. A Foucault pendulum with a length of 9 m has a period of 6 s. What is its frequency?

64. A mass on a spring bobs up and down over a distance of 30 cm from the top to the bottom of its path twice each second. What are its period and amplitude?

65. A spring hanging from the ceiling has an unstretched length of 80 cm. A mass is then suspended at rest from the spring, causing its length to increase to 89 cm. The mass is pulled down an additional 3 cm and released. What is the amplitude of the resulting oscillation?

66. A mass oscillates up and down on a vertical spring with an amplitude of 4 cm and a period of 2 s. What total distance does the mass travel in 10 s?

67. What is the period of a 0.4-kg mass suspended from a spring with a spring constant of 40 N/m?

68. A boy with a mass of 50 kg is hanging from a spring with a spring constant of 200 N/m. With what frequency does the boy bounce up and down?

69. By what factor would you have to increase the mass to double the period for a mass on a spring?

70. By what factor would you have to increase the spring constant to triple the frequency for a mass on a spring?

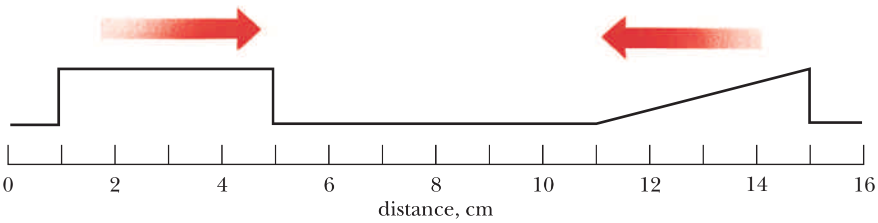
71. A pendulum has a length of 5 m. What is its period?

72. A girl with a mass of 40 kg is swinging from a rope with a length of 2.5 m. What is the frequency of her swinging?

73. The highly idealized wave pulses shown in the figure below at a time equal to zero have the same amplitudes and travel at 1 cm/s. Draw the shape of the rope at 2, 4, 5, and 8 s.

74. Work Exercise 73 but change the rectangular pulse from a crest to a trough.

75. A train consisting of identical 10-m boxcars passes you such that 25 boxcars pass you each minute. Find the speed of the train.



EXERCISE 73

76. You observe that 25 crests of a water wave pass you each minute. If the wavelength is 10 m, what is the speed of the wave?

77. A periodic wave on a string has a wavelength of 25 cm and a frequency of 3 Hz. What is the speed of the wave?

78. If the breakers at a beach are separated by 5 m and hit shore with a frequency of 0.3 Hz, at what speed are they traveling?

79. What is the distance between adjacent crests of ocean waves that have a frequency of 0.2 Hz if the waves have a speed of 3 m/s?

80. Sound waves in iron have a speed of about 5100 m/s. If the waves have a frequency of 400 Hz, what is their wavelength?

81. For sound waves, which travel at 343 m/s in air at room temperature, what frequency corresponds to a wavelength of 1 m?

82. What is the period of waves on a rope if their wavelength is 0.8 m and their speed is 2 m/s?

83. A rope is tied between two posts separated by 3 m. What possible wavelengths will produce standing waves on the rope?

84. A 3-m-long rope is tied to a thin string so that one end is essentially free. What possible wavelengths will produce standing waves on this rope?

85. What is the fundamental frequency on a 6-m rope that is tied at both ends if the speed of the waves is 18 m/s?

86. Tweety Bird hops up and down at a frequency of 0.5 Hz on a power line at the midpoint between the poles, which are separated by 20 m. Assuming Tweety is exciting the fundamental standing wave, find the speed of transverse waves on the power line.

16

Sound and Music

▶Sounds are all around us. Some are pleasant, and some irritate and distract us. How do the sounds of music differ from those sounds we call noise, and why do musicians use different-sized instruments?

(See page 346 for the answer to this question.)

Jake Rajs/Stone/Getty

A military marching band at a ticker-tape parade in New York City.

WHEN we think of sound, we generally think of signals traveling through the air to our ears. But sound is more than this. Sound also travels through other media. For instance, old-time Westerns showed cowboys and Indians listening for the "iron horse" by putting their ears to the rails or the ground; two rocks clapped together underwater are easily heard by swimmers below the surface; a fetus inside a mother's womb can be examined with ultrasound; and the voices of people talking in the next room are often heard through the walls. Sounds can be soothing and musical, but they can also be irritating or even painful.

Sound is a wave phenomenon. For example, we talk of the pitch, or frequency, of sounds, which is definitely a wave characteristic. But what other evidence do we have? The conclusive evidence is that sound exhibits superposition, something that we know distinguishes waves from particles.

Sound

A vibrating object produces disturbances in the surrounding air. When the surface moves outward, the air molecules are pushed away, creating a *compression*. When its surface moves in the opposite direction, a partial void, or *rarefaction*, is created near the object. Pressure differences cause the air molecules to rush back into the region only to get pushed out again. Thus, the air molecules vibrate back and forth near the object's surface as illustrated by the tuning fork in Figure 16-1. The compressions and rarefactions travel away from the vibrating surface as a sound wave. Because the vibrations of the air molecules are along the same direction as the motion of the wave, sound is a longitudinal wave.

sound is a longitudinal wave ▶

As with all waves, it is energy—not mass—that is transported. The individual molecules of the air are not moving from one place to another; they simply vibrate back and forth. It is the disturbance that moves across the room when you talk. This disturbance, or wave, moves with a certain speed.

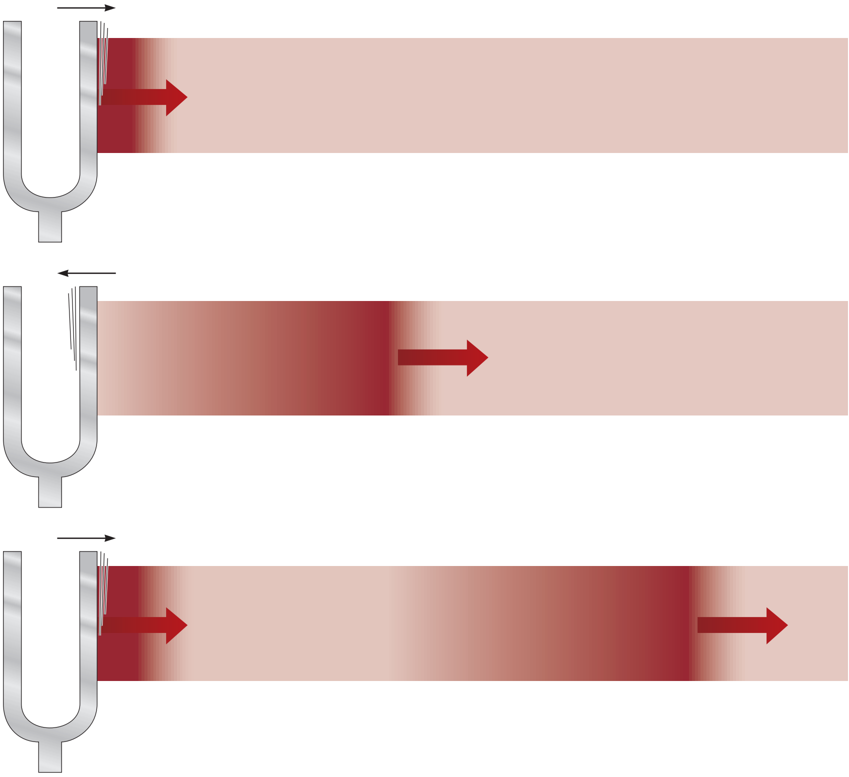


Figure 16-1 Sound is a longitudinal wave in which the air molecules vibrate along the direction the wave is traveling, producing compressions and rarefactions that travel through the air.

Speed of Sound

Extended presentation available in the *Problem Solving* supplement

Echoes demonstrate that sound waves reflect off surfaces and that they move with a finite speed. You can use this phenomenon to measure the speed of sound in air. If you know the distance to the reflecting surface and the time it takes for the echo to return, you can calculate the speed of sound. The speed of sound is 343 meters per second (1125 feet per second), or 1235 kilometers per hour (767 miles per hour), at room temperature.

speed of sound = 343 m/s

Experiments have shown that the speed of sound does not depend on the pressure, but it does depend on the temperature and the type of gas. Sound is slower at lower temperatures. At the altitude of jet airplanes, where the temperature is typically −40°C (−40°F), the speed of sound drops to 310 meters per second, or 1020 kilometers per hour (690 miles per hour). The speed is higher for gases with molecules that have smaller masses. The speed of sound in pure helium at room temperature is three times that in air.

FLAWED REASONING

Heidi and Russell are discussing the speed of sound in air:

Heidi: "A sound wave travels through air via collisions of air molecules. If the air is compressed, the molecules are closer together. The sound wave should speed up because each molecule does not need to travel as far to collide with its neighbor."

Russell: "The textbook claims that the speed of sound in air depends only on the temperature of the air, not on its pressure."

Find the error in Heidi's reasoning.

ANSWER Imagine a relay race in which every runner runs *exactly* 5 mph. If one of the teams has twice as many runners (spaced half as far apart), the race still ends in a tie. Each runner on the smaller team runs twice as far before passing the baton, but there are half as many baton transfers. Similarly, each molecule in the compressed gas reaches its neighbor in less time, but more neighbors are involved. The air temperature determines the average molecular speed, and this determines how quickly the wave travels.

Knowing the speed of sound in air allows you to calculate your distance from a lightning bolt. Because the speed of light is very fast, the time light takes to travel from the lightning to you is negligible. Therefore, the time delay between the arrival of the light flash and the sound of the thunder is essentially all due to the time it takes the sound to travel the distance. Given the speed of sound, we can use the definition of speed from Chapter 2 to calculate that sound takes approximately 3 seconds to travel 1 kilometer (5 seconds for a mile).

Sound waves also travel in other media. The speed of sound in water is about 1500 meters per second, much faster than in air. The speed of sound in solids is usually quite a bit higher—as high as 5000 meters per second—and the sound is quite a bit louder. Hearing the sounds of an approaching train through the rails works well because the sound moving through the rail does not spread out like sound waves in air and experiences less scattering along the path. You can easily experience this phenomenon in the classroom. Have a friend scratch one end of a meter stick while you hold the other end next to your ear. The effect is striking.

Sonar (*so*und *na*vigation *r*anging) uses the echoes of sound waves in water to determine the distances to underwater objects. These sonar devices emit sound pulses and measure the time required for the echo to return. Sonar is used to determine the depth of the water, search for schools of fish, and locate submarines.

Are You On the Bus?

Q: If it takes the thunder 9 seconds to reach you, how far away is the lightning?

A: Because it takes the thunder 3 seconds to travel a kilometer, the lightning bolt must have been 3 kilometers away (a little less than 2 miles).

Hearing Sounds

Sound perception is a complex phenomenon whose study involves a wide variety of sciences, including physiology, psychology, and the branch of physics called acoustics. Experiments with human hearing, for example, clearly show that our perception is often different from the simple interpretations of the measurements taken by instruments. Our perception of pitch depends mainly on frequency but is also affected by other properties, such as loudness. And loudness depends on the amplitude of the wave (as well as the response of our ears). When our ear–brain systems tell us that one sound is twice as loud as another, instruments show that the power output is nearly eight times as great.

Our ears intercept sound waves from the air and transmit their vibrations through internal bone structures to special hairlike sensors. The ear canal acts as a resonator, greatly amplifying frequencies near 3000 hertz. This amplified sound wave moves the eardrum, which is located at the end of the ear canal, as shown in Figure 16-2. The eardrum is connected to three small bones in the middle ear. When sound reaches the middle ear, it has been transformed from a wave in air to a mechanical wave in the bones. These bones then move a smaller oval window inside the ear. The leverage advantage of the bone structure and the concentration of the pressure vibration onto a smaller window further amplify the sound, increasing our ability to hear faint sounds. The final transformation of mechanical sound waves to nerve impulses takes place in the inner ear. The pressure vibrations in the fluid of the inner ear resonate with different hairlike sensors, depending on the frequency of the sound.

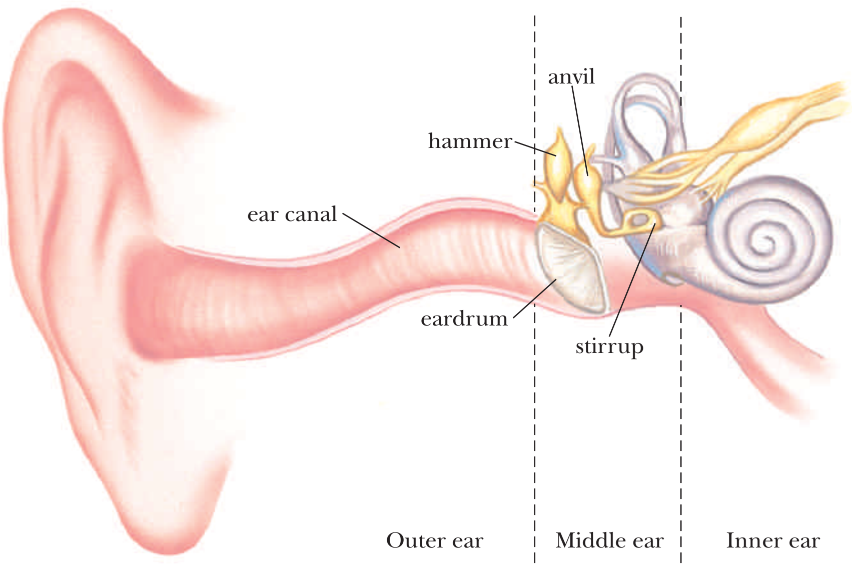


Figure 16-2 The structure of the human ear.

The range of frequencies that we can hear clearly depends on the resonant structures within our ears. When a frequency is too high or too low, the sound wave is not amplified like those within the audible range. The audible range is normally from 20 hertz to 20,000 hertz, although it varies with age and the individual. The sensitivity of our ears varies over this range, with low sensitivity occurring at both ends of the range. As we get older, our ability to hear higher frequencies decreases. However, modern digital hearing aids can be individually programmed to compensate for a person's hearing loss at each frequency.

Courtesy of SENSO Digital Hearing Aid by Widex

Digital hearing aid.

The Recipe of Sounds

Suppose you are in a windowless room but can hear sounds from the outside. You would have no trouble identifying most of the sounds you hear. A bird sounds different from a foghorn, a trumpet different from a baritone. Why is it that you can recognize these different sources of sound? They may be pro-

Everyday Physics | *Animal Hearing*

Animals have different ranges of sensitivity for hearing than humans. Dogs and bats, for example, have hearing ranges that extend to ultrasonic frequencies—frequencies above those that we can hear.

Unlike humans, most animals use their hearing as an aid in gathering food and escaping danger. Bats squeak at ultrasonic frequencies and detect the echoes from small flying insects (Figure A); robins cock their heads in early spring as they listen for the very faint sounds of worms in the ground, and owls have two different types of ears providing them with binocular hearing for finding mice moving through grassy fields.

Animal researchers have found that homing pigeons and elephants hear very low-frequency sounds—infrasonic frequencies. Low-frequency sounds do not attenuate as rapidly and therefore travel much farther than high-frequency sounds. In the case of pigeons, hearing the sounds of skyscrapers swaying in the wind in distant cities may provide them with navigational bearings. There is evidence that elephants communicate with each other over distances of miles using a low-frequency rumble. Figure B displays the frequency ranges that some animals can hear.

1. Are there frequencies that humans can hear, but dogs cannot? Explain.
2. Which animal can hear the greater range in frequencies, a mouse or an elephant? Give evidence for your answer from the graph in Figure B.

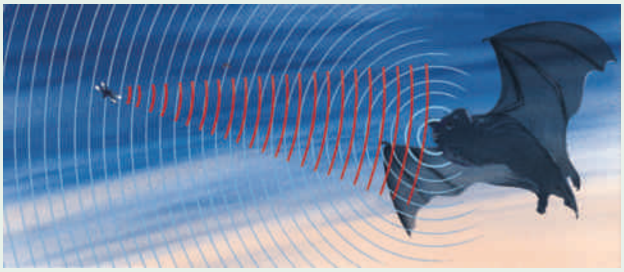


Figure A The bat emits sound waves (shown in blue) that reflect off the insect and return (shown in red) to the bat, giving away the insect's position.

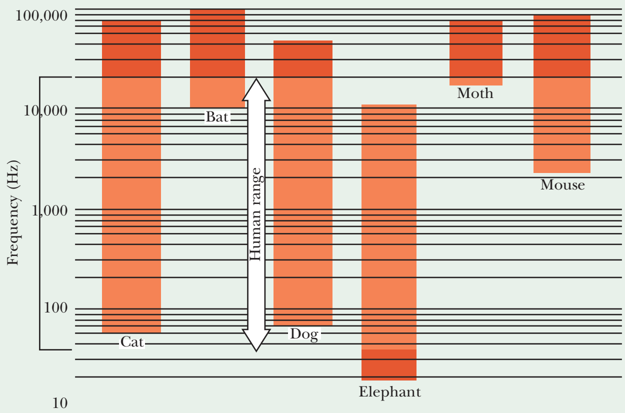
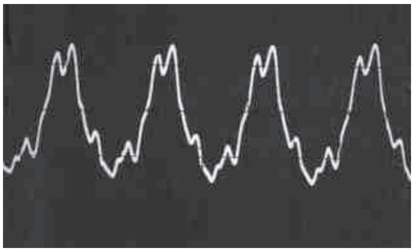
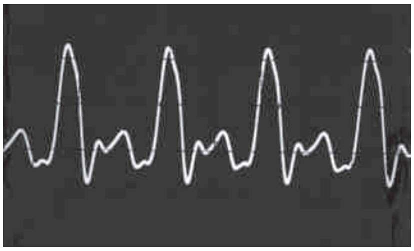


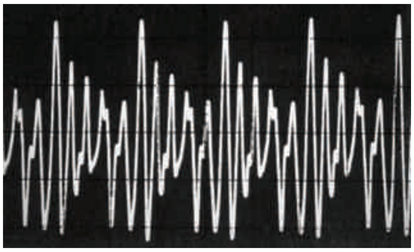
Figure B The frequency ranges of hearing for some animals.



(a)



(b)



(c)

Figure 16-3 Wave patterns of (a) a violin, (b) a trumpet, and (c) a bassoon. These photographs show only general characteristics; the details depend on the placement of the microphone.

ducing different notes; a bird sings much higher than a foghorn. They may make different melodies; some birds have a rather complex sequence of notes in their call, whereas the foghorn produces one continuous note.

What if all the sound makers outside your room are restricted to a single note? Do you suppose that you could still identify the different sources? Most likely. When each source produces the note, it is accompanied by other higher, resonant frequencies. You actually hear a superposition of these frequencies. The reason you can distinguish among the sources is that each sound has a unique combination of intensities of the various harmonics (Chapter 15). One sound may be composed of a strong fundamental frequency and weaker higher harmonics, and another may have a particularly strong second harmonic. Each sound has a particular recipe of resonant frequencies that combine to make the total sound.

The vibrating strings in a piano, violin, guitar, or banjo have their own combinations of the fundamental frequency and higher harmonics. The relative intensities of these harmonics depend primarily on the way the string was initially vibrated and on the vibrational characteristics of the body of the instrument. The initial part of the sound is called the *attack,* and its character is determined, in part, by how the sound is produced. Bowing, for example, produces a sound different from plucking or striking. The manner in which the various components of the sound decay also differs from one instrument to another. The wave patterns produced by various sounds are illustrated in Figure 16-3.

Our voices have the same individual character. We recognize different voices because of their particular recipe of harmonics. Research has shown that this recipe changes under emotional stress. Some people have suggested that these changes in the recipe of sound of a voice can be used in lie detection. Electronic devices can decompose the composite waveform—the superposition of all the harmonics—and evaluate the relative strength of each harmonic. This technique spots changes in higher harmonics that are virtually undetectable in the composite waveform.

How does music differ from other sounds? This question is difficult because what is music to one person may be noise to you. In general, music can be defined as that collection of periodic sounds that is pleasing to the ear.

Most cultures divide the totality of musical frequencies into groups known as octaves. A note in one octave has twice the frequency of the corresponding note in the next lower octave. For example, the pitches labeled C in ascending adjacent octaves have frequencies of 262, 524, and 1048 hertz, respectively. In Western cultures most music is based on a scale that divides the octave into 12 steps. It may seem strange that an octave spans 12 notes, because the word *octave* derives from the Latin word meaning "eight." An octave contains 7 natural notes and 5 sharps and flats. Indian and Chinese music have different divisions within their octaves.

Through the years, people have created instruments to produce sounds that were pleasing to them. Nearly all of these instruments involve the production of standing waves. Although there is an enormous variety of instruments, most of them can be classified as string, wind, or percussion instruments.

Stringed Instruments

When a string vibrates, it moves the air around it, producing sound waves. Because the string is quite thin, not much air is moved, and consequently the sound is weak. In acoustic stringed instruments, this lack of volume is solved by mounting the vibrating string on a larger body. The vibrations are transmitted to the larger body, which can move more air and produce a louder

Everyday Physics | *Loudest and Softest Sounds*

The intensity of sound is measured in terms of the sound energy that crosses 1 square meter of area in 1 second and is measured by a sound level meter. At a reference value of 1000 hertz, the faintest sounds that can be heard by the human ear have intensities a little less than one-trillionth (10^{-12}) of a watt per square meter. This results from a variation in pressure of about 0.3 billionth of an atmosphere and corresponds to a displacement of the air molecules of approximately one-billionth of a centimeter, which is less than the diameter of a molecule! The ear is a very sensitive instrument.

On the loud side, the ear cannot tolerate sound intensities much greater than 1 watt per square meter without experiencing pain. This corresponds to a variation in pressure of about one-thousandth of an atmosphere, 1 million times as big as for sounds that can be barely heard. The displacements of the molecules are also 1 million times as large.


Erwin C. "Bud" Nielsen/Visuals Unlimited

A source of sound that is transmitting at 100 watts in a spherically symmetric pattern is painful to our ears at a distance of 1 meter and theoretically audible at about 3000 kilometers (2000 miles!), if we assume no losses in moving through the air. Psychoacoustic scientists report that when the intensity of sound is increased about eight times, people report a doubling in the loudness of the sound.

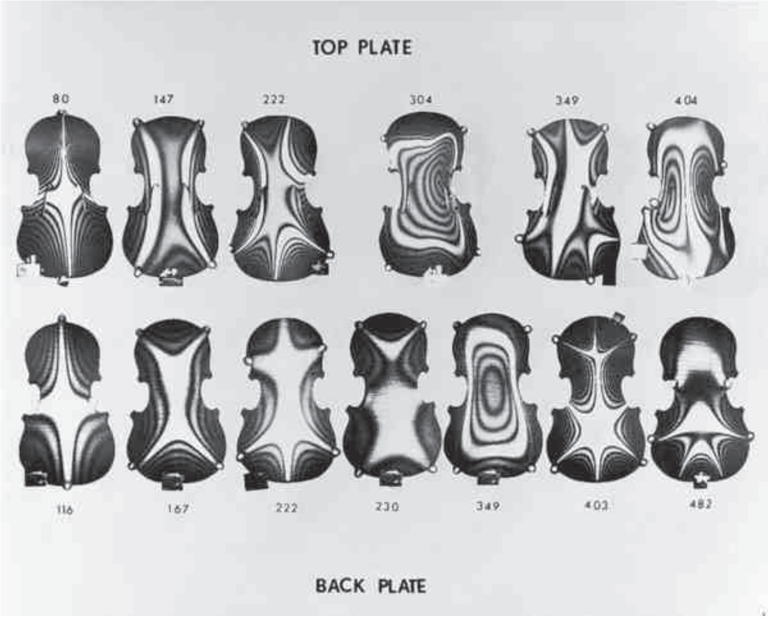
Because of the large range of sensitivity of the ear as well as our perceptual scale, the scale scientists use to distinguish different sound levels is based on multiples of 10. A sound that has 10 times as much intensity as another sound has a level of 10 decibels, or 10 dB (pronounced "dee bee"), higher than the first sound. Because every increase of 10 decibels corresponds to a factor-of-10 increase in intensity, an increase of 20 decibels means that the intensity increases by $10 \times 10 = 100$ times. Sound-level meters use this same nonlinear scale. The table gives some representative values of sound intensity.

A sound level of 85 decibels is safe for unlimited exposure. However, a sound level of 100 decibels is safe for only 2 hours, and an increase to 110 decibels reduces the safe period to 30 minutes. Note that a typical rock concert has an intensity 10 times as large as this. It is also interesting to note that 75% of hearing losses are due to exposure to loud noises and not due to aging. Earplugs typically reduce sound levels by 20–30 decibels and should be used whenever sound levels exceed safe limits.

1. By what factor do earplugs reduce the intensity of sound if they reduce the sound level by 30 dB?
2. The sound intensity from a typical rock concert (not the Vienna Boy's Choir) is greater than the sound intensity of a screaming baby by what factor?

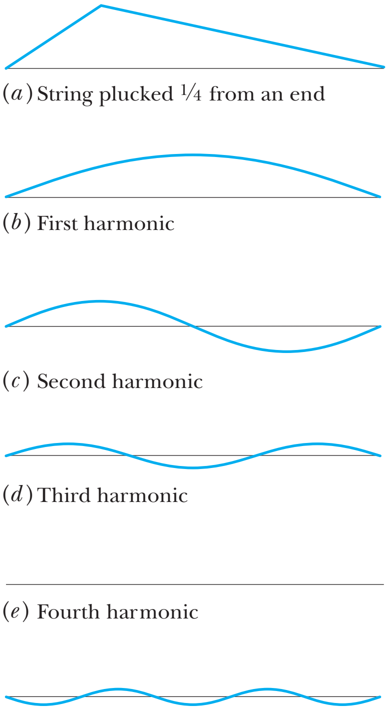
Decibel Levels for Some Common Sounds

Source	Decibels	Energy Relative to Threshold	Sensation
Nearby jet taking off	150	1 quadrillion	
Jackhammer	130	10 trillion	
Rock concert, automobile horn	120	1 trillion	Pain
Police siren, video arcade	110	100 billion	
Power lawn mower, loud music	100	10 billion	
Screaming baby	90	1 billion	Endangers hearing
Traffic on a busy street, alarm clock	80	100 million	Noisy
Vacuum cleaner	70	10 million	
Conversation	60	1 million	
Library	40	10,000	Quiet
Whisper	30	1000	Very quiet
Breathing	10	10	Just audible
	0	1	Hearing threshold



Catgut Acoustical Society, Inc.

Figure 16-4 The vibrations in the body of the violin are made visible by holographic techniques.



(*a*) String plucked 1/4 from an end

(*b*) First harmonic

(*c*) Second harmonic

(*d*) Third harmonic

(*e*) Fourth harmonic

(*f*) Fifth harmonic

Figure 16-5 The shape of the plucked string (a) is the superposition of the first five harmonics (b–f). Note that the fourth harmonic does not contribute to this particular shape.

sound. Figure 16-4 shows a variety of vibrations produced in a violin's body by its vibrating strings. The design of the instrument produces variations in the instrument's vibrational patterns and thus changes the character of the sound produced. There was something special about the way Antonio Stradivari made his violins that made their sound more pleasing than others.

Modern electric guitars do not use the body of the instrument to transmit the vibrations of the strings to the air. The motions of the vibrating strings are converted to oscillating electric signals by pickup coils mounted under the strings. This signal is then amplified and sent to the speakers.

Whether acoustic or electric, plucking a guitar string creates vibrations. This initial distortion causes waves to travel in both directions along the string and reflect back and forth from the fixed ends. The initial shape of the string is equivalent to a unique superposition of many harmonic waves. Figure 16-5 shows the shape of a plucked string at the moment of release and the contributions of the first five harmonics. (For this particular shape, the fourth harmonic is zero everywhere and does not contribute.)

These harmonic waves travel back and forth on the string, creating standing waves with nodes at the two ends of the string. The first three standing waves are shown in Figure 16-6. Because the distance between nodes is one-half the wavelength, the wavelength of the fundamental, or first harmonic (a), is twice as long as the string, or $\lambda_1 \times 2L$. The wavelength of the second harmonic (b) is $\lambda_2 = L$, the length of the string, and therefore half as long as that of the fundamental. The wavelength of the third harmonic (c) is $\lambda_3 = \frac{2}{3}L$, one-third of the fundamental wavelength.

harmonic wavelengths $\lambda_1, \frac{\lambda_1}{2}, \frac{\lambda_1}{3}, \ldots$ ►

Q: What is the wavelength corresponding to the third harmonic on a 60-centimeter-long wire?

A: It must be one-third the length of the fundamental wavelength. Because the fundamental wavelength is twice the length of the string, we obtain 2(60 centimeters)/3 = 40 centimeters.

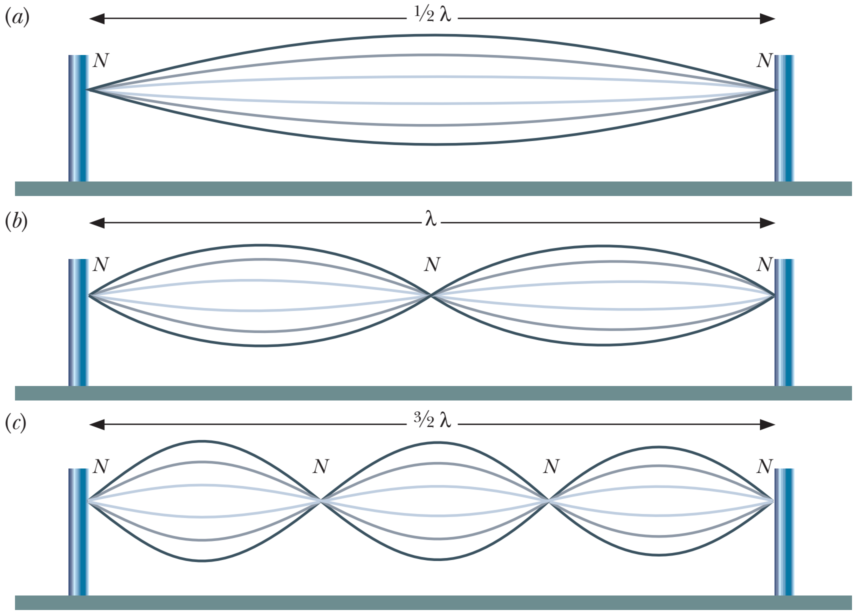


Figure 16-6 The first three standing waves on a guitar string. The wavelengths of these harmonics are $2L$, L, and $\frac{2}{3}L$, respectively.

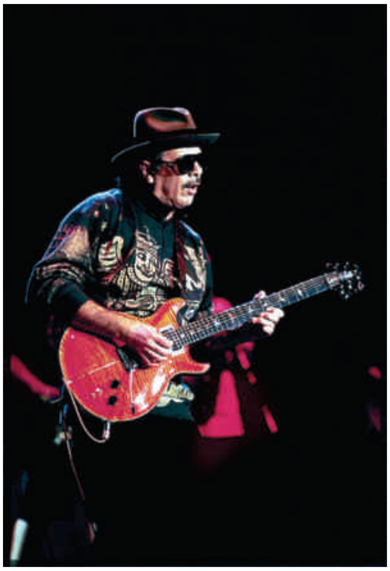
The frequencies of the various harmonics can be obtained from the relationship for the wave's speed, $v = \lambda f$. As long as the vibrations are small, the speeds of the different waves on the plucked string are all the same. Therefore, $v = \lambda_1 f_1 = \lambda_2 f_2 = \lambda_3 f_3 = \ldots$. Because the second harmonic has half the wavelength, it must have twice the frequency. The third harmonic has one-third the wavelength and three times the frequency. The harmonic frequencies are whole-number multiples of the fundamental frequency.

◀ harmonic frequencies $f_1, 2f_1, 3f_1, \ldots$

There is a simple way to verify that more than one standing wave is present on a vibrating string. By lightly touching the string in certain places, particular standing waves can be damped out and thus make others more obvious. For example, suppose you touch the center of the vibrating guitar string. We can see from Figure 16-6(a) that the fundamental will be damped out because it has an antinode at the middle. The second harmonic, however, has a node at the middle [Figure 16-6(b)], so it is unaffected by your touch. The third harmonic is damped because it also has an antinode at the middle of the string. In fact, all odd-numbered harmonics are damped out, and all even-numbered ones remain.

When you do this experiment, you hear a shift in the lowest frequency. With the initial pluck, the fundamental is the most prominent frequency. After touching the middle of the string, the fundamental is gone, so the lowest frequency is now that of the second harmonic, a frequency twice the original frequency. In musical terms we say that the note shifts upward by one octave. You can find the higher harmonics by plucking the string again and gently touching it $\frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots$ of the way from one end.

Music, of course, consists of many notes. The string we have been discussing can play only one note at a time because the fundamental frequency determines the musical note. Instruments must be able to play many different frequencies to be useful. To play other notes, we must have more strings or a simple way of changing the vibrational conditions on the string. Most stringed instruments use similar methods for achieving different notes. Pianos, harps, and harpsichords have many strings. Striking or plucking different strings produces different notes.



Guitarist Carlos Santana changes notes by pushing the strings against the frets to shorten the lengths that vibrate.

Are You On the Bus?

Q: What are the different ways to increase the fundamental frequency of a note played on a guitar string?

A: To increase the fundamental frequency of a note on a guitar string, you can increase the tension, finger the string, or use a string with less mass per unit length.

Figure 16-7 The note played by this one-string bass is changed by varying the tension (and thus the wave speed) in the string.

A guitar usually has six strings, some more massive than others to extend the range of frequencies. Still, only six strings aren't enough to produce many interesting songs. A guitarist must be able to change the note produced by each string. By fingering the string the guitarist shortens the vibrating portion of the string, creating new conditions for standing waves. The new fundamental wavelength is now twice this *shortened* length and therefore smaller than before. This smaller wavelength produces a higher frequency.

The guitarist can change the speed of the waves by changing the tension of the string. An increase in the tension increases the speed and therefore increases the frequency. It takes too long to change the tension in the middle of a song, so changes in tension are only used to tune the instruments. An exception is the "washtub bass" shown in Figure 16-7.

Wind Instruments

Wind instruments—such as clarinets, trumpets, and organ pipes—are essentially containers for vibrating columns of air. Each has an open end that transmits the sound and a method for exciting the air column. With the exception of the organ pipe, wind instruments also have a method for altering the fundamental frequency. In many ways wind instruments are analogous to stringed instruments. A spectrum of initial waves is created by a disturbance. The instrument governs the standing waves that are generated; all other frequencies are quickly damped out. The sound we hear is a combination of the frequencies of these standing waves.

Figure 16-8 A drawing of a longitudinal sound wave. The dark regions represent compressions (high density); the lighter regions represent rarefactions (low density).

But there are three main differences between wind and string instruments. First, unlike vibrating strings, the vibrational characteristics of air cannot be altered to change the speed of the waves. Only the length of the vibrating air column can be changed. Second, a string has a node at each end, but there is an antinode near the open end of the wind instrument. At an antinode the vibration of the air molecules is a maximum. A node occurs at the closed end, where there is no vibration of the air molecules. Finally, the disturbance in wind instruments produces longitudinal waves in the air column instead of the transverse waves of the stringed instruments. A longitudinal wave at a single moment is represented in Figure 16-8.

Drawing longitudinal standing waves is difficult because the movement of the air molecules is along the length of the pipe. If we draw them in this manner for even one period, crests and troughs overlap and produce a confusing drawing. When illustrating longitudinal waves in a windpipe, we normally draw them as if the air movements were transverse. This should not be overly confusing as long as you remember that these drawings are basically graphs of the displacements of the air molecules versus position. In Figure 16-9 two curves are drawn to represent the range of displacement of the standing waves. Note that the curves meet at the closed end of the pipe. There is a node at that spot, indicating that the air molecules near the wall do not move.

Consider a closed organ pipe—one that is closed at one end and open at the other. The largest wavelength that produces a node at the closed end and an antinode at the open end is four times the length of the pipe, or $4L$. This is

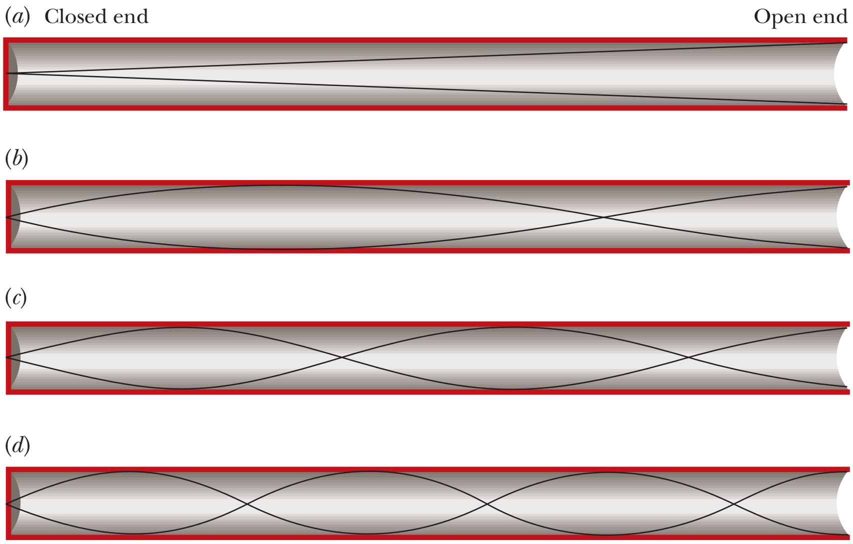


Figure 16-9 Graphs of the first four standing waves in a closed organ pipe. Note that there is a node at the closed end and an antinode at the open end.

illustrated in Figure 16-9(a). The next-smaller wavelength [Figure 16-9(b)] is four-thirds the length of the pipe, or $\frac{4}{3}L$; and the wavelength in Figure 16-9(c) is four-fifths the length of the pipe, or $\frac{4}{5}L$.

As with the stringed instruments, we can generate a relationship among the various frequencies by examining the relationship between the wave's speed, its wavelength, and its frequency: $v = \lambda f$. The possible wavelengths are 1, $\frac{1}{3}$, $\frac{1}{5}$, $\frac{1}{7}$, . . . times the fundamental wavelength. Because the wave speeds are constant, the corresponding frequencies are 1, 3, 5, 7, . . . times the fundamental frequency.

Q: What is the wavelength corresponding to the fifth harmonic in a 50-centimeter-long closed organ pipe?

A: It must be one-fifth the length of the fundamental wavelength. Because the fundamental wavelength is four times the length of the pipe, we obtain 4(50 centimeters)/5 = 40 centimeters.

The closed organ pipe does not have resonant frequencies that are 2, 4, 6, . . . times the fundamental frequency; the even-numbered harmonics are missing because these frequencies would not produce a node at the closed end and an antinode at the open end. On the other hand, the open organ pipe (one that is open at both ends) has all harmonics.

Percussion Instruments

Percussion instruments are characterized by their lack of the harmonic structure of the string and wind instruments. Percussionists employ a wide range of instruments, including drums, cymbals, bells, triangles, gongs, and xylophones. Although all of these resonate at a variety of frequencies, the higher frequencies are not whole-number multiples of the lowest frequency.

Each of the percussion instruments behaves in a different way. The restoring force in a drum is provided by the tension in the drumhead. The two-dimensional standing-wave patterns are produced by the reflection of transverse waves from the edges of the drum and are characterized by nodal lines. These

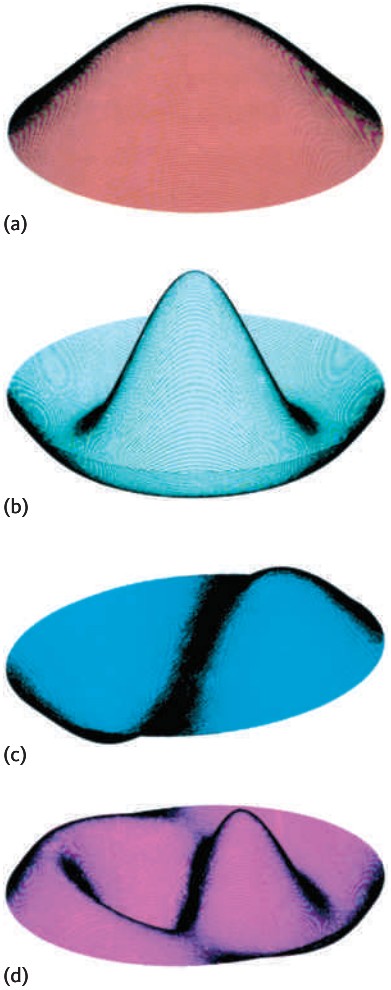


Figure 16-10 A circular drumhead can vibrate in many different modes with nonharmonic frequencies.

WORKING IT OUT *Organ Pipe*

An organ pipe is found to produce resonances at 150 Hz and 250 Hz but at no frequency in between. Is this an open pipe or a closed pipe? What is the pipe's length?

There is no formula to answer the first part of this question. What we have to ask is whether this appears to be part of a series of all multiples of a single frequency or just the odd multiples. The highest frequency that goes evenly into both values is 50 Hz. An open pipe with a 50-Hz fundamental would produce 100 Hz, 150 Hz, 200 Hz, 250 Hz, and so on. Even though it would produce the two required frequencies, it would also produce a resonance at 200 Hz, which we know doesn't exist in this pipe. A closed pipe would produce 50 Hz, 150 Hz, 250 Hz, 350 Hz, and so on, as required. This is therefore a closed organ pipe with a fundamental frequency of 50 Hz. To find the tube's length, we begin by using the wave equation to find the wavelength:

$$\lambda = v/f$$

$$\lambda = \frac{343\ \text{m/s}}{50\ \text{Hz}} = 6.86\ \text{m}$$

At the fundamental frequency, only one-quarter of a wavelength fits in the pipe, so the length of the organ pipe must be

$$L = \lambda/4 = 1.7\ \text{m}$$

nodal lines are the two-dimensional analogs of the nodal points in vibrating strings.

The nodal lines for a circular membrane are either along a diameter or circles about the center; there is always a nodal line around the edge of the drumhead. Figure 16-10(a) shows the fundamental mode, in which the center of the drumhead moves up and down symmetrically. Figure 16-10(c) shows a mode with an antinodal line along a diameter. When one-half of the drumhead is up, the other half is down. This mode has a frequency 1.593 times that of the fundamental frequency. The second symmetric mode is shown in Figure 16-10(b) and has a frequency 2.295 times the fundamental frequency.

Beats

When we listen to two steady sounds with nearly equal frequencies, we hear a periodic variation in the volume. This effect is known as **beats** and should not be confused with the rhythm of the music that you may dance to; beats are the result of the superposition of the two waves. Because the two waves have different frequencies, there are times when they are in step and add together and times when they are out of step and cancel. The result is a periodic cancellation and reinforcement of the waves that is heard as a periodic variation in the loudness of the sound.

This is illustrated by the drawings in Figure 16-11. It is important to realize that these drawings do not represent strobe pictures of the waves. The horizontal line represents time, not position. The drawings represent the variation in the amplitude of the sound at a single location—for instance, as it reaches one of your ears. Figures 16-11(a) and (b) show each wave by itself, and Figure 16-11(c) shows the superposition of these two waves. Your ears hear a frequency that is the average of the two frequencies, and that varies in amplitude with a beat frequency.

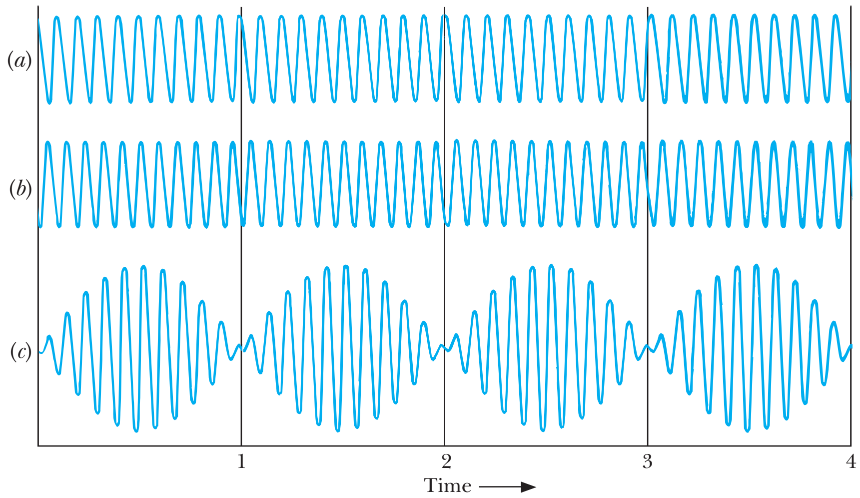


Figure 16-11 The superposition of two sound waves of different frequencies produces a sound wave (c) that varies in amplitude.

We can obtain the beat frequency by examining Figure 16-11. If the time between vertical lines is 1 second, wave (a) has a frequency of 10 hertz and wave (b) 11 hertz. At the beginning of a second, a crest from wave (a) cancels a trough from wave (b), and the sound level is zero. Wave (a) has a lower frequency and continually falls behind. At the end of 1 second, it has fallen an entire cycle behind, and the waves once again cancel. The difference of 1 hertz in their frequencies shows up as a variation in the sound level that has a frequency of 1 hertz. This same process is valid for any frequencies that differ by 1 hertz. For example, the beat frequency produced when two sound waves of 407 hertz and 408 hertz are combined is also 1 hertz.

If the frequencies differ by 2 hertz, it takes only $\frac{1}{2}$ second for the lower-frequency wave to fall one cycle behind. Therefore, this pattern happens two times per second, or with a beat frequency of 2 hertz. This reasoning can be generalized to show that the beat frequency is equal to the difference in the two frequencies.

◀ beat frequency $= \Delta f$

Piano tuners employ this beat phenomenon when tuning pianos. The tuner produces the desired frequency by striking a tuning fork and then adjusts the piano wire's tension until the beats disappear. Modern electronics have now made it possible for a tone-deaf person to make a living tuning pianos.

WORKING IT OUT *Beats*

You play an unknown tuning fork alongside a 440-Hz tuning fork. The resulting sound has a beat frequency of 3 Hz. You then play the unknown alongside a 445-Hz tuning fork and hear a beat frequency of 8 Hz. What is the frequency of the unknown tuning fork?

We can reason the solution based on the idea that the beat frequency is the difference in the two source frequencies. From the first experiment we know that the unknown tuning fork must be different from 440 Hz by 3 Hz. The choices are 437 Hz and 443 Hz. Clearly, the first experiment does not provide enough information to answer the question. The second experiment tells us that the unknown tuning fork must be different from 445 Hz by 8 Hz. The choices are 437 Hz and 453 Hz. The choice that is consistent with both experiments is 437 Hz.

Q: Why don't we hear beats when adjacent keys on a piano are hit at the same time?

A: The beats do exist, but the beat frequency is too high for us to notice.

Are You On the Bus?

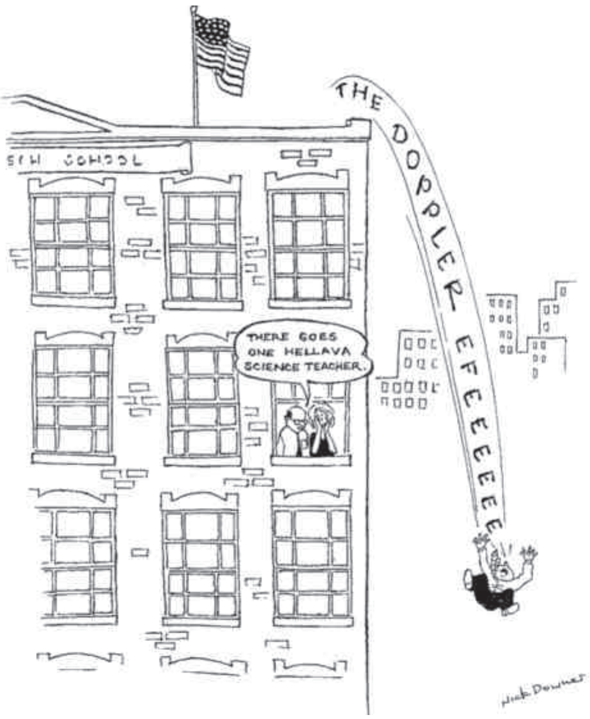
Doppler Effect

What we hear is not necessarily the sound that was originally produced, even when only a single source is involved. Recall the sound of a car as it passes you. This variation in sound is especially noticeable with race cars and some motorcycles because they emit distinctive roars. Imagine standing by the side of a road as a car passes you with its horn blasting constantly. Two things about the horn's sound change. First, as the car approaches, its sound gets progressively louder; as it leaves, the sound gets quieter. The volume changes simply because the wave spreads out as it moves away from the source. If you are far from the source, the energy of the sound waves intercepted by your ears each second is smaller.

The second change in the sound of the horn may not be as obvious. The frequency *that you hear* is not the same as the frequency that is actually emitted by the horn. The frequency you hear is higher as the car approaches you and lower as the car recedes from you. This shift in frequency due to the motion is called the **Doppler effect,** after the Austrian physicist and mathematician Christian Doppler.

The pitch of the sound we hear is determined by the frequency with which crests (or troughs) hit our ears. Our ears are sensitive to the frequency of a wave, not to the wavelength. Figure 16-12 shows a two-dimensional drawing of sound waves leaving a tuning fork. A small portion of the sound is intercepted by the ear. If the tuning fork and the ear are stationary relative to each other, the frequency heard is the same as that emitted.

When the tuning fork is moving, the listener hears a different frequency. Because the tuning fork moves during the time between the generation of one crest and the next, the sound waves crowd together in the forward direction and spread out in the backward direction, as shown in Figure 16-13. If the tun-



Q: If you were flying a model airplane on a wire so that it traveled in circles about you, would you hear a Doppler shift?

A: Because the airplane is not moving toward or away from you, you would not hear a Doppler shift. However, someone standing off to the side watching you would hear a Doppler shift.

ing fork moves toward the listener, the ear detects the sound waves that are crowded together. The frequency with which the waves hit the ear is therefore higher than was actually emitted by the source. Similarly, if the tuning fork moves away from the listener, the observed frequency is lower.

Notice, however, that in both Figures 16-12 and 16-13 the spacings do not change with the separation of the source and receiver. Because the shift does not depend on the distance, the Doppler-shifted frequency does not change as the source gets closer or farther away.

We will see in later chapters that the Doppler effect also occurs for light. The observation that the light from distant galaxies is shifted toward lower frequencies (their colors are *redshifted*) tells us that they are moving away from our Galaxy and that the universe is expanding.

The Doppler effect also occurs when the receiver moves and the source is stationary. If the listener moves toward a stationary tuning fork, her ear intercepts wave crests at a higher rate, and she hears a higher frequency. If she recedes from the tuning fork, the crests must continually catch up with her ear, which therefore intercepts the crests at a reduced rate, and she consequently hears a lower frequency. Notice that the shift in the frequency is still constant over time as long as the velocities are constant. Of course, the loudness of the signal decreases as the distance to the tuning fork increases.

When a wave bounces off a moving object, it experiences a similar Doppler shift in frequency due to successive crests (or troughs) having longer or shorter distances to travel before reflecting. By monitoring these frequency shifts, we can determine the speed of objects toward or away from the original source. Because the Doppler effect occurs for all kinds of waves, this technique is used in many situations, such as catching speeding motorists using radar (an electromagnetic wave) and monitoring the movement of dolphins using sound waves in water.

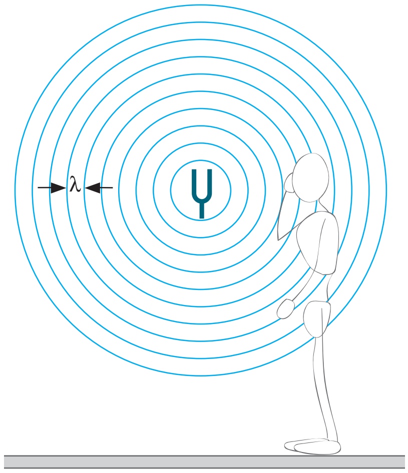


Figure 16-12 If the source and the ear are stationary relative to each other, the ear hears the same frequency as emitted by the source.

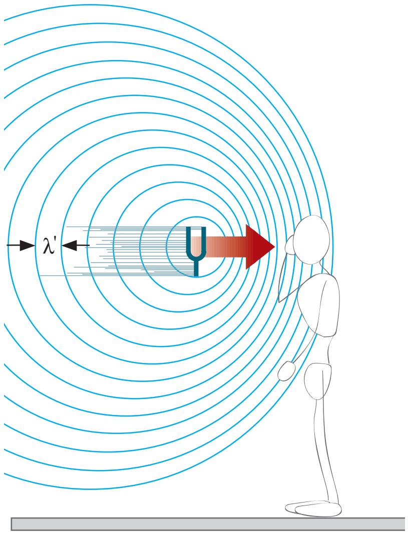


Figure 16-13 If the source of sound moves toward the right, the waves bunch up on the right-hand side and spread out on the left-hand side. The person hears a frequency that is higher than that emitted by the source.

FLAWED REASONING

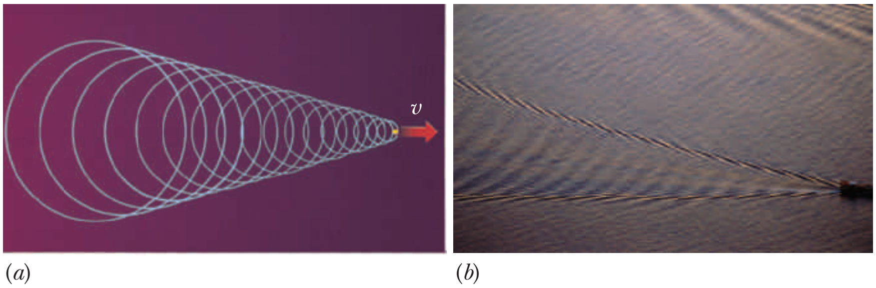
The following question appears on the final exam: "You and a friend are standing a city block apart. An ambulance with its siren blaring drives down the street toward you and away from your friend. The ambulance passes you and continues down the street at a constant speed. Is the frequency that you hear higher than, lower than, or the same as the frequency your friend hears?"

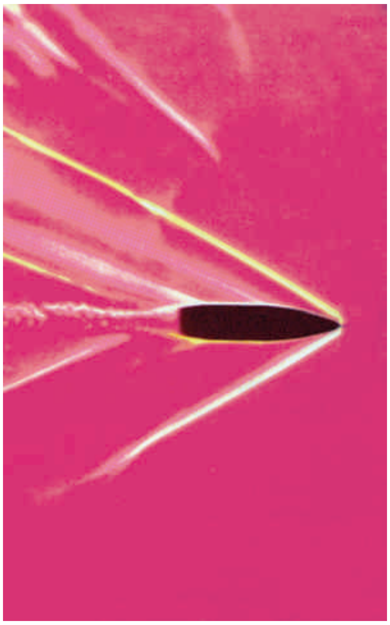
Matt gives the following answer to this question: "According to the Doppler shift, as the siren gets farther away, the wavelengths get farther apart. The longer the wavelength, the lower the frequency. I will therefore hear a higher frequency than my friend."

What is wrong with Matt's reasoning, and what is the correct answer to the exam question?

ANSWER The Doppler shift does not depend on the distance to the source. The shift in frequency depends only on the velocity of the source relative to the observer. The ambulance is moving away from both of you with the same speed. Therefore, both of you will hear the same lower frequency, but the siren will be louder to you.

Figure 16-14 (a) When the source moves faster than the speed of sound, the waves form a cone known as the shock wave. (b) The presence of the bow wave (analogous to a shock wave) indicates that the canoe is traveling faster than the speed of the water waves.





A stroboscopic photograph of a bullet traveling through the hot air above a candle. The formation of the shock wave tells us that the bullet was traveling faster than the speed of sound.

Shock Waves

When a source of waves moves faster than the speed of the waves in the medium, the next crest is generated in front of the leading edge of the previous crest. This causes the expanding waves to superimpose and form the conical pattern shown in Figure 16-14(a). The amplitude along the cone's edge can become very large because the waves add together with their crests lined up. The edge of the cone is known as a **shock wave** because it arrives suddenly and with a large amplitude.

Shock waves are common in many media. When speedboats go much faster than the speed of the water waves, they create shock waves commonly known as wakes. The Concorde, a supersonic plane, travels much faster than the speed of sound in air and therefore produces shock waves. Some people are concerned about the effects of the sonic boom when the edge of the pressure cone reaches the ground.

The return of a space shuttle to Earth produces a double sonic boom. The nose produces one boom, and the engine housings near the rear of the spacecraft produce the other. Listen for this the next time you watch a television broadcast of a shuttle returning.

Are You On the Bus?

Q: Would you expect a spacecraft traveling to the Moon to produce a shock wave during its entire trip?

A: No. Because there is no air in most of the space between Earth and the Moon, the spacecraft would not produce any sound.

Summary

Sound is a longitudinal wave that travels through a variety of media. The speed of sound is 343 meters per second in air at room temperature, four times as fast in water, and more than ten times as fast in solids. In a gas the speed of sound depends on the temperature (slower at lower temperatures) and the type of gas (faster in molecules with less mass).

Ears detect sound waves and send electric signals to the brain. The range of frequencies that can be detected depends on the resonant structures in the animals' ears. The audible range in humans is roughly 20–20,000 hertz, although it varies with age and the individual.

You can recognize different sources of sound because each source produces a unique combination of intensities of the various harmonics—its own recipe of harmonics.

Music generally differs from other sounds in its periodicity. Nearly all musical instruments—string, wind, or percussion—involve the production of stand-

Everyday Physics | *Breaking the Sound Barrier*

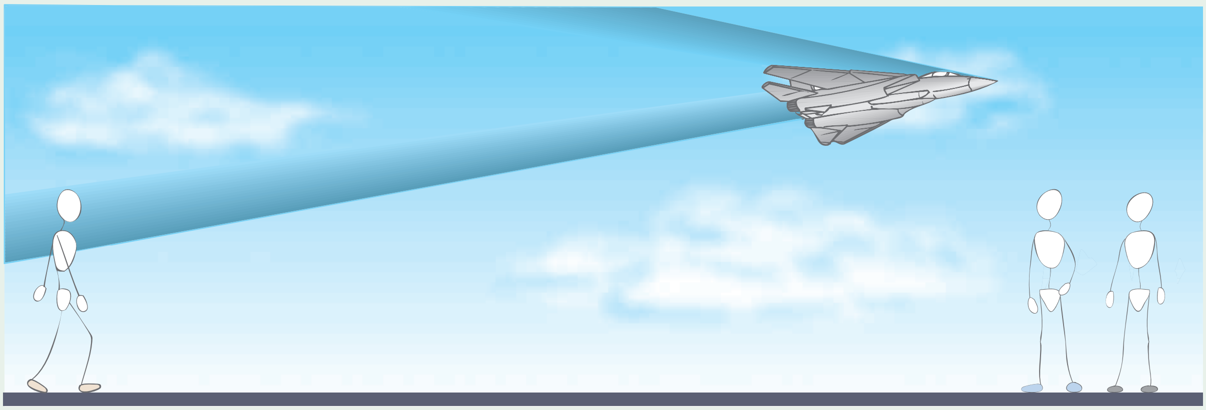
A common misunderstanding is that the sonic boom occurs when the aircraft "breaks" the sound barrier. The boom actually continues as long as the plane flies at supersonic speeds. However, at a given location, the loud noise is heard only when the edge of the pressure cone reaches the ear, as illustrated in the figure. This short, loud noise results from the accumulated superposition of the wave crests generated by the plane.

The first person to break the sound barrier was Chuck Yeager on October 14, 1947. He flew a Bell XS-1 rocket plane at 12,800 meters over Edwards Air Force Base in California at a speed of 1080 kilometers per hour (670 miles per hour). This speed was also known as Mach 1.015 because it was 1.015 times the speed of sound at this altitude on that day. The supersonic Concorde cruises at speeds up to Mach 2.2 (1450 miles per hour), and the Russian MIG-25B has been tracked on radar at speeds up to Mach 3.2 (2110 miles per hour).

1. If a MIG-25B speeds up from Mach 2.0 to Mach 3.0, does the angle defining the shape of the trailing pressure cone get larger or smaller? Explain.
2. What is the reading on your car's speedometer (in mph) if you are traveling at Mach 1.6? (Don't try this!)

Courtesy of U.S. Navy

This F/A-18 Hornet is transitioning from subsonic to supersonic flight. The atmospheric conditions were just right for the shock waves produced by the cockpit and the wings/fuselage to produce condensation.



A plane traveling faster than the speed of sound produces a shock wave that is heard after the plane passes overhead. The couple on the right has not yet heard the sonic boom or any other sound from the plane.

ing waves. In many of these instruments, the standing-wave frequencies, or harmonics, are whole-number multiples of the fundamental frequency. The frequency of the second harmonic is twice the fundamental frequency. The string's actual shape is a superposition of all the standing waves and looks quite different from those of the individual harmonics.

Changing the standing-wave conditions generally produces different notes. The relationship $v = \lambda f$ tells us that if the speed or the wavelength is changed, the frequency must change. A smaller wavelength gives a higher frequency when the speed is held constant.

Two steady sounds with nearly equal frequencies superimpose to form beats, a variation in the loudness of the sound that has a frequency equal to the difference in the frequencies.

The frequency of the sound from a moving object is shifted in frequency according to the Doppler effect. The frequency is shifted upward as the emitting object approaches and downward as it recedes from the receiver. The shift in the frequency is constant as long as the velocity is constant. The Doppler effect occurs for all kinds of waves.

When a source of waves moves faster than the speed of the waves in the medium, a conical shock wave is formed. When the edge of the cone created by supersonic aircraft reaches us, we hear a sonic boom.

CHAPTER 16 *Revisited*

By studying the sounds produced by instruments such as guitars, pianos, and organs, we learn that musical sounds often have combinations of frequencies that are whole-number multiples of the lowest frequency. The size of the body of the instrument determines the range of frequencies that it amplifies.

Key Terms

beats A variation in the amplitude resulting from the superposition of two waves that have nearly the same frequencies. The frequency of the variation is equal to the difference in the two frequencies.

Doppler effect A change in the frequency of a periodic wave due to the motion of the observer, the source, or both.

shock wave The characteristic cone-shaped wave front that is produced whenever an object travels faster than the speed of the waves in the surrounding medium.

Questions and exercises are paired so that most odd-numbered are followed by a similar even-numbered.
Blue-numbered questions and exercises are answered in Appendix B.
indicates more challenging questions and exercises.
WebAssign Many Conceptual Questions and Exercises for this chapter may be assigned online at WebAssign.

Conceptual Questions

1. What is the evidence that sound is a wave phenomenon?
2. Why is it impossible to send a longitudinal wave down a string or a transverse wave through the air?
3. Which of the following has the greatest effect on the speed of sound in air: amplitude, frequency, wavelength, or temperature?
4. Does a 220-hertz sound wave move faster than, slower than, or at the same speed as a 440-hertz sound wave? Explain.
5. Is the speed of sound faster in air or water?
6. Why does the voice of a clown who has inhaled helium sound so high pitched?

7. Why do you see the lightning before you hear the thunder?

8. When the time interval between seeing the lightning and hearing the thunder is short, the thunder is loud. Why?

9. Even though you may be far away from an orchestra, the trumpet and the trombone do not sound "out of step" with each other. What does this tell us about sound waves?

10. You visit the Grand Canyon on a particularly hot day and, like all good tourists, you yell, "Hello!" Would you expect to hear the echo sooner or later than on a cooler day? Why?

11. If earplugs are advertised to reduce sound levels by 20 decibels, by how much do they reduce the intensity of the sound?

12. A friend tells you that a real estate developer has reassured her that the increased traffic due to the new mall on her street will increase the average traffic noise a mere 20%, from 50 decibels to 60 decibels. How would you advise her?

13. Is it possible for a 440-hertz sound wave to be louder than an 880-hertz sound wave?

14. Which of the following properties of a sound wave determines its loudness (or intensity): wavelength, speed, amplitude, or frequency?

15. Which of the following properties of a sound wave determines its pitch: wavelength, speed, amplitude, or frequency?

16. Which is longer—the wavelength of infrasound or ultrasound?

17. The speed of sound is much greater in water than in air. If you were to jump into a pool holding a 500-hertz buzzer, would the frequency of the sound you hear while underwater be greater than, less than, or equal to 500 hertz? Explain your reasoning.

18. For sound waves in room-temperature air, is it possible to change the wavelength of a sound without changing the frequency? Explain your reasoning.

19. What frequency is two octaves higher on the musical scale than 100 hertz?

20. Which harmonic is one octave higher in frequency than the fundamental?

Courtesy of C.F. Martin & Co., Nazareth, PA

21. How can you increase the speed of the waves on a guitar string?

22. Why is the low E string on a guitar made of thicker wire than the high E string?

23. How does the wavelength of the fundamental standing wave on a violin string compare with the length of the string?

24. Two standing waves are created on the same guitar string. Will the frequency of the one with the shorter wavelength be higher than, lower than, or the same as the frequency of the one with the longer wavelength? Explain.

25. On what two places on a guitar string of length *L* could you place your finger on an antinode to damp out the second harmonic? Would the fourth harmonic be damped out if you placed your finger at either of these locations?

26. A string vibrates with a fundamental frequency of 100 hertz. You lightly place your finger at a point one-quarter of the way from one end to damp out various harmonics. What is the lowest frequency that will still resonate on the string?

27. Does plucking the string harder affect the fundamental frequency of a guitar string? Explain.

28. Does increasing the tension of the string affect the wavelength of the fundamental standing wave on a guitar string?

29. Does the second or the third harmonic on a guitar string have traveling waves with the higher speed, or are the speeds the same?

30. You can tune the frequency of a guitar string without changing its length. To tune a flute, however, you have to slightly adjust the length of the tube. How do you account for this difference?

31. Would you expect to find nodes or antinodes at the ends of a guitar string? Explain.

32. Would you expect to find a node or an antinode at each end of an organ pipe that is closed at one end?

33. How does the wavelength of the fundamental standing wave for an organ pipe that is open at both ends compare to the length of the pipe?

Courtesy of Henry Leap

34. How does the wavelength of the fundamental standing wave for an organ pipe that is closed at one end and open at the other compare to the length of the pipe?

35. How does the fundamental wavelength of an organ pipe that is open at both ends change as you close one of the ends?

36. How does the fundamental frequency of an organ pipe that is open at both ends change as you close one of the ends?

37. You have an organ pipe that resonates at frequencies of 300, 450, and 600 hertz but nothing in between. It may resonate at lower and higher frequencies as well. Is the pipe open at both ends or open at one end and closed at the other? How can you tell?

38. You have an organ pipe that resonates at frequencies of 500, 700, and 900 hertz but nothing in between. It may resonate at lower and higher frequencies as well. Is the pipe open at both ends or open at one end and closed at the other? How can you tell?

39. What happens to the fundamental frequency of an organ pipe as the temperature in the room decreases? Explain.

40. What would happen to the fundamental wavelength of an organ pipe if you fill it with helium instead of air? What would happen to the fundamental frequency? Explain.

41. The fundamental wavelength for a bar on a xylophone that is suspended at a point midway between the ends and the middle of the bar is _____ the length of the bar.

42. What is the fundamental wavelength for standing waves on a rod held at its center?

43. Electronic signal generators can produce pure frequencies with great precision. Why then is it so difficult to make an electronic piano that sounds like an acoustic piano?

44. Why does middle C played on an oboe sound different from middle C played on a piano?

45. As the frequencies of two waves get farther apart, what happens to the beat frequency?

46. As the frequencies of two waves get closer together, what happens to the beat frequency?

47. You are singing along with Celine Dion on the car radio as she sings a long steady note. You know that you are a little off-key because you are hearing beats at 2 hertz. Are you singing a little too high or a little too low, or can't you tell? Explain.

48. You have two pairs of tuning forks, one pair at 345 and 348 hertz and the other pair at 489 and 491 hertz. Which pair of tuning forks, when sounded together, produces the higher beat frequency?

49. The two wires corresponding to one key on a piano are out of tune. If we decrease the tension of the wire producing the lower frequency, will the beat frequency increase, decrease, or stay the same? Why?

50. You have two tuning forks that, when played together, produce beats at a frequency of 4 hertz. You place a small piece of putty on one of the tuning forks, which lowers its frequency, and find that the beat frequency increases. Did you place the putty on the tuning fork with the lower or the higher initial frequency? Explain.

51. Describe the sounds you would hear if a train passed you with its whistle blowing.

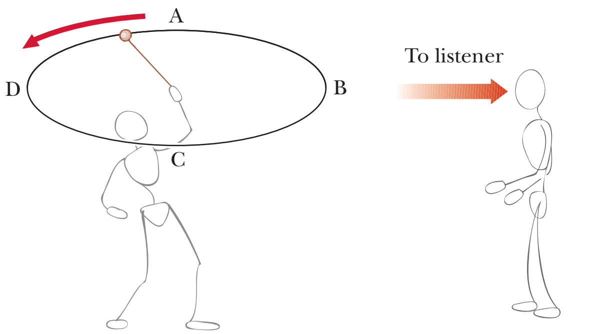
52. A tuning fork rings with a frequency of 400 hertz. As the tuning fork moves away from you, would the frequency you hear be lower than, equal to, or higher than 400 hertz? Explain your reasoning.

53. Which of the following properties of the wave does not change in the Doppler effect: wavelength, speed, or frequency? Explain.

54. An automobile sounding its horn is moving toward you at a constant speed. How does the frequency you hear compare with that heard by the driver?

55. A police car with a fixed-frequency siren drives straight toward you. It slows down gradually and stops right next to you. As the car is slowing down, what do you hear happening to the frequency and the loudness of the siren?

56. You tie a battery-powered, fixed-frequency buzzer to the end of a string and swing it around in a large horizontal circle as shown. Your friend stands well off to the right and listens to the tone. At which of the labeled locations will the buzzer be when your friend hears the highest frequency? (Assume that the time it takes the sound to travel from the buzzer to your friend's ear can be ignored.)



57. Billy and Elaine are holding electric buzzers that sound at slightly different frequencies. When they are both standing still, the two buzzers produce a beat frequency of 4 hertz. Elaine begins to run away, and Billy hears the beat frequency gradually increase. Whose buzzer has the higher frequency? Explain.

58. If a wind is blowing from a siren toward your ear, does the frequency of the siren change? Why?

59. Explain why a sonic boom sounds much like an explosion.

60. What does the observation of a wake tell you about the speed of a boat compared with the speed of water waves?

61. A jet that has been traveling faster than the speed of sound for several miles passes overhead. Describe what you would hear from the moment the plane is directly overhead until it is several miles away.

62. A jet is flying along close to the ground at just below the speed of sound. Describe how the pilot could "break the sound barrier" without flying faster.

Exercises

63. The musical note "middle C" has a frequency of 262 Hz. What is its period of vibration?

64. What is the period of a tuning fork that is vibrating at 1200 Hz?

65. What is the wavelength of a musical note with a frequency of 1248 Hz?

66. How long is a wavelength of infrasound with a frequency of 4 Hz?

67. What frequency would you need to produce a sound wave in room-temperature air with a wavelength of 1 m?

68. Radio waves travel at the speed of light, which is 3×10^8 m/s. What is the wavelength for FM station WFIZ broadcasting at 97.3 MHz? (The prefix M is "mega" and means 10^6.)

69. What is the longest wavelength that can be heard by a normal ear?

70. What is the shortest wavelength that the average human can hear?

71. You observe that the delay between a lightning flash and the thunder is 8 s. How far away is the lightning?

72. If the "shot heard 'round the world" could actually travel around the world, how long would it take? (Assume that the circumference of Earth is 40,000 km.)

73. If the sonar signal sent straight down from a boat takes 0.8 s to return, how deep is the lake?

74. Your sonic range finder measures the distance to a nearby building at 20 m. The range finder is calibrated for sound traveling 343 m/s, but on this very cold day (−37°C) the speed of sound is only 309 m/s. How far away is the building?

Courtesy of PASCO Scientific

75. If ear protectors can reduce the sound intensity by a factor of 10,000, by how many decibels is the sound level reduced?

76. You go out and spend a bucket of money on a new stereo amplifier that advertises that it can produce maximum volumes 20 dB greater than your current system. How many times as powerful must this amplifier be?

77. If the fundamental frequency of a guitar string is 125 Hz, what harmonic frequencies are possible?

78. You measure all of the possible resonant frequencies for a guitar string in the range 500–1000 Hz and find that it will resonate only at 600, 750, and 900 Hz. What is the fundamental frequency for this string?

79. What is the fundamental frequency for a 50-cm banjo string if the speed of waves on the string is 470 m/s?

80. If the fundamental frequency of a 60-cm-long guitar string is 500 Hz, what is the speed of the traveling waves?

81. What harmonic frequencies are possible in a closed organ pipe that is 1.2 m long?

82. What harmonic frequencies are possible in an open organ pipe that is 1.2 m long?

83. What length of closed organ pipe is required to produce the note B_4 with a frequency of 493.88 Hz?

84. How long is an open organ pipe with a fundamental frequency of 493.88 Hz?

85. You have an organ pipe that resonates at frequencies of 375, 450, and 525 Hz but nothing in between. It may resonate at lower and higher frequencies as well. What is the fundamental frequency for this pipe?

86. You have an organ pipe that resonates at frequencies of 500, 700, and 900 Hz but nothing in between. It may resonate at lower and higher frequencies as well. What is the fundamental frequency for this pipe?

87. A tuning fork has been damaged, and its frequency slightly changed. What could its altered frequency be if it produces two beats per second with a tuning fork that is known to vibrate at 262 Hz?

88. You have a tuning fork of unknown frequency. When you ring it alongside a tuning fork with known frequency of 360 Hz, you hear beats at a frequency of 2 Hz. When you ring it alongside a tuning fork with known frequency of 355 Hz, you hear beats at a frequency of 3 Hz. What is the unknown frequency?

The Big Picture

The Mystery of Light


Jerry Irwin/Photo Researchers, Inc.

Rainbow over a farm in Pennsylvania.

Throughout history our knowledge, attitudes, and values have been reflected in our sciences and in our arts. The most obvious example of this parallelism is light. Light and one of its characteristics, color, have followed parallel trends in science and art. In the Middle Ages, for example, light was sacred and mysterious; artists placed it in some heavenly plane. When St. John the Evangelist said, "God is light," his words reflected the belief that the presence of light in a place of worship was a sign of the presence of the Holy Spirit. Beautiful, brightly colored leaded-glass windows became the norm in the great churches being built at the time.

Light, as seen in these early works, had a surrealistic behavior. It wasn't until the Renaissance that artists studied the effects of light on objects, how it illuminates some sides of an object but not others. During the late 1600s, when light was shown to take a finite time to travel through space, paintings were just beginning to show objects casting shadows on the ground.

Still the scientific questions remained: What is light? Does light behave as a wave or as a stream of particles? It was known that light travels from the Sun to Earth and that space is a very good vacuum. If light is a wave, how can it travel through a vacuum where there is nothing to wave? The history of this debate about the wave or particle nature of light goes back at least to Newton's time. Newton, perhaps influenced by his successful work on apples and moons, believed that light was a stream of very tiny particles. His contemporary, Dutch scientist Christiaan Huygens, believed that light was a wave. The controversy was not resolved until 100 years later. (However, the question of the nature of light reemerged during the 20th century with the development of quantum mechanics.)

And what is the origin of the colors? Prior to Newton's work with clear glass prisms, color was thought to be something emanating from objects. A ruby's redness was only due to the characteristics of the ruby. As we will see, this is partially correct; the character of the light shining on the ruby also plays a role.

The next time you look at a star, think about the starlight that reaches your eyes. It probably began its journey hundreds or thousands of years ago, in many cases long before civilization began on Earth. When it arrives at your eye, its journey is over. The light is absorbed in your retina, producing an electrical signal that travels to your brain. However, this light is carrying a much more complex message than simply the location of the star in the heavens. With the exception of a few artificial satellites within the solar system and a variety of cosmic particles striking Earth, our entire knowledge of the cosmos comes from the information carried by light.

We continue to build our physics world view by studying this mysterious phenomenon in the next three chapters; then we will use our newly gained knowledge of light to probe the structure of matter at the atomic level . . . and beyond.

A prism spreads light out into its component colors.

Bill Kamin/Visuals Unlimited

17 Light

▶When a magician thrusts a sword through the lovely assistant or spooks sit next to you as you traverse the haunted mansion's dark innards, the explanation is, "It's done with mirrors." But how does light produce these illusions that are so convincing?

(See page 370 for the answer to this question.)

John C. Muegge/Visuals Unlimited

The reflection of light at the surface of the water acts like a mirror to produce a virtual image.

ALTHOUGH the phenomenon of light is so common to our everyday experience and has played such a central role in the histories of religion, art, and science, it is actually quite elusive. Even the act of seeing has confused people. We talk of looking *at* things—of looking *into* a microscope, of sweeping our glance *around* the room—as if seeing were an active process, much like beaming something in the direction of interest. This notion goes back to some early ideas about light in which rays were supposedly emitted by the eyes and found the objects seen. The idea is still common, perhaps perpetuated in part by the Superman stories. In these stories Superman supposedly has the ability to emit powerful X rays from his eyes, enabling him to see through brick walls or, in the modern movie version, through clothes.

Seeing is actually a rather passive activity. What you see depends on the light that enters your eyes and not on some mysterious rays that leave them. The light is emitted whether or not your eyes are there to receive it. You simply point your eyes toward the object and intercept some of the light. In fact, light passing through clean air is invisible. If a flashlight is shined across a room, you don't see the light passing through the air. You only see the light that strikes the wall and bounces back into your eyes. (Sometimes you can see the beam's path because part of the light scatters from dust, fog, or smoke particles in the air and is sent toward your eyes.)

Shadows

One of the earliest studies of light was how it moved through space. By observing shadows and the positions of the light sources and the objects causing the shadows, it is easy to deduce that light travels in straight lines. In drawings illustrating the paths of light, it is convenient to use the idea of **light rays.** Because there are an infinite number of paths, we draw only enough to illustrate the general behavior.

An opaque object illuminated by a point source of light blocks some of the rays from reaching a screen behind it, producing a shadow like the one shown in Figure 17-1. The shadow has the same shape as the cross section of the object but is larger.

Most sources of light are not points but extend over some space. However, we can think of each small portion of the source as a point source casting its own sharp shadow. All these point-source shadows are superimposed on the screen behind the object. The darkest region is where all of the shadows overlap. This is known as the **umbra** (Figure 17-2). Surrounding the umbra is the **penumbra,** where only some of the individual shadows overlap.

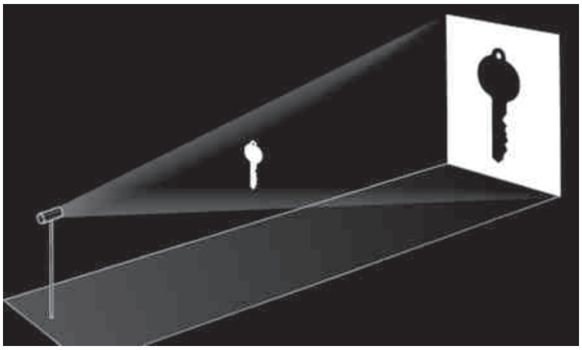


Figure 17-1 The shadow produced by a point source of light is very sharp.

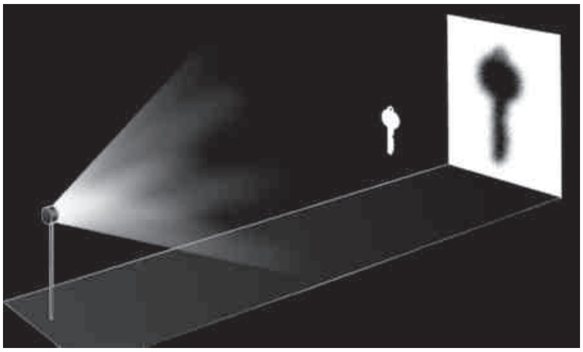


Figure 17-2 The shadow produced by an extended source of light has a dark central umbra surrounded by a lighter penumbra.

Everyday Physics | *Eclipses*

The most spectacular shadows are eclipses, especially total solar eclipses. During a solar eclipse, the Moon's shadow sweeps a path across a portion of Earth, as shown in Figure A. If you are in the path of the umbra, the Sun is totally obscured (Figure B). Observers to the side of the umbra's path but in the path of the penumbra see a partial eclipse.

One of Aristotle's arguments that Earth is a sphere involved the shadow during lunar eclipses (Figure C). Here, Earth's shadow falls onto the face of the Moon. Because (1) the shape of the shadow is always circular and (2) the only solid that always casts a circular shadow is a sphere, Aristotle correctly concluded that Earth must be a sphere.

Few people experience a solar eclipse, although the population on half of Earth can see a lunar eclipse. To see a solar eclipse, observers must be directly in the shadow of the Moon. This only covers a small portion of Earth's surface. During a lunar eclipse, we observe Earth's shadow on the Moon (Figure D). Anyone who can see the Moon can therefore see this eclipse.

1. Explain the difference between the umbra and the penumbra. Use the idea that an extended light source (such as the Sun) can be thought of as a collection of many point sources of light.
2. Why is it more likely that you will observe a total lunar eclipse than a total solar eclipse in your lifetime?
3. Your friend thinks that the phases of the Moon are caused by Earth's shadow on the Moon. What evidence could you present to convince your friend that she is mistaken?

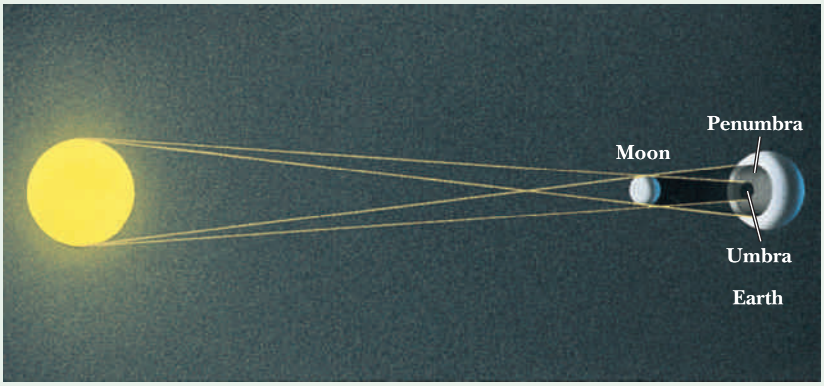


Figure A A total solar eclipse occurs when the umbra of the Moon's shadow falls on Earth.

Jim Anderson, Montana State University, Physics Department

Figure B A total eclipse of the Sun.

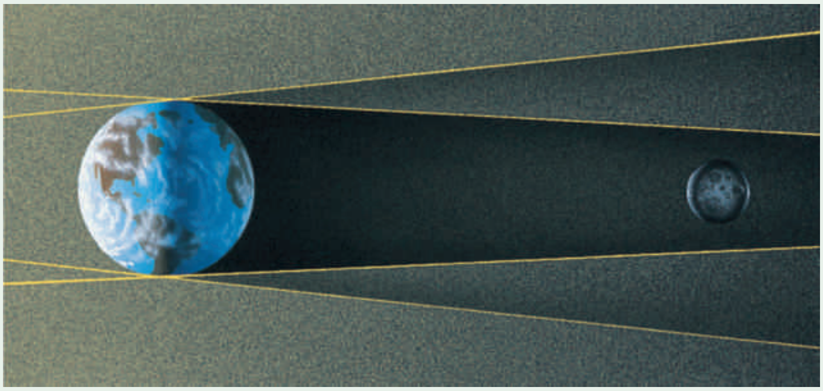


Figure C During a lunar eclipse, the Moon is in Earth's shadow.

Courtesy of Mike Murray

Figure D During a lunar eclipse, we can see part of Earth's circular shadow.

Rather than looking at the shadow, imagine standing in the shadow looking back toward the source. If your eye is located in the umbra, you will not see any portion of the source of light; the object between you and the source blocks out all of the light. If your eye is in the penumbra, you will see part of the light source.

Shadows that we observe on Earth are not totally black and devoid of light. We can, for example, see things in shadows. This is due to the light that scatters into the shadow from the atmosphere or from other objects. On the Moon, however, the shadows are much darker because there is no lunar atmosphere. Astronauts exploring the Moon's surface have to be careful. Stepping into their own shadows can be dangerous because the extreme blackness of the shadow would hide everything within it—sharp rocks, uneven terrain, even a deep hole.

Pinhole Cameras

Extended presentation available in the *Problem Solving* supplement

Light that strikes most objects leaves in a great many directions. (This must be true because we can see the object from many different viewing directions.) If we place a piece of photographic film in front of a vase of flowers, as in Figure 17-3(a), the film will be completely exposed, leaving no record of the scene. Light from one part of the vase hits the film at the same place as light from many other parts of the vase and flowers. Every spot on the film receives light from virtually every spot that faces the film.

We can get an image by controlling which light rays hit the film. A screen with a small hole in it is placed between the vase and the film, as shown in Figure 17-3(b). Now only the light from a small portion of the vase reaches a given region of the film. Making the hole smaller, as in Figure 17-3(c), further reduces the portion of the vase exposed to a given spot on the film. If the hole is made small enough, a recognizable image of the vase is formed on the film. This technique can be used to make a pinhole camera by enclosing the film in a light-tight box with a hole that can be opened and shut. The photograph in Figure 17-4 was taken with a pinhole camera made from a shoebox.

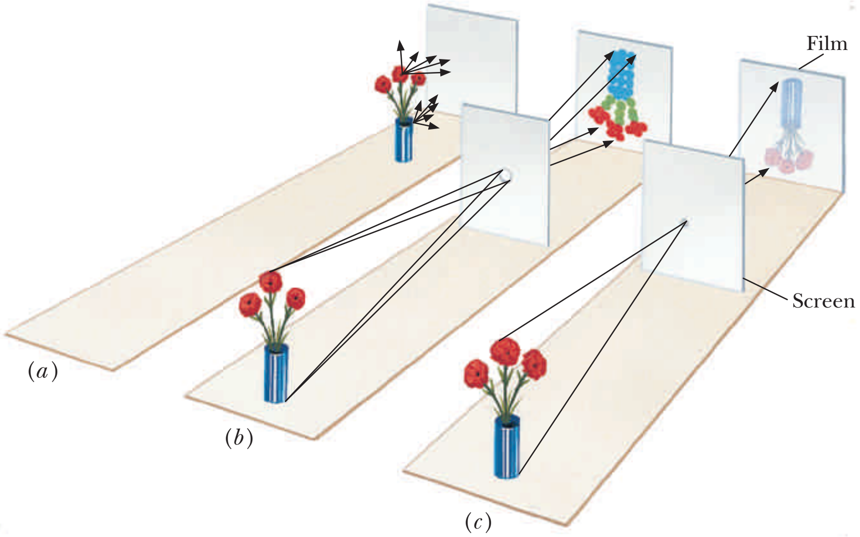


Figure 17-3 (a) Each part of the film receives light from many parts of the vase and flowers, and no image is recorded. (b) The screen restricts the light so that each part of the film receives light from a small portion of the scene. (c) Reducing the size of the hole produces a sharper but dimmer image.

Figure 17-4 A photograph made with a pinhole camera. The exposure time was 2 seconds.

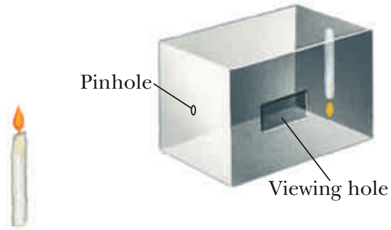


Figure 17-5 A pinhole in one end of the box produces an inverted image on the opposite end.

The problem with pinhole cameras is the very small amount of light that reaches the film. Exposure times generally must be long, requiring that the scene be relatively static. We will see in Chapter 18 that this problem was eventually overcome by using a lens.

Pinhole cameras were used before the invention of film. If you add another hole to the box, as illustrated in Figure 17-5, you can see the image on the back wall. During solar eclipses, some observers use this process to watch the partial phases of the eclipse safely. On a grander scale, an entire room can be converted into a pinhole camera known as a *camera obscura*. Renaissance artists used smaller versions to help them draw landscapes and portraits; they traced the images formed on the back wall.

Reflections

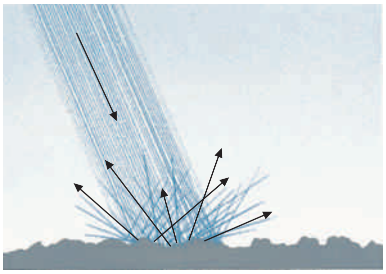


Figure 17-6 A beam of light striking a rough surface is scattered in many directions.

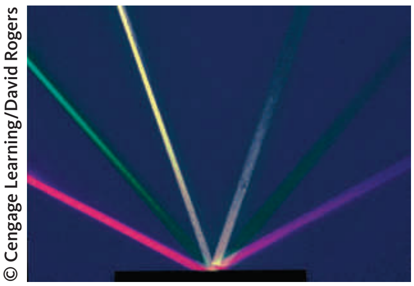
Some things emit their own light—a candle, a lightbulb, and the Sun, to mention a few. But we see most objects because they reflect some of the light that hits them. The incident light, or the light striking the object, is scattered in many directions by the relatively rough surface of the object, a process known as **diffuse reflection.** If the surface of the object is very smooth, a light beam reflects off it much as a ball rebounds from a wall. Presumably then, when light hits rougher surfaces, the same thing happens. But with rough surfaces, different portions of the incident light reflect in many directions, as shown in Figure 17-6.

By looking at the reflection of a thin beam of light (a good approximation to a "ray" of light) on a smooth surface, we can discover a rule of nature. Figure 17-7(a) shows the reflections of light beams hitting a mirror at three different angles. Clearly, the angle between the incident and the reflected ray is different in each situation. However, if we examine only one case at a time, we notice that the angles that the incident and reflected rays make with the surface of the mirror are equal. If we measure the angles of the incident ray and the reflected ray for many such situations, we discover that the two angles are always equal.

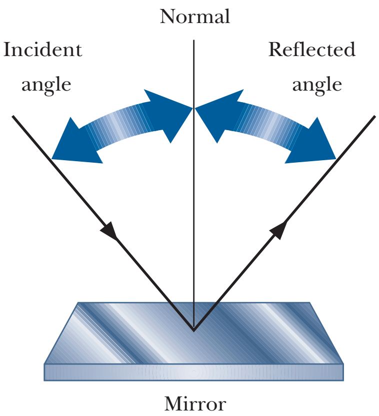
In our illustration of this phenomenon, we used the angles made with the reflecting surface. It is more convenient to consider the angles between the rays and the **normal,** an invisible line perpendicular to the surface that touches the surface where the rays hit [Figure 17-7(b)]. The reflected ray lies in the same plane as the normal and the incident ray. Using these angles we state the **law of reflection:**

law of reflection ▶

The angle of reflection is equal to the angle of incidence.



(a)



(b)

Figure 17-7 (a) Reflections of three thin beams of light hitting a mirror at different angles from the left. (b) For each one the angle of reflection is equal to the angle of incidence.

Q: Assume that you are stranded on an island. Where would you aim a mirror to signal a searching aircraft with sunlight?

A: The normal to the mirror would have to be directed at the point midway between the Sun and the aircraft. Luckily, this is easier to do than it may seem.

Flat Mirrors

When we look at smooth reflecting surfaces, we don't see light rays; we see images. We can see how images are produced by looking at the paths taken by light rays. In Figure 17-8 we locate the image Q of a single point P in front of a flat mirror. Light leaves the point P in all directions; some of the light strikes the mirror at point A, then reflects and travels in the direction of B. An eye at B would receive the light coming along the direction AB.

Remember that seeing is a passive activity. Our eye–brain system records only the direction from which the light arrives. We do not know from how far away the light originated, but we do know that it came from someplace along the line from B to C. However, we can say the same thing about the light that arrives from another direction, say, at E. The eye perceives this light as coming from someplace along the line EF. Because the only place that lies on both lines is Q, our brain says that the light originated at point Q. This is the location of the image. We see an image of P located behind the mirror at point Q. After the light reflects from the mirror, it has all of the properties it would have had if the object had actually been at Q.

The image has a definite location in space. Figure 17-8 shows that the point Q is located the same distance behind the mirror as the point P is in front of the mirror. A straight line drawn between P and Q is normal to the mirror. These observations allow us to locate the image quickly. For example, in Figure 17-9 we can locate the image of the pencil by first locating the image of its tip and then the image of its eraser. The size of the image is the same as that of the object (although it is farther away from you and looks smaller). Note that the pencil doesn't have to be directly in front of the mirror. However, the mirror must be between the entire image and the observer's eyes, as illustrated in Figure 17-10.

A magician's trick illustrates the realism of an image by presenting the audience with a live, talking head on a table, as in Figure 17-11(a). Like many other illusions, the trick lives up to the cliché that "it's all smoke and mirrors." But where is the person's body? Your eye–brain system is tricked by the images. You think that the table is an ordinary one with legs and a wall behind it. You "know" this because you can see the wall between the table's legs. Figure

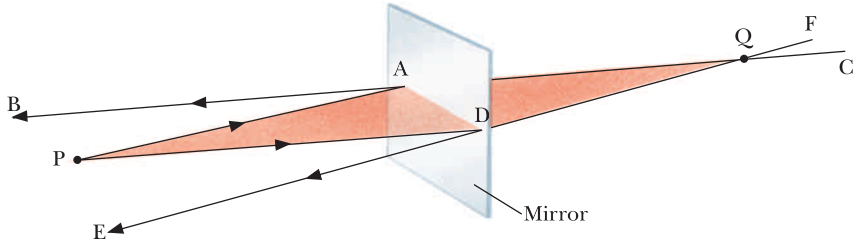


Figure 17-8 A ray diagram showing that the location of the image Q is determined by the crossing of the two lines of sight. The image Q is as far behind the mirror as the object P is in front.

Figure 17-9 The image of a pencil formed by a flat mirror is located behind the mirror.

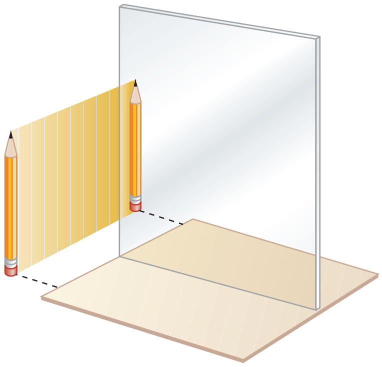


Figure 17-10 A flat mirror forms an image of an object even if the object is not located directly in front of the mirror.

Gerald F. Wheeler (both)

(a)

(b)

Figure 17-11 (a) The woman's head appears to sit on the table. How is this done? (b) The same scene with both mirrors removed.

FLAWED REASONING

Cassandra complains to her mother: "Physics is hard. I am supposed to light a match in front of a mirror and find the location of the flame's image. When I move my head to the left, the image of the flame is located on the left-hand side of the mirror. When I move my head to the right, it is located on the right-hand side of the mirror. How can I find the location of something that keeps moving?"

Help Cassandra's mother convince her that physics is easy.

ANSWER Suppose Cassandra looks at a tree through her kitchen window. By moving her head back and forth, she can align the tree with the left-hand side or the right-hand side of the window. The tree, however, is not moving. The mirror is like the kitchen window, and the flame is like the tree. The light that reflects from the mirror behaves as if it came directly from the image. The image is not on the mirror's surface; it is located behind the mirror and does not move. Cassandra should think of the mirror as a window through which she can view this "image world."

17-11(b) shows the set-up with the mirrors removed. The table has mirrors between its legs so that the walls you see under the table are really images of the side walls.

Multiple Reflections

When a light beam reflects from two or more mirrors, we get interesting new optical effects from the multiple reflections. If the two mirrors are directly opposite each other, such as in some barbershops and hair salons, we get an infinite number of images, with successive images being farther and farther away from the object. To see how this works, remember that the light appearing to come from an image has the same properties as if it actually came from an object *at the location of the image*. Each mirror forms images of everything in front of it, *including* the images formed by the other mirror. Each of the mirrors forms an image of the object. The light from these images forms an additional set of images behind the opposing mirrors, and on and on, as illustrated in Figure 17-12.

Q: Would the opposing mirrors in Figure 17-12 allow you to see the back of your head?

A: Yes, provided the mirrors are tilted a bit so that your head doesn't get in the way of your view. The second image in the mirror in front of you will be of the back of your head.

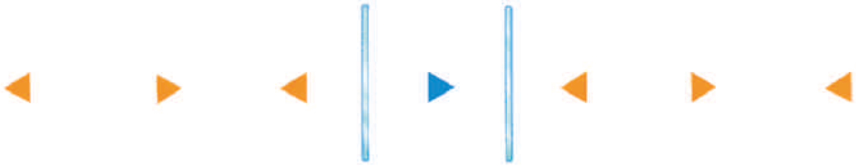


Figure 17-12 Two opposing parallel mirrors form an infinite number of images. Notice the reversals in the orientations of the triangles.

WORKING IT OUT Full-Length Mirror

Jonathan needs to buy a "full-length" mirror to mount on his wall. He is extremely thrifty (cheap) and wants to buy the shortest mirror that will allow him to see a reflection of his entire body. If Jonathan is 6 feet tall, how short can the mirror be and how high should it be mounted on the wall? Does your answer depend on how far from the mirror Jonathan will be standing?

If Jonathan stands some distance d in front of the mirror, then an image of Jonathan, exactly the same size, will be formed a distance d behind the mirror, as shown in Figure 17-12. Jonathan can see his toes "in the mirror" because light travels from his toes, bounces off the mirror, and enters his eyes. This light appears to travel in a straight line path from the toes of the image of Jonathan behind the mirror. The mirror is mounted on the wall, halfway between Jonathan and his image, so the light coming from his toes will bounce off the mirror at a height halfway between the floor and his eyes. A similar argument can be used to claim that the light from the top of his head bounces off the mirror at a height halfway between his eyes and the top of his head. The shortest "full-length" mirror that would do the job would be 3 feet tall, and should be mounted on the wall with the top of the mirror at the same height as Jonathan's forehead (halfway between his eyes and the top of his head). Our reasoning is independent of the distance d, so it does not matter how far from the mirror Jonathan will be standing.

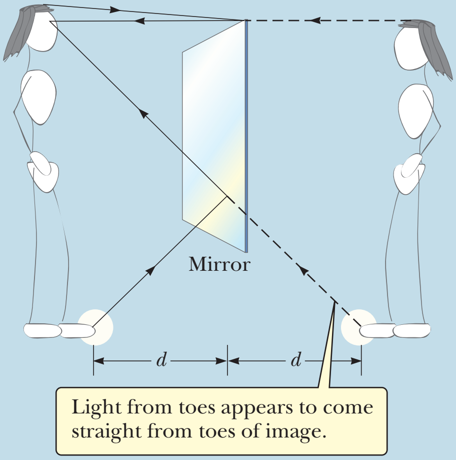
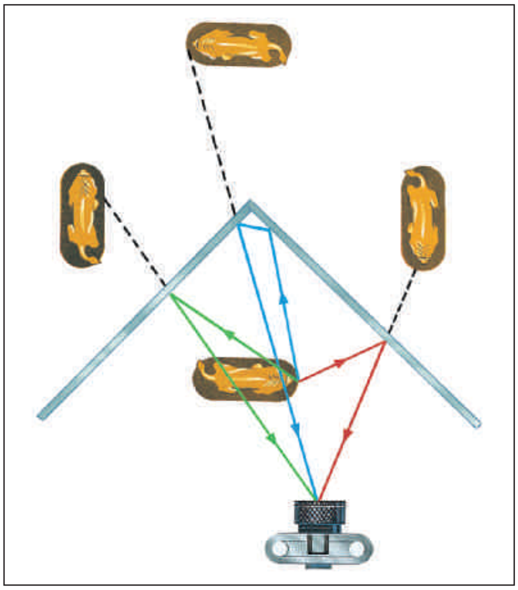


Figure 17-13 A "full-length" mirror need only be half the height of the person using it.

(a)



(b)

Figure 17-14 (a) Two mirrors at right angles to each other produce three images of the lion. (b) The red, green, and blue lines show sample paths taken by rays that enter the camera lens.

In Figure 17-14(a) we have placed a figurine of a lion in front of two mirrors that form a right angle. Figure 17-14(b) shows the paths taken by the rays that entered the camera. Notice that there are three images; each mirror forms

Figure 17-15 When the mirrors form an angle of 60 degrees, five images of the lion are produced.

one image, and then each mirror forms an image of these images. (Remember that a mirror does not have to actually extend between the image and the object. However, it is often useful to imagine that each mirror is extended.) If the angle between the mirrors is precisely 90 degrees, these latter two images overlap to form a single image beyond the corner.

If the angle between the mirrors is made smaller, the overlapping images beyond the corner separate. When the angle between the mirrors reaches 60 degrees, we once again get overlapping images beyond the corner and a total of five images, as shown in Figure 17-15.

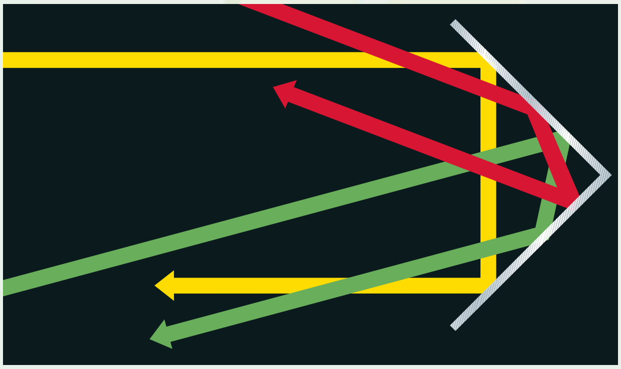
Curved Mirrors

Fun-house mirrors are interesting because of the distortions they produce. The distortions are not caused by a failure of the law of reflection but result from the curvature of the mirrors. Some distortions are desirable. If the distortion is a magnification of the object, we can see more detail by looking at the image.

Cosmetics mirrors and some rearview mirrors on cars are simple curved mirrors that don't produce bizarre distortions but do change the image size. A cosmetics mirror uses the concave side—the reflecting surface is on the inside of the sphere—to generate a magnified image of your face. The con-

Everyday Physics | *Retroreflectors*

An interesting consequence of having two mirrors at right angles is that an incoming ray (in a plane perpendicular to both mirrors) is reflected back parallel to itself, as shown by the three rays in the figure. This works for all rays if we add a third mirror to form a "corner" reflector, such as by putting mirrors on the ceiling and the two walls in the corner of a room. These *retroreflectors* are used in the construction of bicycle reflectors so that the light from a car's headlights is reflected back to the car driver and not off to the side as it would be with a single mirror. Examination of many reflectors reveals a surface covered with holes in the shape of the corners of cubes. Reflectors used on clothing and the surfaces of stop signs are often covered with a layer of reflective beads. The surfaces in the regions between the beads work like corner reflectors.



Each of the three rays is reflected back parallel to itself.

An outer-space application of retroreflectors involves an experiment to accurately measure the distance to the Moon. The Apollo astronauts placed panels of retroreflectors on the Moon to allow scientists on Earth to bounce a laser beam off the Moon and receive the reflected signal back on Earth. Ordinary mirrors would not have worked because the astronauts could not have aimed them well enough to send the beam back to Earth. Additionally, the Moon's wobbly rotation would send the reflected beam in many directions. Because retroreflectors work for all incident angles, they could be simply laid on the Moon's surface. We now know that the Earth–Moon distance increases by 3 to 4 centimeters per year.

1. Explain in detail the process used to measure the distance between Earth and the Moon.
2. Is the orbital speed of the Moon increasing, decreasing, or staying the same? Explain your answer using the concept of conservation of angular momentum, discussed in Chapter 8.

Are You On the Bus?

Q: Sometimes mirrors are installed in stores to inhibit shoplifting. Are these concave or convex mirrors?

A: The mirrors must be convex to provide a wide view of the store.

vex reflecting surface—the outside of the sphere—always produces a smaller image but has a bigger field of view. Convex mirrors are quite often used on cars and trucks and on "blind" street corners because they provide a wide-angle view.

Figure 17-16 shows the essential geometry for a concave spherical mirror; the reflecting surface is the inside of a portion of a sphere. The line passing through the center of the sphere C and the center of the mirror M is known as the **optic axis.** Light rays parallel to the optic axis are reflected back through a common point F called the **focal point.** The focal point is located halfway between the mirror and the center of the sphere. The distance from the mirror to the focal point is known as the **focal length** and is equal to one-half the radius *R* of the sphere.

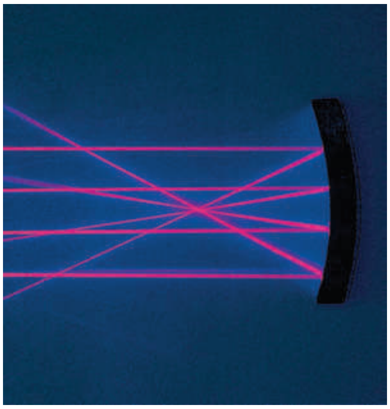
Concave mirrors can be used to focus light. Some solar collectors use them to concentrate sunlight from a large area onto a smaller heating element. Because the Sun is very far away, its rays are essentially parallel, and a concave spherical mirror can focus the sunlight at the focal point. A cylindrical mirror focuses sunlight to a line instead of a point. A pipe containing a fluid can be placed along this line to carry away the thermal energy. If the heat energy is used to generate electricity, the higher temperature makes the process more efficient.

Light rays are reversible. The law of reflection is still valid when the incident and reflected rays are reversed. So the shape that focuses parallel rays to a point will take rays from that point and send them out as parallel rays. This idea is used in automobile headlights. The bulb is placed near the focal point of the mirror, producing a nearly parallel beam.

The optic axis of a convex mirror also passes through the center of the sphere C, the focal point F, and the center of the mirror M, but in this case, the focal point and the center of the sphere are on the back side of the mirror. (Again, the focal point is halfway between the center of the sphere and the center of the mirror.) Rays parallel to the optic axis are now reflected *as if they came from* the focal point behind the mirror, as shown in Figure 17-17.

The curved surfaces of fun-house mirrors produce interesting images.

Convex mirrors allow clerks to watch for potential shoplifters.



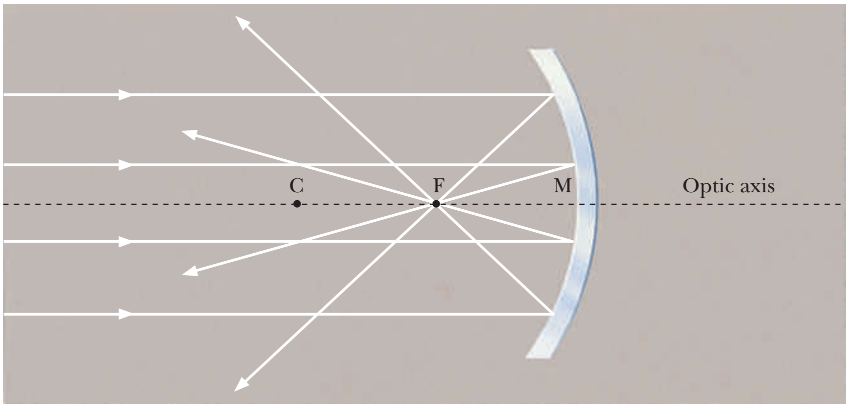
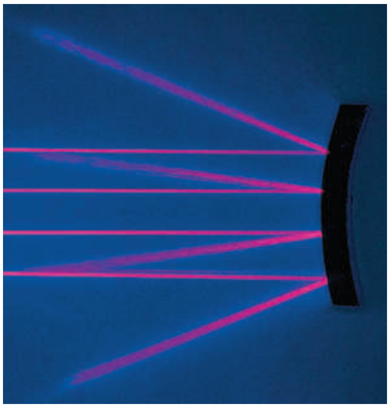


Figure 17-16 The focal point F of a spherical concave mirror lies along the optic axis midway between the center of the sphere C and the center of the mirror M. Rays parallel to the optic axis are focused at the focal point F.

Cylindrical mirrors are used to concentrate sunlight at this solar farm.

Mark E. Gibson/Visuals Unlimited



© Cengage Learning/David Rogers

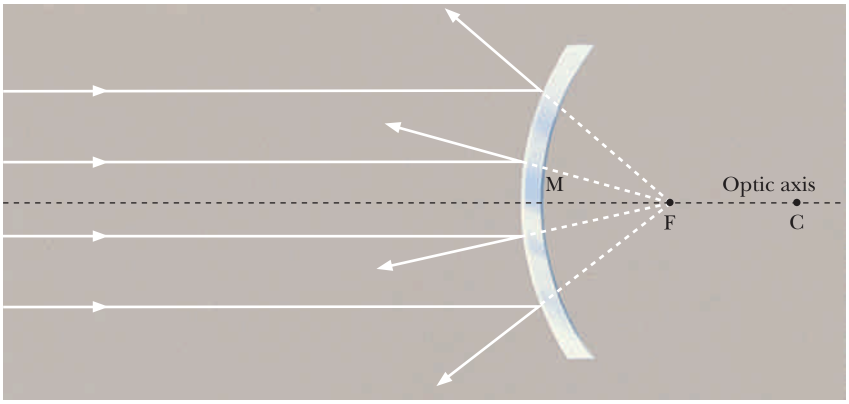


Figure 17-17 The focal point F of a spherical convex mirror lies along the optic axis midway between the center of the sphere C and the center of the mirror M. Rays parallel to the optic axis are reflected as if they come from the focal point F.

© Cengage Learning/David Rogers

Figure 17-18 An inverted image is formed when the head is located outside the focal point of a concave mirror. Note that both the face and the floor tiles are inverted.

Images Produced by Mirrors

A concave mirror can form two different types of image, depending on how close the object is to the mirror. Imagine walking toward a large concave mirror. At a large distance from the mirror, you will see an image of your face that is inverted and reduced in size (Figure 17-18). This image is formed by light from your face reflecting from the mirror and converging to form an image in front of the mirror. This image is known as a **real image,** because the rays reflect from the mirror and converge to form the image. The rays then diverge, behaving as if your face were actually at the image location.

As you walk toward the mirror, the image of your face moves toward you and gets bigger. As you pass the center point of the mirror, the image moves behind you. When your face is closer to the mirror than the focal point, the image is similar to that in a flat mirror except that it is magnified (Figure 17-19). As with the flat mirror, the rays diverge after reflecting, and the image is located behind the mirror's surface. In this case there can be no light at the location of the image because it is formed behind the mirror. The light only appears to come from the location of the image. This second type of image is called a **virtual image.** As you continue approaching the mirror, the image moves closer and gets smaller.

The essential difference between a real and a virtual image is whether the light actually comes from the image location or only appears to come from there. If the rays diverge upon reflecting, they never come together to form a real image. They will, however, appear to originate from a common location behind the mirror. Reflected rays that converge to form a real image can be seen on a piece of paper placed at the image location because the light actually converges at that location. However, if you put a piece of paper at the location of a virtual image, you get nothing because there is no light there.

Figure 17-19 A magnified, erect image is formed when the head is located inside the focal point of a concave mirror.

Locating the Images

A simple way of locating an image without measuring any angles is by looking at a few special rays. Light leaves each point on the object in all directions; rays that strike the mirror form an image. Although any of these rays can be used to locate the image of a point, three are easy to draw and are therefore useful in locating the image. Because all rays from a given point on the object are focused at the same place for a real image (or appear to come from the same point for a virtual image), we need only find the intersection of any two of them. The third one can be drawn as a check. (In actual drawings these three rays do not always meet at a point. However, they give a pretty good location for the image if the object is small enough that the special rays strike the mirror near the optic axis.)

The three rays that are useful in these ray diagrams are shown in Figure 17-20. The easiest ray to draw is the red one lying along a radius of the sphere. It strikes the mirror normal to the surface and reflects back on itself. Another easy ray is the blue one that approaches the mirror parallel to the optic axis. It is reflected back toward the focal point. The third ray (shown in green) is a reverse version of the second one; a ray passing through the focal point reflects back parallel to the optic axis. (This ray does not actually need to pass through the focal point. If the object is closer than the focal point, the ray still lies along the line from the focal point to the mirror.)

If the mirror is small, some of these special rays may not strike the actual surface of the mirror. For the purposes of the ray diagram, we extend the mirror because a larger mirror with the same focal length would produce the same image. In fact, the mirror could be so small that none of the three easily drawn rays hits the mirror.

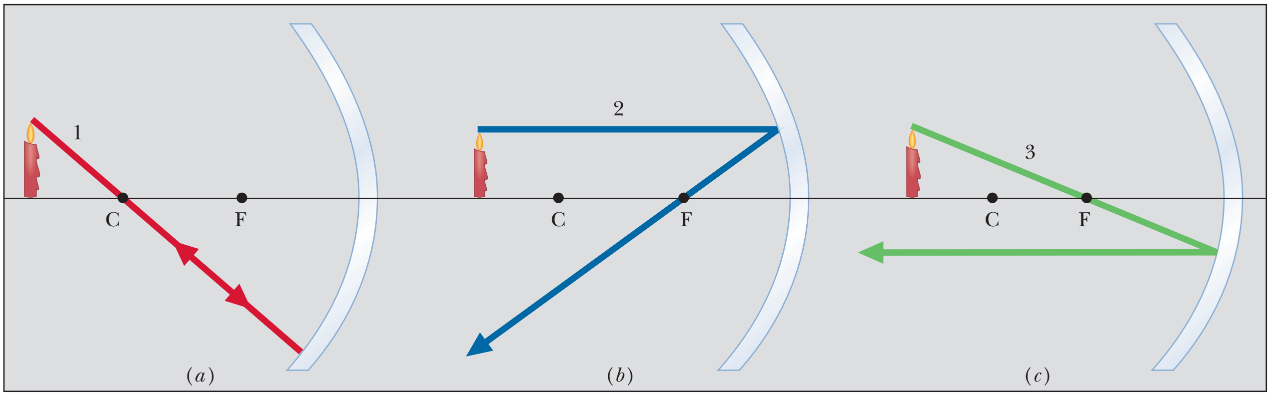


Figure 17-20 The three light rays used in drawing ray diagrams for concave mirrors.

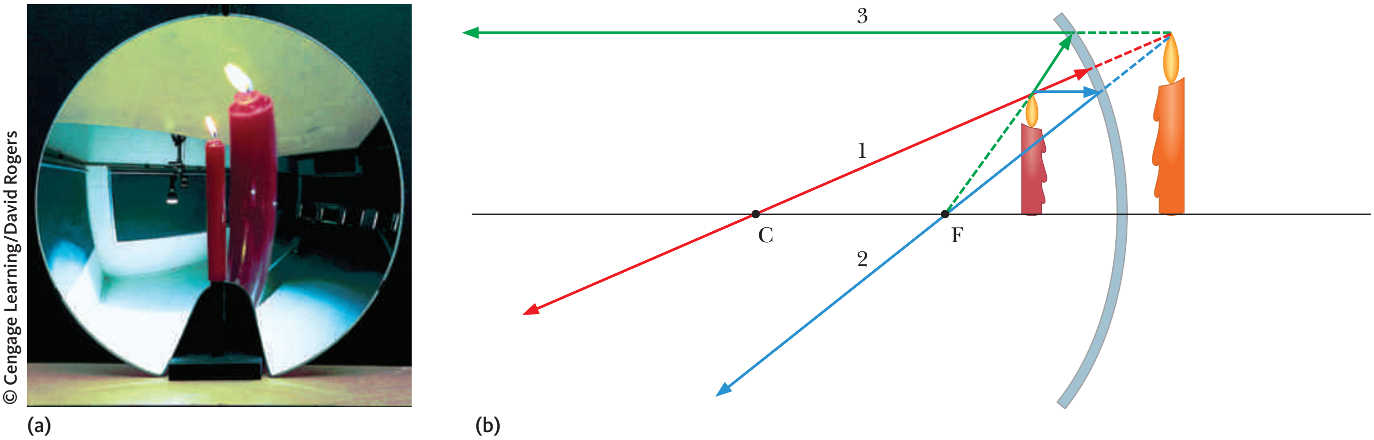


Figure 17-21 The image of a candle inside the focal point of a concave spherical mirror is virtual, erect, and magnified.

The descriptions of these rays can be abbreviated as follows:

rays for curved mirrors ▶

1. Along radius—back on itself
2. Parallel to optic axis—through focal point
3. Through focal point—parallel to optic axis

To illustrate the use of these ray diagrams, consider an object located inside the focal point of a concave mirror [Figure 17-21(a)]. Figure 17-21(b) shows the three special rays that we use to locate the image of the tip of the candle. The rays are color-coded to correspond to the descriptions given earlier. Because the base of the candle is on the optic axis, we know that the image of the base is also on the optic axis. So finding the location of the tip of the candle gives the image location, orientation, and magnification. For the case illustrated in Figure 17-21, we can see that the rays intersect behind the mirror, forming a virtual image that is erect and magnified.

As the candle is moved away from the mirror, the image size and the distance of the image behind the mirror increase. As the candle approaches the focal point, the image becomes infinitely large and infinitely far away. You can verify this by drawing a ray diagram.

When the candle is beyond the focal point, a real image is formed [Figure 17-22(a)]. The ray diagram in Figure 17-22(b) shows that the reflected rays do not diverge as in the previous case but come together, or converge. These rays actually cross at some point in front of the mirror to form a real image.

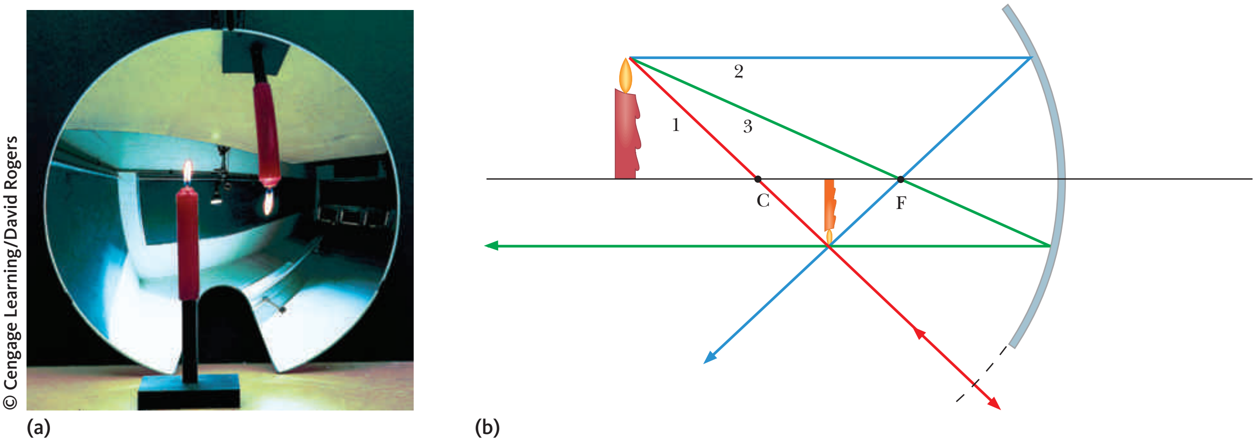
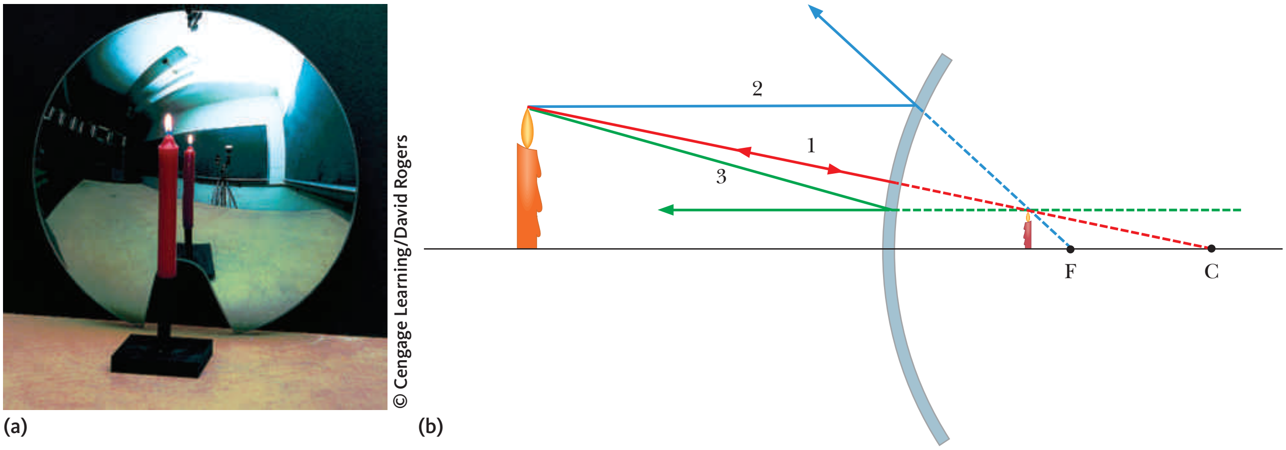


Figure 17-22 The image of a candle outside the focal point of a concave spherical mirror is real and inverted and may be larger or smaller than the object.

Q: Where would you place an object to get an image at the same location?

A: At the center of the sphere. This can be checked with a ray diagram.



© Cengage Learning/David Rogers

(a) (b)

Figure 17-23 The image of a candle in front of a convex spherical mirror is always virtual, erect, and reduced in size. Ray 2 reflects as if it came from the focal point, and ray 3 starts toward the focal point.

Fortunately, locating images formed by convex mirrors is the same process used for images formed by concave ones. A ray diagram showing how to locate the image is given in Figure 17-23(b). The same three rays are used. You must only remember that the focal point is now behind the mirror. With ray diagrams you can verify that images formed by convex mirrors are always erect, virtual, and reduced in size.

Speed of Light

Besides traveling in a straight line in a vacuum, light moves at a very high speed. In fact, people originally thought that the speed of light was infinite—that it took no time to travel from one place to another. It was clear to these observers that light travels much faster than sound; lightning striking a distant mountain is seen long before the thunder is heard.

Because of light's great speed, early attempts to measure the speed of light failed. Galileo made one interesting attempt when he tried to measure the speed of light by sending a light signal to an assistant on a nearby mountain. The assistant was instructed to uncover a lantern upon receiving Galileo's signal. Galileo measured the time that elapsed between sending his signal and receiving that of his assistant. Knowing the distance to the mountain, he was able to calculate the speed of light. Upon repeating the experiment with a more distant mountain, however, he found the same elapsed time! Had the speed of light increased? No. Galileo correctly concluded that the elapsed time was due to his assistant's reaction time. Therefore, the time it took light to travel the distance was either zero or much smaller than he was able to measure.

About 40 years later, in 1675, Danish astronomer Ole Roemer made observations of the moons of Jupiter that showed that light had a finite speed. Roemer found that the period of revolution of a moon around Jupiter was shorter during the part of the year when Earth approached Jupiter and longer when Earth receded from Jupiter. He expected the period to be constant like

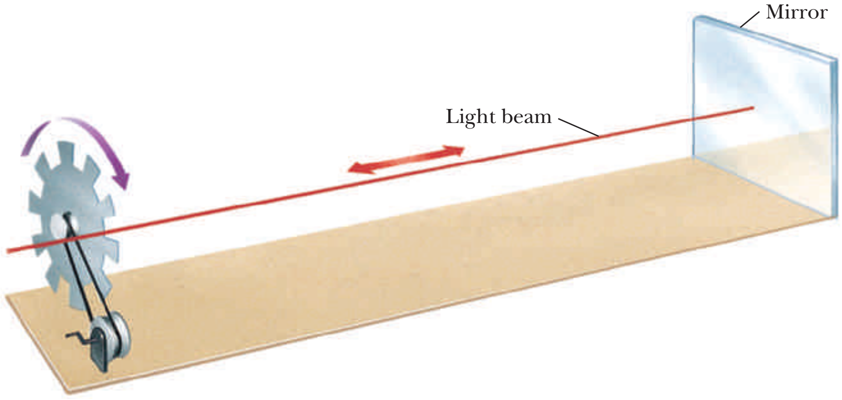


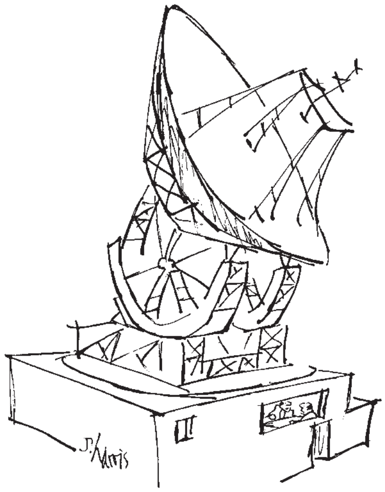
Figure 17-24 A schematic drawing of Fizeau's apparatus for measuring the speed of light.

that of our Moon about Earth. He correctly concluded that the period of the Jovian moon was constant and that the variations observed on Earth were due to the varying distance between Earth and Jupiter. When Earth approaches Jupiter, the light emitted at the beginning of the period must travel farther than that emitted at the end of the period. Therefore, the light emitted at the end of the period arrives sooner than it would if Earth and Jupiter remained the same distance apart. This difference in distance makes the period appear shorter. Once the radius of Earth's orbit was determined, the speed of light could be calculated.

speed of light ▶

French physicist Hippolyte Fizeau performed the first nonastronomical measurement of the speed of light in 1849. He sent light through the gaps in the teeth of a rotating gear to a distant mirror. The mirror was oriented to send the light directly back (Figure 17-24). At moderate speeds of rotation, the returning light would strike a tooth. But at a certain rotational speed, the light would pass through the next gap. Knowing the speed of the gear and the distance to the mirror, Fizeau was able to calculate the speed of light to a reasonable accuracy.

The speed of light has been measured many times. As the methods improved, the uncertainty in its value became less than 1 meter per second. In 1983 an international commission set the speed of light in a vacuum to exactly 299,792,458 meters per second and used it with atomic clocks to define the length of the meter. The speed of light is usually rounded off to 3×10^8 meters per second (186,000 miles per second). If light traveled in circles, it could go around Earth's equator 7.5 times in 1 second. It's no wonder that early thinkers thought light traveled at an infinite speed. Although the speed of light is finite, it can be considered infinite for many everyday experiences.



"As I understand it, they want an immediate answer. Only trouble is, the message was sent out 3 million years ago."

Are You On the Bus?

Q: Given that radio waves travel at the speed of light, would you expect there to be longer than normal pauses in a conversation between two people talking via radio from the two U.S. coasts? How about in a conversation with astronauts on the Moon?

A: Because light can travel around the world 7.5 times in 1 second, we would expect it to travel across the United States and back in a small part (about $\frac{1}{30}$) of a second. Therefore, the pauses would seem normal. The round-trip time for the signal to the Moon is a little less than 3 seconds; this would produce noticeable pauses.

Color

Color is all around us and adds beauty to our lives. Color and the way we perceive it is an area of research involving many disciplines, including physics, chemistry, physiology, and psychology. One example of our body's role in detecting color is the fact that there is no such thing as white light. What we perceive as white is really the summation of many different colors reaching our eyes. The color of an object is determined by the color, or colors, of the light that enters our eyes and the way that this is interpreted by our brains.

Besides the additive effect within our brain, there is also the issue of what reaches our eyes. If you have ever tried to match the color of two pieces of clothing under different kinds of lighting, you know it can be difficult. They may match very well in the store but be quite different in sunlight. The colors we perceive are determined by two factors: the color present in the illuminating light and the colors reflected by the object. A red sweater is red because the pigments in the dye absorb all colors except red. When viewed under white light, the sweater looks red. If the illuminating light does not contain red, the sweater will appear black because all the light is absorbed. A brightly colored box and dice look different under different illuminating lights (Figure 17-25).

Although most lights give off all colors, the colors do not have the same relative intensities found in sunlight. Fluorescent lights are often brighter in the blue region and therefore highlight blues. At the risk of taking some of the romance out of candlelit dinners, we note that your date's warm glow is due to the yellow-red light from the candles and may have nothing to do with your date's feelings toward you.

Often an object that appears to be a single color reflects several different colors. Although different colors may enter our eyes, we do not see each of these colors. Our eye–brain systems process the information, and we perceive a single color sensation at each location. The color perceived may appear to have nothing in common with the component colors. For instance, if an object reflects red and green, it will appear yellow. This behavior is in sharp contrast with the sense of hearing. Our ear–brain combination can hear and distinguish many different sounds coming from the same place at the same time.

Placing colored filters in front of spotlights or slide projectors and allowing the colored beams from each one to overlap demonstrates the additive effects of color. One combination of light beams that produces most of the colors that we perceive is red, green, and blue. Figure 17-26 illustrates the colors that are seen on the screen. For instance, red and green yield yellow, blue and green yield cyan (a bluish green), and red and blue yield magenta (a reddish purple). All three colors together produce white!

Two colors that produce white light when added together are called **complementary colors.** This process is illustrated by positive and negative color

(a)

(b)

Figure 17-25 (a) A brightly colored box and dice illuminated with yellow light. (b) The same box and dice illuminated with white light.

Q: What color would you expect to see if you remove some of the blue light from white light?

A: This leaves red and green behind, creating a yellow color.

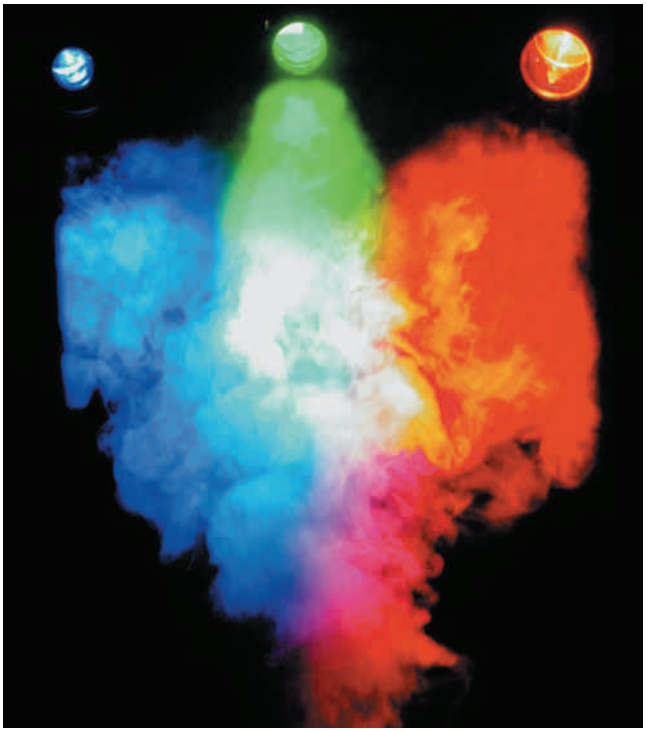


Figure 17-26 The overlap of colored lights produces new colors.

Figure 17-27 The colors in the positive image on the film are the complements of those in the negative image.

film of the same scene; the colors on one are the complements of those on the other, as illustrated in Figure 17-27.

By using dimmer switches to vary the brightness of each beam, we can generate a wide range of colors. This process is the basis of color television. Exam-

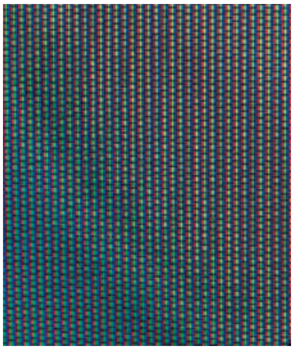


Figure 17-28 The array of colors on your television comes from the addition of three basic colors—red, green, and blue.

ine a color TV screen with a magnifying glass and you will see that it is covered with arrays of red, green, and blue dots or lines like those in Figure 17-28.

It may seem strange to find that mixing certain colors yields white; from childhood we have learned that mixing many different-colored paints together does not yield white, but rather a dark brownish color. Mixing paints is different from mixing colored lights. Mixing paints is a *subtractive* process, whereas with light beams you are *adding* colors. When white light strikes a red object, the pigment in the object subtracts all colors except red and reflects the red back to the viewer. Likewise, a red filter subtracts all colors except the red that passes through it. Each additional color pigment or filter subtracts out more colors from the incident light. This suggests that the primary colors of light—red, green, and blue—are not the best colors to use as primary colors of paint. Red paint absorbs both green and blue. Magenta paint, on the other hand, reflects both blue and red, absorbing only one primary color of light, green. Likewise, cyan paint absorbs only red and reflects both blue and green, and yellow paint absorbs only blue light and reflects both red and green. Indeed, most color printing is done with four colors of ink—cyan, magenta, yellow, and black. (Although black could be created by combining the other three colors, using a separate black improves image quality.)

FLAWED REASONING

What is wrong with the following statement? "If I take a green banana into a dark room and shine red light on it, it should appear yellow—that is, ripe."

ANSWER It is true that overlapping green and red lights produce yellow. However, green bananas are not a source of green light. Green bananas absorb all colors of light except green, so they appear green under white light. However, if we shine only red light on the bananas, they will absorb this red light and appear black (overripe).

This brief coverage of color perception allows us to answer the questions, Why is the sky blue? and, Why is the Sun yellow? The Sun radiates light that is essentially white. Because it appears yellow, we can assume that some of the complementary color has somehow been removed. The complement of yellow is blue—the color of the sky. The molecules in the atmosphere are more effective in scattering blue light than red light. As the sunlight passes through the atmosphere, more and more of the blue end of the spectrum is removed, leaving the transmitted light with a yellowish color (Figure 17-29). When we look away from the Sun, the sky has a bluish cast because more of the blue light is scattered into our eyes. This effect is enhanced when the rays have a

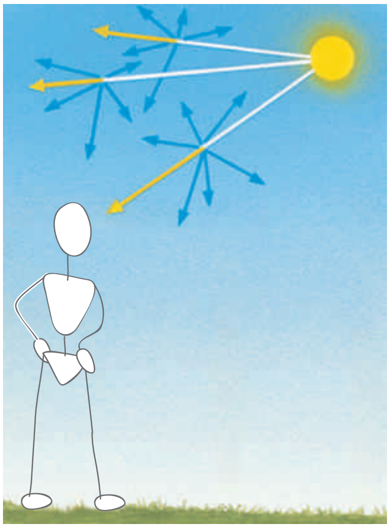


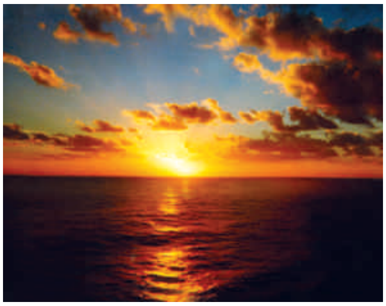
Figure 17-29 The sky is blue and the Sun yellow because air scatters more blue light than red light.

longer path through the atmosphere: the Sun turns redder near sunrise and sunset. The redness also increases with increased numbers of particles in the air (such as dust). The additional dust in the air during harvest time produces the spectacular harvest moons. Although much less romantic, the same effect results from the air pollution near urban industrial sites.

These ideas also account for the color of water. Because water absorbs red light more than the other colors, the water takes on the color that is complementary to red—that is, cyan. Consequently, underwater photographs taken without artificial lighting look bluish green.

Q: If red light were scattered more than blue light, what color would the Sun and sky appear?

A: The sky would appear red because of the scattered light, and the Sun would appear cyan because of the removal of more of the red end of the spectrum.



Courtesy of American Association of Physics Teachers. Photo by Rhian E. Vanderburg.

The sun appears redder as it sets because the sunlight passes through more of the atmosphere.

Summary

Light is seen only when it enters our eyes. Because we know that light travels in straight lines, we can use rays to understand the formation of shadows and images.

When light reflects from smooth surfaces, it obeys the law of reflection, which states that the angles the incident and reflected rays make with the normal to the surface are equal. The reflected ray lies in the same plane as the normal and the incident ray.

Mirrors produce real and virtual images. Light converges to form real images that can be projected, whereas light only appears to come from virtual images. The virtual image formed by a flat mirror is located on the normal to the mirror that passes through the object. The image is located the same distance behind the mirror as the object is in front of the mirror. The sizes of the image and the object are the same.

Images formed by spherical mirrors can be located by drawing three special rays: (1) along the radius—back on itself; (2) parallel to the optic axis—through the focal point; and (3) through the focal point—parallel to the optic axis. The focal point is located halfway between the surface and the center of the sphere.

Light travels through a vacuum at 299,792,458 meters per second.

The additive effects of color mean that adding red and green lights yields yellow, blue and green lights yield cyan, and red and blue lights yield magenta. All three colors produce white. Mixing paints is a subtractive process, whereas mixing light beams is an additive process. The sky is blue and the Sun is yellow because the molecules in the atmosphere preferentially scatter blue light, leaving its complement.

CHAPTER 17 *Revisited*

When light reflects from a smooth mirror or clean piece of glass, its direction is changed, giving the viewer false information about the location of the source of the light. For instance, reflections of the walls of a box make the box appear to be empty when, in fact, a rabbit is hiding behind the mirrors. This is the basis of many visual illusions.

Key Terms

complementary color For lights, two colors that combine to form white.

diffuse reflection The reflection of rays from a rough surface. The reflected rays do not leave at fixed angles.

focal length The distance from a mirror to its focal point.

focal point The location at which a mirror focuses rays parallel to the optic axis or from which such rays appear to diverge.

law of reflection The angle of reflection (measured relative to the normal to the surface) is equal to the angle of incidence. The incident ray, the reflected ray, and the normal all lie in the same plane.

light ray A line that represents the path of light in a given direction.

normal A line perpendicular to a surface or curve.

optic axis A line passing through the center of a curved mirror and the center of the sphere from which the mirror is made.

penumbra The transition region between the darkest shadow and full brightness. Only part of the light from the source reaches this region.

real image An image formed by the convergence of light.

umbra The darkest part of a shadow where no light from the source reaches.

virtual image The image formed when light only appears to come from the location of the image.

Questions and exercises are paired so that most odd-numbered are followed by a similar even-numbered.

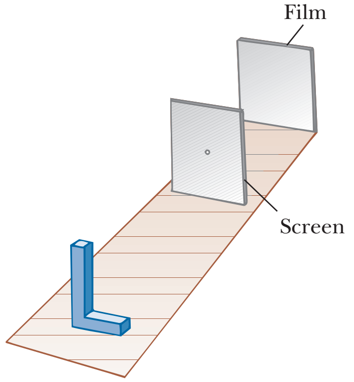
Blue-numbered questions and exercises are answered in Appendix B.

🌶 indicates more challenging questions and exercises.

WebAssign Many Conceptual Questions and Exercises for this chapter may be assigned online at WebAssign.

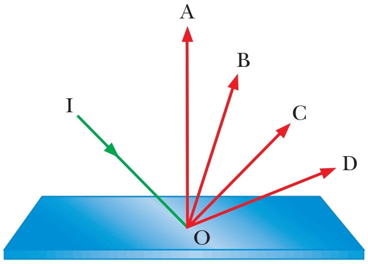
Conceptual Questions

1. A professor shines a light beam across the front of a lecture hall. Why can you see the light on the wall but not in the air?
2. If you shine a laser pointer across the room, you see only a red spot on the far wall. However, you can see the path of the beam if you create a dust cloud with a pair of chalk erasers. Explain how the cloud allows you to see the beam.
3. You place a plane mirror flat on the floor ahead of you and shine a laser beam toward its center. Why do you not see a red dot on the face of the mirror?
4. If you spread a fine layer of dust on the mirror in Question 3, you suddenly see a red dot on the face of the mirror. Why does this happen?
5. Which of the following will cast a shadow that has an umbra but no penumbra: the Sun, a lightbulb, a campfire, or a point source of light? Explain.
6. You hold your hand 3 feet above the ground and look at the shadow cast by the Sun. You repeat this inside using the light from a 2-foot by 4-foot fluorescent light box in the ceiling. In which case will the penumbra be more pronounced? Which of these two sources is acting more like a point source? Explain.
7. You are in a dark room with a single incandescent 60-watt bulb in the center of the ceiling. You hold a book directly beneath the bulb and begin lowering it toward the floor. As the book is lowered, what happens to the size of the umbra?
8. Repeat Question 7 using a 2-foot by 4-foot fluorescent light box in place of the incandescent bulb.
9. 🌶 Under what conditions will the shadow of a ball on a screen not have an umbra? What does this have to do with the observation that some solar eclipses are not total for any observer on Earth?
10. 🌶 During some solar eclipses, the angular size of the Moon is smaller than that of the Sun. What would observers on Earth see if they stood directly in line with the Sun and Moon?
11. What effect does enlarging the hole in a pinhole camera have on the image?
12. What happens to the image produced by a pinhole camera when you move the back wall of the camera closer to the pinhole?
13. You form the letter L using two fluorescent light tubes in front of the opening of a pinhole camera as shown in the following figure. Sketch the image formed on the film.

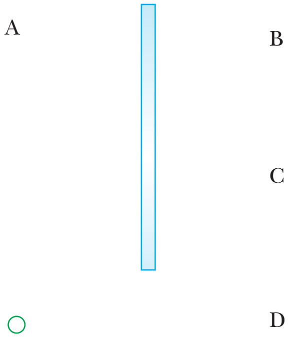


14. You are sitting under a large maple tree on a sunny day. You notice that the light filtering through the leaves does not have the sharp lines you would expect from maple leaves. Instead, the pattern consists of many small round circles of light. Use the concept of a pinhole camera to explain this.

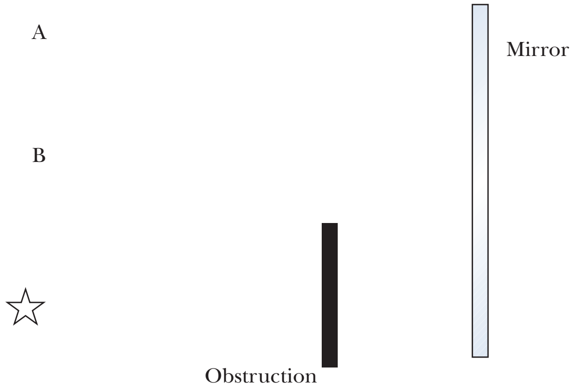
15. When the incident ray IO reflects from the mirror in the following figure, the reflected ray lies along the line O______.



16. Which letter corresponds to the location of the image of the object O in the following figure?



17. In the following figure, observers at locations A and B are attempting to see the image of the star in the plane mirror. An obstruction is placed in front of the mirror as shown. Which observers, if either, can see the star's image?

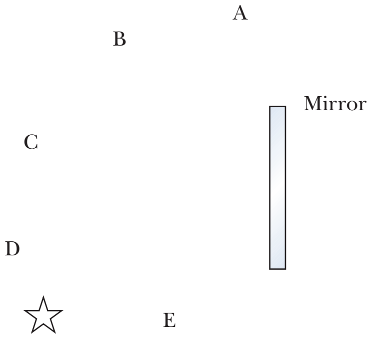


18. In talking about Question 17, Miguel claims, "The star still has an image directly behind the mirror because the light doesn't have to travel from the object to the image." Antonio counters, "The obstruction prevents light from going directly to the mirror, so no image can be formed." Jocelyn argues, "There has to be an image, it just can't be directly behind the obstruction." Which student(s), if any, do you agree with?

19. If a 0.7-meter-tall child stands 0.6 meter in front of a vertical plane mirror, how tall will the image of the child be?

20. How do the size and location of your image change as you walk away from a flat mirror?

21. At which of the lettered locations in the following figure would an observer be able to see the image of the star in the mirror?



22. Where is the image of the arrow shown in the following figure located? Shade in the region where an observer could see the entire image.

23. What is the magnification of a flat mirror? What is its focal length?

24. If you walk toward a flat mirror at a speed of 1.2 meters per second, at what speed do you see your image moving toward you?

25. How does the height of the shortest mirror in which a woman can see her entire body compare with her height? Does your answer depend on how far she stands from the mirror?

26. The word *AMBULANCE* is often written backward on the front of the vehicle so that it can be read correctly in a rearview mirror. Why do we have to switch the left and right but not the up and down?

27. You are standing in a room that has large plane mirrors on opposite walls. Why do the images produced appear to get progressively smaller? Are these images real or virtual?

28. How many images would be formed by two mirrors that form an angle of 45 degrees?

29. Why are the back surfaces of automobile headlights curved?

30. If rays of light parallel to the optic axis converge to a point after leaving the mirror, what kind of mirror is it?

31. Most of us find that we really have to strain our eyes to focus on objects located close to our noses. You hold two mirrors 1 foot in front of your face. One is a plane mirror, and the other is a concave mirror with a 3-inch focal length. In which case are you more likely to have to strain your eyes to see the image of your nose?

32. Can the image produced by a convex mirror ever be larger than the object? Explain.

33. What type of mirror would you use to produce a magnified image of your face?

34. The image produced by a convex mirror is always closer to the mirror than the object. Then why do the convex mirrors used on cars and trucks often have the warning "Caution: Objects Are Closer Than They Appear" printed on them?

35. What are the size and location of the image of your face when your face is very close to a concave mirror? How do the size and location change as you move away from the mirror?

36. If you hold your face very close to a convex mirror, what are the size and location of the image of your face? How do the size and location change as you move away from the mirror?

37. What is the fundamental difference between a real image and a virtual one?

38. Can both real and virtual images be photographed? Explain.

39. You hold a small lightbulb directly in front of a convex mirror. Is it possible for two rays leaving the lightbulb to intersect after reflecting from the mirror? Is the bulb's image real or virtual? Explain.

40. You hold a small lightbulb directly in front of a concave mirror beyond the mirror's focal point. Is it possible for two rays leaving the lightbulb to intersect after reflecting from the mirror? Is the bulb's image real or virtual? Explain.

41. A searchlight uses a concave mirror to produce a parallel beam. Where is the bulb located?

42. What happens to the location of the real image produced by a concave mirror if you move the object to the original location of this image?

43. Why does the arrival of the sound from a bass drum in a distant band not correspond to the blow of the drummer?

44. Astronomers claim that looking at distant objects is the same as looking back in time. In what sense is this true?

45. The Sojourner rover that explored the surface of Mars as part of NASA's Pathfinder mission had to make decisions on its own rather than be driven by remote control from Earth. Why?

NASA

46. Without asking, how could you tell whether you were talking to astronauts on the Moon or on Mars?

47. What color is produced by the overlap of a blue spotlight and a red spotlight?

48. A substance is known to reflect red and blue light. What color would it have when it is illuminated by white light? By red light?

49. A surface appears yellow under white light. How will it appear under red light? Under green light? Under blue light?

50. An actress wears a blue dress. How could you use spotlights to make the dress appear to be black?

51. A Crest toothpaste tube viewed under white light has a red C on a white background. What would you see if you used red light?

© Cengage Learning/George Semple

52. When you place a blue filter in the light from a projector, it produces a blue spot on the wall. If you use a red filter, you get a red spot. What would you see on the wall if you passed the light through both filters at once?

53. When you mix red and green light from separate projectors, you get a yellow spot on the wall. However, if you mix red and green paint, you get a muddy brown color. How do you account for this difference?

54. What color do you expect to get if you mix magenta and cyan paints?

55. What is the complementary color to green?

56. What is the complementary color to cyan?

57. If you removed all of the red light from white light, what color would you see?

58. A lens for a spotlight is coated so that it does not transmit yellow light. If the light source is white, what color is the spot?

Blue-numbered answered in Appendix B = more challenging questions

59. If the atmosphere primarily scattered green light instead of blue light, what color would the sky and Sun appear?

60. How would the color of sunlight change if the atmosphere were much more dense?

Exercises

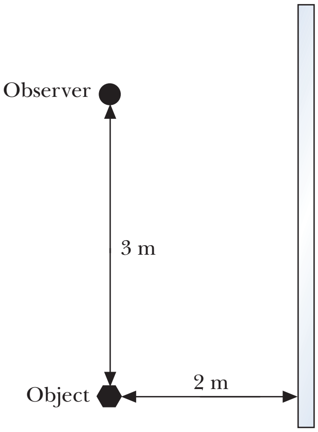
61. A 4-cm-diameter ball is located 40 cm from a point source and 80 cm from a wall. What is the size of the shadow on the wall?

62. A 4-cm-diameter ball is located 50 cm in front of a pinhole camera. If the film is located 10 cm from the pinhole, what is the size of the image on the film?

63. Use a ruler and a protractor to verify that the image produced by a flat mirror is as far behind the mirror as the object is in front.

64. George's eyes are 60 in. from the floor. His belt buckle is 36 in. from the floor. Determine the maximum distance from the floor that the bottom of a plane mirror can be placed such that George can see the belt buckle's image in the mirror. (*Hint:* You can verify that it does not matter how far George stands from the mirror.)

65. An object and an observer are located 2 m in front of a plane mirror, as shown in the following figure. If the observer is 3 m from the object, find the distance between the observer and the location of the object's image.



66. In Exercise 65 find the distance that the light travels from the object to the observer.

67. In the figure associated with Exercise 65, light leaves the object, reflects from the mirror, and reaches the observer. Use a protractor to find the angle of reflection.

68. Use a compass and a protractor to verify that the three rules for drawing ray diagrams for spherical mirrors satisfy the law of reflection.

69. What is the radius of the spherical surface that would produce a mirror with a focal length of 5 m?

70. A telescope mirror is part of a sphere with a radius of 3 m. What is the focal length of the mirror?

71. An object is located three times the focal length from a concave spherical mirror. Draw a ray diagram to locate its image. Is the image real or virtual, erect or inverted, magnified or reduced in size? Explain.

72. An object is located midway between the focal point and the center of a concave spherical mirror. Draw a ray diagram to locate its image. Is the image real or virtual, erect or inverted, magnified or reduced in size? Explain.

73. A 6-cm-tall object is placed 60 cm from a concave mirror with a focal length of 20 cm. Draw a ray diagram to find the location and size of the image.

74. How would your answer to Exercise 73 change if the same object were 120 cm from a concave mirror with a focal length of 40 cm?

75. Draw a ray diagram to locate the image of a 10-cm-tall object located 90 cm from a convex mirror with a focal length of 45 cm.

76. How would your answer to Exercise 75 change if the same object were 30 cm from a concave mirror with a focal length of 15 cm?

77. A convex mirror has a focal length of 60 cm. Draw a ray diagram to find the location and magnification of the image of an object located 30 cm from the mirror.

78. Repeat Exercise 77 for a concave mirror.

79. If you place an object 40 cm in front of a concave spherical mirror with a focal length of 20 cm, where will the image be located?

80. You have a concave spherical mirror with a focal length of 30 cm. Where could you place a candle to make it appear to burn at both ends?

81. If Galileo and his assistant were 15 km apart, how long would it take light to make the round-trip? How does this time compare with reaction times of about 0.2 s?

82. Suppose Galileo, in the experiment described in Exercise 81, had assumed that the entire 0.2-s delay was due to the travel time of light rather than to his assistant's reaction time. What value would he have calculated for the speed of light?

83. Approximately how long would it take a telegraph signal to cross the United States from the East Coast to the West Coast? (Telegraph signals travel at about the speed of light.)

84. Mars and Earth orbit the Sun at radii of 228 million km and 150 million km, respectively. When, in the future, your friend from Mars calls you on the phone and you answer, "Hello," what are the minimum and maximum times you will have to wait for your friend to reply?

85. How far does light travel in 1 year? This distance is known as a light-year and is a commonly used length in astronomy.

86. How far does light travel in 1 nanosecond—that is, in one-billionth of a second?

18

Refraction of Light

▶When light travels from one transparent material to another—say, from air to water—some interesting visual effects are created. Fish in an aquarium look bigger, and a tree that has fallen into a lake looks bent. Is it ever possible that the light cannot travel from one transparent material to another?

(See page 395 for the answer to this question.)

Refraction of light in the atmosphere causes the oval shape of the setting Sun.

WHEN light strikes a transparent material, it usually changes direction. This change accounts for many interesting effects such as the apparent distortion of objects and the beauty of an afternoon rainbow. This bending of light that occurs at the surface of a transparent object is called **refraction.**

Refraction can be studied by looking at the paths the light takes as the incident angle is varied, as shown in Figure 18-1. As in reflection, the angles are measured with respect to the normal to the surface. In this case the normal is extended into the material, and the angle of refraction is measured with respect to the extended normal. The amount of bending is zero when the angle of incidence is zero; that is, light incident along the normal to the surface is not bent. As the angle of incidence increases relative to the normal, the amount of bending increases; the angle of refraction differs more and more from the angle of incidence.

▶ Extended presentation available in the *Problem Solving* supplement

Index of Refraction

The amount of bending that occurs when light enters the material depends on the incident angle and an optical property of the material called the **index of refraction.** (We will refine the definition of the index of refraction in the next chapter.) A mathematical relationship can be written that predicts the refracted angle given the incident angle and the type of material. This rule, called *Snell's law,* is not as simple as the rule for reflection because it involves trigonometry. A simpler way to express the relationship is to construct a graph of the experimental data. Of course, although graphs are easier to use, they often have the disadvantage of being less general. In this case a graph has to be made for each substance. The graph in Figure 18-2 gives the angle of refraction in air, water, and glass for each angle of incidence in a vacuum. Although the curves for water and glass have similar shapes, light is refracted more on entering glass than water.

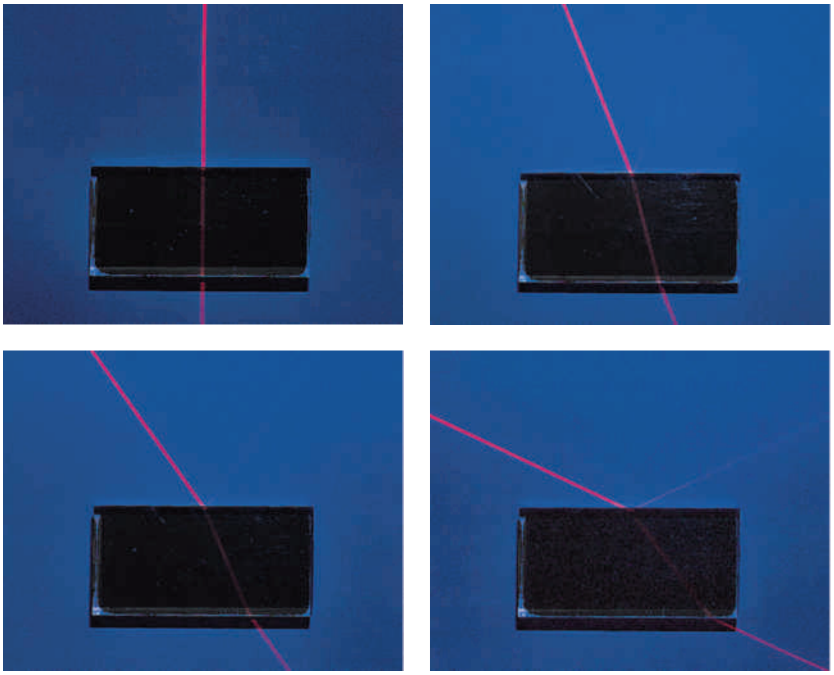


Figure 18-1 The amount of refraction depends on the angle of incidence. Notice that some of the incident light is reflected and some is refracted.

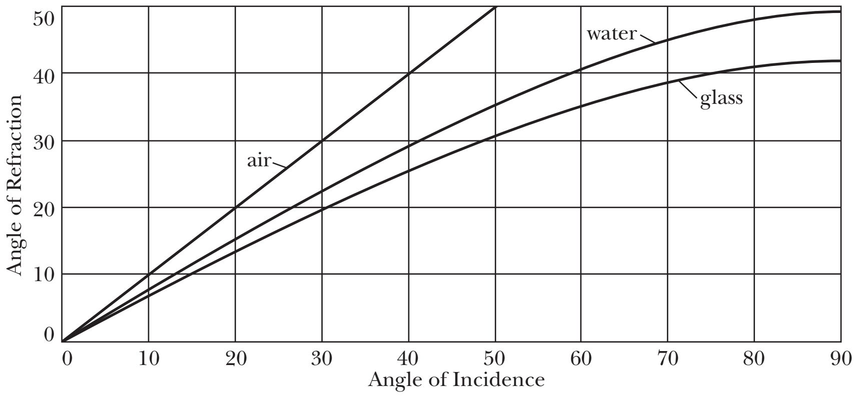


Figure 18-2 This graph shows the relationship between the angle of refraction and the angle of incidence for light entering air, water, and glass from a vacuum.

If no refraction takes place, the index of refraction is equal to 1. You can see from the graph that very little bending occurs when light goes from a vacuum into air; the index of refraction of air is only slightly greater than 1. Because the index of refraction of air is very close to 1, air and a vacuum are nearly equivalent. Therefore, we will use the graph in Figure 18-2 for light entering water or glass from either air or a vacuum. The index of refraction for water is 1.33; for different kinds of glass, it varies from 1.5 to 1.9. The curve for glass on the graph is drawn for an index of 1.5. The index of refraction for diamond is 2.42. A larger index of refraction means more bending for a given angle of incidence. For example, the graph indicates that light incident at 50 degrees (50°) has an angle of refraction of 31 degrees in glass and 35 degrees in water. Thus, the light is bent 19 degrees going into glass (index = 1.5) and only 15 degrees going into water (index = 1.33).

Q: What is the angle of refraction for light incident on glass at 30 degrees? How much does the ray bend?

A: The graph in Figure 18-2 gives an angle of refraction of approximately 20 degrees. Therefore, the ray bends 30 degrees − 20 degrees = 10 degrees from its original direction.

Light entering a transparent material from air bends *toward* the normal. What happens if light originates in the material and exits into the air? Experiments show that the paths of light rays are reversible. The photographs in Figure 18-1 can be interpreted as light inside the glass passing upward into the air. (If this were really the case, however, there would also be a faint reflected beam in the glass.) This example shows that when light moves from a material with a higher index of refraction to one with a lower index, the light leaving the material is bent *away from* the normal. Because of the reversibility of the rays, you can still use the graph in Figure 18-2 to find the angle of refraction; simply reverse the labels on the two axes.

Q: If a ray of light in water strikes the surface at an angle of incidence of 40 degrees, at what angle does it enter the air?

A: Locate the 40-degree angle on the *vertical axis* of the graph in Figure 18-2 and move sideways until you encounter the curve for water. Then, moving straight down to the horizontal axis, we obtain an angle of 58 degrees.

© Cengage Learning/David Rogers

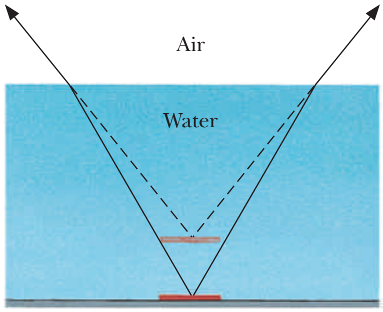
Figure 18-3 A straight pencil appears to be bent at the surface of the water.

© Cengage Learning/David Rogers (both)

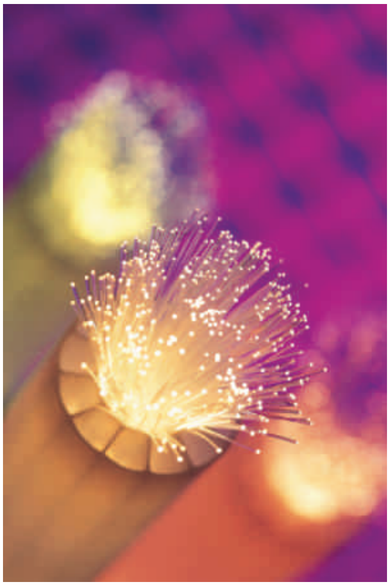
(a)

(b)

Figure 18-4 A coin underwater (a) appears closer than an identical coin in air (b). (c) Some of the rays that produce the virtual image.



(c)



Dennis O'Clair/Stone/Getty Images

Strands of glass optical fibers are used to carry voice, video, and data signals in telecommunication networks. Typical fibers have diameters of 60 micrometers.

Another consequence of this reversibility is that light passing through a pane of glass that has parallel surfaces continues in its original direction after emerging. The glass has the effect of shifting the light sideways, as shown in Figure 18-1.

The refraction of light produces interesting optical effects. A straight object partially in water appears bent at the surface. The photograph of a pencil in Figure 18-3 illustrates this effect. Looking from the top, we see that the portion of the pencil in the water appears to be higher than it actually is.

This phenomenon can also be seen in the photographs of identical coins, one underwater [Figure 18-4(a)] and the other in air [Figure 18-4(b)]. Even though the coins are the same distance from the camera, the one underwater appears closer and larger. The drawing in Figure 18-4(c) shows some of the rays that produce this illusion. This effect also makes fish appear larger—although never as large as the unlucky fisherman would like you to believe.

Let's examine the reason for the coin's appearing larger when it is in the water. Is it because the image is closer, or is the image itself bigger? It is fairly straightforward to see that the increase in size is due to the image being closer. To see that the image hasn't increased in size, we need to remind ourselves that rays normal to the surface are not refracted. Therefore, if we use vertical rays to locate the images of all points on the rim of the coin, each image will be directly above the corresponding point on the rim. This means that the image has the same size as the coin.

Q: If you keep your stamp collection under thick pieces of glass for protection, will the stamps appear to have their normal sizes?

A: No. Just like the coin in water, the stamps appear to be closer and are therefore apparently larger in size.

Total Internal Reflection

In some situations, light can't pass between two substances even if they are both transparent. This occurs at large incident angles when the light strikes a material with a lower index of refraction, such as from glass into air, as shown at the lower surface in Figure 18-5. At small angles of incidence, both reflection and refraction take place. The refracted angle is larger than the incident angle as shown in Figure 18-5(a). As the incident angle increases, the refracted angle increases even faster. At a particular incident angle, the refracted angle reaches 90 degrees. Beyond this incident angle—called the **critical angle**—the light no longer leaves the material; the light is totally reflected as shown in Figure 18-5(b). This is called **total internal reflection.**

The critical angle can be found experimentally by increasing the incident angle and watching for the disappearance of the emerging ray. Because the graph in Figure 18-2 works for both directions, we can find the critical angle by looking for the angle of refraction for an incident angle of 90 degrees. The intersection of the curve with the right-hand edge of Figure 18-2 indicates that the critical angle for our glass is about 42 degrees. The critical angle for diamond is only 24 degrees.

This total internal reflection has many applications. For example, a 45-degree right prism can act as a mirror, as shown in Figure 18-6. If the incident angle of 45 degrees is greater than the critical angle, when the light beam hits the back surface, the beam is totally reflected. This reflecting surface has many advantages over ordinary mirrors. It doesn't have to be silvered, it is easier to protect than an external surface, and it is also more efficient for reflecting light.

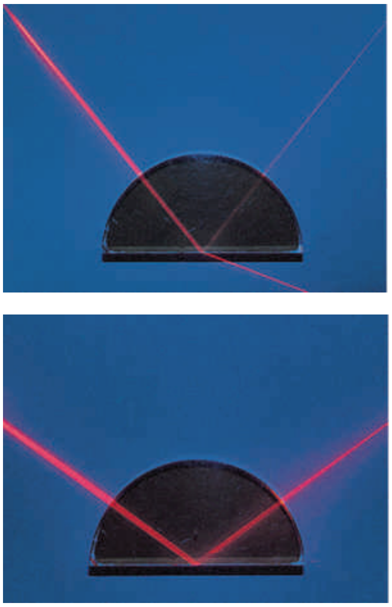


Figure 18-5 (a) Light traveling from glass into air at the lower surface bends away from the normal. (b) When the incident angle is larger than some critical angle, the light is totally reflected. None of the light passes through the surface.

Q: What is the critical angle for water?

A: The graph in Figure 18-2 shows that the angle of refraction in water never exceeds 49 degrees, so this is the critical angle.

Another application of this principle is to "pipe" light through long narrow fibers of solid plastic or glass, as shown in Figure 18-7. Light enters the fiber from one end. Once inside, the light doesn't escape out the side because the angle of incidence is always greater than the critical angle. The rays finally exit at the end of the fiber because there the incident angles are smaller than the critical angle. Fiber-optic applications are found in photography, medicine, telephone transmissions, and even decorative room lighting.

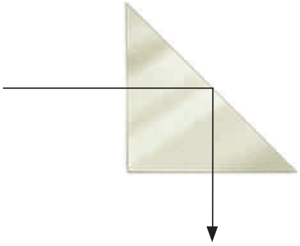


Figure 18-6 A prism acts as a flat mirror when the light is totally internally reflected.

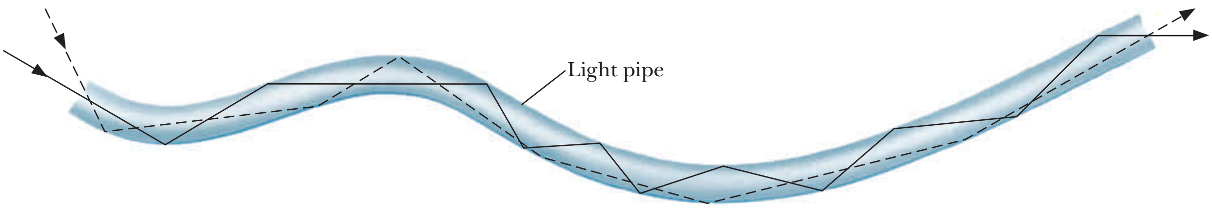


Figure 18-7 Light may be "piped" through solid plastic or glass rods using total internal reflection.

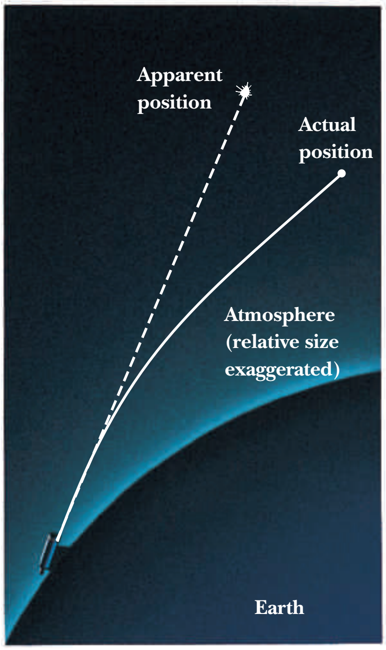


Figure 18-8 Atmospheric refraction changes the apparent positions of celestial objects, making them appear higher in the sky.

Atmospheric Refraction

We live at the bottom of an ocean of air. Light that reaches us travels through this air and is modified by it. Earth's atmosphere is not uniform. Under most conditions the atmosphere's density decreases with increasing altitude. As you may guess, the index of refraction depends on the density of a gas because the less dense the gas, the more like a vacuum it becomes. We therefore conclude that the index of refraction of the atmosphere gradually decreases the higher we go.

Refraction occurs whenever there is any change in the index of refraction. When there is an abrupt change, as at the surface of glass, the change in the direction of the light is abrupt. But when the change is gradual, the path of a light ray is a gentle curve. The gradual increase in the index of refraction as light travels into the lower atmosphere means that light from celestial objects such as the Sun, Moon, and stars bends toward the vertical. Figure 18-8 shows that this phenomenon makes the object appear higher in the sky than its actual position. Astronomers must correct for atmospheric refraction to get accurate positions of celestial objects.

This shift in position is zero when the object is directly overhead and increases as it moves toward the horizon. Atmospheric refraction is large enough that you can see the Sun and Moon before they rise and after they set. Of course, without knowing where the Sun and Moon should be, you are not able to detect this shift in position. You can, however, see distortions in their shapes when they are near the horizon, as shown in the photographs in Figure 18-9. Because the amount of refraction is larger closer to the horizon, the apparent change in position of the bottom of the Moon is larger than the change at the top. This results in a shortening of the diameter of the Moon in the vertical direction and gives the Moon an elliptical appearance.

There are other changes in the atmosphere's index of refraction. Because of the atmosphere's continual motion, there are momentary changes in the density of local regions. Stars get their twinkle from this variation. As the air moves, the index of refraction along the path of the star's light changes, and the star appears to change position slightly and to vary in brightness and color—that is, to twinkle. Planets do not twinkle as much because they are close enough to Earth to appear as tiny disks. Light from different parts of the disk averages out to produce a steadier image.

Images not available due to copyright restrictions

Dispersion

Although the ancients knew that jewels produced brilliant colors when sunlight shone on them, they were wrong about the origin of the colors. They thought the colors were part of the jewel. Newton used a prism to show that the colors don't come from jewels but rather from light itself—that the colors are already present in sunlight. When sunlight passes through a prism, the light refracts and is split up into a spectrum of colors ranging from red to violet, a phenomenon known as **dispersion** (Figure 18-10). To eliminate the idea that the prism somehow produced the colors, Newton did two experiments. He took one of the colors from a prism and passed it through a second prism, demonstrating that no new colors were produced. He also recombined the colors and obtained white light. His experiments showed conclusively that white light is a combination of all colors. The prism just spreads them out so that the individual colors can be seen.

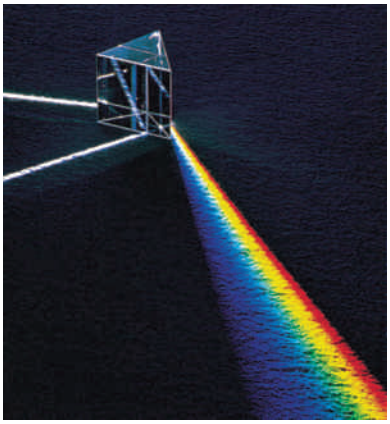

David Parker/Science Photo Library/ Photo Researchers, Inc.

Figure 18-10 A prism separates white light into the colors of the rainbow.

The name ROY G. BIV is a handy mnemonic for remembering the order of the colors produced by a prism or those in the rainbow: red, orange, yellow, green, blue, indigo, and violet. (Indigo is included mostly for the mnemonic; people can seldom distinguish it from blue or violet.)

The light changes direction as it passes through the prism because of refraction at the faces of the prism. Dispersion tells us that the colors have slightly different indexes of refraction in glass. Violet light is refracted more than red and therefore has a larger index. ("Blue bends better" is an easy way of remembering this.) The brilliance of a diamond is due to the small critical angle for internal reflection and the separation of the colors due to the high amount of dispersion.

Rainbows

Sometimes after a rain shower, you get to see one of nature's most beautiful demonstrations of dispersion, a rainbow. Part of its appeal must be that it appears to come from thin air. There seems to be nothing there but empty sky.

In fact, rainbows result from the dispersion of sunlight by water droplets in the atmosphere. The dispersion that occurs as the light enters and leaves the droplet separates the colors that compose sunlight. You can verify this by making your own rainbow. Turn your back to the Sun and spray a fine mist of water from your garden hose in the direction opposite the Sun. Each color forms


David Cavagnaro/Visuals Unlimited

A rainbow formed in the spray from a sprinkler hose.

Patrick J. Endres/Visuals Unlimited

Figure 18-11 A rainbow's magic is that it seems to appear out of thin air. Notice the secondary rainbow on the right.

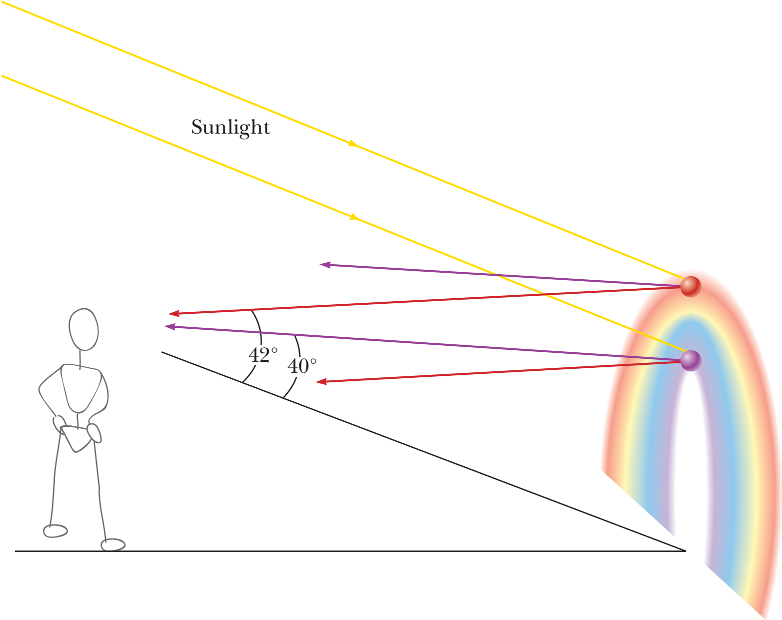


Figure 18-13 The color of each water droplet forming the rainbow depends on the viewing angle.

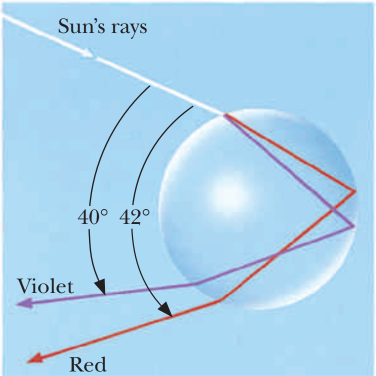


Figure 18-12 Dispersion of sunlight in a water droplet separates the sunlight into a spectrum of colors.

part of a circle about the point directly opposite the Sun (Figure 18-11). The angle to each of the droplets along the circle of a given color is the same. Red light forms the outer circle and violet light the inner one. Figure 18-12 shows the paths of the red and violet light. The other colors are spread out between these two according to the mnemonic ROY G. BIV. Each droplet disperses all colors. Your eyes, however, are only in position to see one color coming from a particular droplet. For instance, if the droplet is located such that a line from the Sun to the droplet and a line from your eyes to the droplet form an angle of 42 degrees, the droplet appears red (Figure 18-13). If this angle is 40 degrees, the droplet appears violet. Intermediate angles yield other colors.

Whether or not you believe there is a pot of gold at the end of the rainbow, you will never be able to get there to find out. As you move, the rainbow "moves." In your new position, different droplets produce the light you see as the rainbow.

FLAWED REASONING

A friend calls you at 8:00 a.m. and tells you to go outside and observe a beautiful rainbow in the east. **Would you hire this friend as a hiking guide?**

ANSWER Your friend has serious compass issues. The Sun comes up in the east. You see rainbows by looking away from the Sun. Indeed, the center of the rainbow will lie along a line passing through the Sun and your head. Therefore, at 8:00 a.m. you will see the rainbow in the west.

If you are willing to get wet, it is possible to see a complete circular rainbow. Near noon on a sunny day, spray the space around you with a fine mist. Looking down, you will find yourself in the center of a rainbow. A circular rainbow can sometimes be seen from an airplane.

If viewing conditions are good, you can see a secondary rainbow that is fainter and larger than the first (Figure 18-11). It is centered on the same point, but the colors appear in reverse order. This rainbow is produced by light that reflects twice inside the droplets.

Q: If you see a rainbow from an airplane, where do you expect to see the shadow of the airplane?

A: Because the center of the rainbow is always directly opposite the Sun, the shadow of the airplane will be at the center of the rainbow.

Halos

Sometimes a large halo can be seen surrounding the Sun or Moon. These halos and other effects, such as *sun dogs* and various arcs, are caused by the refraction of light by ice crystals in the atmosphere.

Atmospheric ice crystals have the shape of hexagonal prisms. Each one looks like a slice from a wooden pencil that has a hexagonal cross section. Light hitting the crystal is scattered in many different directions, depending on the angle of incidence and which face it enters and exits. Light entering and exiting alternate faces, as shown in Figure 18-14, has a minimum angle of scatter of 22 degrees. Although light is scattered at other angles, most of the light concentrates near this angle.

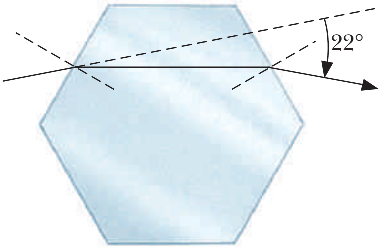


Figure 18-14 Light passing through alternate surfaces of a hexagonal ice crystal changes direction by at least 22 degrees.

To see a ray of light that has been scattered by 22 degrees, you must look in a direction 22 degrees away from the Sun. Light scattering this way from crystals randomly oriented in the atmosphere forms a 22-degree halo around the Sun, as shown in Figure 18-15. The random nature of the orientations ensures that at any place along the halo there will be crystals that scatter light into your eyes. Dispersion in the ice crystals produces the colors in the halo.

Occasionally one also sees "ghost" suns located on each side of the Sun at the same height as the Sun, as seen in Figure 18-15. Ice crystals that have vertically oriented axes produce these *sun dogs*. These crystals can refract light into your eyes only when they are located along or just outside the halo's circle at the same altitude as the Sun.

An even larger but dimmer halo at 46 degrees exists but is less frequently seen. It is formed by light passing through one end and one side of the crystals. Other effects are produced by light scattering through other combinations of faces in crystals with particular orientations.

Courtesy of Robert Greenler

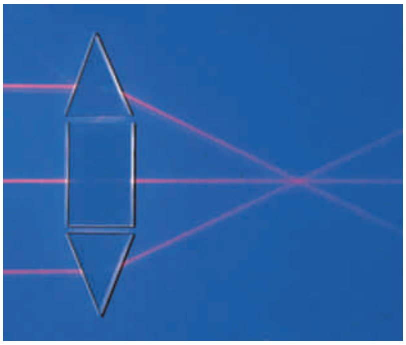
Figure 18-15 A photograph of the 22-degree halo and its associated sun dogs taken with a fisheye lens. A building was used to block out the Sun's direct rays.

Lenses

When light enters a material with entrance and exit surfaces that are not parallel, unlike a pane of glass, the direction of the light beam changes. Two prisms and a rectangular block can be used to focus light, as shown in Figure 18-16. However, most other rays passing through this combination would not be focused at the same point. The focusing can be improved by using a larger number of blocks or by shaping a piece of glass to form a lens.

We see the world through lenses. This is true even for those of us who don't wear glasses, because the lenses in our eyes focus images on our retinas. Other lenses extend our view of the universe—microscopes for the very small and telescopes for the very distant.

Although many lens shapes exist, they can all be put into one of two groups: those that converge light and those that diverge light. If the lens is thicker at



© Cengage Learning/David Rogers

Figure 18-16 Two prisms and a rectangular block form a primitive lens.

its center than at its edge, as in Figure 18-17(a), it is a converging lens. If it is thinner at the center, as in Figure 18-17(b), it is a diverging lens.

Are You On the Bus?

Q: Lenses in eyeglasses are made with one convex surface and one concave surface. How can you tell if the lenses are converging or diverging?

A: Check to see if they are thicker at the center than at the edges. If they are thicker at the center, they are converging.

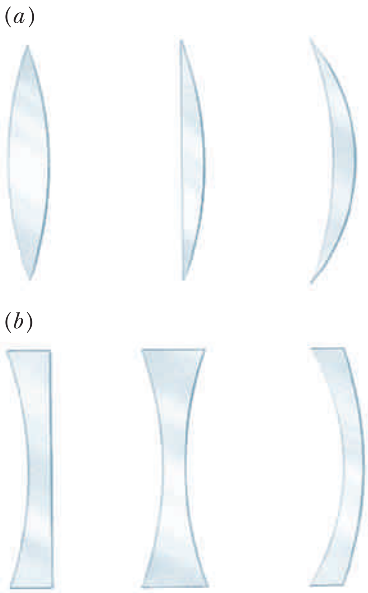


Figure 18-17 (a) Converging lenses are thicker at the middle. (b) Diverging lenses are thinner at the middle.

Lenses have two focal points—one on each side. A converging lens focuses incoming light that is parallel to its **optic axis** at a point on the other side of the lens known as the *principal* **focal point** (Figure 18-18). The distance from the center of the lens to the focal point is called the **focal length.** We can find the other focal point by reversing the direction of the light and bringing it in from the right-hand side of the lens. The light then focuses at a point on the left-hand side of the lens that we refer to as the "other" focal point in drawing ray diagrams.

For a diverging lens, incoming light that is parallel to the optic axis appears to diverge from a point on the same side of the lens (Figure 18-19). This point is known as the *principal focal point,* and the focal point on the other side is known as the "other" one. You can show by experiment that the two focal points are the same distance from the center of the lens if the lens is thin. A lens is considered to be thin if its thickness is very much less than its focal length.

The shorter the focal length, the "stronger" the lens; that is, the lens focuses light parallel to the optic axis at a point closer to the lens.

Images Produced by Lenses

The same ray-diagramming techniques used for curved mirrors in the previous chapter will help us locate the images formed by lenses. Again, three of the rays are easily drawn without measuring angles. The intersection of any two determines the location of the image. Figure 18-20 shows the three rays.

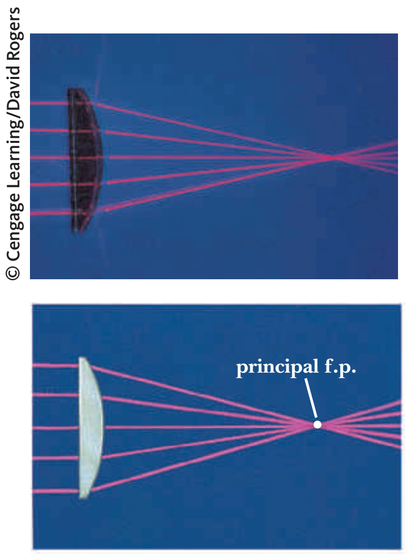


Figure 18-18 Light parallel to the optic axis of a converging lens is focused at the principal focal point.

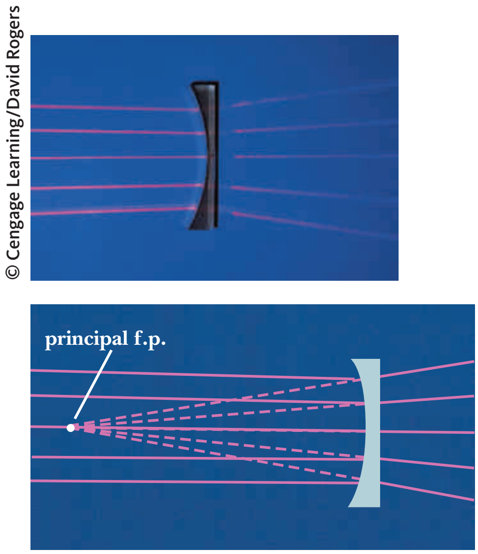


Figure 18-19 Light parallel to the optic axis of a diverging lens appears to come from the principal focal point.

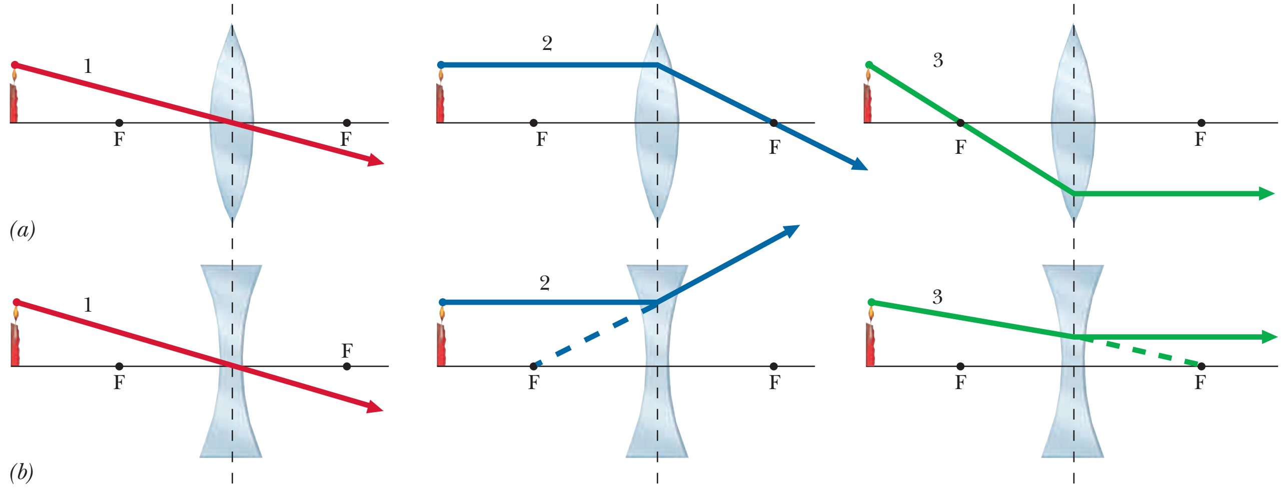


Figure 18-20 The three rays used in drawing ray diagrams for (a) converging and (b) diverging lenses.

First, a ray passing through the center of the lens continues without deflection. Second, for a converging lens, a ray parallel to the optic axis passes through the principal focal point. Third, a ray coming from the direction of the other focal point leaves the lens parallel to the optic axis [Figure 18-20(a)]. (The optic axis passes through the center of the lens and both focal points.) Notice that the second and third rays are opposites of each other. For a diverging lens, the second ray comes in parallel to the optic axis and leaves as if it came from the principal focal point, and the third ray heads toward the other focal point and leaves parallel to the optic axis [Figure 18-20(b)].

These rays are similar to the ones used for mirrors. There are two main differences: the first ray passes through the center of the lens and not the center of the sphere as it did for mirrors, and there are now two focal points instead of one. We can still give abbreviated versions of these rules (the words in parentheses refer to diverging lenses).

1. Through center—continues
2. Parallel to optic axis—through (from) principal focal point
3. Through (toward) other focal point—parallel to optic axis

◀ rays for lenses

These rules assume that the lens is thin. The first rule neglects the offset that takes place when a light ray passes through parallel surfaces of glass at other than normal incidence (see Figure 18-1). For the purposes of drawing these rays, the bending of the light is assumed to take place at a plane perpendicular to the optic axis and through the center of the lens. A vertical dashed line indicates this plane.

We can apply these rays to locate the image of a candle that is located on the optic axis outside the focal point of a converging lens. The ray diagram in Figure 18-21 shows that the image is located on the other side of the lens and is real and inverted. (See the previous chapter for a discussion of the types of image.) Whether the image is magnified depends on how far it is from the focal point. As the candle is moved away from the lens, the image moves closer to the principal focal point and gets smaller.

If the candle is moved inside the focal point, as illustrated in Figure 18-22, the image appears on the same side of the lens. This is the arrangement that is used when a converging lens is used as a magnifying glass. The lens is positioned such that the object is inside the focal point, producing an image that is virtual, erect, and magnified.

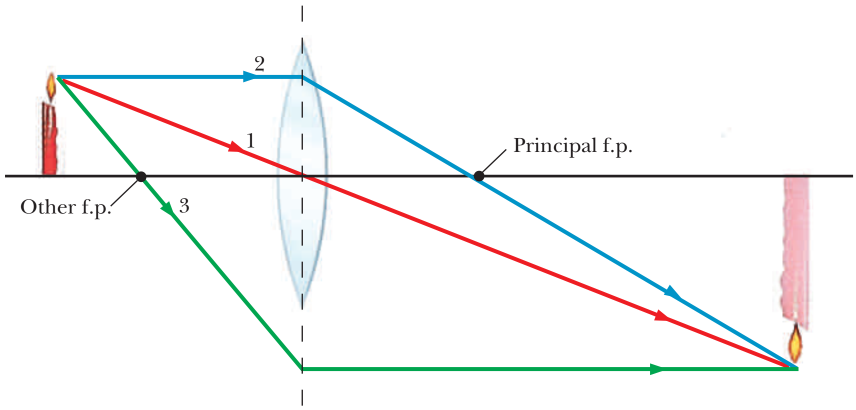


Figure 18-21 The image formed by a candle located outside the focal point of a converging lens is real and inverted and may be magnified or reduced in size.

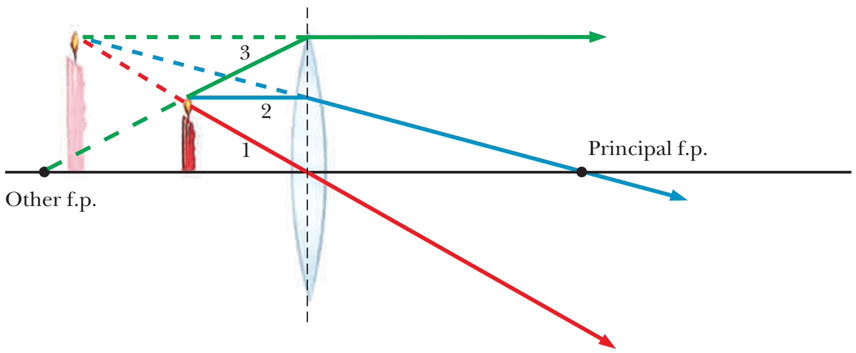


Figure 18-22 The image formed by a candle located inside the focal length of a converging lens is virtual, erect, and magnified.

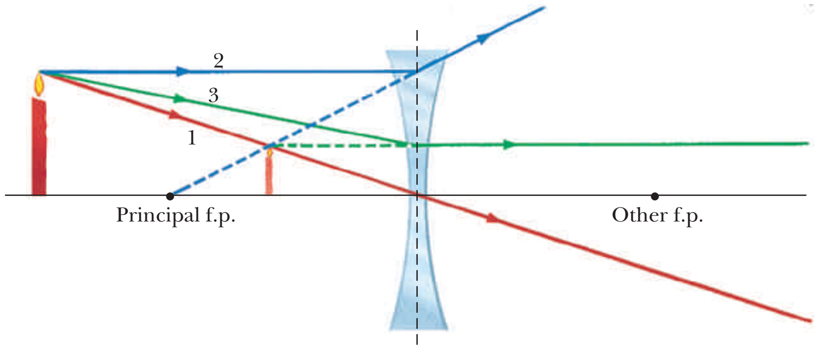


Figure 18-23 The image formed by a diverging lens is virtual, erect, and reduced in size.

A diverging lens always produces a virtual image, as shown in Figure 18-23. The image changes location and size as the object is moved, but the image remains erect and virtual.

Q: Is the lens used in a slide projector converging or diverging?

A: It must be converging because it forms a real image on the screen.

Notice that one of the rays in Figure 18-21 does not pass through the lens. This isn't a problem because there are many other rays that do pass through the lens to form the image. Ray diagramming is just a geometric construction that allows you to locate images, a process that can be illustrated with an illu-

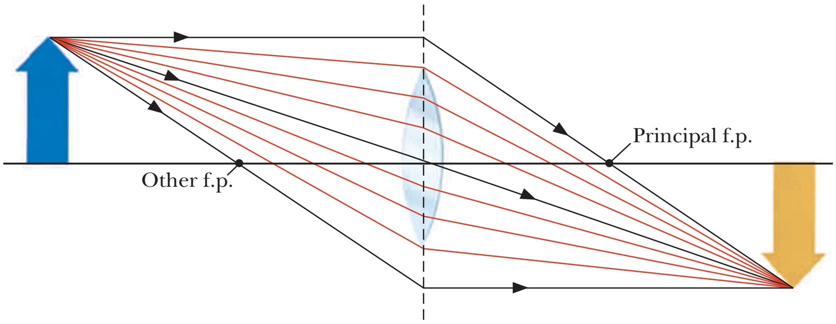


Figure 18-24 The orange rays form the image. The black ones are used because they are easy to draw.

minated arrow and a large-diameter lens, as drawn in Figure 18-24. A piece of paper at the image's location allows the image to be easily seen. If the lens is then covered with a piece of cardboard with a hole in it, the image is still in the same location, is the same size, and is in focus. The light rays from the arrow that form the image are those that pass through the hole. The image is not as bright because less light now forms the image. The orange lines illustrate the paths of some of the other rays.

FLAWED REASONING

The following question appears on the final exam: "Three long light filaments are used to make a letter Y that is placed in front of a large converging lens such that it creates a real image on the other side of the lens. The meeting point of the three filaments lies on the optic axis of the lens. A piece of cardboard is then used to cover up the bottom half of the lens. Describe what happens to the image of the Y."

Three students give their answers:

Jacob: "The cardboard will block the light from the lower filament, so the image will appear as a letter V."

Emily: "The real image formed by a converging lens is inverted. The image would now appear to be an upside-down letter V."

Michael: "The image is inverted, so the light from the lower filament must pass through the top half of the lens and the light from the upper two filaments will be blocked by the cardboard. The image will appear as the letter I."

All three students have answered incorrectly. **Find the flaws in their reasoning.**

ANSWER A point source of light sends light to all parts of the lens's surface. This light converges at a single point on the other side of the lens (the image location). Covering half the lens blocks half the light, but the other half still forms an image at the same location. The three long filaments can be thought of as a collection of many point sources. They still form the same image (an upside-down Y). The image will be dimmer because half the light is blocked.

Cameras

We saw in the last chapter that pinhole cameras produce sharp images if the pinhole is very small. The amount of light striking the film, however, is quite small. Very long exposure times are needed, which means that the objects in the scene must be stationary. The amount of light reaching the film can be substantially increased (and the exposure time substantially reduced) by using a converging lens instead of a pinhole.

The essential features of a simple camera are shown in Figure 18-25. This camera has a single lens at a fixed distance from the film. The distance is chosen so that the real images of faraway objects are formed at, or at least near, the film. These cameras are usually not very good for taking close-up shots, such as portraits, because the images are formed beyond the film and are therefore out of focus at the film. More expensive cameras have an adjustment that moves the lens relative to the film to position (focus) the image on the film.

Q: If the focal length of the lens in a simple camera is 50 millimeters, how far is it from the lens to the film for a subject that is very far from the camera?

A: If the objects are effectively at infinity, the light from each point will be focused at a distance equal to the focal length. Thus, the film should be about 50 millimeters from the center of the lens.

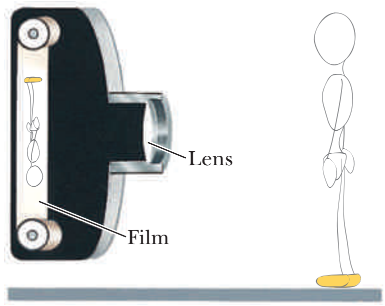


Figure 18-25 The essential features of a simple camera.

Ideally, all light striking the lens from a given point on the object should be focused to a given point on the film. However, real lenses have a number of defects, or **aberrations,** so that light is not focused to a point but is spread out over some region of space.

A lens cannot focus light from a white object to a sharp point because of dispersion. A converging lens focuses violet light at a point closer to the lens than it does red light. This chromatic aberration produces images with colored fringes. Because the effect is reversed for diverging lenses and the amount of dispersion varies with material, lens designers minimize chromatic aberration by combining converging and diverging lenses made of different types of glass.

A spherical lens (or a spherical mirror, for that matter) does not focus all light parallel to the optic axis to a sharp point. Light farther from the optic axis is focused at a point closer to the lens than light near the optic axis. Using a combination of lenses usually corrects this spherical aberration; using a diaphragm to decrease the effective diameter of the lens also reduces it. Although this sharpens the image, it also reduces the amount of light striking the film. New techniques for reducing spherical aberration by grinding lenses with nonspherical surfaces and by making lenses in which the index of refraction of the glass changes with the distance from the optic axis have been developed.

Our Eyes

Gerald F. Wheeler

The retina of a human eye.

Leonardo da Vinci stated in the 15th century that the lens of an eye forms an image inside the eye that is transmitted to the brain. He believed that this image must be upright. It was a century before it was shown that he was half right: the lens forms an image inside the eye, but the image is upside down. The inverted nature of the image was demonstrated by removing the back of an excised animal eye and viewing the image. The inverted world received by our retinas is interpreted as right-side up by our eye–brain system. The essential features of this remarkable optical instrument include the cornea, the lens, and some fluids, which act collectively as a converging lens to form real, inverted images on the retina (Figure 18-26).

When you look at a distant object, nearby objects are out of focus. Only distant objects form sharp images on the surface of the retina. The nearby objects form images that would be behind the retina, and the images on the retina are therefore fuzzy. This phenomenon occurs because the locations of images of objects at various distances depend on the distances between the lens and the objects and on the focal length of the lens. The lens in the eye changes its shape and thus its focal length to accommodate the different distances.

Opticians measure the strength of lenses in diopters. The lens strength in *diopters* is equal to the reciprocal of the focal length measured in meters. For example, a lens with a focal length of 0.2 meter is a 5-diopter lens. In this case a larger diopter value means that the lens is stronger. Converging lenses have positive diopters, and diverging lenses have negative diopters. Diopters have the advantage that two lenses placed together have a diopter value equal to the sum of the two individual ones.

In the relaxed eye of a young adult who does not wear corrective lenses, all the transparent materials have a total "power" of +60 diopters. Most of the refraction (+40 diopters) is due to the outer element of the eye, the cornea, but the relaxed lens contributes +20 diopters.

The eye can vary the strength of the lens from a relaxed value of +20 diopters to a maximum of +24 diopters. When the relaxed eye views a distant object, the +60 diopters produce an image at 1.7 centimeters (0.7 inch), which is the distance to the retina in a normal eye. The additional +4 diopters allow the eye to view objects as near as 25 centimeters (10 inches) and still produce sharp images on the retina.

The ability of the eye to vary the focal length of the lens decreases with age as the elasticity of the lens decreases. A 10-year-old eye may be able to focus as close as 7 centimeters (+74 diopters), but a 60-year-old eye may not be able to focus any closer than 200 centimeters ($6\frac{1}{2}$ feet). An older person often wears bifocals when the eyes lose their ability to vary the focal length.

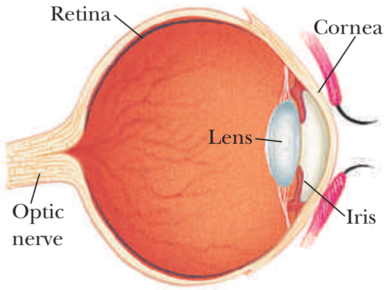


Figure 18-26 Schematic drawing of the human eye.

	Sphere	Cylinder	Axis
R	−6.50	+3.25	089
L	−5.75	+2.75	074
R	+2.00		
L	+2.00	Bifocals	

The prescription for your author's eyeglasses. The spherical and cylindrical corrections are given in diopters. Axis is the number of degrees the axis of the cylinder is rotated from the vertical. The bifocal correction is added to the others.

WORKING IT OUT *Diopters*

A converging lens of focal length 25 cm is placed next to a diverging lens of length 20 cm. What is the effective focal length for this combination? Is it diverging or converging?

A lens with a shorter focal length is more effective in bending the light; it is a "stronger" lens. The strength of the lens is therefore given by the inverse of the focal length, measured in diopters. The strength of the converging lens is

$$d_1 = \frac{1}{0.25\ \text{m}} = +4\ \text{diopters}$$

The strength of the diverging lens is

$$d_2 = -\frac{1}{0.20\ \text{m}} = -5\ \text{diopters}$$

where the negative sign indicates that it is spreading the light rather than collecting it. The combined strength of the two lenses is given by the sum of the diopters:

$$d_{total} = d_1 + d_2 = +4\ \text{diopters} - 5\ \text{diopters} = -1\ \text{diopter}$$

The effective focal length is the inverse of d_{total}, or −1 m. The two lenses combined could be replaced by a single diverging lens with focal length of 1 m.

The amount of light entering the eye is regulated by the size of the pupil. As with the ear, the range of intensities that can be viewed by the eye is very large. From the faintest star that can be seen on a dark, clear night to bright sunlight is a range of intensity of approximately 10^{10}.

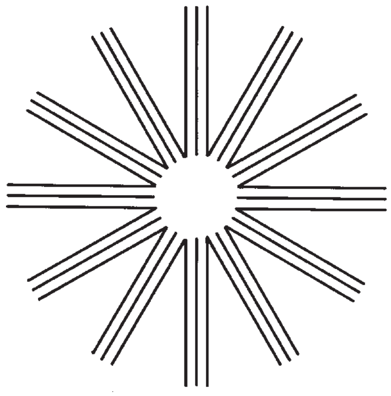


Figure 18-27 A test pattern for astigmatism. If you see some lines blurred while other lines are sharp and dark, you have some astigmatism.

Another common visual defect is *astigmatism.* When some of the refracting surfaces are not spherical, the image of a point is spread out into a line. Use the pattern in Figure 18-27 to check for astigmatism in your eyes. Lines along the direction in which images of point sources are spread remain sharp and dark, but the others become blurred. Are your two eyes the same?

Magnifiers

It has been known since the early 17th century that refraction could bend light to magnify objects. The invention of the telescope and microscope produced images of regions of the universe that until then had been unexplored. Galileo used the newly discovered telescope to see Jupiter's moons and the details of our Moon's surface. English scientist Robert Hooke spent hours peering into another unexplored world with the aid of the new microscope.

The size of the image on the retina depends on the object's physical size and on its distance away. The image of a dime held at arm's length is much larger than that of the Moon. What really matters is the angular size of the object—that is, the angle formed by lines from your eye to opposite sides of the object. The angular size of an object can be greatly increased by bringing it closer to your eye. However, if you bring it closer than about 25 centimeters (10 inches), your eye can no longer focus on it, and its image is blurred. You can get both an increased angular size and a sharp image by using a converging lens as a magnifying glass. When the object is located just inside the focal point of the lens, the image is virtual and erect and has nearly the same angular size as the object. Moreover, as shown in Figure 18-28, the image is now far enough away that the eye can focus on it and see it clearly.

An even higher magnification can be achieved by using two converging lenses to form a compound microscope, as shown in Figure 18-29. The object is located just outside the focal point of the objective lens. This lens forms a real image that is magnified in size. The eyepiece then works like a magnifying glass to further increase the angular size of this image.

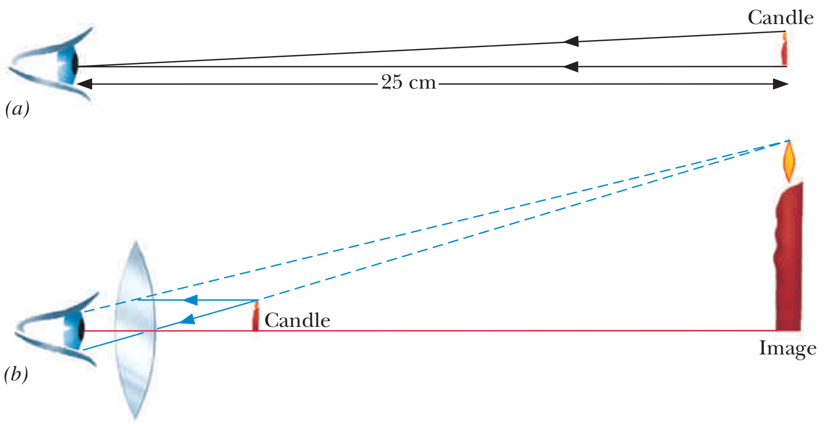


Figure 18-28 A candle viewed (a) at a distance of 25 centimeters and (b) through a magnifying glass.

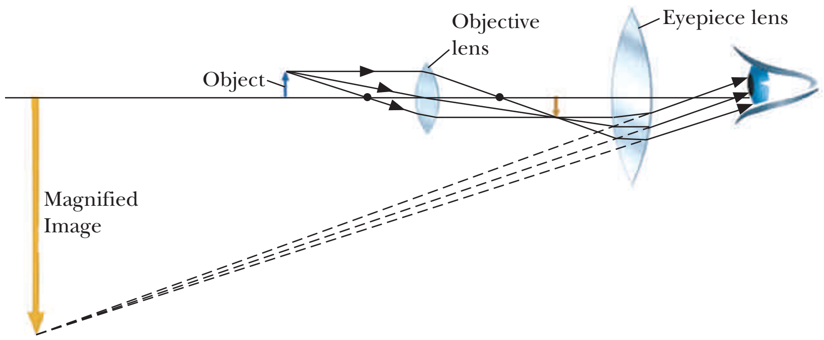


Figure 18-29 Schematic of a compound microscope.

Everyday Physics | *Eyeglasses*

Our optical system is quite amazing. The lens in our eye can change its shape, altering its focal length to place the image on the retina. Sometimes, however, the eye is too long or too short, and the images are formed in front of or behind the retina. When the eye is too long, the images of distant objects are formed in front of the retina, as shown in Figure A. Such a person has *myopia* (nearsightedness) and can see things that are close but has trouble seeing distant objects. When the eye is too short, the person has *hyperopia* (farsightedness) and has trouble seeing close objects. Distant objects can be imaged on the retina, but close objects form images behind the retina (Figure B).

Our knowledge of light and refraction allows us to devise instruments—eyeglasses—that correct these deficiencies. The nearsighted person wears glasses with diverging lenses to see distant objects (Figure A). The farsighted person's sight is corrected with converging lenses (Figure B).

Even people with perfect vision early in life lose some of the lens's range, particularly the ability to shorten the focal length. These people have trouble creating a focused image on their retina for objects that are close. Many older people wear bifocals when their eyes lose the ability to shorten the focal length. The upper portion of the lens is used for distant viewing; the lower portion is used for close work or reading. When people work at intermediate distances such as looking at computer screens, they sometimes wear trifocals.

The difficulty of making glasses to correct vision is increased when a person has astigmatism. This occurs when some of the refracting surfaces are not spherical and the image of a point is spread out into a line. This visual defect is corrected by adding a cylindrical curvature to the spherical curvature of the lens. You can check your (or a friend's) glasses for correction for astigmatism by looking through them in the normal way but holding the glasses at a distance from your head. When you rotate the lens about a horizontal axis, you will see background distortion if the lens has an astigmatic correction.

Many people wear contact lenses to correct their vision. The use of contact lenses has created some interesting challenges in correcting for astigmatism and the need to wear bifocals. When correcting for astigmatism, the cylindrical correction must have the correct orientation. Contact lenses can be weighted at one place on the edge to keep the lens oriented correctly.

People who need bifocals can sometimes be fitted with different corrections in each eye; one eye is used for close vision and the other is used for distant vision. The eye–brain system switches from one eye to the other as the situation demands.

Finally, a medical procedure uses laser technology to correct nearsightedness. In this procedure, the laser is used to make radial cuts in the cornea of the eye to reduce its curvature.

1. Explain how diverging lenses can be used to correct for nearsightedness.
2. Why do your contact lenses need to be weighted more heavily on one side if you have astigmatism?

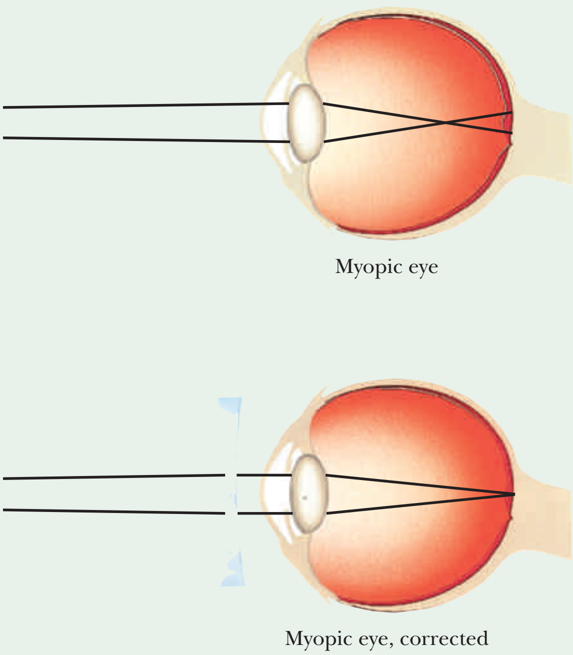


Figure A The myopic eye forms the images of distant objects in front of the retina. This is corrected with diverging lenses.

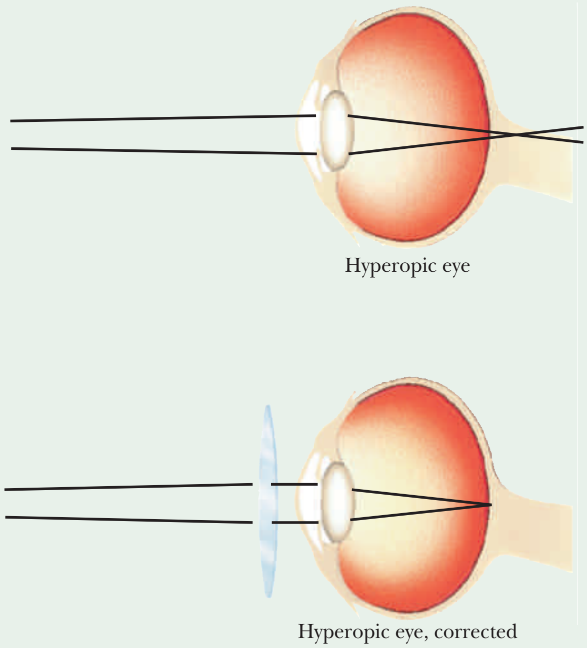


Figure B The hyperopic eye forms images of close objects in back of the retina. This is corrected with converging lenses.

Telescopes

There are many varieties of telescope. A simple one using two converging lenses is known as a **refracting telescope,** or refractor. Figure 18-30 shows that this type of telescope has the same construction as a compound microscope except that now the object is far beyond the focal point of the objective lens. Like the microscope, the refractor's objective lens produces a real, inverted image. Although the image is much smaller than the object, it is much closer to the eye. The eyepiece acts as a magnifying glass to greatly increase the angular size of the image. The magnification of a telescope is equal to the ratio of the focal lengths of the objective lens and the eyepiece. To get high magnification, the focal length of the objective lens needs to be quite long.

Binoculars were designed to provide a long path length in a relatively short instrument. The diagram in Figure 18-31 shows that this is accomplished by using the internal reflections in two prisms to fold the path.

Large-diameter telescopes are desirable because they gather a lot of light, allowing us to see very faint objects or to shorten the exposure time for taking pictures. The problem, however, is making a large-diameter glass lens. It is difficult, if not impossible, to make a piece of glass of good enough quality. Also, a lens of this diameter is so thick that it sags under its own weight. Therefore, most large telescopes are constructed with concave mirrors as objectives and

Figure 18-30 Schematic of an astronomical telescope.

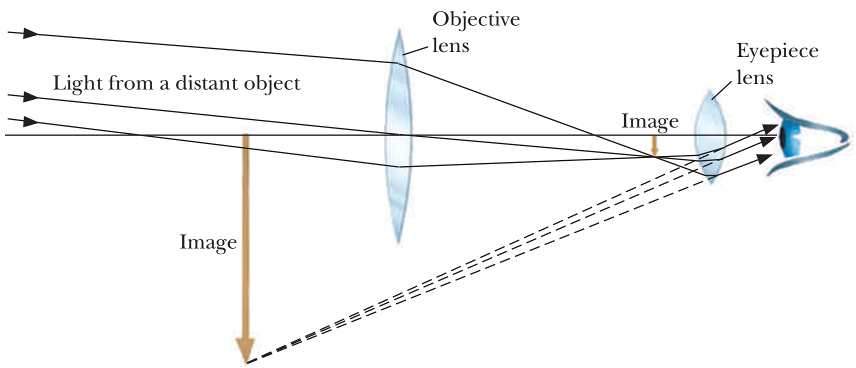
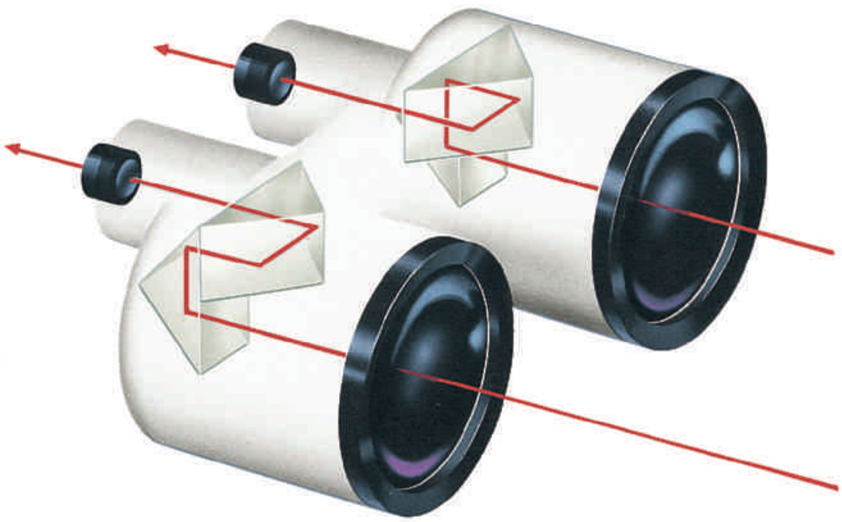


Figure 18-31 Schematic of prism binoculars.



are known as **reflecting telescopes,** or reflectors. The use of a concave mirror to focus the incoming light has several advantages: the construction of a mirror requires grinding and polishing only one surface rather than two; a mirror can be supported from behind; and, finally, mirrors do not have the problem of chromatic aberration. Figure 18-32 illustrates several designs for reflecting telescopes.

The world's largest refractor has a diameter of 1 meter (40 inches), whereas the largest reflector has a diameter of 6 meters (236 inches). This is just about the limit for a telescope with a single objective mirror; the costs and manufacturing difficulties are not worth the gains. Telescope makers have recently built telescopes in which the images from many smaller mirrors are combined to increase the light-gathering capabilities.

Summary

When light strikes a transparent material, part of it reflects and part refracts. The amount of refraction depends on the incident angle and the index of refraction of the material. Light entering a material of higher index of refraction bends toward the normal. Because the refraction of light is a reversible process, light entering a material with a smaller index of refraction bends away from the normal. For light in a material with a larger index of refraction, total internal reflection occurs whenever the angle of incidence exceeds the critical angle.

The refraction of light at flat surfaces causes objects in or behind materials of higher indexes of refraction to appear closer, and therefore larger. The apparent locations of celestial objects are changed by refraction in the atmosphere.

White light is separated into a spectrum of colors because the colors have different indexes of refraction, a phenomenon known as dispersion. Rainbows

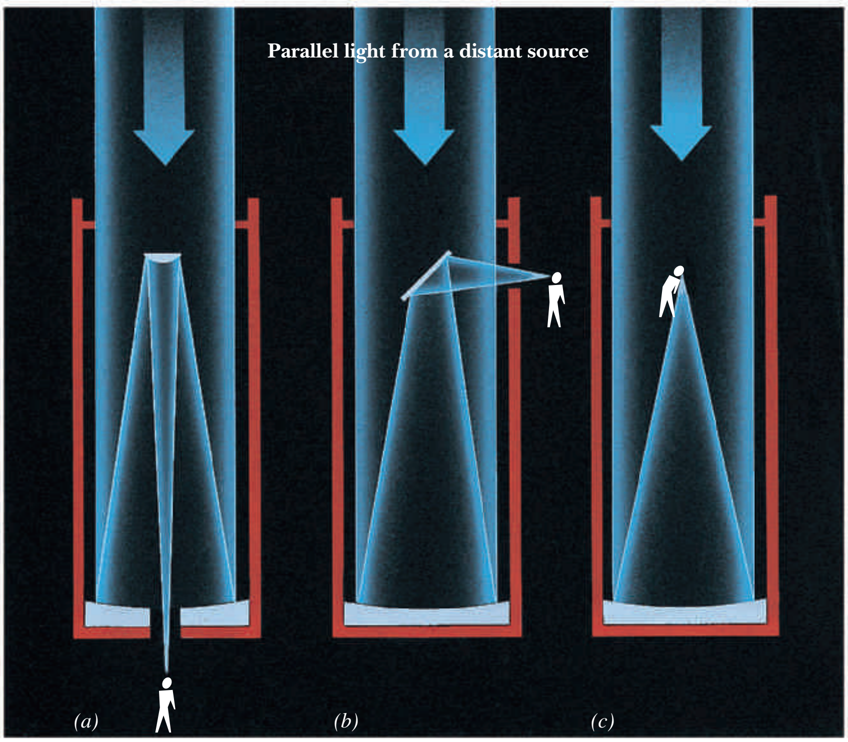


Figure 18-32 Schematics of (a) a Cassegrain reflector, (b) a Newtonian reflector, and (c) a prime-focus telescope.

Everyday Physics | *The Hubble Space Telescope*

At a cost of $1.6 billion, the Hubble Space Telescope was placed in an orbit 575 kilometers (357 miles) above Earth by the space shuttle *Discovery* in April 1990. Its mission is to provide astronomers with observations of the universe without the disturbances caused by Earth's atmosphere. The atmosphere absorbs most of the radiation reaching us from space, except for two broad bands in the radio region and around the visible region. A variety of experiments were planned that ranged from viewing distant, faint objects to accurately measuring the positions of stars. The design specifications indicated that the Hubble Space Telescope would be able to see objects seven times as far away as what could be observed from Earth's surface.

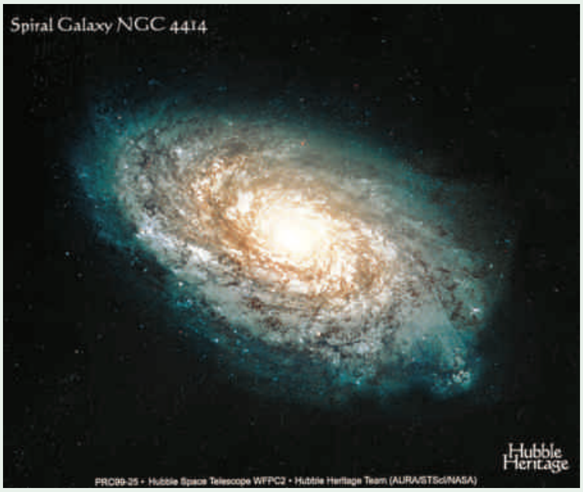
However, these experiments were seriously hampered by a defect in the telescope's primary mirror. Although the error in its shape was only 0.002 millimeter (about one-fortieth the thickness of a human hair) at the edge of the 2.4-meter-diameter mirror, it caused light from the edge of the mirror to focus 38 millimeters beyond light from the center. This left the telescope with an optical defect—spherical aberration—that created fuzzy halos around images of stars, and blurred images of extended objects such as galaxies and giant clouds of gas and dust. Only 15% of the light from a star was focused into the central spot compared to the design value of 70%. However, because this type of aberration is well known, computer enhancement was used to sharpen images of brighter objects, producing some rather remarkable views and some very good science, but the technique did not work for faint objects.

Fortunately, the Hubble Space Telescope was designed to be serviced by the space shuttles, and new optics were designed to compensate for the error. In December 1993 the space shuttle *Endeavour* docked with the Hubble Space Telescope to repair the optics and replace a number of mechanical and electrical components that had failed or required scheduled replacement. The repairs required five spacewalks involving two astronauts each. The total repair mission had a cost of $700 million. The repaired Hubble Space Telescope is able to see much farther into space.

1. Why is a space-based telescope superior to an Earth-based one?
2. In what ways did spherical aberration affect the initial images formed by the Hubble Space Telescope?

NASA

The Hubble Space Telescope.



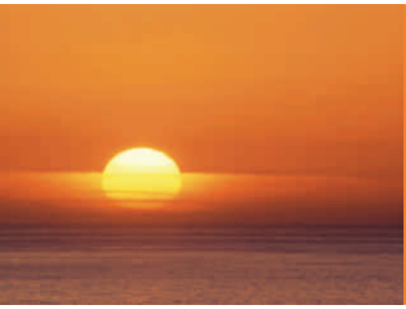
NASA

Careful study of images of this spiral galaxy (NGC 4414) taken by the Hubble Space Telescope allowed astronomers to determine that it is 60 million light-years from Earth.

are formed by dispersion in water droplets. Each color forms part of a circle about the point directly opposite the Sun. Halos are caused by the refraction of sunlight in ice crystals.

Ray diagrams can be used to locate the images formed by lenses. The rays are summarized by the following rules: (1) through center—continues; (2) parallel to optic axis—through (from) principal focal point; and (3) through (toward) other focal point—parallel to optic axis.

Cameras and our eyes contain converging lenses that produce real, inverted images. Converging lenses can be used as magnifiers of objects located inside the focal points. Lenses can be combined to make microscopes and telescopes.



CHAPTER 18 *Revisited*

The most obvious consequence of the passage of light between two transparent materials is that the direction of the light changes at the interface. This may produce virtual images that are closer, making the fish look bigger and the tree look bent. It is possible for light to get "trapped" if it is in the material with the larger index of refraction and if the angle of incidence is larger than the critical angle.

Key Terms

aberration A defect in a mirror or lens causing light rays from a single point to fail to focus at a single point in space.

critical angle The minimum angle of incidence for total internal reflection to occur.

dispersion The spreading of light into a spectrum of colors.

focal length The distance from the center of a lens to its focal point.

focal point The location at which a lens focuses rays parallel to the optic axis or from which such rays appear to diverge.

index of refraction An optical property of a substance that determines how much light bends on entering or leaving it.

optic axis A line passing through the center of a lens and both focal points.

reflecting telescope A type of telescope using a mirror as the objective.

refracting telescope A type of telescope using a lens as the objective.

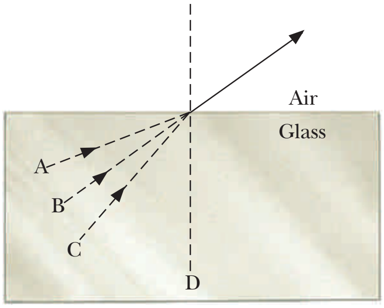
refraction The bending of light that occurs at the interface between transparent media.

total internal reflection A phenomenon that occurs when the angle of incidence of light traveling from a material with a higher index of refraction into one with a lower index of refraction exceeds the critical angle.

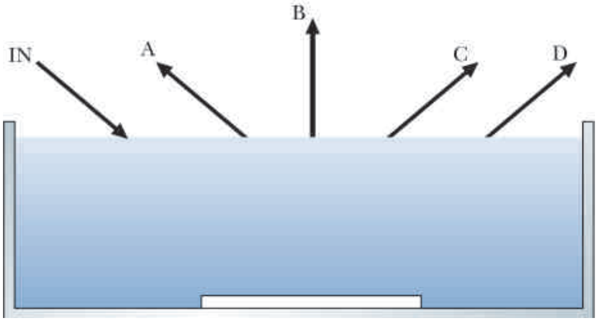
Questions and exercises are paired so that most odd-numbered are followed by a similar even-numbered.
Blue-numbered questions and exercises are answered in Appendix B.
indicates more challenging questions and exercises.
WebAssign **Many Conceptual Questions and Exercises for this chapter may be assigned online at WebAssign.**

Conceptual Questions

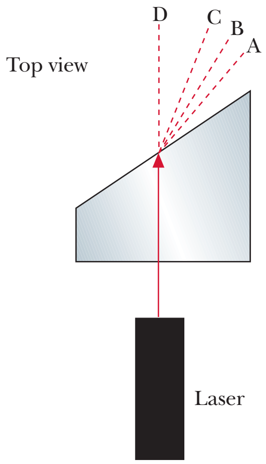
1. A narrow beam of light emerges from a block of glass in the direction shown in the following figure. Which arrow best represents the path of the beam within the glass?



2. A mirror is lying on the bottom of a fish tank that is filled with water. If IN represents a light ray incident on the top of the water, which possibility in the following figure best represents the outgoing ray?



3. You place a waterproof laser and a glass prism flat on the bottom of an empty aquarium, as shown in the following figure. The light leaving the prism follows path B. If you filled the aquarium with water, which path would the light leaving the prism now follow?



4. Figure 18-2 shows the refraction curves for air, water, and glass. If we were to draw the curve for diamond, would it appear above or below the curve for glass? Why?
5. Why do clear streams look so shallow?
6. Suppose you are lying on the bottom of a swimming pool looking up at a ball that is suspended 1 meter above the surface of the water. Does the ball appear to be closer, farther away, or still 1 meter above the surface? Explain.
7. Is the critical angle greater at a water–air surface or a glass–air surface?
8. For what range of incident angles is light totally reflected at a water–air surface?
9. Telephone companies are using "light pipes" to carry telephone signals between various locations. Why does the light stay in the pipe?
10. There is a limit to how much a fiber-optic cable can be bent before light "leaks" out because bending the pipe allows light to strike the surface at angles less than the critical angle. If you were laying fiber-optic cable under water instead of in air, would this be a greater or lesser problem? Why?
11. The distance from Earth to Mars varies from 48 million miles to 141 million miles as the two planets orbit the Sun. At which distance would Mars appear to twinkle more? Why?
12. We observe stars twinkling when viewed from Earth, but astronauts in a space shuttle do not observe this. Why?
13. As you look toward the west, you see two stars one above the other with a 5-degree separation. As the two stars move closer to the western horizon, will their apparent separation increase, decrease, or stay the same? Why?
14. In the absence of an atmosphere, a star moves across the sky from horizon to horizon at a constant speed. How does the star appear to move in the presence of an atmosphere?
15. How is the time of sunrise affected by atmospheric refraction?
16. How does the presence of an atmosphere affect the length of day and night?
17. How does the refraction of light in the atmosphere affect the appearance of the Sun or Moon as it approaches the horizon?
18. During a total lunar eclipse, the Moon lies entirely within Earth's umbra and yet is still faintly visible. If Earth lacked an atmosphere, the Moon would not be visible at

all. Explain how Earth's atmosphere allows the Moon to be seen during an eclipse.

19. You are trying to spearfish from a boat, and you spy a fish about 2 meters from the boat. Should you aim high, low, or directly at the fish? Why?

20. If you were going to send a beam of light to the Moon when it is just above the horizon, would you aim high, low, or directly at the Moon? Explain.

21. Does a beam of white light experience dispersion as a result of reflecting from a mirror? Explain.

22. Why is a diamond more brilliant than a clear piece of glass having the same shape?

23. A fiber-optic cable is used to transmit yellow light. At one sharp bend, the incident angle is exactly at the critical angle for yellow light. If the pipe were bent any more, the yellow light would "leak" out. If the cable were used with blue light, would this bend be a problem? What if it were used with red light? Explain.

24. While transmitting white light down a fiber-optic cable, you bend the cable too much in one place, and some of the light "leaks" out. Which is the first color of light to "leak" out?

25. If your line of sight to a water droplet makes an angle of 41 degrees with the direction of the sunlight, what color would the raindrop appear to be?

26. You are looking at a rainbow from the ground floor of an apartment building and notice that a kite is hovering right in the green portion. If you were to go up to the second floor, would you be likely to see the kite hovering in the red portion or the blue portion of the rainbow? Explain.

27. Why is a shadow of your head always in the center of a rainbow?

28. If you were flying in an airplane and saw a rainbow at noon, what shape would it be, and where would you be looking?

29. At what time of day might you expect to see the top of a rainbow set below the horizon? In what direction would you look to see it?

30. At what time of day and in what direction would you look to see the top of a rainbow rise above the horizon?

31. To produce a hologram, a narrow beam of laser light must be spread out enough to expose the surface of a piece of film. What type of lens would accomplish this? Explain your reasoning.

32. What type of lens would be helpful in starting a campfire? Why?

33. You find a converging lens in the storeroom and wish to determine its focal length. Describe how you could use two lasers to accomplish this.

34. You place a laser at the principal focus of a converging lens and aim the beam toward any part of the lens. Describe the beam's path after passing through the lens.

35. You find a diverging lens in the storeroom and wish to determine its focal length. Describe how you could use two lasers to accomplish this.

36. You wish to use a diverging lens to redirect the light from six lasers to produce beams that are parallel to the optic axis of the lens. How should you aim the lasers to accomplish this?

37. What type of lens would you use to construct an overhead projector? Explain your reasoning.

38. Can a prism be used to form an image? Explain.

39. Where does a ray arriving parallel to the optic axis of a converging lens go after passing through the lens?

40. What path does a ray take after passing through a converging lens if it passes through a focal point before it enters the lens?

41. A converging lens is used to form a sharp image of a candle. What effect does covering the upper half of the lens with paper have on the image?

42. How does covering all but the center of a lens affect the image of an object?

43. Two converging lenses with identical shapes are made from glasses with different indexes of refraction. Which one has the shorter focal length? Why?

44. You are building a device in which the dimensions of a diverging lens cannot be changed. On testing the device, however, you discover that the focal length of your lens is too short. You decide to fix this problem by grinding a lens of identical shape from different glass. Should you use a glass with a larger or a smaller index of refraction? Why?

45. Consider the image of a candle as shown in Figure 18-21. Explain why you would not be able to see the image if your eye were located to the left of the image location.

46. Consider the image of a candle as shown in Figure 18-21. To clearly see the image, should you locate your eye at the image location or to the right of the image location? Explain.

47. What kind of image is formed on the retina of an eye?

48. How might you convince a friend that the image formed by a camera is a real image?

49. What is the purpose of the pupil in the eye?

50. What is the purpose of a diaphragm in a camera?

51. Why does a telescope that uses a mirror to focus the light not exhibit chromatic aberration whereas a telescope that uses a lens to focus the light does?

52. You measure the focal length of a converging lens by finding the crossing point of parallel beams of red laser light. If you measure the focal length using lasers that emit green light, would the measured value be greater than, less than, or the same as before? Why?

53. Are reading glasses used by older people converging or diverging? Explain.

54. Without glasses, the light entering our eyes comes directly from the object. With glasses, the light comes from the image formed by the glasses. Is this image real or virtual? Explain.

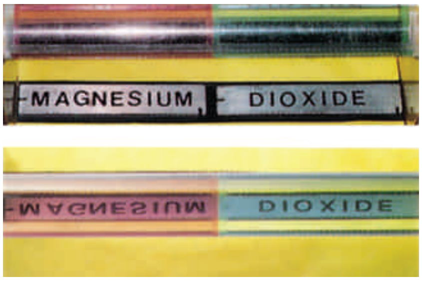
55. When a person has cataract surgery, the lens of the eye is removed and replaced with a plastic one. Would you expect this lens to be converging or diverging? Explain.

56. Stamp and coin collectors often wear special glasses that allow them to see the details of the stamps and coins. Are the lenses in these glasses converging or diverging? Why?

57. The ray diagram for a magnifying glass is shown in Figure 18-22. As the object is moved toward the focal point, the direction of emerging ray 2 does not change, whereas the direction of emerging ray 1 does. By looking at where these rays now intersect, determine whether the image gets larger or smaller as the object is moved closer to the focal point.

58. If the magnifying lens in Figure 18-22 were replaced with a lens of shorter focal length, the direction of emerging ray 1 would not change, whereas the direction of emerging ray 2 would. By looking at where these rays would now intersect, determine whether the magnification would be increased or decreased.

59. The figure shows the words MAGNESIUM DIOXIDE viewed through a solid plastic rod. Why does MAGNESIUM appear upside down, while DIOXIDE appears right-side up?



60. When a single converging lens is used to focus white light, the image has a colored fringe due to chromatic aberration. Describe the changes in the color of the fringe as a screen is moved through the focal point.

Exercises

61. If light in air is incident at 30°, at what angle is it refracted in water? In glass?

62. Light in air is incident on a surface at an angle of 60°. What is its angle of refraction in glass? In water?

63. You are spearfishing in waist-deep water when you spot a fish that appears to be down at a 45° angle. You recognize that the light coming from the fish to your eye has been refracted and you must therefore aim at some angle below the apparent direction to the fish. What is this angle?

64. Light from the bottom of a swimming pool is incident on the surface at an angle of 30°. What is the angle of refraction?

65. Use Figure 18-2 to estimate the critical angle for glass with an index of refraction of 1.6.

66. Use Figure 18-2 to estimate the angle of refraction for light in air incident at 50° to the surface of glass with an index of refraction of 1.6.

67. A prism made of glass with an index of refraction of 1.5 has the shape of an equilateral triangle. A light ray is incident on one face at an angle of 48°. Use a protractor and Figure 18-2 to find the path through the prism and out an adjacent side. What is the exit angle?

68. A prism made of plastic with an index of refraction of 1.33 has the shape of a cube. A light ray is incident on one face at an angle of 70°. Use a protractor and Figure 18-2 to find the path through the prism and out an adjacent side. What is the exit angle?

69. You are scuba diving below a fishing pier. You look up and see a fishing pole that appears to be 2 m above the surface of the water. Use a ray diagram to show that the pole is actually closer to the surface of the water.

70. Using data from Figure 18-2 to determine the exact angles, redraw Figure 18-4(c) to locate the image of something sitting on the bottom of a pond. Use your scale diagram and a ruler to show that an object in water appears to be about $\frac{3}{4}$ as deep as it actually is. (Note that the index of refraction for water is $\frac{4}{3}$.)

71. An object is located midway between the other focal point and the center of a converging lens. Draw a ray diagram showing how you locate the image. Estimate the magnification of the image from your diagram.

72. Use a ray diagram to find the location and magnification of the image of an object located three focal lengths from a converging lens.

73. Use a ray diagram to locate the image of an arrow placed 60 cm from a diverging lens with a focal length of 30 cm.

74. The focal length of a converging lens is 30 cm. Use a ray diagram to locate the image of an object placed 60 cm from the center of this lens.

75. Over what range of positions can an object be located so that the image produced by a converging lens is real and smaller than the object?

76. Over what range of positions can an object be located so that the image produced by a converging lens is real and magnified?

77. In Figure 18-21 a weakly diverging lens is inserted at the principal focal point of the converging lens. Use a ray diagram to show that this results in the real image being shifted to the right.

78. Use ray diagrams to show that a diverging lens cannot produce a magnified image of an object.

79. Draw a ray diagram to locate the image of an object placed inside the focal point of a diverging lens. Is the image real or virtual? Erect or inverted? Magnified or reduced in size? Explain.

80. What is the location of the image of an object placed at the focal point of a diverging lens? Is the image real or virtual? Erect or inverted? Magnified or reduced in size? Explain.

81. How many diopters are there for a converging lens with a focal length of 0.4 m?

82. If a lens has a focal length of 25 cm, how many diopters does it have?

83. A converging lens of focal length 20 cm is placed next to a converging lens of focal length 50 cm. What is the effective focal length for this combination?

84. What focal-length lens would you need to place next to a converging lens of focal length 25 cm to create an effective focal length of 20 cm for the combination?

85. A converging lens of focal length 50 cm is placed next to a diverging lens of length 25 cm. What is the effective focal length for this combination? Is it diverging or converging?

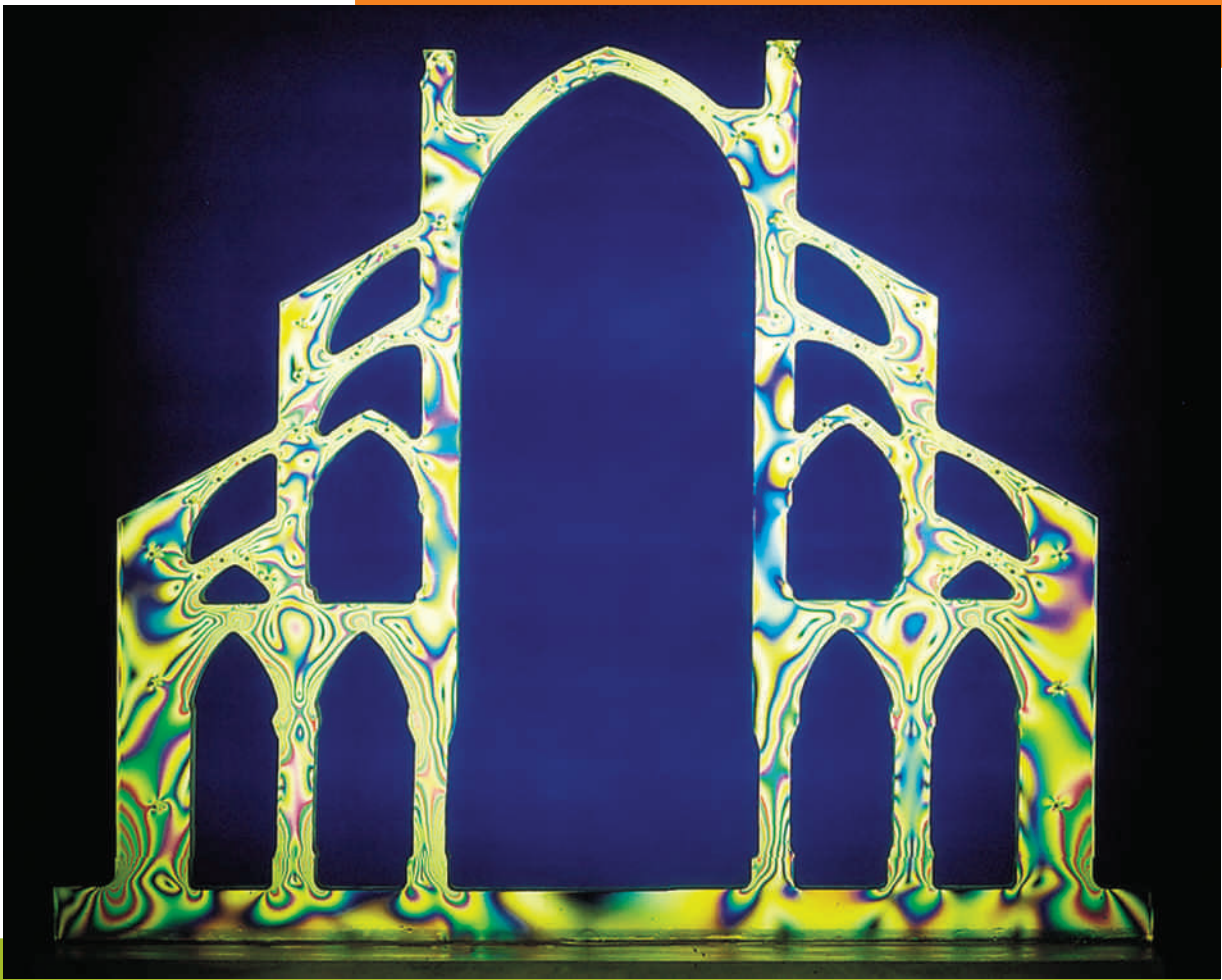
86. You ordered a converging lens of focal length 2 m, but the company delivered a lens whose focal length was only 1 m. You don't have time to wait for a replacement, so you decide to correct the problem by placing a diverging lens next to the lens they sent you. What must the focal length be for this diverging lens?

19

A Model for Light

▶ Light is one of our most common everyday phenomena, and our knowledge of light is responsible for our understanding of the beautiful effects in soap bubbles and for technological advances into extending human perception. But what is light?

(See page 416 for the answer to this question.)



Robert Mark, Princeton University

Viewing a plastic model of a structure through crossed pieces of Polaroid filter shows the stress patterns in the structure. This photograph shows a model of the flying buttresses of the Notre Dame cathedral.

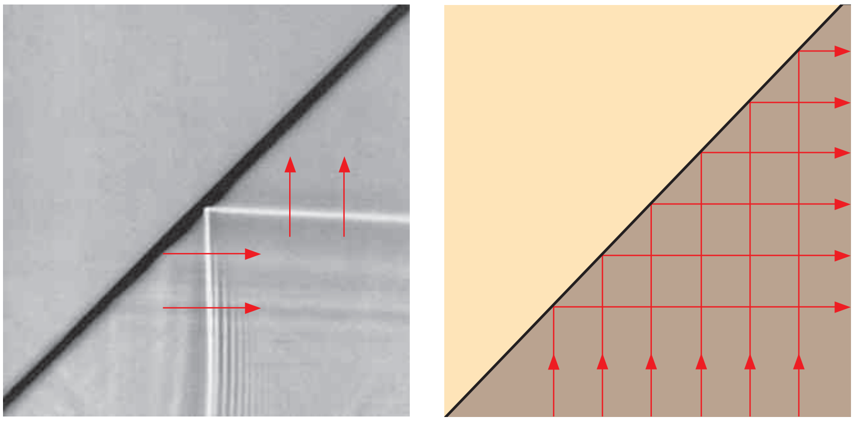
IN the previous two chapters, we learned a great deal about light by simply observing how it behaves, but we did not ask, "What is light?" This question is easy to ask, but the answer is more difficult to give. For example, we can't just say, "Let's look!" What we "see" is the stimulation caused by light entering our eyes, not light itself. To understand what light is, we need to look for analogies, to ask ourselves what things behave like light. In effect, we are building a model of a phenomenon that we can't observe directly. This same problem occurs quite often in physics because many components of nature are not directly observable. In the case of light, we need a model that accounts for the properties that we studied in the previous two chapters. When we look at the world around us, we see that there are two candidates: *particles* and *waves*.

So the question becomes, does light behave as if it were a stream of particles or a series of waves? Newton thought that light was a stream of particles, but other prominent scientists of his time thought that it behaved like waves. Because of Newton's great reputation, his particle model of light was the accepted theory during the 18th century. However, many of the early observations could be explained by a particle model *or* a wave model. Scientists continually looked for new observations that could distinguish between the two theories. We will examine the experimental evidence to see whether light behaves like particles or waves. The process of deciding is as important as the answer.

Reflection

We know from such common experiences as echoes and billiards that both waves and particles can bounce off barriers. But this fact is not enough. Our model for the behavior of light must agree with our conclusion that the angle of reflection is equal to the angle of incidence. A particle model for light can account for this if the reflecting surface is frictionless and perfectly elastic. A wave model of light also has no problem accounting for the law of reflection. The photograph in Figure 19-1 (a) shows a straight wave pulse on the surface of water striking a smooth, straight barrier. The wave was initially moving toward the top of the picture and is being reflected toward the right. We see that the angle between the incident wave and the barrier is equal to that between the reflected wave and the barrier. In the corresponding ray diagram [Figure 19-1 (b)], we draw the rays perpendicular to the straight wave fronts—that is, in the direction the wave is moving.

Observing that light reflects from surfaces gives us no clues as to its true nature; both particles and waves obey the law of reflection.



(*a*) (*b*)

Figure 19-1 (a) The horizontal white line is a straight water-wave pulse moving toward the top of the picture. The part that has already hit the black barrier is reflected toward the right. (b) The corresponding ray diagram has rays perpendicular to the wave fronts and shows that the angle of reflection is equal to the angle of incidence.

▶ Extended presentation available in the *Problem Solving* supplement

Refraction

We learned a great deal in the previous chapter about the behavior of light when it passes through boundaries between different transparent materials. Newton thought that the particles of light experienced a force as they passed from air into a transparent material. This inferred force would occur only at the surface, act perpendicular to the surface, and be directed into the material. This force would cause the particles to bend toward the normal, as shown in Figure 19-2. In this scheme the light particles would also experience this force on leaving the material. Because the force acts into the material, it now has the effect of bending the particles away from the normal. Furthermore, when Newton calculated the dependence of the amount of **refraction** on the angle of incidence, his answer agreed with the curves in Figure 18-2. So the model gives not only the correct qualitative results but also the correct quantitative results.

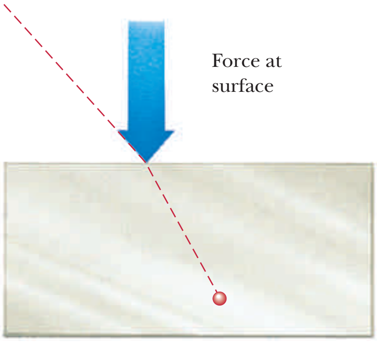


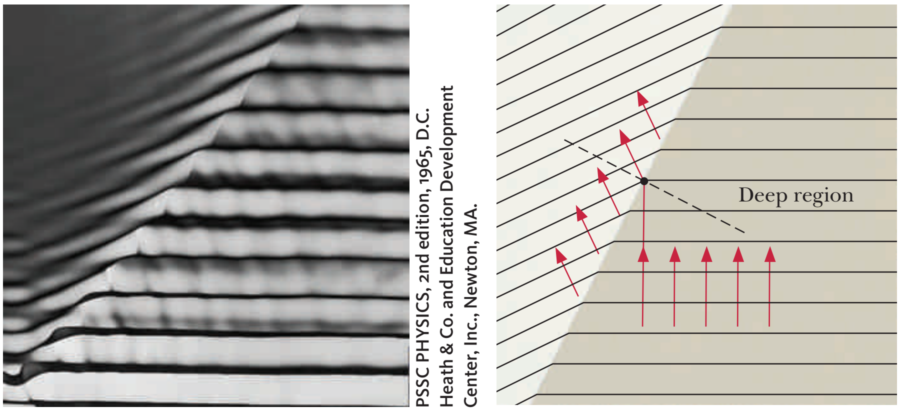
Figure 19-2 A force perpendicular to the surface would cause light particles to be refracted.

Water waves also refract. The photograph in Figure 19-3 shows the refraction of water waves. The boundary that runs diagonally across the photograph separates the shallow region on the left from the deeper region on the right. Once again, the numerical relationship between the angle of refraction and the angle of incidence is in agreement with the law of refraction.

There is an important difference between the wave and particle predictions; the speeds of the particles and waves do not change in the same way. Figures 19-2 and 19-3 both correspond to light entering a substance with a higher index of refraction. In the particle model, particles bend toward the normal because they speed up. Thus, the particle model predicts that the speed of light should be faster in substances with higher indexes of refraction.

In waves the opposite is true. Because the crests in Figure 19-3 are continuous across the boundary, we know that the frequency of the waves doesn't change. However, the wavelength does change; it is shorter in the shallow region to the left of the boundary. Because $v = \lambda f$ (Chapter 15), a decrease in the wavelength means a decrease in the speed of the wave. This means that the speed of the waves in the shallow region is smaller, and therefore the wave model predicts that the speed of light should be slower in substances with higher indexes of refraction.

Because the two models predict opposite results, we have a way of testing them; the speed of light can be measured in various materials to see which model agrees with the results. The speed of light in a material substance was not measured until 1862, almost two centuries after the development of the two theories. French physicist Jean Foucault measured the speed of light in



(a) *(b)*

Figure 19-3 (a) Water waves refract when passing from deep to shallow water. (b) The corresponding ray diagram shows that the waves refract toward the normal.

air and water and found the speed in water to be less. This dealt a severe blow to the particle model of light and consequently caused a modification of the physics world view.

◀ light travels slower in materials

American physicist Albert Michelson improved on Foucault's measurements and found a ratio of 1.33 for the speed of light c in a vacuum to the speed of light v in water. This value is equal to the **index of refraction** n of water as predicted by the wave model, and thus $n = c/v$. Because the speed of light in a vacuum is the maximum speed, the indexes of refraction of substances must be greater than 1. This gives us a way of determining the speed of light in any material once we know its index of refraction. The speed of light in a substance is equal to its speed in a vacuum divided by the index of refraction, $v = c/n$.

◀ $n = c/v$

In the previous chapter, we discovered that different colors have slightly different indexes of refraction within a material, which results in **dispersion.** Because the index of refraction is related to the speed, we can now conclude that different colors must have different speeds in a material.

Q: Does red or blue light have the slower speed in glass?

A: Because blue light is refracted more than red, it has the higher index of refraction. Therefore, blue light has a slower speed.

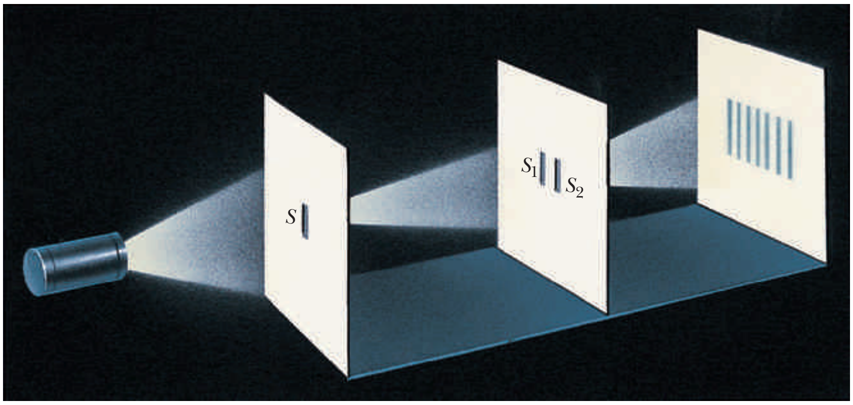
Interference

Although the speed of light in materials provided definitive support early on for the fact that light behaves like a wave, there was other supporting evidence in the debate. Newton's main adversary in this debate was largely unsuccessful at convincing others of the importance of the evidence. We now examine some of the other properties of waves that differ from those of particles and show that light exhibits these properties.

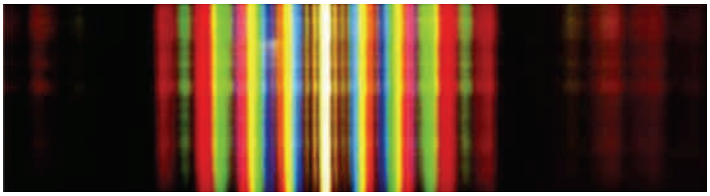
We begin with interference. The interference of two sources of water waves is shown in the photograph and drawing in Figure 15-20. Attempts to duplicate these results with two lightbulbs, however, fail. But if light is a wave phenomenon, it should exhibit such interference effects. Light from the two sources does superimpose to form an interference pattern, but the pattern isn't stationary. Stationary patterns are produced only when the two sources have the same wavelength and a constant phase difference; that is, when the time interval between the emission of a crest from one source and the emission of a crest from the second source does not change. Because the phase difference between the lightbulbs varies rapidly, the pattern blurs out, and the region looks uniformly illuminated.

The interference of two light sources was first successfully demonstrated by Thomas Young in 1801 (more than a half century before Foucault's work on the speed of light in materials), when he let light from *one* pinhole impinge on two other pinholes [Figure 19-4(a)]. Passing the light through the first pinhole produced light at the second pinholes that was reasonably in phase. The pattern shown in Figure 19-4(b) consists of colored bars formed by light from the two slits interfering to form antinodal regions. These antinodal regions have large amplitudes and appear bright. Each color forms its own set of colored bars. Antinodal regions for the different rainbow colors superimpose to produce the colors we perceive. Modern versions of this experiment are in complete agreement with the wave model.

Figure 19-4 (a) A schematic of Young's experiment that demonstrated the interference of light. (b) The interference pattern produced by white light incident on two slits.



(a)



(b)

Courtesy of Gregory Francis

If the slits are illuminated with a single color, the pattern looks much like the one you would expect to find along the far edge of the ripple tank in Figure 15-20. As the color of the light changes, the pattern on the screen changes size (Figure 19-5). Red light produces the widest pattern, and violet light produces the narrowest one. The sizes produced by other colors vary in the same order as the colors of the rainbow.

Q: What happens to the width of one of these interference patterns if the distance between the two slits is increased?

A: We learned in Chapter 15 that the spacing of the nodal lines depends on the ratio of the wavelength to the slit separation. Therefore, the wider spacing will produce a narrower pattern.

Recall that in looking at the two-source interference patterns in Chapter 15, the nodal lines on each side of the central line were created by a difference in the path lengths from the two sources that was equal to one-half wavelength (Figure 15-21). Increasing the wavelength causes these nodal lines to move farther away from the central line; the pattern widens. Therefore, the shifting of the dark regions on our screen with color can be explained by a change in

Figure 19-5 The two-slit interference pattern produced by red light (a) is wider than that produced by blue light (b).

(a)

(b)

Courtesy of Gregory Francis (both)

wavelength. We conclude that the color of light depends on its wavelength, with red being the longest and violet the shortest.

Measurements of the interference pattern and the separation of the slits can be used to calculate the wavelength of the light. Experiments show that visible light ranges in wavelength from 400 to 750 nanometers (nm), where a nanometer is 10^{-9} meter. It takes more than 1 million wavelengths of visible light to equal 1 meter. Knowing the speed of light to be 3.0×10^8 meters per second, we can use the relationship $c = \lambda f$ to calculate the corresponding frequency range to be roughly $(4.0 \text{ to } 7.5) \times 10^{14}$ hertz.

You can perform Young's two-slit experiment by taping the slits to one end of the cardboard tube from a roll of paper towels. To make the two slits, blacken a microscope slide (or other piece of glass) with a candle flame. Scratch two viewing slits on the blackened side by holding two razor blades tightly together and drawing them across the slide. Cut a slit about 2 millimeters wide in a piece of paper and tape it to the other end of the tube. Aim this end of the tube at a light source and look through the end with the two slits to see the interference pattern (Figure 19-6).

You can vary the wavelength of the light by using pieces of red and blue cellophane as filters. Putting a thin spacer between the razor blades will vary the spacing.

Figure 19-6 A device for viewing two-slit interference.

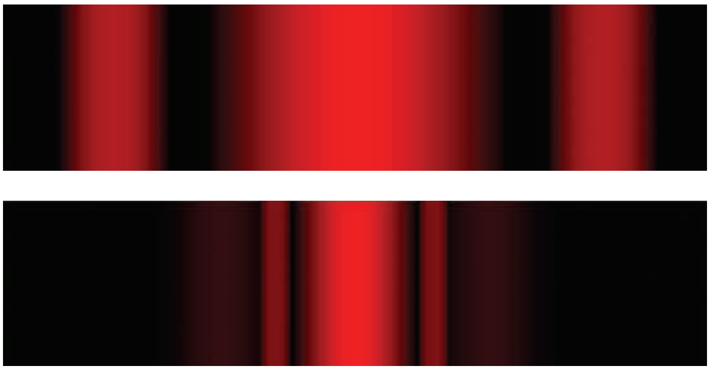
Diffraction

Young's experiment points out another aspect of waves. His interference pattern was possible only because light from one pinhole overlapped that from the other. Light spreads out as it passes through the pinholes. In other words, light exhibits diffraction just like the water waves of Figure 15-23.

The photographs in Figure 19-7 were taken of the diffraction pattern produced by light passing through a narrow slit. The slit was wider in Figure 19-7(b)! But this difference in the patterns makes sense if you examine the photographs of water waves in Figure 15-23. Contrary to what common sense may tell us, the narrower slit produces the wider pattern. These patterns are in agreement with the wave model and contrary to what we would expect if light traveled only in straight lines. In the case of particles, the wider slit would produce the wider pattern, not the narrower one.

Q: Does red light or blue light produce the wider diffraction pattern?

A: The width of the diffraction pattern depends on the ratio of the wavelength to the width of the slit. Because red light has the longer wavelength, it would produce the wider pattern.



(a)

(b)

Courtesy of Gregory Francis (both)

Figure 19-7 Diffraction patterns produced by red light incident on (a) a narrower and (b) a wider slit. The narrower slit produced the wider pattern.

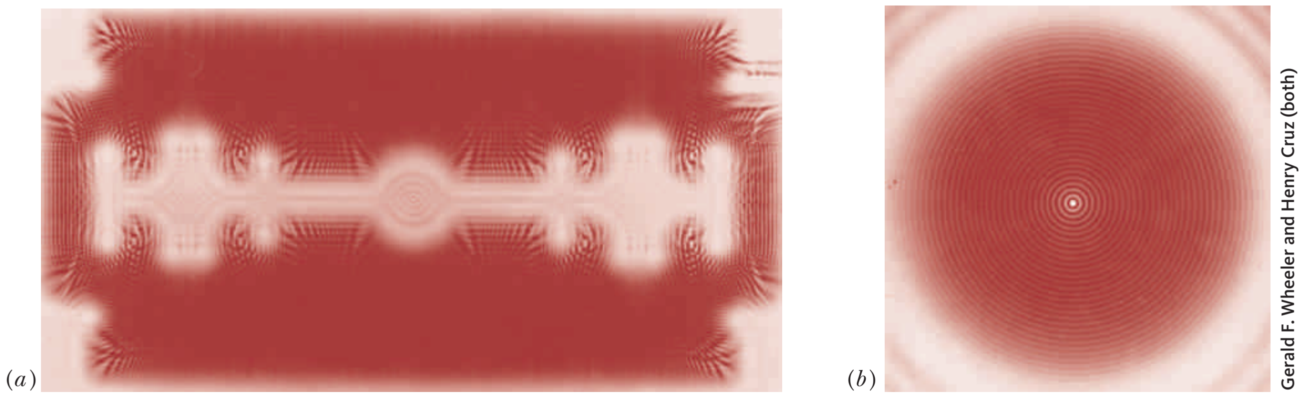


Figure 19-8 Photographs of diffraction patterns in the shadows of (a) a razor blade and (b) a penny.

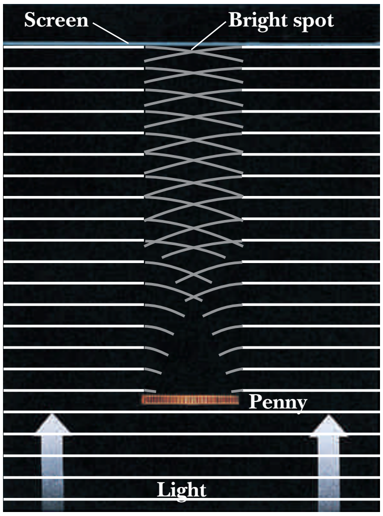


Figure 19-9 Diagram illustrating how wave interference and diffraction produce a bright spot at the center of the penny's shadow.

Water waves also show that diffraction takes place around the edges of barriers. To observe this effect with light, we should look along the edge of the shadow of an object. This effect is usually not observed with light because most shadows are produced by broad sources of light. The resulting shadows have smooth changes from umbra through penumbra to full brightness. Therefore, we should look at shadows produced by point sources to eliminate the penumbra.

The photographs in Figure 19-8 were created by putting photographic paper in the shadows of a razor blade and a penny. We used red light from a laser because the light waves are in phase and can easily be made to approximate a point source. You can see that the edges of the shadows show the effects of the interference of diffracted light. Only a wave model of light can explain these effects.

Notice the center of the shadow of the penny. Even though the penny is solid, the shadow has a bright spot in the center. Can the wave model explain this bright spot? The diagram in Figure 19-9 shows how the light diffracts around the edge of the penny. The light coming from each point on the edge travels the same distance to the center of the shadow. These waves arrive in phase and superimpose to form the bright spot. This is added support for the wave model of light.

It is possible to observe the diffraction of light in everyday situations. Look at a lightbulb through the slit between two fingers held in front of your eyes, as in Figure 19-10. How does the image of the lightbulb change as you slowly change the spacing between your fingers? Notice the dark, vertical lines in the image. What happens to those lines as your fingers move closer together?

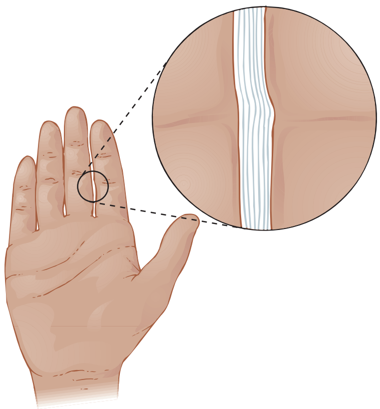


Figure 19-10 Light viewed through the thin slit between two fingers exhibits diffraction effects.

Thin Films

Soap bubbles, oil slicks, and thin films of any transparent material exhibit beautiful arrays of color under certain conditions—an effect not of dispersion but of interference.

Consider a narrow beam of red light incident on a thin film, as shown in Figure 19-11. At the first surface, part of the light is reflected and part is transmitted. The same thing happens to the transmitted part when it attempts to exit the film; part is transmitted and part is reflected, and so on. Considering just the first two rays that leave the film on the incident side, we can see that

Everyday Physics | *Diffraction Limits*

The wave nature of light and the size of the viewing instrument limit how small an object we can see, even with the best instruments. Because of the wave nature of light, there is always some diffraction. The image is spread out over a region in space; the region is larger for longer wavelengths and smaller openings.

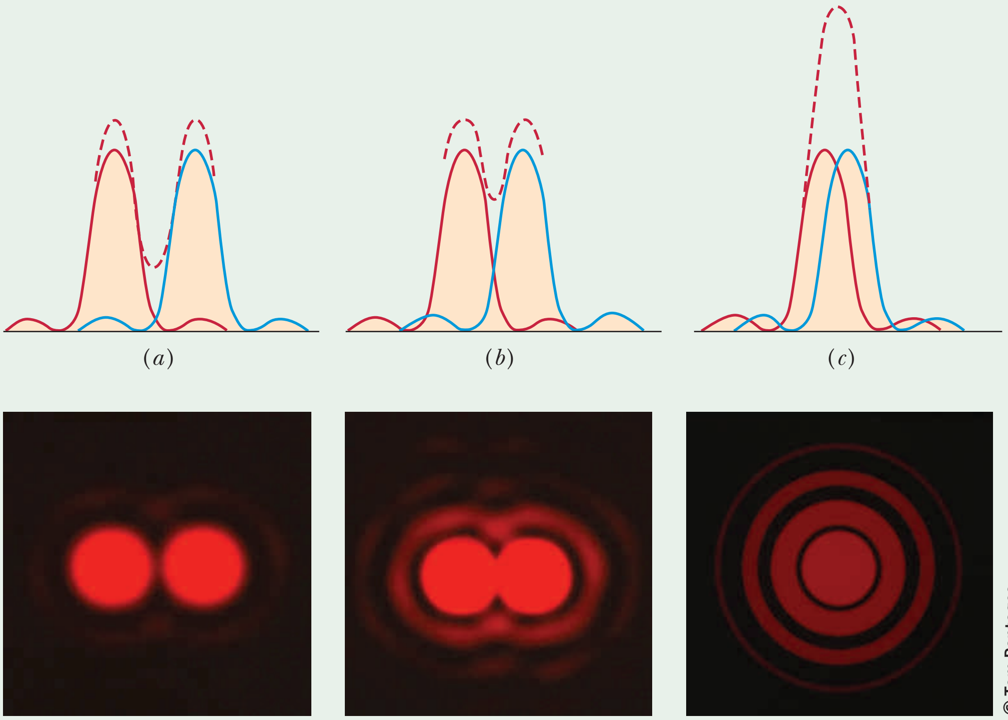
Consider two objects with small *angular* sizes (they could be very big but so far away that they look small) separated by a small angular distance. Each of these will produce a diffraction pattern when its light passes through a small opening such as in our eye or a telescope. The diagram and photograph in (a) correspond to the case in which the diffraction patterns can still be clearly distinguished. In (c) the overlap is so extensive that you cannot resolve the individual images. The limiting case is shown in (b); the separation of the centers of the two patterns is less than the width of the central maximum of either pattern. At this limit the first minimum of each pattern lies on top of the maximum of the other pattern.

This minimum angular separation can be calculated mathematically. The light coming from two point sources can be resolved if the angle (measured in radians) between the objects is greater than 1.22 times the ratio of the wavelength of the light to the diameter of the aperture. Imagine viewing an object from a distance of 25 centimeters, the nominal distance of closest vision. If we use an average pupil size of 5 millimeters and visible light with a wavelength of 500 nanometers, we calculate that you can distinguish a separation of 0.03 millimeter, about the radius of a human hair. A similar calculation tells us that we should be able to distinguish the headlights of an oncoming car at a distance of 10 kilometers.

From this relationship you can see why astronomers want bigger telescopes. With a larger mirror, the resolving angle is smaller, and they can distinguish more detail in distant star clusters. For instance, the resolving angle of a 5-meter telescope in visible light is about 0.02 arc second, where 1 arc second is $\frac{1}{3600}$ degree. In practice, this resolution is never obtained because turbulence in Earth's atmosphere limits the resolution to about 1 arc second. This is one of the major reasons for placing the Hubble Space Telescope in orbit.

Diffraction effects also place lower limits on the sizes of objects that can be examined under an optical microscope because the details to be observed must be separated by more than the diffraction limits set by the microscope.

1. Is it better to use red light or blue light to minimize diffraction effects while photographing tiny objects through a microscope? Why?
2. Why are the diffraction effects of your eyes more important during the day than at night?



© Tom Pantages

Two overlapping diffraction patterns can just barely be resolved (b) if the central maximum of each pattern lies on the first minimum of the other. If the patterns are closer, they appear to be a single object (c).

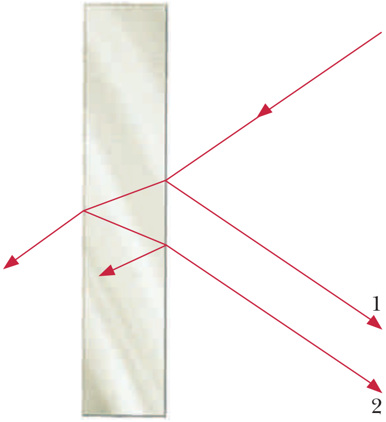


Figure 19-11 Light incident on a thin film is reflected and transmitted at each surface. Rays 1 and 2 interfere to produce more or less light, depending on the thickness of the film and the wavelength of the light in the film.

they interfere with each other. Because of the thickness of the film, ray 2 lags behind ray 1 when they overlap on the incident side of the film. Ray 2 had to travel the extra distance within the film. At a certain thickness, this path difference is such that the crests of one ray line up with the troughs from the other and the two rays cancel. In this case little or no light is reflected. At other thicknesses, the crests of one ray line up with the crests from the other, and the two rays reinforce. In this case light is reflected from the surface.

Because different colors have different wavelengths, a thickness that cancels one color may not be the thickness that cancels another. If the film is held vertically (Figure 19-12), the pull of gravity causes the film to vary in thickness, being very thin at the top and increasing toward the bottom. In fact, when the film breaks, it begins at the top where it has been stretched too thin. If the film varies in thickness and the incident light is white, different colors will be reflected for different thicknesses, producing the many colors observed in soap films.

Notice that the very thin region at the top of the soap film in Figure 19-12 has no light reflecting from the film. At first this seems confusing. Because there is essentially no path difference between the front and back surfaces, we may expect light to be reflected. However, the first ray is inverted when it reflects; that is, crests are turned into troughs, and vice versa. This process is analogous to the inversion that takes place when a wave pulse is reflected from the fixed end of a rope (Chapter 15). Light waves are inverted when they reflect from a material with a higher index of refraction. At the back surface, the rays reflect from air (a lower index of refraction than the soap), and no inversion takes place.

Are You On the Bus?

Q: Why does part of the thin film in Figure 19-12 appear white when white is not one of the rainbow colors?

A: We observe white when most of the colors overlap.

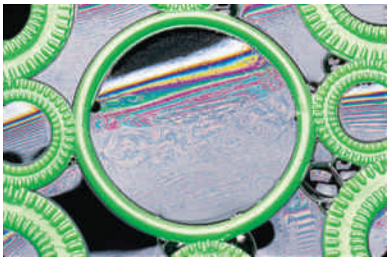


Figure 19-12 Different colors are reflected from different thicknesses of the thin film. The very thin region at the top of the soap film appears black because the reflected light is inverted at the front surface but not at the back surface.

Because the ray reflected from the front surface is inverted while the ray reflected from the back surface is not inverted, a crest from the front surface will overlap a trough from the back surface if the thickness of the film is much smaller than a wavelength. The two reflected rays will cancel, and no light will be reflected. If the light is normal to the surface and the thickness of the film is increased until it is one-quarter of a wavelength thick, the light will now be strongly reflected. The ray that reflects from the back surface must travel an extra one-half wavelength, and these crests and troughs will be delayed so that they now line up with those reflected from the front surface. It is important to note that we are referring to the wavelength of the light *in the film*. Because the frequency of the light is the same in the air and in the film while the speed is reduced, the wavelength in the film is equal to that in the air divided by the index of refraction, $\lambda_f = \lambda/n$.

This colorful phenomenon also occurs after rain has wetted the highways. Oil dropped by cars and trucks floats on top of the puddles. Sunlight reflecting from the top surface of the oil and from the oil–water interface interferes to produce the colors. Again, variations in the thickness of the oil slick produce the array of colors.

Interference also occurs for the transmitted light. Light passing directly through the thin film (Figure 19-13) can interfere constructively or destructively with light that is reflected twice within the film. The effects are complementary to those of the reflected light. When the thickness of the film is chosen to minimize the reflection of a certain color of light, the transmission

WORKING IT OUT Thin Film

You are coating a glass lens of index of refraction 1.5 with a film of material of index of refraction 1.6. You start with the thinnest film possible that creates a strong reflection for 600-nm (orange) light. You gradually increase the film thickness until you again get strong reflection. What is the thickness of the film now? How would your answer change if the glass lens had an index of refraction of 1.7?

The light rays are inverted when they reflect at the front surface of the film (going from index of refraction of 1.0 in air to 1.6 in the film). The rays are not inverted at the back surface of the film (going from index of refraction of 1.6 in the film to 1.5 in the glass). The thinnest possible film that will yield strong reflection is therefore $\lambda/4$, where λ is the wavelength of the light in the film. The next strong reflection will occur when the extra path length in the film is one whole wavelength. This means increasing the thickness of the film by half a wavelength:

$$t = \frac{\lambda}{4} + \frac{\lambda}{2} = \frac{3\lambda}{4} = \frac{3\lambda_{air}}{4n} = \frac{3(600\ \text{nm})}{4(1.6)} = 281\ \text{nm}$$

If the glass lens instead has an index of refraction of 1.7, greater than that of the film, then the light rays will invert at both surfaces. The first strong reflection will now occur when the film is half a wavelength thick. The next strong reflection will again occur when the path length in the film is increased by one whole wavelength. This means increasing the thickness of the film by half a wavelength:

$$t = \frac{\lambda}{2} + \frac{\lambda}{2} = \lambda = \frac{\lambda_{air}}{n} = \frac{600\ \text{nm}}{1.6} = 375\ \text{nm}$$

FLAWED REASONING

A factory worker is coating camera lenses with a special film to filter out red light by reflecting the red light from the surface. The index of refraction of the film is greater than that of the lens. The boss has instructed him to make the film 110 nanometers thick. The worker decides that "more is always better" and that the boss should not be so stingy with the coating material. He therefore makes the coating twice as thick. **Will the customers get a better red filter than they bargained for?**

ANSWER The boss is not being stingy. The coating filters red light if its thickness is one-quarter the wavelength of the red light in the material. This causes the light that reflects from the back of the coating to be in phase (crest lined up with crest) with the light that reflects from the front of the coating. If the thickness of the coating is doubled, these two reflections will be out of phase, and the red light will pass through the lens instead of being reflected.

of that color is maximized. This process is a consequence of the conservation of energy. The light must go someplace.

Some modern office buildings have windows with thin films to reduce the amount of light entering the offices. The visors in the helmets of space suits are coated with a thin film to protect the eyes of the astronauts. Thin films are also important to lens makers because the proper choice of material and thickness allows them to coat lenses so that they do not reflect certain colors or conversely so that they do not transmit certain colors. It is common

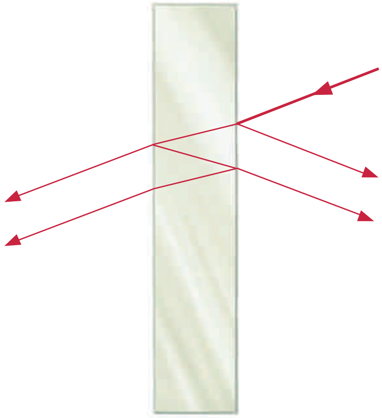


Figure 19-13 The light transmitted through a thin film also displays interference effects.

The colors on the puddle are caused by the interference of light reflected from the top and bottom of a thin oil film floating on the surface.

The thin film on the visor protects the astronaut's eyes.

to coat the lenses in eyeglasses with a thin film to stop the transmission of ultraviolet light. Lenses in high-quality telescopes and binoculars are coated to reduce the reflection (and therefore enhance the transmission) of visible light, increasing the brightness of images.

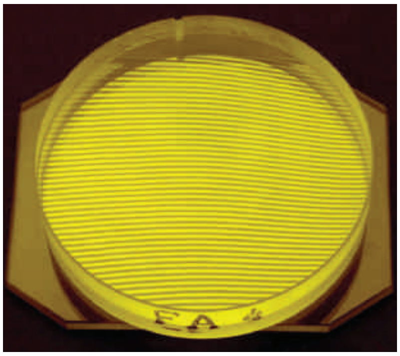
An interesting example of thin-film interference was observed by Newton (even though he did not realize that it supported the wave model of light). When a curved piece of glass such as a watch glass is placed in contact with a flat piece of glass, a thin film of air is formed between the two. An interference pattern is produced when the light reflects from the top and bottom of the air gap. The interference pattern that is formed by a wedge-shaped air gap is shown in Figure 19-14. Such patterns are commonly used to test the surface quality of lenses and mirrors.

Are You On the Bus?

Q: Would you expect to find a dark spot or a bright spot when you look at the reflected light from an air gap of almost zero thickness?

A: The light waves will be inverted when they reflect off the second interface, but not the first one. Therefore, you would see a dark spot in the reflected light and a bright spot in the transmitted light.

Figure 19-14 (a) Interference lines are produced by light reflecting from the top and bottom surfaces of an air wedge between two pieces of glass.



(*a*)

Image not available due to copyright restrictions

Polarization

We have established that light is a wave phenomenon, but we have not discussed whether it is transverse or longitudinal—that is, whether the vibrations take place perpendicular to the direction of travel or along it. We will determine this by examining a property of transverse waves that does not exist for longitudinal waves and then see whether light exhibits this behavior.

Transverse waves traveling along a horizontal rope can be generated so that the rope vibrates in the vertical direction, the horizontal direction, or any direction in between. If the vibrations are in only one direction, the wave is said to be *plane polarized,* or often just **polarized.** This property becomes important when a wave enters a medium in which various directions of polarization are not treated the same. For instance, if our rope passes through a board with a vertical slit cut in it (Figure 19-15), the wave passes through the slit if the vibration is vertical but not if it is horizontal.

What happens if the slit is vertical but the polarization of the wave is someplace between vertical and horizontal? Imagine that you are looking along the length of the rope, and we represent the polarization of the wave by an arrow (a vector) along the direction of vibration. This polarization can be

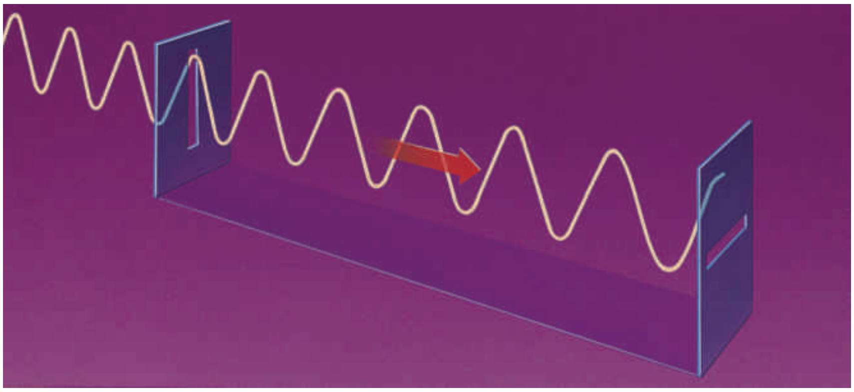


Figure 19-15 Waves with vertical polarization pass through a vertical slit, but not through a horizontal one.

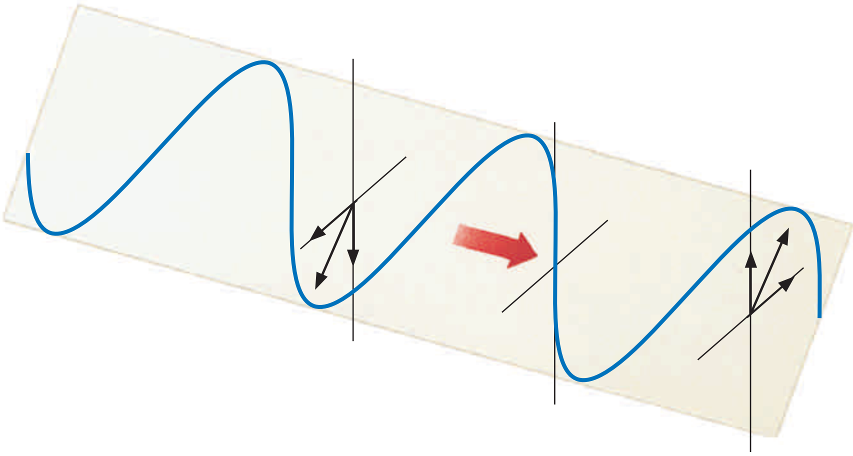


Figure 19-16 The plane-polarized wave can be broken up into two perpendicular component waves: one in the vertical plane and one in the horizontal plane. The displacements of these component waves are shown at two sample locations.

imagined as a superposition of a vertical vibration and a horizontal vibration, as illustrated in Figure 19-16. The vertical slit allows the vertical vibration to pass through while blocking the horizontal vibration. Therefore, the wave has a vertical polarization after it passes through the vertical slit. The amplitude of the transmitted wave is equal to the amplitude of the vertical vibration of the incident wave.

Determining whether light can be polarized is a little difficult because our eyes cannot tell whether light is polarized. Nature, however, provides us with materials that polarize light. The mere existence of polarized light demonstrates that light is a transverse wave. If it were longitudinal, it could not be polarized. Commercially available light-polarizing materials, such as Polaroid filters, consist of long, complex molecules whose long axes are parallel. These molecules pass light waves with polarizations perpendicular to their long axes but absorb those parallel to it.

◀ **light is a transverse wave**

We can use a piece of Polaroid filter (or the lens from Polaroid sunglasses) to analyze various light sources to see whether they are polarized. If the light is polarized, the intensity of the transmitted light will vary as the Polaroid filter is rotated. This simple procedure shows that common light sources such as ordinary incandescent lamps, fluorescent lights, candles, and campfires are unpolarized. However, if we examine the light reflected from the surface of a lake, we find that it is partially polarized in the horizontal direction. Light reflected from nonmetallic surfaces is often partially polarized in the direction parallel to the surface. This is the reason why the axes of polarization of Polaroid sunglasses are in the vertical direction. Boaters know that Polaroid sunglasses remove the glare from the surface of water and allow them to see below the surface.

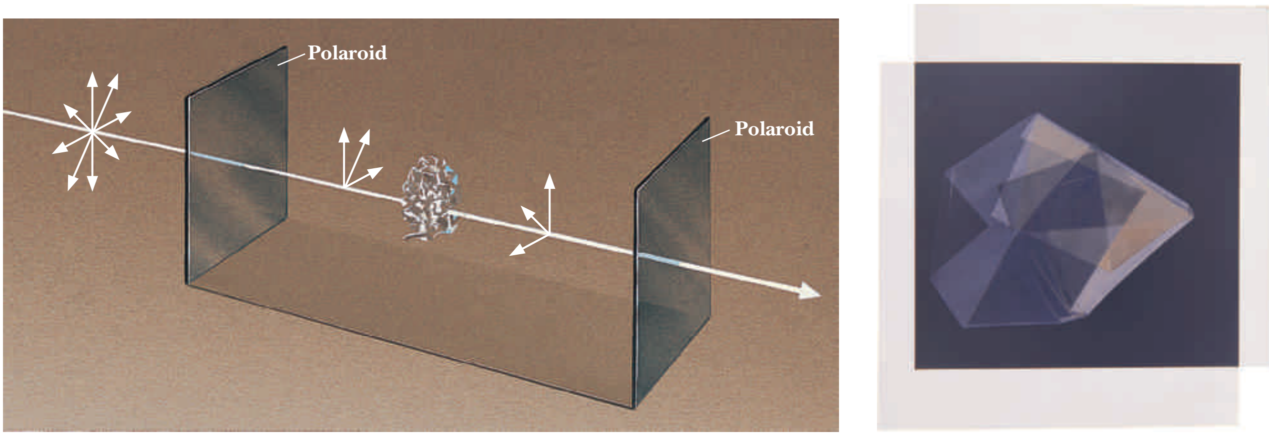
Photographs taken through a plate-glass window (a) with and (b) without a polarizing filter on the camera lens. Notice that the polarizing filter eliminates the glare.

© Cengage Learning/George Semple (both)

(a) *(b)*

Until now we have discussed materials in which one direction passes the light and the other absorbs it. Some materials allow both polarizations to pass through, but the orientations have different speeds. This can have the effect of rotating the plane of polarization. The cellophane on cigarette packages and some types of transparent adhesive tape have this property.

You can make an interesting display by crumpling a cigarette wrapper and looking at it between two Polaroid filters, as in Figure 19-17(a). The Polaroid filter on the side near the light polarizes the light striking the display. When this light is viewed with the second Polaroid filter, the different thicknesses of cellophane have different colors, as shown in Figure 19-17(b). The amount of rotation depends on the thickness of the cellophane and the wavelength of the light. Certain thicknesses rotate certain colors by just the right amount to pass through the second Polaroid filter. Others are partially or completely blocked. The colors of each section change as either Polaroid filter is rotated.



Charles D. Winters

(a) *(b)*

Figure 19-17 (a) Cellophane or transparent tape rotates the plane of polarization. (b) A pattern formed by different thicknesses of cellophane viewed between two pieces of Polaroid filter.

Glasses and plastics under stress rotate the plane of polarization, and the greater the stress the more the rotation. Plastic models of structures, such as cathedrals, bones, or machined parts, can be analyzed to discover where the stress is greatest.

Looking Ahead

We have not finished our quest to understand light. Notice, for example, that although we established that the wave model of light explains all the phenomena we have observed so far, we still have not said what it is that is waving! We know that light travels from the stars to our eyes through incredible distances in a vacuum that is much better than any we can produce on Earth. How is this possible?

We will see in Chapter 23 that understanding the nature of light was the key to our modern understanding of atoms. This reexamination of the model for light occurred at the beginning of the 20th century and began with Einstein asking many of the same questions that we addressed here.

Summary

Developing a theory of light involves building a model from known experiences that can be compared with the behavior of light. Both particle and wave models can account for the law of reflection and the law of refraction. However, only a wave model can correctly account for the speed of light in transparent materials. The index of refraction n is the ratio of the speed of light c in a vacuum (300 million meters per second) to its speed v in the material; that is, $n = c/v$. Dispersion of light indicates that different colors have different speeds in a material.

Other properties of waves that differ from those of particles further support the fact that light is a wave phenomenon. Interference of two coherent light sources produces a stationary pattern. If different colors are used, the pattern changes size, demonstrating that colors have different wavelengths. Red light has the largest wavelength and produces the widest pattern, and violet light produces the narrowest one. Visible light ranges in wavelength from 400 to 750 nanometers.

Light exhibits diffraction, which is the spreading out of a wave as it passes through narrow openings and around the edges of objects. The width of the diffraction pattern depends on the ratio of the wavelength to the size of the opening—the narrower the opening, the wider the pattern. Because red light has the longest visible wavelength, it produces the widest pattern for a given opening.

Thin films of transparent materials exhibit beautiful arrays of colors due to the interference of light rays reflecting from the two surfaces. A ray that reflects from a material with a higher index of refraction is inverted. Different wavelengths are strongly reflected or transmitted at different thicknesses of film.

Light exhibits polarization, demonstrating that it is a transverse wave. Polarizing materials pass light waves with polarizations perpendicular to one axis, but absorb those parallel to it. Common light sources are usually unpolarized. However, light reflected from the surface of a lake or glass is partially polarized parallel to the surface.

Everyday Physics | *Holography*

Wouldn't it be fantastic to have a method for catching waves, preserving the information they carry, and then at some later time playing them back? In fact, high-fidelity equipment routinely records a concert and plays it back so well that the listeners can hardly tell the difference. Similarly, high-fidelity photography should record the light waves coming from a scene and then play them back so well that the image is nearly indistinguishable from the original scene.

Conventional photography, however, does not do this. Although modern chemicals and papers have created images with extremely fine resolution, nobody is likely to mistake a photograph for the real object because a photographic image is two-dimensional, whereas the scene is three-dimensional. This loss of depth means that you can view the scene in the photograph from only one angle, or perspective. You cannot look around objects in the foreground to see objects in the background. The objects in the scene do not move relative to one another as you move your point of view.

Holography is a photographic method that produces a three-dimensional image that has virtually all the optical properties of the scene. This process was conceived by Dennis Gabor in 1947; he received a Nobel Prize in 1971 for this work. Gabor chose the word *hologram* to describe the three-dimensionality of the image by combining the Greek roots *holo* ("complete") and *gram* ("message").

Although Gabor could make a hologram out of a flat transparency, the lack of a proper light source prevented him from making one of an object with depth. With the invention of the laser in 1960, interest in holography was renewed. We will postpone the discussion of laser operation until Chapter 24. At the moment we need only know that it is a device that produces light with a single color and a constant phase relationship.

Figure A shows the essential features of a setup for making holograms. Notice that there is no lens between the object and the film. Light from the object reflects onto all portions of the film. Although this *object beam* carries the information about the object, if this were all that happened, the film would be completely exposed, and virtually no information about the object would be recorded.

Another portion of the beam from the laser (the *reference beam*) illuminates the film directly. Once again, if this were all that happened, the film would not record the scene. However, the light from the laser has a single well-defined wavelength (color) with the crests lined up. Therefore, the light reflected from the object and the light in the reference beam produce an interference pattern that is recorded in the film. This interference pattern contains the three-dimensional information about the scene.

The hologram is viewed by placing it back in a reference beam. The pattern in the film causes the light passing through it to be deflected so that it appears to come from the original scene (Figure B).

Because information from each point in the scene is recorded in all points of the hologram, a hologram contains information from

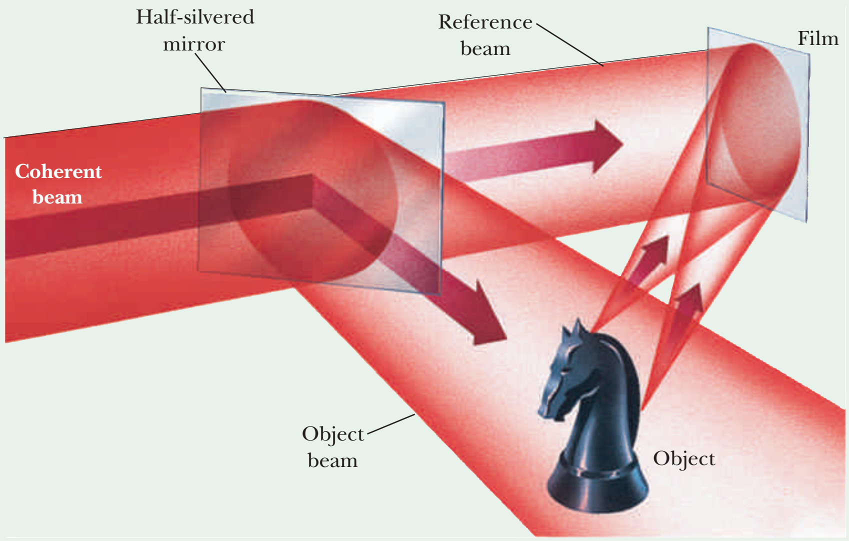


Figure A Making a hologram.

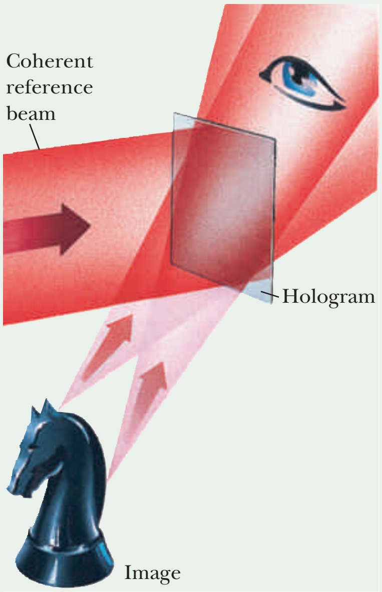


Figure B A hologram is viewed by placing it in the reference beam. The light passing through the film produces a virtual image at the location of the original scene.

all perspectives covered by the film. Therefore, viewing a hologram is just like viewing the original scene through a window. The photographs in Figure C were taken of a single hologram. Note that the relative locations of the chessmen change as the camera position changes. This is what we would see if we were looking at the actual chessboard.

Because information from each point in the scene is recorded in all points of the hologram, a hologram can be broken, and each piece will produce an image of the entire scene. Of course, something must be lost. Each piece will allow the scene to be viewed only from the perspective of that piece. This is analogous to looking at a scene through a window that has been covered except for a small hole.

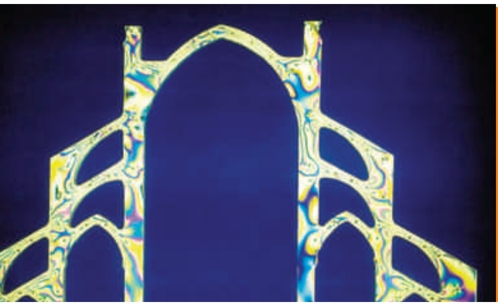
Advances in holography have made it possible to display holograms with ordinary sources of white light instead of the much more expensive lasers. It is now possible to hang holograms in your home as you would paintings. In fact, they have become so inexpensive that they have appeared on magazine covers. Holograms can also be made in the shape of cylinders so that you can walk around the hologram and see all sides of the scene. In some of these, individuals in the scene move as you walk around the hologram. Holographic movies are possible, and holographic television should be possible in the future.

In addition to being a fantastic art medium, holography has found many practical applications. It can be used to measure the three-dimensional wear patterns on the cylinder walls of an engine to an accuracy of a few ten-thousandths of a millimeter (a hundred-thousandth of an inch), perform nondestructive tests of the integrity of machined parts or automobile tires, study the shape and size of snowflakes while they are still in the air, store vast quantities of information for later retrieval, and produce three-dimensional topographical maps.

1. What characteristics of laser light are critical to the creation of a holographic image?
2. A piece of film containing a hologram is cut into four equal parts. How will the hologram observed in one of these parts compare to the original hologram?

Figure C The relative positions of the chessmen in the holographic image change as the point of view changes.

Gerald F. Wheeler (both)



CHAPTER 19 *Revisited*

Light is probably the most fascinating and elusive phenomenon in nature. The answer to the question, "What is light?" changes with the techniques used to examine the question. In this chapter we found good evidence to support the conclusion that light is a wave phenomenon. In future chapters we'll return to this question and delve further into the mystery of light.

Key Terms

dispersion The spreading of light into a spectrum of colors. The variation in the speed of a periodic wave due to its wavelength or frequency.

index of refraction An optical property of a substance that determines how much light bends on entering or leaving it. The index is equal to the ratio of the speed of light in a vacuum to its speed in the substance.

polarized A property of a transverse wave when its vibrations are all in a single plane.

refraction The bending of light that occurs at the interface between two transparent media. It occurs when the speed of light changes.

Questions and exercises are paired so that most odd-numbered are followed by a similar even-numbered.
Blue-numbered questions and exercises are answered in Appendix B.
indicates more challenging questions and exercises.
WebAssign Many Conceptual Questions and Exercises for this chapter may be assigned online at WebAssign.

Conceptual Questions

1. Newton believed that light beams consist of tiny particles. If these beams travel in straight lines, what does that imply about their speed?
2. How does the particle theory of light account for the diffuse reflection of light?
3. Argue that the law of reflection would not hold for particles rebounding from a surface that is not frictionless or not perfectly elastic.
4. When a particle reflects elastically from a smooth surface, the component of the particle's momentum parallel to the surface is conserved while the component of the particle's momentum perpendicular to the surface is reversed. Use this information to argue that Newton's particle theory of light is consistent with the observation that the angle of incidence equals the angle of reflection.
5. If particles incident at 45 degrees from the normal strike a completely elastic surface that has friction, will the angle of reflection (with respect to the normal) be greater than, equal to, or less than 45 degrees? Explain.
6. If particles incident at 45 degrees from the normal strike a frictionless surface that is not completely elastic, will the angle of reflection (with respect to the normal) be greater than, equal to, or less than 45 degrees? Explain.
7. How does Newton's idea of light particles explain the law of refraction?
8. Explain how Newton's idea of light particles predicts that the speed of light in a transparent material will be faster than in a vacuum.
9. Does the wave's frequency or its wavelength remain the same when the wave crosses from one medium into another? Explain.
10. In which region of Figure 19-3(a) (top left or bottom right) are the waves traveling at the higher speed? Explain.
11. Which color of light, red or blue, travels faster in a diamond? Explain your reasoning.
12. Do you expect the speed of light in glass to be slower than, faster than, or the same as that in diamond? Why?
13. What property of a light wave determines its brightness?
14. Does the amplitude of a light wave increase, decrease, or stay the same on reflection from a transparent material? Explain.
15. Starting with the observation that waves that have been bent toward the normal have a shorter wavelength than the incident waves, explain how the wave model for light

predicts that the speed of light in glass will be slower than the speed in a vacuum.

16. Imagine that Newton knew that light travels slower in glass than in air but was unaware of the law of refraction. In what direction would he have predicted light to bend when passing from air into glass?

17. Does total internal reflection result from light trying to pass from a slow medium to a fast medium or from a fast medium to a slow medium? Explain.

18. Different colors of light have different critical angles for total internal reflection. Is the critical angle greater for colors of light that travel faster or slower in the medium? Explain.

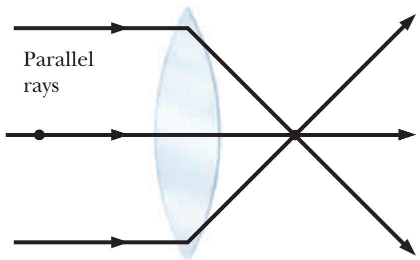
19. What is the physical difference between red and blue light?

20. How does the slow speed of light in diamonds affect their brilliance?

21. Why do we not notice any dispersion when white light passes through a windowpane?

22. What does the dispersion of light tell us about the speeds of various colors of light in a material?

23. Will the converging lens in the following figure focus blue light or red light at a closer distance to the lens? Explain.



Questions 23 and 24.

24. Using blue light, you determine the focal point for the lens in the preceding figure. If you were to shine a green laser beam from this focal point to a point near the top of the lens, would the emerging beam be bent toward or away from the optic axis? Explain.

25. Red light is used to form a two-slit interference pattern on a screen. As the two slits are moved farther apart, does the separation of the bright bands on the screen decrease, increase, or remain the same? Why?

26. What happens to the separation of the bright bands in a two-slit interference pattern if the slits are made narrower but their separation remains the same?

27. Would yellow light or green light produce the wider two-slit interference pattern? Why?

28. We observe that the two-slit interference pattern produced by blue light is narrower than that produced by red light. What does this tell us about red and blue light?

29. What determines whether two light beams with the same wavelength tend to cancel or reinforce each other?

30. Why don't we notice interference patterns when we turn on two lights in a room?

31. If light and sound are both wave phenomena, why can we hear sounds around a corner but cannot see around a corner?

32. Approximately how narrow should a slit be for the diffraction of visible light to be observable?

33. Blue light is used to form a single-slit diffraction pattern on a screen. As the slit is made wider, does the separation of the bright bands on the screen decrease, increase, or remain the same? Explain.

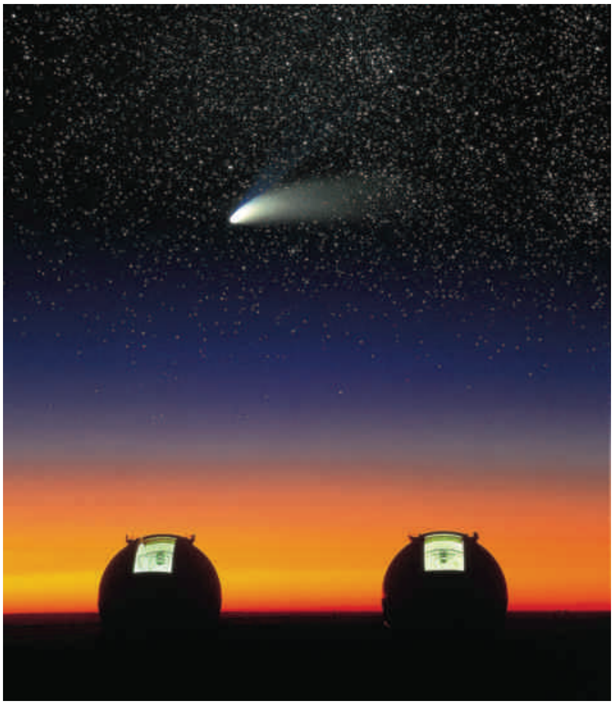
34. Would orange light or blue light produce the wider diffraction pattern? Why?

35. Would a slit with a width of 300 nanometers or of 400 nanometers produce a wider diffraction pattern when illuminated by light of the same wavelength? Why?

36. Which of the following single-slit diffraction experiments would produce the wider diffraction pattern: 800-nanometer light passing through a 500-nanometer-wide slit, or 450-nanometer light passing through a 400-nanometer-wide slit? Why?

37. Why can't an ordinary microscope using visible light be used to observe individual molecules?

38. A common technique used by astronomers for overcoming diffraction limits is to electronically combine the light from more than one telescope. This effectively increases the diameter of the aperture to the distance between the telescopes. If the signals from two 5-meter-diameter telescopes located 100 meters apart were being combined when one of the telescopes stopped functioning, by what factor would the minimum resolvable angle be increased?



David Nunuk/SPL/Photo Researchers, Inc.

39. Will you observe multicolored patterns if you illuminate a thin soap film with monochromatic light? Why?

40. A thin film of oil on top of a bucket of water produces multicolored patterns. However, a bucket full of oil produces no such effect. Explain the difference.

41. Assume that you have the thinnest film that strongly reflects red light. Would you need to make the film thinner or thicker to completely reflect blue light? Why?

42. You are coating glass with a film of higher index of refraction. You make the thinnest film that will produce a strong reflection for a particular monochromatic light source. You then gradually increase the film's thickness until you find another strong reflection. How many times thicker is this film than the original?

43. A glass pane with index of refraction 1.5 is coated with a thin film of a material with index of refraction 1.6. The coating is as thin as possible to produce maximum reflection for blue light. If this same material is used to coat a different kind of glass with index of refraction 1.9, the light reflected from the back surface of the film now experiences an inversion. Does the coating have to be thicker or thinner in this case to produce strong reflection? Explain.

44. The office workers in a skyscraper complain that the morning sun shines too brightly into their work areas. The problem is resolved by applying a thin film to each windowpane. The film has an index of refraction smaller than the glass and is designed to reflect yellow light when applied to the glass. If a sheet of this film is held in front of a yellow spotlight, would any of the light pass through the film? Explain.

Phil Jude/SPL/Photo Researchers, Inc.

45. A thin film in air strongly reflects orange light. Will it still reflect orange light when it is placed in water?

46. A thin, transparent film strongly reflects yellow light in air. What does the film do when it is applied to a glass lens that has a higher index of refraction than the film?

47. If all the labels had come off the sunglasses in the drug store, how could you tell which ones were polarized?

48. Can sound waves be polarized?

49. The digital displays at fuel pumps often use liquid crystal displays (LCDs) to show the price. Because the light from LCDs is polarized, they can often be impossible to read while wearing Polaroid sunglasses. What could you do to read the display without removing your glasses?

50. How could you use Polaroid sunglasses to tell whether light from the sky is polarized?

51. How would you distinguish a hologram from a flat transparency?

52. If each point on a holographic film contains the entire image, what is gained by making the hologram larger?

53. What kind of light is required to make a hologram of a three-dimensional object?

54. What kind of light is required to display a hologram?

55. To gather enough light to expose the film, long time exposures are often necessary to make holograms of inanimate objects. Why is a very powerful laser required to make a hologram of a person's face?

56. Which of the following phenomena does not show a difference between the wave theory and particle theory of light: reflection, refraction, interference, diffraction, or polarization?

Exercises

57. What is the speed of light in glass with an index of refraction of 1.6?

58. What is the speed of light in water?

59. The speed of light in diamond is 1.24×10^8 m/s. What is the index of refraction for diamond?

60. Zircon is sometimes used to make fake diamonds. What is its index of refraction if the speed of light in zircon is 1.6×10^8 m/s?

61. If it takes light 5 ns (1 nanosecond $= 10^{-9}$ s) to travel 1 m in an optical cable, what is the index of refraction of the cable?

62. If an optical cable has an index of refraction of 1.5, how long will it take a signal to travel between two points on opposite coasts of the United States separated by a distance of 5000 km?

63. The index of refraction for red light in material X is measured at 1.80. Blue light travels 5×10^6 m/s slower than red light in this material. What is the index of refraction for blue light in material X?

64. For crown glass, the index of refraction for violet light is 1.532 and the index of refraction for red light is 1.515. How much faster is red light than violet light in this medium?

65. What is the wavelength of the radio signal emitted by an AM station broadcasting at 1420 kHz? Radio waves travel at the speed of light.

66. What is the wavelength of light that has a frequency of 5×10^{14} Hz?

67. The red light from a helium–neon laser has a wavelength of 633 nm. What is its frequency?

68. What is the frequency of the yellow light with a wavelength of 590 nm that is emitted by sodium lamps?

69. What is the wavelength of the red light from a helium–neon laser when it is in glass with an index of refraction of 1.6? The wavelength in a vacuum is 633 nm.

70. A transparent material is known to have an index of refraction equal to 1.9. What is the wavelength of light in this material if it has a wavelength of 650 nm in a vacuum?

71. Light from a sodium lamp with a wavelength in a vacuum of 590 nm enters diamond in which the speed of light is 1.24×10^8 m/s. What is the wavelength of this light in diamond?

72. What is the wavelength of light in water if it has a frequency of 6.6×10^{14} Hz?

73. For distant objects, the angular size in degrees can be approximated as $57° \times w/d$, where w is the width of the object and d is its distance. What is the angular separation of the headlights on a car 10 km away if the headlights are 1.2 m apart?

74. The minimum angular separation in arc seconds ($\frac{1}{3600}$ degree) is found by first finding the ratio of the wavelength of light to the diameter of the aperture and then multiplying by 2.5×10^5. Using visible light with a wavelength of 550 nm, calculate the minimum angular separation for an eye with a pupil size of 5 mm.

75. Using the information in Exercise 74, find the theoretical resolution of a telescope with a 10-m-diameter mirror for visible light at 550 nm.

76. What is the theoretical resolution of a radio telescope with a 10-m-diameter collecting dish for radio waves with a wavelength of 21 cm? (See Exercise 74.)

© Photodisc Green/Getty Images

77. What is the thinnest soap film that will strongly reflect light with a wavelength of 400 nm in the film?

78. What is the thinnest soap film that will strongly reflect red light from a helium–neon laser? The wavelength of this light is 633 nm in air and 470 nm in soapy water.

79. You are coating a glass lens of index of refraction 1.6 with a film of material of index of refraction 1.7. You start with the thinnest film possible that creates a strong reflection for 500-nm light. You gradually increase the film thickness until you again get strong reflection. What is the thickness of the film now?

80. Repeat Exercise 79 for a glass lens of index of refraction 1.8.

The Big Picture

An Electrical and Magnetic World


Lief Skoogfors/The Franklin Institute Science Museum/Woodfin Camp & Assoc

The hairs on this girl's head repel each other because they have the same charge.

If our common experiences are any guide, electrical and magnetic effects appear insignificant. They don't seem to be a common property of all matter, nor do they usually seem very strong. Magnets are used to lift a couple of paper clips, and static electricity lifts bits of paper.

If we ignore the electricity supplied by both power companies and lightning, electricity seems only to give us occasional annoying shocks when we scuff our feet on rugs or slide across vinyl seat covers. Furthermore, the effects are intermittent; shocks usually occur only in winter or when it is especially dry. Even then, they occur only with a few materials: we feel a shock when we touch metal doorknobs but not when we touch wooden doors.

Magnetic properties seem more permanent: a piece of metal that is magnetic in the winter is still magnetic in the summer. But again, these properties don't seem to be universal. Only a few materials are magnetic, and the magnetic forces don't seem very strong.

Most people held these attitudes in Newton's time. The investigation of electrical and magnetic properties of matter was not considered important. Gravity reigned supreme as the governing force in the universe. All objects attract each other via the gravitational force. It holds objects to Earth's surface, it holds our life-supporting atmosphere around Earth, and it even controls the motions of the planets. But, in fact, gravity is the *weakest* force. It is the electric and magnetic forces that control the structure of matter on the atomic level and therefore the structure of matter in the macroscopic world.

We are electrical creatures. What tastes good or bad, what's medicine and what's poison, and the contractions of our muscles or our vocal cords are all controlled by electrical interactions. Even our thoughts are due to electrical signals. Magnets and magnetism seem more removed from our daily lives, but images of our brains can be taken using magnetic techniques. As we will learn in Chapter 22, the seemingly separate phenomena of electricity and magnetism are in fact intimately connected.

As we have seen, the age of Newton, Galileo, and Huygens ushered in a new way of thinking about and seeing everyday phenomena. The one belief they all shared was that the universe is composed of atoms. However, atoms and the void between them implied some mysterious action-at-a-distance, and no one could adequately account for many optical, electrical, and magnetic phenomena. Waves and atoms did not seem to mix well with the concept of a void, and by 1800 it was clear that some new theory was required.

The next three chapters follow a path similar to the one followed by the early pioneers in this area. We first look at electricity and magnetism separately. Then we look for connections between the two. As we develop the connections, you will find yourself returning to components of the physics world view that we developed in earlier chapters. All of this discussion will set the stage for the study of the basic building block of nature, the atom.

20 Electricity

▶We see electrical effects when sparks occur. On dry days, we may feel a shock when we touch a metal doorknob after walking across a carpet. If we look closely, we can see a spark jump between our hand and the doorknob. Although they may surprise us, the sparks do not hurt us. On the other hand, nature produces long and dangerous sparks during lightning storms. What determines the length of the sparks?

(See page 440 for the answer to this question.)

Jean-Loup Charmet/SPL/Photo Researchers, Inc.

A lightning storm produces very high voltages, creating large and dangerous electric currents.

THE early Greeks knew that amber—a fossilized tree sap currently used in jewelry—had the interesting ability to attract bits of fiber and hair after it was rubbed with fur. This was one way of recognizing an object that was electrified. In modern terminology we say the object is **charged.** This doesn't explain what **charge** is, but is a handy way of referring to this condition.

In 1600 English scientist William Gilbert published a pioneering work, *De Magnete*, in which he pointed out that this electrical effect was not an isolated property of amber but a much more general property of matter. Materials such as gems, glass, and sealing wax could also be charged. Rubbing together two objects made from different materials was the most common way of charging an object. In fact, both objects become charged.

After Gilbert, many experimenters joined in the activity of investigating electricity. However, the question of what happens when an object gains or loses this electrical property remained unanswered. A modern response may include words such as *electron* or *proton*. But we have to be careful that we are not simply substituting new words or phrases for old ones. To answer this question fully and thus expand our world view, we need to carefully examine our electrical world.

Electrical Properties

In an effort to explain electricity, Gilbert proposed the existence of an electric fluid in certain types of objects. He suggested that rubbing an object removed some of this fluid, leaving it in the region surrounding the object. Bits of fiber were attracted to the object by the "draft" of the fluid returning to the object. Although many other electrical phenomena could not be explained with this idea, it was the beginning of attempts to model the underlying, invisible processes that caused the electrical effects.

Little progress was made for more than a century. In the 1730s it was shown that charge from one object would be transferred to a distant object if metal wires connected them but not if silk threads connected them. Materials that can transfer charge are known as **conductors;** those that cannot are called nonconductors, or **insulators.** It was discovered that metals, human bodies, moisture, and a few other substances are conductors.

The discovery that moisture is a conductor explains why electrical effects vary from day to day. You normally experience bigger shocks in the winter when the humidity is naturally low. On more humid days, any charge that you get by scuffing your feet is quickly dissipated by the moisture in the air surrounding your body.

Many charged objects have to be suspended from insulators such as silk threads or plastic bases or they quickly lose their charge to Earth, or ground. In fact, we speak of **grounding** an object to ensure that it is not charged. Other objects hold their charges without being insulated. When we charge a plastic rod by rubbing it with a cloth, the charge stays on the rod even if we hold it. But our bodies are conductors. Why doesn't the charge flow to ground through our bodies? It stays on the rod because the rod is an insulator; charge generated at one end remains there. The charge can be removed by moving our hands along the charged end. As we touch the regions that are charged, the charges flow through our bodies to ground.

A metal rod cannot be charged by holding it in our hands and rubbing it with a cloth because metal conducts the charge to our hands. A metal rod can be charged if it is mounted on an insulating stand or if we hold it with an insulating glove; that is, the rod must be insulated from its surroundings.

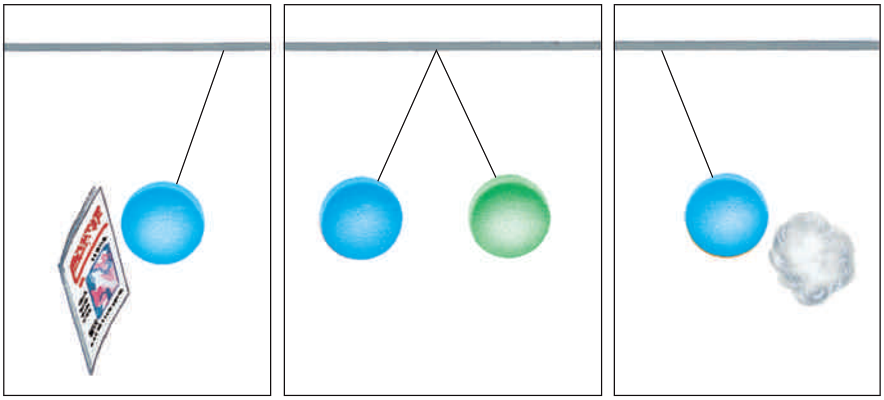
Fuel trucks are grounded to remove the charge built up by the fuel flowing through the hoses.

Even the flow of liquid through pipes is enough of a rubbing action to charge the liquid. Fuel trucks tend to build up charge as they dispense fuel. If this charge becomes large enough, sparking may occur and cause a fire or explosion. To avoid this danger, a conducting wire is connected from the truck to the ground to allow the charge to flow into the ground.

Two Kinds of Charge

It was also the early 1730s before there was any mention that charged objects could repel one another. Electricity, like gravity, was believed to be only attractive. It may seem strange to us now because both the attractive and repulsive aspects of electricity are easy to demonstrate. If you comb your hair, the comb becomes charged and can be used to attract small bits of paper. After contacting the comb, some of these bits are then repelled by the comb.

This phenomenon can be investigated more carefully using balloons and pieces of wool. If we rub a balloon with a piece of wool, the balloon becomes charged; it attracts small bits of paper and sticks to walls or ceilings. If we suspend this balloon by a thread and bring uncharged objects near it, the balloon is attracted to the objects [Figure 20-1(a)]. Everything seems to be an attractive effect.



(*a*) (*b*) (*c*)

Figure 20-1 A charged balloon is (a) attracted by uncharged objects, (b) repelled by an identically charged balloon, and (c) attracted by the wool used in charging it.

Charging another balloon in the same way demonstrates the new effect: the two balloons repel one another [Figure 20-1(b)]. Because we believe that any two objects prepared in the same way are charged in a like manner, we are led to the idea that like-charged objects repel one another.

Whenever we charge an object by rubbing it with another, both objects become charged. If we examine the pieces of wool, we find that they are also charged: they each attract bits of paper.

Are You On the Bus?

Q: Will the two pieces of wool attract or repel each other?

A: Because they have been charged in a like manner, they will repel each other.

The piece of wool and the balloon, however, attract each other after being rubbed together [Figure 20-1(c)]. If they had the same kind of charge, they would repel. We are therefore led to the idea that there must be two different kinds of charge and that the two kinds attract each other. These experiments can be summarized by stating the following:

Like charges repel; unlike charges attract.

Conservation of Charge

Like Gilbert, Benjamin Franklin believed that electricity was a single fluid and that an excess of this fluid caused one kind of charged state, whereas a deficiency caused the other. Because he could not tell which was which, he arbitrarily named one kind of charge *positive* and the other kind *negative*. By convention the charge on a glass rod rubbed with silk or plastic film is positive (Figure 20-2), whereas that on an amber or rubber rod rubbed with wool or fur is negative.

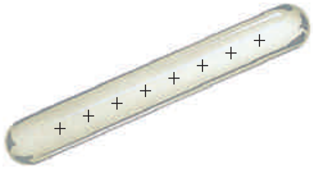


Figure 20-2 The charge on the glass rod is called positive. This is indicated by a few plus signs.

At first glance these names seem to have no advantage over other possible choices, such as black and white or yin and yang. They were, however, more significant. Franklin's use of the fluid model led him to predict that charge should be conserved. The amount of electric fluid should remain the same; it is just transferred from one object to another. If you start with two uncharged objects and rub them together, the amount of excess fluid on one is equal to the deficiency on the other. In other words, the positive and negative charges are equal. Using his system of positive and negative numbers, we can add them and see that the total charge remains zero.

We no longer believe in Franklin's fluid model; the fluid model was abandoned as it could not account for later experimental observations. However, Franklin's idea about the **conservation of charge** has been verified to a high precision. It is one of the fundamental laws of physics. In its generalized form, it can be stated as follows:

In an isolated system the total charge is conserved.

◀ conservation of charge

In our modern physics world view, all objects are composed of negatively charged electrons, positively charged protons, and uncharged neutrons. The electron's charge and the proton's charge have the same size. An object is uncharged (or neutral) because it has equal amounts of positive and negative charges, not because it contains no charges. For example, atoms are electrically neutral because they have equal numbers of electrons and protons.

Biosketch | *Franklin: The American Newton*

When Benjamin Franklin (1706–1790) arrived in Paris as Revolutionary America's ambassador, he was known to the sympathetic French as a famous scientist who "stole the scepter from tyrants and lightning from the gods." Precisely because he was the "American Newton," he was the most effective representative the young nation-in-the-making could have sent.

Benjamin Franklin

Born in Boston, Franklin made his fortune and reputation in Philadelphia—then the largest English-speaking city outside of London. He was a one-man chamber of commerce who greeted newcomers, brought promising young people into his circle of friends, and helped organize everything from the first philosophical society in the country to a street-lighting district, a fire department, and a mental hospital. His publications on electricity brought him membership in the Royal Philosophical Society of London. (Newton had been president just 50 years earlier.) He invented lightning rods and used them on his own house to great effect.

Franklin believed that electricity was a fluid that pervaded all bodies in varying quantities. A body would seek to maintain electrical equilibrium. If there was a deficit, the body would be in a negative state, and electricity would flow into the body. If there was a surplus, electricity would flow out. He also believed that unlike gravitation, which is only attractive, electrified matter could repel as well as attract other electrified matter. It could also attract nonelectrified matter.

Like others of his era, he was also interested in heat, illumination, weather, and other aspects of science. Like Michael Faraday a bit later, Franklin was not a mathematician and so could not give definitive form to his arguments. Even so, this well-known patriot was the best known of early American scientists and a fascinating human being.

—Pierce C. Mullen, historian and author

Sources: Carl Van Doren, *Benjamin Franklin* (New York: Viking, 1938); *The Autobiography of Benjamin Franklin* (New York: Dover Thrift, 1966).

Positively charged objects may have an excess of positive charges or a deficiency of negative charges—that is, an excess of protons or a deficiency of electrons. We simply call them positively charged because the electrical effects are the same in both situations.

The modern view easily accounts for the conservation of charge when charging objects. The rubbing simply results in the transfer of electrons from one object to the other; whatever one object loses, the other gains.

Q: If you remove one electron from a neutral quarter and one electron from a neutral Buick, which (if either) has the greater net charge?

A: In either case, each proton is paired with an electron except for one. The net charge on either object would just be the charge of a single proton.

Induced Attractions

Attraction is more common than repulsion because charged objects can attract *uncharged* objects. How do we explain the observation that charged objects attract uncharged objects? Consider a positively charged rod and an uncharged metal ball. As the rod is brought near the ball, the rod's positive charges attract the negative charges and repel the positive charges in the ball. Because the charges in the ball are free to move, this results in an excess of negative charges on the near side and an excess of positive charges on the far side of the ball (Figure 20-3). Because charge is conserved, the excess negative charge on one side is equal to the excess positive on the other. Experi-

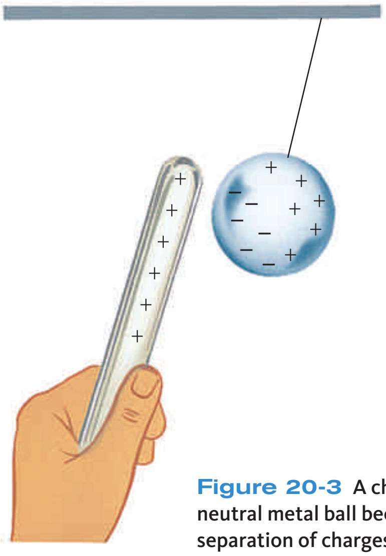


Figure 20-3 A charged rod attracts a neutral metal ball because of the induced separation of charges.

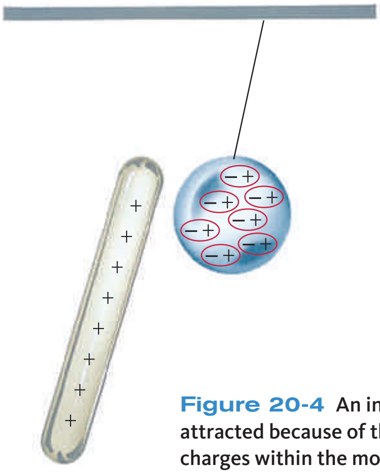


Figure 20-4 An insulating object is attracted because of the separation of charges within the molecules.

ments with balloons show that the electric force varies with distance; the force gets weaker as the balloons are moved farther apart. Thus, the ball's negative charge is attracted to the rod more than its positive charge is repelled. These induced charges result in a net attraction of the uncharged ball toward the rod.

Experiments show that if the ball is made of an insulating material, the attraction still occurs. In insulators the charges are not free to move across the object, but there can be motion on the molecular level. Although the molecules are uncharged, the presence of the charged rod may induce a separation of charge within the molecule. Other molecules are naturally more positive on one end and more negative on the other. The positive rod rotates these polar molecules so that their negative ends are closer to the rod, as shown in Figure 20-4. This once again results in a net attractive force.

A particularly graphic example of the attraction of an uncharged insulator is illustrated in Figure 20-5. A long two-by-four piece of lumber is balanced on

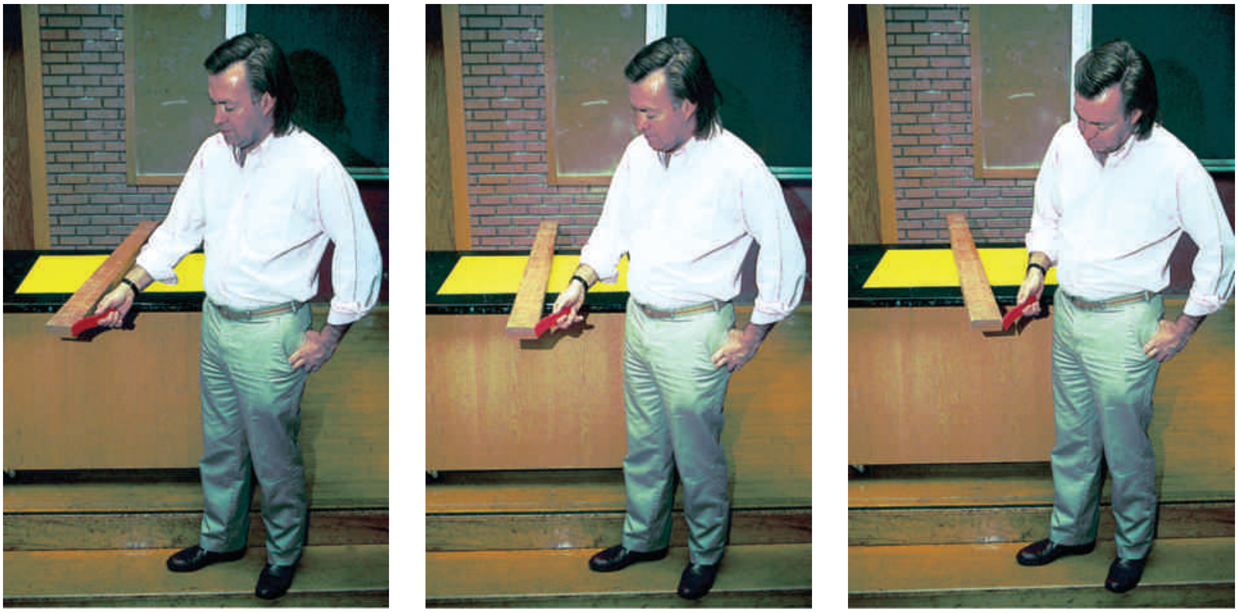


Figure 20-5 A standard two-by-four piece of lumber balanced on a curved dish can be rotated by the attractive force of a charged rod.

a watch glass so that the board is free to rotate. When a charged rod is brought near one end, the resulting attractive force produces a torque that rotates the board.

Q: Assuming that the two-by-four was attracted by a positively charged rod in the previous example, what direction would it rotate if a negatively charged rod was used?

A: It would rotate in the same direction. The interaction between any charged object and an uncharged object is always attractive.

We can now return to the charged comb that attracted and then repelled bits of paper. Initially, the bits of paper are uncharged. They are attracted to the comb by the induced charge. When the bits of paper touch the comb, they acquire some of the charge on the comb and are repelled because the comb and the paper have the same sign of charge and like charges repel one another.

FLAWED REASONING

Your teacher demonstrates that the south pole of a bar magnet is attracted to a charged glass rod. She then asks your class to predict what will happen when the charged glass rod is brought near the north pole of the magnet. The best student in the class answers, “The glass rod is positively charged, so the south pole of the magnet must be negatively charged. I know that a north pole will attract a south pole, so north poles of magnets must be positively charged. I predict that the glass rod will repel the north pole of the magnet.”

Your teacher then performs the experiment and shows that the glass rod *attracts* the north pole. She then shows that a negatively charged rubber rod also attracts both poles of the magnet, demonstrating that the magnet must be electrically neutral. **Clearly the best student in the class made a mistake in his reasoning. What is it?**

ANSWER An electrical *attraction* between two objects only indicates that at least one of the objects is charged. The class already knew that the glass rod was charged, so the first demonstration teaches nothing about the charge (if any) on the magnet. If the charged rod had *repelled* one of the poles of the magnet, it would indicate that the magnet must be charged. However, bar magnets are normally not charged and are attracted by any charged rod.

The Electroscope

In most experiments we transfer so few charges that the total attraction or repulsion is small compared with the pull of gravity. Detecting that an object is charged becomes difficult unless it is very light—like bits of paper or a balloon. We get around this difficulty with a device called an *electroscope,* which gives easily observable results when it is charged. By bringing the object in question near the electroscope, we can deduce whether it is charged by the effect it has on the electroscope.

The essential features of an electroscope are a metal rod with a metal ball on top and two very light metal foils attached to the bottom. Figure 20-6 shows a homemade electroscope constructed from a chemical flask. The glass enclosure protects the very light foils from air currents and insulates the foils and rod from the surroundings.

Figure 20-6 An electroscope made from a flask, a metal rod, and two pieces of thin metal foil.

If the electroscope isn't charged, the foils hang straight down under the influence of gravity. When we touch the ball of the electroscope with a charged rod, some of the rod's excess charge is shared with the electroscope. The charges flow from the rod onto the electroscope because the charges on the rod repel each other and therefore distribute themselves over the largest region possible, which includes the ball, the metal rod, and the foils. Because the foils have the same kind of charge, they repel each other and swing apart (Figure 20-7). The separation of the foils indicates that the electroscope is charged; the amount of separation is a rough indication of the amount of excess charge. This happens when the electroscope is touched with a positive rod as well as a negative one.

Figure 20-7 When an electroscope is charged by direct contact, it has the same charge as the rod.

When a charged object is touched to an electroscope, the electroscope takes on the same kind of charge as the object. However, the electroscope only shows the presence of charge; it does not indicate the sign. We can determine the sign by slowly bringing a rod with a known positive or negative charge toward the ball of the charged electroscope and watching the motion of the foils.

For example, as a positive rod nears the electroscope, it induces a redistribution of charge in the electroscope. It repels the positive charge and attracts the negative charge, causing an increase in positive charge (or a decrease in negative charge) on the foils. If the foils were originally positive, the extra positive charge would cause them to move farther apart. If the foils were originally negative, the extra positive charge would cancel part of the negative charge and the foils would move closer together.

Are You On the Bus?

Q: When a negative rod is slowly brought near an electroscope that is initially charged, the foils move closer together. What is the charge on the electroscope?

A: The negative rod repels negative charges on the electroscope's ball, increasing their concentration on the foils. Because this negative charge reduces the charge on the foils, they must have been positive initially.

A charged rod will also separate the foils on an uncharged electroscope. As a positive rod is brought near an uncharged electroscope, some of the negative charges in the electroscope are attracted to the ball on top; the ball has a net negative charge, and the foils have a net positive charge, as shown in Figure 20-8(a). Note that the net charge on the electroscope is still zero because we did not touch it with the rod. We have only moved the charges around in the electroscope by bringing the charged rod near the ball. When the charged rod is removed, the charges redistribute themselves and the foils fall. An analogous phenomenon happens with a negative rod.

Imagine that while the positive rod is held near the electroscope you touch the ball with your finger, as shown in Figure 20-8(b). Negative charges are attracted by the large positive charge on the rod and travel from the ground through your body to the electroscope. If you first remove your finger and then the rod, the electroscope is now charged and the foils repel each other. This is known as charging the electroscope by induction. The electroscope acquires a charge *opposite* to that on the rod.

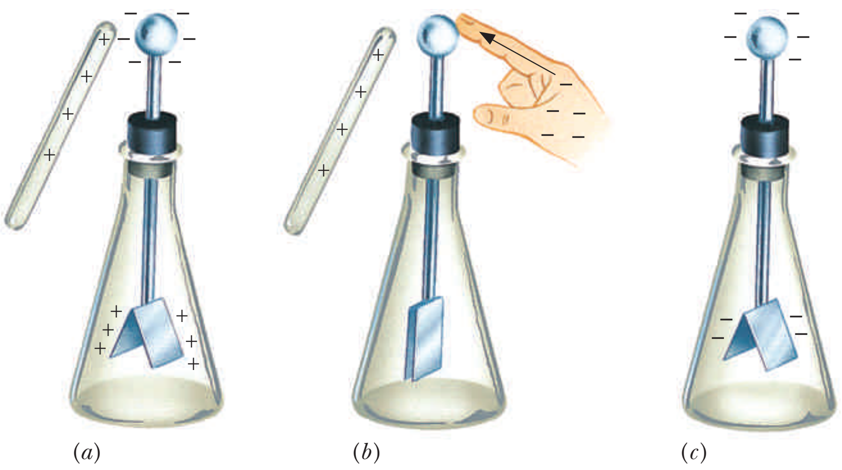


Figure 20-8 (a) The positively charged rod brought near the electroscope attracts negative charges to the ball, leaving the foils with an excess positive charge. (b) Touching the ball with a finger allows negative charges to flow into the electroscope, (c) leaving it with a net negative charge.

The Electric Force

Extended presentation available in the *Problem Solving* supplement

Simple observations of the attraction or repulsion of two charged objects indicate that the size of the electric force depends on distance. For instance, a charged object has more effect on an electroscope as it gets nearer. But we need to be more precise. How does the size of this force vary as the separation between two charged objects changes? And how does it vary as the amount of charge on the objects varies?

In 1785 French physicist Charles Coulomb measured the changes in the electric force as he varied the distance between two objects and the charges on them. He verified that if the distance between two charged objects is doubled (without changing the charges), the electric force on each object is reduced to one-fourth the initial value. If the distance is tripled, the force is reduced to one-ninth, and so on. This type of behavior is known as an **inverse-square relationship**—*inverse* because the force gets smaller as the distance gets larger, *square* because the force changes by the square of the factor by which the distance changes.

Coulomb also showed that reducing the charge on one of the objects by one-half reduced the electric force to one-half its original value. Reducing the charge on each by one-half reduced the force to one-fourth the original value. This means that the force is proportional to the product of the two charges.

Q: Coulomb could not measure charge directly. What technique could he have used to reduce the charge to one-half its original value?

A: He could have used identical conducting spheres and put the charged sphere into contact with a neutral one. By symmetry each sphere would have one-half the original charge.

These two effects are combined into a single relationship known as Coulomb's law:

Coulomb's law

$$F = k\frac{q_1 q_2}{r^2}$$

In this equation, q_1 and q_2 represent the amount of charge on objects 1 and 2, r is the distance between their centers, and k is a constant (known as Coulomb's constant) whose value depends on the units chosen for force, charge, and distance.

Each object feels the force due to the other. The forces are vectors and act along the line between the centers of the two objects. The force on each object is directed toward the other if the charges have opposite signs and away from each other if the charges have the same sign (Figure 20-9). Because the two forces are due to the interaction between the two objects, the forces are an action–reaction pair. According to Newton's third law, the forces are equal in magnitude, point in opposite directions, and act on different objects.

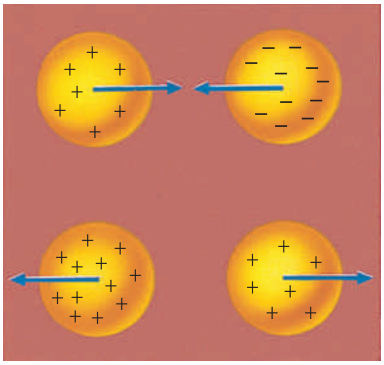


Figure 20-9 The forces on two charged objects are equal in size and opposite in direction in accordance with Newton's third law.

Because the existence of an elementary, fundamental charge was not known until the 20th century, the unit of electric charge, the **coulomb** (C), was chosen for convenience in use with electric circuits. (We will formally define the coulomb later.) Using the coulomb as the unit of charge, the value of Coulomb's constant is determined by experiment to be

Coulomb's constant

$$k = 9 \times 10^9 \frac{\mathrm{N \cdot m^2}}{\mathrm{C^2}}$$

The coulomb is a tremendously large unit for the situations we have been discussing. For instance, the force between two spheres, each having 1 coulomb of charge and separated by 1 meter, is

$$F = k\frac{q_1 q_2}{r^2} = \left(9 \times 10^9 \frac{\mathrm{N \cdot m^2}}{\mathrm{C^2}}\right) \frac{(1\ \mathrm{C})(1\ \mathrm{C})}{(1\ \mathrm{m})^2} = 9 \times 10^9\ \mathrm{N}$$

This is a force of 1 million tons!

The charges that we have been discussing are much less. When you scuff your shoes on a carpet, the charges transferred are about a millionth of a coulomb. But even this charge is very large compared with the charge on a single electron or proton. The size of the charge on an electron or a proton is now known to be 1.6×10^{-19} coulomb. Conversely, it would take the charge on 6.25×10^{18} protons (or electrons) to equal a charge of 1 coulomb.

FLAWED REASONING

The following question appears on the final exam: "Two small, identical metal balls are hung next to each other from silk threads. Both balls are given the same net charge, such that they repel each other, as shown in Figure 20-10. If the net charge on ball 2 is reduced by half, will the angle that ball 1 makes with the vertical be smaller than, the same as, or larger than the angle that ball 2 makes with the vertical?"

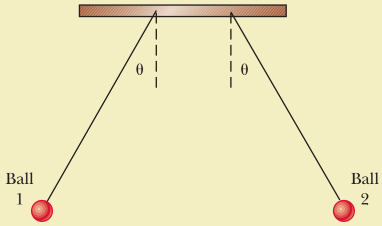


Figure 20-10

Peter gives the following answer: "Ball 2 will still hang at the same angle as before. The charge on ball 1 has not changed, so it exerts the same force on ball 2. Ball 2 has half the charge, so it exerts half the force on ball 1. Therefore, ball 1 will hang at a smaller angle than ball 2." **Should Peter get full credit for this problem?**

ANSWER Coulomb's law automatically satisfies Newton's third law. The force exerted on ball 1 by ball 2 must be equal in magnitude (and opposite in direction) to the force exerted on ball 2 by ball 1. Reducing the charge on ball 2 by half will reduce the force felt by *each* ball by half. Both balls hang at the same, smaller angle.

Electricity and Gravity

The mathematical form of Coulomb's law is the same as that of the universal law of gravitation discussed in Chapter 5. Therefore, the drawing in Figure 5-2 illustrating the dependence of the inverse-square law on distance also holds for electricity if Earth is replaced by a charged object. There are, however, several important differences between gravity and electricity that govern the roles of these two forces in the universe.

There are two kinds of electric charge—positive and negative. Opposite charges attract each other and like charges repel each other. There is, however, only one kind of mass. Negative mass does *not* exist, and the gravitational force is never repulsive. Because of the induced charge distribution possible with two kinds of charge, conducting materials can shield a region from all

Courtesy of Prof. Clint Sprott, University of Wisconsin-Madison

Professor Sprott is protected from the high-voltage discharge by the metallic cage.

external electric forces. Suppose we were to build a large metal room. If this room were placed near a large electric charge, the charges in the metal walls would redistribute in such a way as to cancel the effect of the external charge for all locations in the room. The cancellation works only because there are two kinds of electric charge. This result is not possible for gravitational forces; a gravitation-free chamber cannot be created because there is no negative mass. Antigravity spaceships, although they make good material for science fiction writers, are not possible.

Another difference between electricity and gravity lies in the behavior of objects in the respective fields. The gravitational force is proportional to an object's mass, and therefore objects of all sizes and compositions have the same acceleration in a gravitational field. For example, all objects have the same free-fall acceleration near Earth's surface. In contrast, the electric force is proportional to the charge on an object and *not* to the object's mass. Therefore, charged objects have different accelerations in an electric field. For example, if we were to release a proton and an electron near a charged object, the electron would have a much larger acceleration. The forces on the electron and the proton are the same size, but the electron's mass is much smaller, resulting in a much larger acceleration. In this case, the accelerations would be in opposite directions because the charges have opposite signs.

Finally, all charges occur as integral multiples of one fixed size. Experiments have shown that the charges on electrons and protons are the same size. The charges on other ordinary subatomic particles are either this same size or integral multiples of this size; the charges can be 2 times or 3 times the charge on the proton but never 1.5 times this charge. Mass also occurs in lumps, but there are many different sizes—for instance, the masses of the electron, the proton, and the neutron. These can combine in many different ways, and the total masses are not multiples of a single unit mass.

A consequence of these differences is the different roles these two forces play in our lives. The electric force is the dominant force in the atomic world; it determines the properties of atoms and molecules. On the other hand, the gravitational force dominates on the macroscopic scale of people, planets, and

© Royalty-Free/Corbis

The electric force dominates on the atomic scale, such as in large crystals.

NASA

The gravitational force dominates in the solar system and beyond, such as in this galaxy.

galaxies. The reason electricity and gravity switch roles in these two domains is that macroscopic objects are essentially uncharged; that is, there are approximately equal numbers of positive and negative charges on most large objects. Even though the total attractive force between all the positive charges in Earth and all the negative charges in the Moon is tremendous, it is equal and opposite to the repulsive force between the negative charges in Earth and the negative charges in the Moon. The net electric force is essentially zero because of this cancellation. However, because there is only one kind of mass, there is nothing to cancel the much weaker gravitational force between the atoms in Earth and atoms in the Moon. So as we move from people to planets and eventually to galaxies, the gravitational force becomes the dominant force.

WORKING IT OUT *Gravitational and Electric Forces*

As a numerical example, let's calculate the sizes of the electric and gravitational forces between an electron and a proton in a hydrogen atom. We start by assuming that the electron and proton are separated by a distance of 5.29×10^{-11} m and use the known masses and charges of the electron and proton. (These are given on the back end pages of this text.)

$$F_g = G\frac{m_1 m_2}{r^2} = \left(6.67 \times 10^{-11}\,\frac{\mathrm{N \cdot m^2}}{\mathrm{kg^2}}\right)\frac{(9.11 \times 10^{-31}\,\mathrm{kg})(1.673 \times 10^{-27}\,\mathrm{kg})}{(5.29 \times 10^{-11}\,\mathrm{m})^2}$$

$$= 3.63 \times 10^{-47}\,\mathrm{N}$$

$$F_e = k\frac{q_1 q_2}{r^2} = \left(9 \times 10^{9}\,\frac{\mathrm{N \cdot m^2}}{\mathrm{C^2}}\right)\frac{(1.6 \times 10^{-19}\,\mathrm{C})^2}{(5.29 \times 10^{-11}\,\mathrm{m})^2} = 8.23 \times 10^{-8}\,\mathrm{N}$$

Therefore, the electric force is more than 10^{39} times as strong as the gravitational force. (That's 1000 trillion trillion trillion times as strong!) Because the two forces change in the same way as the separation of the electron and proton changes, the separation does not matter and the electric force is always this much stronger.

The Electric Field

Implicitly, we have assumed the force between two charges to be the result of some kind of direct interaction—sort of an action-at-a-distance interaction. This type of interaction is a little unsettling because there is no direct pushing or pulling mechanism in the intervening space. Electrical effects are evident even when there is a vacuum between the charges.

In many cases, it proves to be both conceptually and computationally simpler to separate the electrical interaction that one object feels as the result of another into two distinct steps. First, one of the objects generates, by virtue of its charge, an **electric field** at every point in space. Second, another object interacts, by virtue of its charge, with the electric field to experience the force. The electric field is created by electric charges and exerts forces on other electric charges. This method is analogous to what we did with the gravitational force near the end of Chapter 5.

If this were the only purpose of the field idea, it would play a minor role in our physics world view. In fact, it probably seems as if we are trading one unsettling idea for another. However, as we continue our studies, we will find that the electric field takes on an identity of its own. As we will learn in Chapter 22, electric and magnetic fields can travel through space as waves.

We define the electric field **E** at every point in space as the force exerted on a unit positive charge placed at the point. This is equivalent to the way that the gravitational field was defined, with the unit mass replaced by a unit positive charge.

◀ electric field = force on unit positive charge

Because force is a vector quantity, the electric field is a vector field; it has a size and a direction at each point in space. You could imagine the space around a positive charge as a "porcupine" of little arrows pointing outward, as shown in the diagrams in Figure 20-11. The arrows farther from the charge would be shorter to indicate that the force is weaker there.

Q: What does the electric field surrounding a negative charge look like?

A: An electric field surrounding a negative charge looks just like the electric field surrounding a positive charge except that all of the arrows are reversed.

The values for an actual electric field can be measured with a test charge. The unit of charge that we have been using is 1 coulomb. This is a very large amount of charge, and if we actually used 1 coulomb as our test charge, it would most likely move the charges that generated the field, thus disturbing the field. Therefore, we use a much smaller charge, such as 1 microcoulomb, and obtain the size of the field by dividing the measured force **F** by the size q of the test charge:

$$\boldsymbol{E} = \frac{\boldsymbol{F}}{q}$$

◀ electric field $= \dfrac{\text{force}}{\text{charge}}$

Notice that the units of electric field are newtons per coulomb (N/C).

If we know the sizes and locations of the charges creating the electric field, we can also calculate the value of the field at any point of interest by assuming that we place a 1-coulomb charge at the location and calculating the force on this charge using Coulomb's law. In doing this, we can take advantage of the fact that each charge acts independently; the effects simply add. This means that we calculate the contribution of each charge to the field and then add these contributions vectorially.

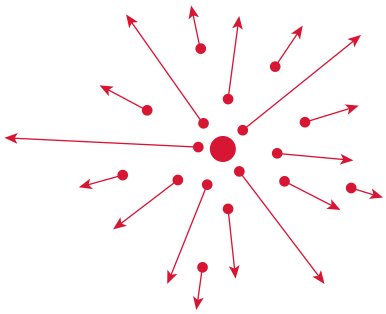


Figure 20-11 Sample electric field vectors around a positive source charge.

WORKING IT OUT *Electric Field*

What is the electric field at a distance of 2 m from a point charge of 3 μC?

By definition, the electric field at a point in space is equal to the force on a unit positive test charge placed at that point. In simple cases its magnitude can be obtained by calculating the force on a test charge q and then dividing by q.

$$E = \frac{F}{q} = \frac{1}{q}k\frac{Qq}{r^2} = k\frac{Q}{r^2}$$

We see that the value of the electric field depends on the source charge Q that creates the field, and is independent of the value of the test charge q that experiences the field. The magnitude of the electric field is

$$E = k\frac{Q}{r^2} = \left(9 \times 10^9\,\frac{\text{N}\cdot\text{m}^2}{\text{C}^2}\right)\frac{3 \times 10^{-6}\,\text{C}}{(2\text{ m})^2} = 6750\,\frac{\text{N}}{\text{C}}$$

Because the electric field is a vector, we must give its direction to completely specify it. The field points directly away from the positive source charge because that is the direction of the force on a unit positive test charge placed at the field point.

Once we know the value of the electric field at any point, we can calculate the force that any charge q would experience if placed at that point:

electric force = charge × electric field ▶

$$\boldsymbol{F} = q\boldsymbol{E}$$

This is read as, "The force on an object is equal to the net charge q on the object times the electric field **E** at the location of the object."

As an example, let's assume that we have generated a uniform electric field that points downward and has a size of 1000 newtons per coulomb. If we place an object in this field that has a positive charge of 1 microcoulomb, the object will experience a downward force of

$$F = qE = (10^{-6}\ \mathrm{C})\left(10^{3}\ \frac{\mathrm{N}}{\mathrm{C}}\right) = 10^{3}\ \mathrm{N}$$

If we change the charge on the object, calculating the new force is easy; we do not have to deal with the charges that produced the electric field.

Electric Field Lines

If we are only interested in what is going on at a single point, the electric field vector representation is helpful. However, it becomes cumbersome if we are interested in a region of space because each point in space may have a different electric field vector associated with it. To draw this, we would need to draw a different vector (possibly different in both magnitude and direction) at each point, and these would tend to overlap! To deal with this, we introduce an alternative representation using **electric field lines.**

Imagine a region containing charged particles that are fixed in place and create an electric field at every point in the region. We now draw an electric field line. Find the direction of the electric field at a starting point and take a small step in the direction of this vector. At the new point, again find the direction of the electric field and take a small step in this direction. Continuing this process creates a series of points that we connect with a smooth line. This line is an electric field line. The final step is to put a small arrow on the line to indicate the direction of travel. Starting at a new point in the region leads to a new electric field line. If we do this from enough different starting points, we can get a sample of lines throughout the region, as shown in Figure 20-12.

This sounds like a long, involved process—and it is. The good news is that we can often learn a great deal through qualitative sketches of electric field lines, which are much easier than doing the detailed calculations. We can use our intuition to draw the electric field lines surrounding an isolated positive source charge. A positive test charge would be repelled directly away from the positive source charge, so the electric field lines should start on the source charge and continue radially outward to infinity, as shown in Figure 20-13. The electric field vector at any point in space is then tangent to the electric field line that would pass through that point. Notice that this is consistent with the drawing in Figure 20-11 of the electric field vectors for a positive source charge. (If we are interested in a point that does not lie on one of the electric field lines that have been drawn, we can find the approximate direction of a line that passes through the point by looking at the surrounding lines.)

Notice that the electric field lines in Figure 20-13 start at locations equally spaced around the source charge. When the electric field lines are drawn this way, they give us visual information about the magnitude of the electric field vector at any point. Notice that the field lines are close together near the source charge, where the electric field is strong, and the field lines are far

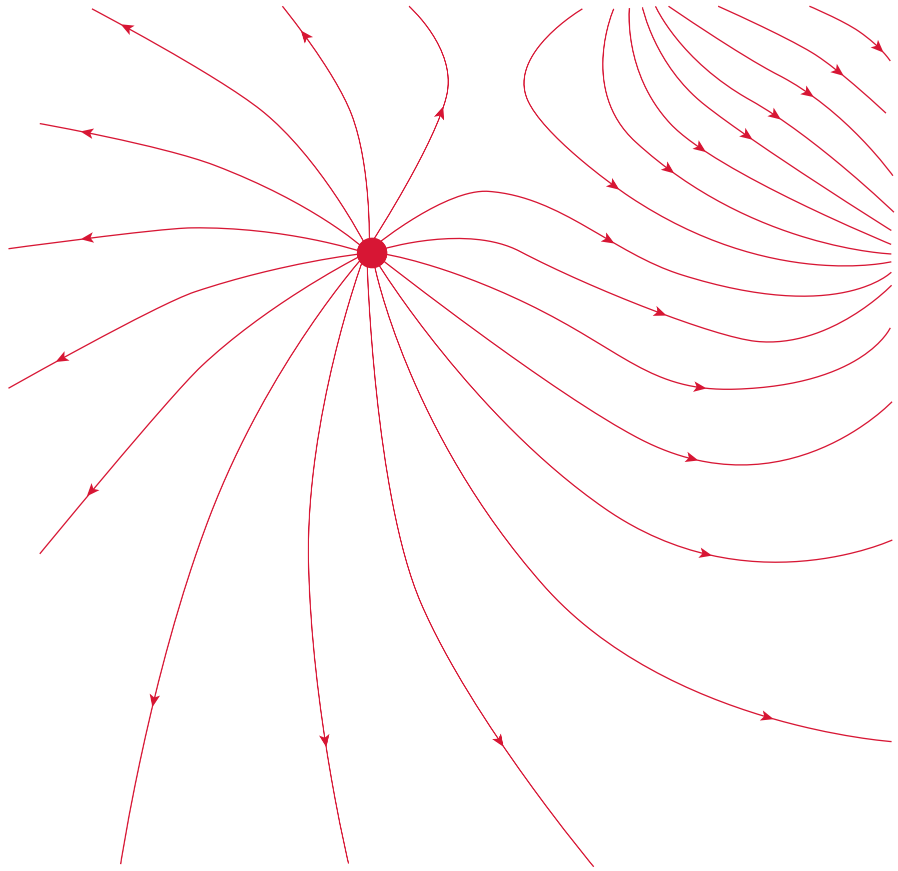


Figure 20-12 Electric field lines represent the total electric field in a region of space.

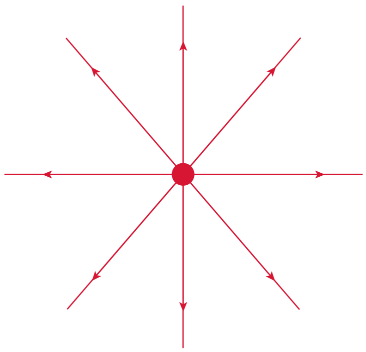


Figure 20-13 Electric field lines due to an isolated positive source charge.

from each other at locations away from the source charge, where the electric field is weak. Indeed, any point in space that is 3 centimeters from the source charge should have the same size electric field. The spherical symmetry of our field lines ensures that the spacing between adjacent field lines is the same for all these points. In general, the strength of the electric field (that is, the length of the vector) is greater in regions where the electric field lines are closer together. Another way to say this is that the electric field is proportional to the density of electric field lines.

Q: How would the diagrams in Figure 20-13 change if the positive source charge were replaced by a negative source charge?

A: The arrows in the drawing would point in opposite directions.

When more than one source charge is present in a region, the field lines represent the total electric field in the region due to all the source charges. At locations very close to one of the source charges, the electric field lines should still be radially symmetric about that source charge (as its contribution will dominate). The number of field lines originating on a positive source charge or ending on a negative source charge should be proportional to the magnitude of the charge. This is because the electric field strength at a location 1 centimeter from a +2-coulomb source charge should be twice as big as the electric field strength 1 centimeter from a +1-coulomb source charge. Figure

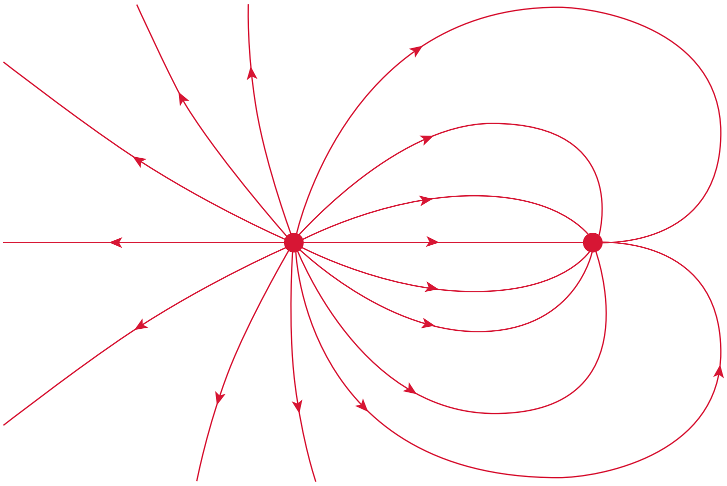


Figure 20-14 The electric field lines indicate that the magnitude of the positive charge is twice that of the negative charge.

20-14 shows the electric field lines around two source charges, one positive and one negative.

Are You On the Bus?

Q: If the positive source charge in Figure 20-14 has a charge of +6 microcoulombs, what is the magnitude of the charge on the negative source charge?

A: The number of electric field lines originating on a positive charge or ending on a negative charge is proportional to the size of the electric charge. Twelve field lines leave the positive source charge, but only six of those field lines end on the negative source charge. The negative charge must be half the size of the positive charge. It must have a charge of −3 microcoulombs.

Notice that electric field lines always begin on positive charges and end on negative charges. If the region you are considering contains more positive than negative charges, some lines will leave the region. If the region contains more negative charges, some lines will come in from outside the region.

Are You On the Bus?

Q: Why do electric field lines never cross?

A: The electric field lines represent the total electric field in a region due to all source charges. The electric field vector at any point in space is tangent to the electric field line through that point. If two electric field lines crossed at any point, the electric field at that point would have to be tangent to both electric field lines. It would have to point in two directions at a single location. This is not possible.

A common situation in which electric field lines provide physical insight is the case of charged parallel metal plates. If electrons are taken from one metal plate and placed on the other, both plates end up with the same amount of excess electric charge (positive on one plate and negative on the other). The charge will spread out on the facing surfaces of the plates until the charge density is uniform. Electric field lines originate on positive charge and end on negative charge, so the uniform charge distribution on the plates dictates that the electric field lines between the plates be parallel to each other, perpendicular to the plates, and uniformly spaced, as shown in Figure 20-15.

These electric field lines are very simple, and they represent an electric field that is uniform in strength everywhere between the two plates. The simplicity of the electric field lines predicts a result that is highly nonintuitive: if you place a small positive test charge at a location halfway between the plates, it experiences the same electric field (and hence the same electric force) as it would very close to the positive plate (or, in fact, anywhere else between the plates). It is possible to show that this is indeed the case mathematically, but it is much easier to use the concept of electric field lines.

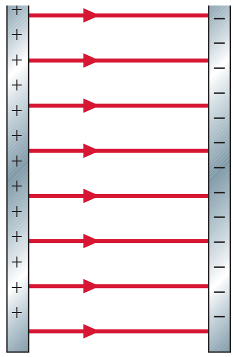


Figure 20-15 The electric field lines between charged metal plates indicate that the electric field is everywhere uniform between the plates.

Electric Potential

We must do work to lift a 1-kilogram block in Earth's gravitational field, and this work increases the gravitational potential energy of the block. When the block is released, it falls and the gravitational potential energy is changed to kinetic energy. Lifting a 5-kilogram block the same distance requires five times the work and the block gains five times the gravitational potential energy. In other words, knowing the energy that is gained by the 1-kilogram block allows us to calculate the energy that would be gained by any other mass moved between the same two points.

In the same way, work is required to move a charged particle in an electric field, and this work changes the **electric potential energy** of the particle. When we release the particle, this electric potential energy can be converted to kinetic energy. Therefore, we define the electric potential energy the same way we did for gravity. The electric potential energy of a charged object is equal to the work done in bringing the object from some zero reference location to the object's location.

◀ electric potential energy = work from reference point

As with gravitational potential energy, the value of the electric potential energy does not depend on the path, but it does depend on the reference location, the location of the object, and the charge on the object (Figure 20-16).

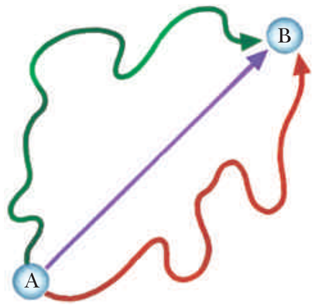


Figure 20-16 The work performed in taking a charge from A to B does not depend on the path.

As with gravitational potential energy, the actual value of the electric potential energy is not important in physical problems; it is only the difference in energy between points that matters. If 10 joules of work is required to move a charged object from point A to point B, the electric potential energy of the object at point B is 10 joules higher than at point A. If point A is the zero reference point, the electric potential energy of the object at point B is 10 joules.

Because objects with different charges have different electric potential energies at a given point, it is often more convenient to talk about the energy available due to the electric field without reference to a specific charged object. The **electric potential** V at each point in an electric field is defined as the electric potential energy EPE divided by the object's charge q:

$$V = \frac{EPE}{q}$$

Notice that it doesn't matter which charged object we use to define the electric potential. This quantity is numerically equal to the work required to bring a positive test charge of 1 coulomb from the zero reference point to the specified point. The units for electric potential are joules per coulomb (J/C), a combination known as a **volt** (V). Because of this, we often speak of the electric potential as a *voltage.*

◀ electric potential $= \dfrac{\text{electric PE}}{\text{charge}}$

Defining the electric potential allows us to obtain the electric potential energy for any charged object by multiplying the potential by the charge. Once again, it is only the *potential difference* that matters. For instance, a 12-volt battery has an electric potential difference of 12 volts between its two terminals. This means that 1 coulomb of charge moving from one terminal to the other would gain or lose (1 coulomb)(12 volts) = 12 joules of energy.

Q: How much work is required to move 3 coulombs of positive charge from the negative terminal of a 12-volt battery to the positive terminal?

A: Each coulomb requires 12 joules of work, so 3 coulombs require 36 joules: $W = qV = (3 \text{ coulombs})(12 \text{ volts}) = 36 \text{ joules}$.

If the potential difference, or voltage, between two points is high and the points are close together, the electric field can be strong enough to tear electrons from molecules in the air. The electrons are pulled one way, and the remaining positive ions are pulled the other way. Because electrons have much less mass than ions, they quickly accelerate to high velocities. As they accelerate they collide with other molecules, knocking additional electrons loose. This forms a cascade of electrons that we call a *spark*. Dry air breaks down in this fashion when the electric field reaches about 30,000 volts per centimeter.

Summary

Objects become electrically charged or uncharged by transferring charges. An uncharged object has equal numbers of positive and negative charges. Positively charged objects may have an excess of positive charges or a deficiency of negative charges. In an isolated system, the total charge is conserved. Charges can flow through conductors such as metal wires but not through insulators such as silk threads.

In electricity, unlike gravity, there are two different types of charge, and the direction of the force depends on the relative types of the charges. Like charges repel; unlike charges attract. A charged object can attract an uncharged object because of an induced separation of charges.

The electric force has the same mathematical form as the gravitational force. The electric force, however, can be repulsive as well as attractive. Electric charge comes in whole-number multiples of one fixed size. The electric charges on electrons and protons have the same size but are opposite in sign.

Electric charges are surrounded by an electric field that is equal to the force experienced by a unit positive charge. This electric field is a vector field with a size and a direction at each point in space. The units for electric field are newtons per coulomb. Electric field lines represent the electric field in a region of space. The electric field at any location points in a direction tangent to the field line at that location, and the strength of the electric field is proportional to the local density of electric field lines.

The electric potential energy of a charged object in an electric field is equal to the work done in bringing the object from some zero reference location. Only the differences in electric potential energy matter in physical situations. The electric potential equals the electric potential energy divided by the object's charge. This quantity is numerically equal to the work required to bring a positive test charge of 1 coulomb from the zero reference point. The units for electric potential are joules per coulomb, or volts.

CHAPTER 20 *Revisited*

Sparks occur when the electric field becomes strong enough to pull electrons from atoms. These electrons are accelerated by the electric field, obtaining sufficient kinetic energy to knock electrons from other atoms. This avalanche of electrons produces the visible sparks that we see. Electric fields of 30,000 volts per centimeter are strong enough to break down atoms in dry air. Because the electric field is proportional to the electric potential difference and inversely proportional to the distance, it takes large voltages to produce long sparks.

Everyday Physics | *Lightning*

Everybody has seen lightning (Figure A) and has probably been frightened by this spectacular display of energy at some time or another. Lightning is unpredictable and seems to occur randomly and instantaneously. Everything but the thunder is usually over in less than half a second. And yet the effects are not inconsequential: each year lightning causes about $80 million of damage in the United States. On the average, lightning kills 85 people a year and injures another 250 in the United States alone.

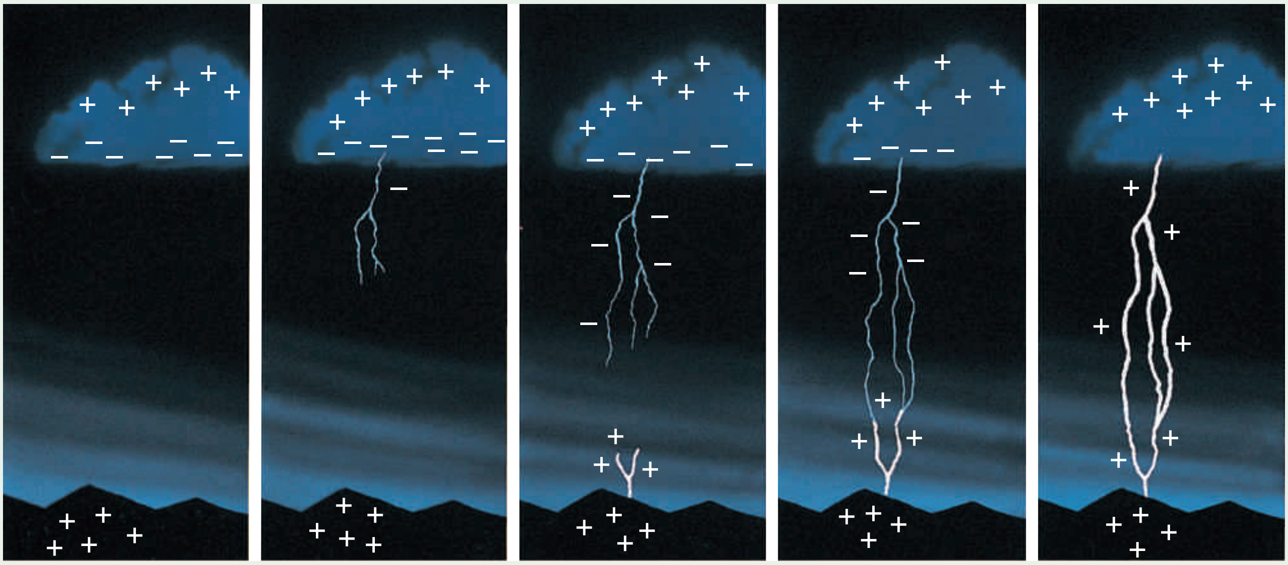
We still have a lot to learn about lightning. For instance, we don't know what causes the initial buildup of charge, what determines the paths the lightning takes, or what triggers the lightning bolts.

We do know that during a storm, clouds develop a separation of electric charge, with the tops of the clouds positively charged and the bottoms negatively charged, as shown in Figure B. The production of a lightning bolt begins when the negative charge on the bottom of the cloud gets large enough to overcome air's resistance to the flow of electricity (Figure C) and electrons begin flowing toward Earth along a zigzag, forked path at about 60 miles per second! As the electrons flow downward, they collide with air molecules and ionize them, producing more free electrons. Even though the current may be as large as 1000 amperes, this is not the lightning bolt we see.

Meanwhile, as the electrons approach the ground, the ground becomes more and more positively charged because of the repulsion of electrons in the ground. This positively charged region moves up through any conducting objects on the ground—houses, trees, and people—into the air (Figure D). When the downward-moving electrons meet the upward-flowing positive regions at an altitude of 100 meters or so, they form a complete circuit (Figure E), and the lightning begins. In less than 1 millisecond, up to 1 billion trillion electrons may reach the ground; the current may be as high as 200,000 amperes. Although the flow of charge is downward, the point of contact between the cloud charge and the ground charge is flowing *upward* at about 50,000 miles per second. (Observers report that the lightning bolt is moving downward because the pathway that is lit is the initial forked pathway coming from the cloud.) The upward surge heats the air to 50,000°F. One meter of the lightning's path may shine as bright as 1 million 100-watt lightbulbs (Figure F). This rapid heating of the air along the lightning's path also produces a shock wave that we hear as the thunder. What we see as a single lightning bolt is usually several bolts in rapid succession along the same path.

Image not available due to copyright restrictions

1. People near a lightning strike report that their hair begins to stand on end, warning of the impending strike. What causes this?
2. What process causes the loud noise that we call thunder?



(B) (C) (D) (E) (F)

Figures B–F The development of a lightning bolt.

Key Terms

charge A property of elementary particles that determines the strength of its electric force with other particles possessing charge. Measured in coulombs, or in integral multiples of the charge on the proton.

charged Possessing a net negative or positive charge.

conductor A material that allows the passage of electric charge. Metals are good conductors.

conservation of charge In an isolated system, the total charge is conserved.

coulomb The SI unit of electric charge; the charge of 6.25×10^{18} protons.

electric field The space surrounding a charged object, where each location is assigned a value equal to the force experienced by one unit of positive charge placed at that location. Electric field is measured in newtons per coulomb.

electric field lines A representation of the electric field in a region of space. The electric field is tangent to the field line at any point, and its magnitude is proportional to the local density of field lines.

electric potential The electric potential energy divided by the object's charge; the work done in bringing a positive test charge of 1 coulomb from the zero reference location to a particular point in space. Electric potential is measured in joules per coulomb.

electric potential energy The work done in bringing a charged object from some zero reference location to a particular point in space.

grounding Establishing an electrical connection to Earth in order to neutralize an object.

insulator A material that does not allow the passage of electric charge. Ceramics are good insulators.

inverse-square relationship A relationship in which a quantity is related to the reciprocal of the square of a second quantity; the electric force is proportional to the inverse square of the distance from the charge.

volt The SI unit of electric potential, 1 joule per coulomb.

Questions and exercises are paired so that most odd-numbered are followed by a similar even-numbered.
Blue-numbered questions and exercises are answered in Appendix B.
indicates more challenging questions and exercises.
WebAssign Many Conceptual Questions and Exercises for this chapter may be assigned online at WebAssign.

Conceptual Questions

1. A handheld glass rod can be charged by rubbing it with silk or a plastic bag while holding it in your hands. Would you conclude from this that glass is a conductor or an insulator? Why?
2. Inexperienced physics teachers often demonstrate the use of the electroscope by touching it with a charged glass rod at a single point. More experienced teachers typically drag the length of the rod across the top of the electroscope to increase the desired effect. Why does this help?
3. A neutral metal sphere is hanging from the ceiling by a thin insulating thread. Which of two equally charged rods, one metal and one rubber, would transfer the greater amount of charge to the metal sphere when touched to the sphere? Why?
4. Could you use fur to charge a metal rod that is held in your bare hand? Explain.
5. Why is it easier to charge a balloon on a dry day than on a humid day?
6. Why is it easier to demonstrate electrostatic phenomena in Fairbanks, Alaska, than in Honolulu, Hawaii?
7. Before an aircraft is fueled from a truck, a wire from the truck is attached to the aircraft. Why?
8. Why do you sometimes feel a shock when you touch a metal doorknob after crossing a carpeted floor?
9. If one rubber rod is charged with fur and another with plastic, they attract each other. How would you explain this?
10. You use a piece of silk to charge a glass rod. Does this leave the silk with a positive or a negative charge? Explain.
11. When you rub balloons with wool and place them near bits of paper, you find three categories of behavior. Balloons attract wool and bits of paper but repel other balloons. Pieces of wool attract balloons and bits of paper but repel other pieces of wool. Bits of paper are attracted to both the balloons and the pieces of wool but do not react with other bits of paper. When we have three distinct behaviors, why do we only need two types of charge in our model?
12. How would you know if you discovered a third kind of charge? What other objects would you expect the third kind of charge to attract and repel?
13. Your classmate claims that magnets are really just charged rods with opposite charges at the two ends. What experi-

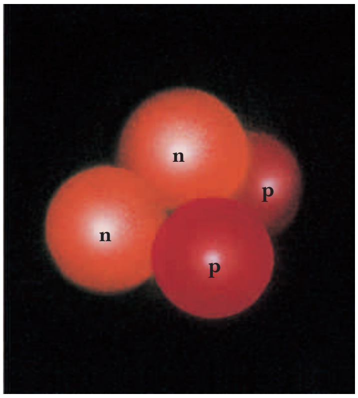
ment could you perform to convince your classmate otherwise?

14. You find that the north pole of a bar magnet is attracted to a charged glass rod. You then discover that the south pole of the magnet is also attracted to the same rod. Will the north pole be attracted or repelled or experience no force if brought near a charged rubber rod? Explain.

15. You have three small balls, each hanging from an insulating thread. You find that balls 1 and 2 repel one another and that balls 2 and 3 repel one another. Will balls 1 and 3 attract or repel one another? Explain.

16. You have three small balls, each hanging from an insulating thread. You find that balls 1 and 2 attract one another and that balls 2 and 3 repel one another. Can any of the balls be neutral? If not, why not? If so, which one?

17. The nucleus of a helium atom contains two protons and two uncharged neutrons. If two electrons surround the nucleus, what is the total charge of the atom?



18. In the modern view of electricity, does an object with a negative charge have an excess of positive charge, an excess of negative charge, a deficiency of negative charge, or a deficiency of positive charge? Explain.

19. Describe how a charged balloon sticks to a wall.

20. When a charged comb is brought near bits of paper, the bits are first attracted to the comb and then repelled after the bits come in contact with the comb. Describe how the charge on the comb and on the bits of paper changes during this process.

21. Why are neutral objects attracted to both negatively and positively charged objects?

22. Describe how a negatively charged rod attracts (a) an uncharged conducting object and (b) an uncharged insulating object.

23. How could you use an electroscope of known positive charge to determine the sign of a charged rod?

24. How could you use a negatively charged rod to determine whether a charged electroscope has a negative or a positive charge without changing the charge on the electroscope?

25. Describe how you would use a positively charged rod to give an electroscope a negative charge.

26. Describe how you would use a positively charged rod to put a positive charge on an electroscope.

27. An electroscope initially has a net negative charge. Why do the foils come together and stay together when a human hand touches the electroscope?

28. An electroscope is initially given a net positive charge. You find that as you bring your hand near the top of the electroscope without touching it, the foils come slightly closer together. When you remove your hand, they move back apart. How do you account for this?

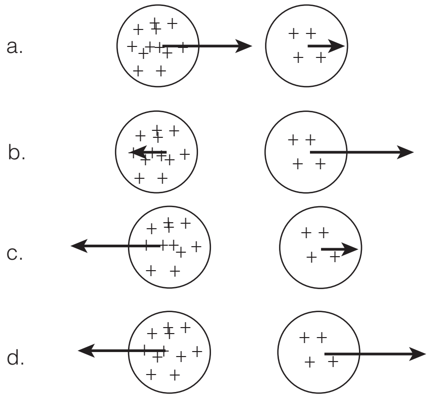
29. You have two metal spheres on insulating stands. You bring them each in turn near the ball of a positively charged electroscope. In each case the deflection of the leaves is reduced slightly. You then bring them each in turn near the ball of a negatively charged electroscope and find that sphere A causes the deflection to increase, whereas sphere B causes it to decrease. What can you say about the net charge on each of the two metal spheres?

30. You hold a charged rubber rod near the ball of an electroscope and then briefly touch the ball with your finger. After removing the rubber rod, you bring a charged glass rod near the ball. Will the deflection of the leaves increase, decrease, or stay the same? Why?

31. When Coulomb was developing his law, he did not have an instrument for measuring charge. And yet he was able to obtain spheres with $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$ of some original charge. How might he have used a set of identical spheres to do this?

32. You have three identical metal spheres on insulating stands. The spheres hold charges $Q_A = -2q$, $Q_B = -q$, and $Q_C = 4q$. First, sphere A is brought into contact with sphere C and separated. Second, sphere C is brought into contact with sphere B and separated. What is the resulting charge on sphere B?

33. Two uniformly charged insulating spheres are held atop fixed posts. The charge on one of the spheres is three times the charge on the other. Which diagram correctly represents the magnitude and direction of the electric forces on the two spheres? Explain your answer.



34. You have two positively charged objects. Object B has twice the charge of object A. When they are brought close together, object A experiences an electric force of 10 newtons. Is the electric force experienced by object B greater than, equal to, or less than 10 newtons? Explain.

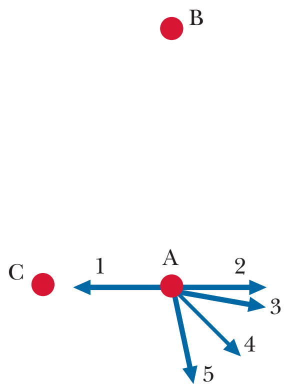
35. Two charged objects are very far from any other charges. If the distance between them is cut to one-third its original value, what happens to the electric force between them?

36. Assume that you have two identically charged objects separated by a certain distance. How would the force change if the objects were twice as far away from each other?

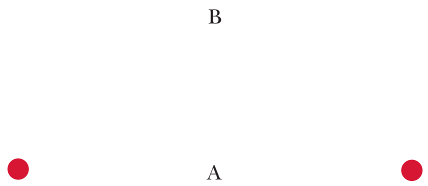
37. Two charged objects are very far from any other charges. If the charge on each of them is tripled, what happens to the electric force between them?

38. Assume that you have two identically charged objects separated by a certain distance. How would the force change if one object had twice the charge and the other object kept the same charge?

39. Three identical charges are arranged as shown in the following figure. The distance from A to B is twice the distance from A to C. Which of the vectors best represents the force on charge A due to B and C? Justify your answer.



40. Two identical charges are fixed at the locations indicated in the following figure. A third charge, identical to the other two, is placed first at position A and then position B. For each case, determine the direction of the net electrical force on the third charge. If the force is zero, state that explicitly.



41. Although the formulas for electric and gravitational forces have the same form, there are differences in the two. What are some of these differences?

42. What are the similarities between electric and gravitational forces?

43. When we approach another person, we are not aware of the gravitational and electric forces between us. What are the reasons in each case?

44. Even though electric forces between electrons and protons are much stronger than gravitational forces, gravitational forces determine the motions in the Solar System. Why?

45. Why are the accelerations of all charged objects near a charged sphere not the same?

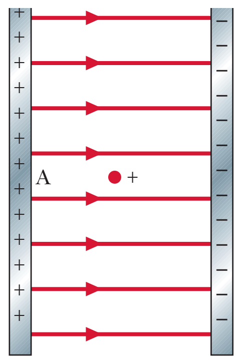
46. In this chapter we compared the electric and gravitational forces between an electron and a proton. Why is the result valid for all separations?

47. A metal sphere with a charge of +2 coulombs experiences an electric force of 20 newtons directed to the left. If the charge on the sphere is increased to 6 coulombs, what force will it experience?

48. A metal sphere with a charge of +4 coulombs experiences an electric force of 50 newtons directed to the left. If the charge on the sphere is changed to −8 coulombs, what force will it experience?

49. The electric field lines between two parallel charged plates are everywhere perpendicular to the plates and parallel to each other, as shown in the following figure. A small positively charged particle placed at the midpoint between the plates experiences an electric force of magnitude F. If the charged particle were instead placed at location A, close to the positive plate, would the force on the particle be greater than, equal to, or less than F? Use the concept of field lines to justify your answer.

50. For the situation described in Question 49, the charged particle is released from rest at the midpoint and strikes the negative plate at speed v. If the particle had instead



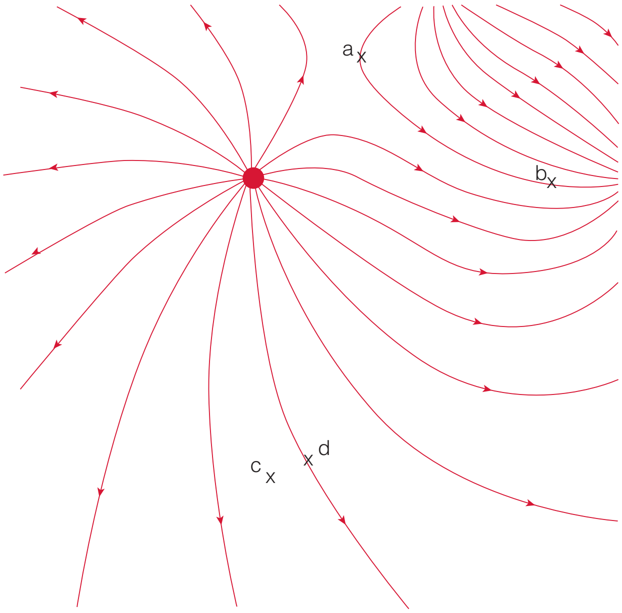
been released from rest at point A, would it have struck the negative plate at a speed greater than, equal to, or less than v? Explain your reasoning.

51. An electron and a proton are released in a region of space where the electric field is vertically upward. How do the electric forces on the electron and proton compare?

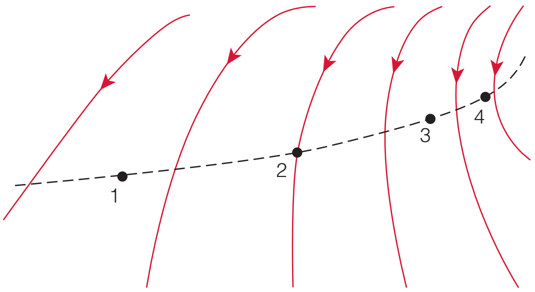
52. How do the accelerations of the electron and proton in Question 51 compare?

53. A portion of an electric field is shown in the following figure. Four locations within the field are marked with **x**'s and labeled **a, b, c,** and **d.** Draw the vectors representing

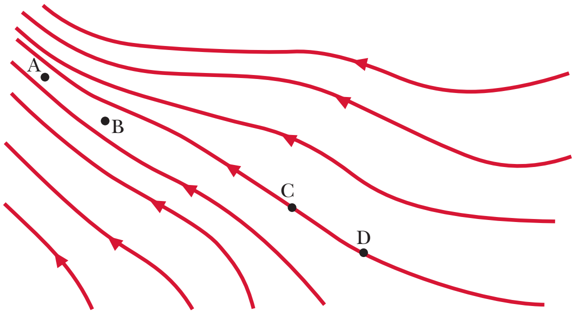
the force on a positive test charge when placed at each of these locations. Explain your reasoning.



54. The field lines (solid lines) shown in the following figure represent the electric field in a certain region of space. A charged particle is released with an initial velocity and follows the trajectory from points 1 to 4, as shown by the dashed line. Draw vectors to represent the force on the charged particle at each location. What is the sign of the charged particle? Explain how you can tell.



55. The following figure shows the electric field lines in a region of empty space. Is the electric field in the region between A and B greater than, equal to, or less than the electric field in the region between C and D? Explain.



56. In the figure described in Question 55, points A and B and points C and D are the same distance apart. Is the potential difference between A and B greater than, equal to, or less than the potential difference between C and D? Explain.

57. A proton is observed moving at speed v_o at point A in space, where the electric potential is 750 volts. It moves to point B, where the electric potential is 950 volts. Is its speed at point B greater than, equal to, or less than v_o? Explain.

58. An electron is observed moving at speed v_o at point A in space, where the electric potential is 750 volts. It moves to point B, where the electric potential is 950 volts. Is its speed at point B greater than, equal to, or less than v_o? Explain.

59. How is the value of the electric potential at each point in space defined?

60. What is the definition of the value of the electric potential energy of a charged object at each point in space?

61. A proton is released from rest in a uniform electric field. Does the proton's electric potential energy increase or decrease? Does the proton move toward a location with a higher or lower electric potential? Explain.

62. An electron is released from rest in a uniform electric field. Does the electron's electric potential energy increase or decrease? Does the electron move toward a location with a higher or lower electric potential? Explain.

63. A chlorine ion contains 17 protons, 18 neutrons, and 18 electrons. What is the net charge of the ion?

64. What is the net charge of an iodine ion containing 53 protons, 74 neutrons, and 54 electrons?

Exercises

65. The nucleus of a certain type of uranium atom contains 92 protons and 143 neutrons. What is the total charge of the nucleus?

66. How many electrons does it take to have a total charge of -1 C?

67. What is the electric force of attraction between charges of $+3$ C and -6 C separated by a distance of 2 m?

68. By what distance must two charges of $+2$ C and $+8$ C be separated so that the repulsive force between them is 3.6×10^{10} N?

69. The nucleus of lithium contains three protons and three neutrons. When two electrons are removed from the neutral lithium atom, the remaining electron has an average distance from the nucleus of 0.018 nm. What is the force between the electron and the nucleus at this separation?

70. In an ionic solid such as ordinary table salt (NaCl), an electron is transferred from one atom to the other. If a distance of 0.1 nm separates the atoms, how strong is the electric force between them?

71. How much stronger is the electric force between two protons than the gravitational force between them?

72. Calculate the ratio of the electric force to the gravitational force between two electrons.

73. A 5-mC charge experiences a force of 2 N directed north. What is the electric field (magnitude and direction) at the location of the charge? Find the force on a −20-mC charge that replaces the 5-mC charge.

74. A −15-mC charge experiences a force of 25 N directed west. What is the electric field (magnitude and direction) at the location of the charge? If this charge is removed and not replaced, what is the electric field at this location?

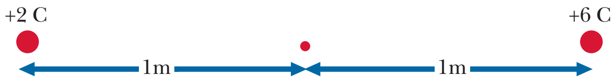
75. What is the electric field at a distance of 3 cm from 4 mC of negative charge?

76. What is the electric field 2 m from 8 C of positive charge?

77. What is the electric field at a distance of 0.2 nm from a carbon nucleus containing six protons and six neutrons?

78. What is the electric field 5.3×10^{-11} m from a proton?

79. What is the electric field midway between charges of 2 C and 6 C separated by 2 m?



80. What is the electric field midway between an electron and a proton separated by 0.2 nm?

81. What is the force on a proton located in an electric field of 3000 N/C? What is the proton's acceleration?

82. How would the values obtained in Exercise 81 change if the proton were replaced by an electron?

83. The electric potential energy of an object at point A is known to be 50 J. If it is released from rest at A, it gains 30 J of kinetic energy as it moves to point B. What is its potential energy at B?

84. If the object in Exercise 83 had a charge of +2 C, what would be the potential difference between points A and B?

85. How much work does a 12-V battery do in pushing 2 mC of charge through a circuit containing one lightbulb?

86. Points A and B each have an electric potential of +9 V. How much work is required to take 3 mC of charge from A to B?

87. What was your electric potential relative to a metal pipe if a spark jumped 1.3 cm (0.5 in.) through dry air from your finger to the pipe?

88. How far can a spark jump in dry air if the electric potential difference is 6×10^5 V? (This is why high-voltage sources are surrounded by a vacuum or an insulating fluid.)

21 Electric Current

▶We live in an electrical world. It is hard to imagine a world without electricity. How does the movement of electrons produce the light that allows us to work and play indoors and at night?

(See page 462 for the answer to this question.)

This night scene of Boathouse Row in Philadelphia is an example of the artistic use of lighting.

NEW phenomena often enter our world view accidentally, a result of someone with an open mind making an observation or conducting an experiment. Initially, these phenomena are often just curiosities, new challenges to our understanding of the world. Many of them give us new ways to do things and eventually find widespread use as our understanding of them increases. In some cases, however, a whole new technology is born. The electrical properties we studied in Chapter 20 were such curiosities. In the early days of investigation into these phenomena, little use was made of electricity other than building devices to amaze and shock friends. In the past 200 years, electricity has developed into a major technological presence in our society. Today, the United States, with 5% of the world's population, uses about one-fourth of the electricity generated in the world. This means that the average U.S. citizen uses five times the electricity the average world citizen does, including those in developed countries. When electricity usage of U.S. citizens is compared with that of citizens in third-world countries, the per capita usage is very much larger.

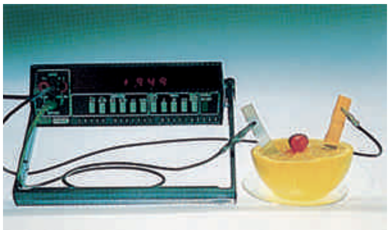
An important advance in understanding electricity came with the development of batteries. Batteries could make electric charges flow continuously as a current rather than in short bursts. This development eventually led to modern electric circuits that have had an obvious impact on our lives. Electric currents can produce heat and light, run motors and radios, operate sound systems and computers, and so on. The list is nearly endless.

An Accidental Discovery

Near the end of the 18th century, Italian anatomist Luigi Galvani announced that an excised frog's leg twitched when he touched it. It was well known from firsthand experiences with static electricity that twitches in humans could result from electric charges. But this was just a leg! Galvani was convinced that he had discovered the secret life force that he believed existed within all animals. Galvani found that a twitch occurred when he touched the frog's leg with a metal object but not when he touched it with an insulator. Furthermore, the metal had to be different from the metal hook holding the other end of the leg. It took two *different* metals to produce the twitch.

Another Italian scientist, Alessandro Volta, heard of these results and performed many experiments with sensitive electroscopes, searching for evidence that electric charge resides in animal tissue. He eventually convinced himself that the electricity was not in the frog's leg but was the result of touching it with the two metals.

Volta tried many different materials and discovered that an electric potential difference was produced whenever two different metals were joined. Some combinations of metals produced larger potential differences than others. One of his demonstrations was to place his tongue between pieces of silver and zinc that were held together at one end. The flow of electricity caused his tongue to tingle. You may have felt a similar sensation when a piece of metal or metal foil touched a metal filling in your tooth. A more detailed explanation of this process would take us beyond the scope of this book, but it is enough to realize that some chemical process moves charges from one metal to the other, creating a potential difference.



A modern version of Volta's demonstration that zinc and copper rods inserted into a grapefruit produce a potential difference. (The cherry is for decorative purposes only.)

Batteries

Volta used his discovery that two dissimilar metals produce an electric potential difference to make the first battery. He made a single cell by putting a

Are You On the Bus?

Q: How many $1\frac{1}{2}$-volt cells does it take to make a 9-volt battery?

A: Because the voltages of the cells add, it takes six cells.

piece of paper that had been soaked in a salt solution between pieces of silver and zinc. He then stacked these cells in a pile like the one shown in Figure 21-1. By doing this he was able to produce a larger potential difference. The potential difference of all the cells together was equal to the sum of the potential differences due to the individual cells.

A cell's potential difference depends on the choice of metals. Those commonly used in batteries have potential differences of $1\frac{1}{2}$ or 2 volts. Putting lead and lead oxide plates (known as electrodes) into dilute sulfuric acid (the electrolyte) makes a simple 2-volt cell. Connecting six of these cells together makes a 12-volt car battery. The positive electrode of one cell is connected to the negative electrode of the next. The connectors sticking out of the first and last cells (the terminals) are usually marked with a + or − sign to indicate their excess charges.

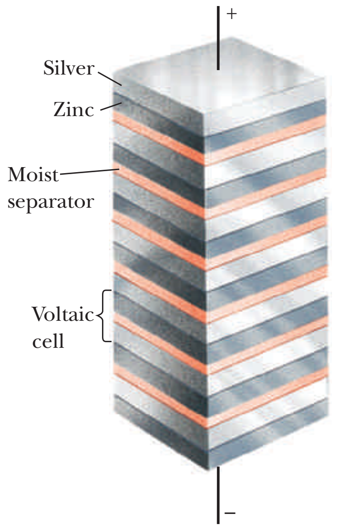


Figure 21-1 A schematic of Volta's battery composed of zinc and silver plates with moist paper between them.

Batteries that can be recharged are called *storage batteries.* This recharging is usually accomplished with an electric charger with the negative terminals of the charger and battery connected to each other and the positive terminals connected to each other. The charger forces the current to run backward through the battery, reversing the chemical reactions and refreshing the battery. This electric energy is stored as chemical energy in the battery for later use.

Nonrechargeable flashlight batteries (sometimes called *dry cells*) are constructed with a carbon rod down the center, as shown in Figure 21-2. Carbon is a conductor and replaces one of the metals. A moist paste containing the electrolyte surrounds the rod. The other electrode is the zinc can surrounding the paste. The cell is then covered with an insulating material. These $1\frac{1}{2}$-volt cells are often used end to end (both pointing in the same direction) to provide the 3 volts used in many flashlights.

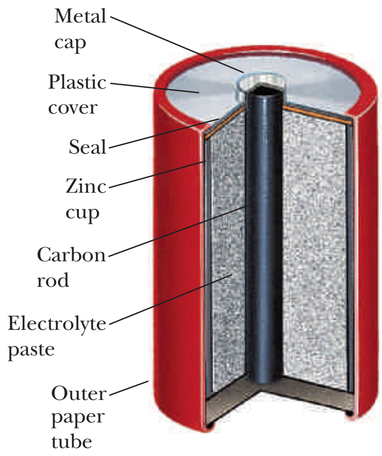


Figure 21-2 Schematic of a standard flashlight-type dry cell.

The voltage produced by an individual cell depends on the materials used and not on its size. The size determines the total amount of chemicals used and therefore the total amount of charge that can be transferred. We have seen that the voltage can be increased by placing cells end to end in a row, an arrangement called **series.** Cells (or batteries) can also be placed side by side, or **parallel,** as shown in Figure 21-3. The cells must all point in the same direction. This parallel arrangement does not increase the voltage but does increase the effective size of the battery. The larger amount of chemicals means that the batteries will last longer before they run down.

Are You On the Bus?

Q: When might you wish to connect batteries in parallel?

A: This is especially useful when the batteries are difficult to change, as in remote radio transmitters on mountaintops.

Household electricity has a number of differences from the electricity produced by batteries. The electricity in homes is supplied at a much higher voltage, typically 120 volts. (For historical reasons, household voltage is often referred to as "one-ten," but modern national standards in the United States and Canada specify that the nominal voltage be 120 volts, plus or minus 5%.) This electricity is supplied by huge electric generators in dams or power plants

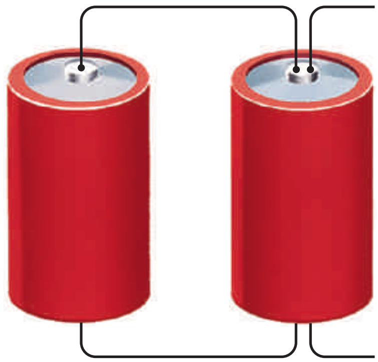


Figure 21-3 Two cells in parallel have the same voltage as a single cell but last twice as long.

Courtesy of Vector Products, Inc.

A car battery can be recharged with a battery charger.

burning coal, oil, gas, or nuclear fuel. (We will discuss generators in the next chapter.) In household circuits the flow of charge (current) reverses direction 120 times a second. This is called *60-cycle alternating current* (ac) as opposed to the *direct current* (dc) supplied by batteries that flows in a single direction.

Because most of the uses of electricity we will discuss in this chapter depend only on the flow of charge and not on its direction, we will simplify our discussions by referring to flashlight batteries. Another advantage of using batteries is that they have low voltages, so you can experiment with some of the ideas mentioned here without fear of being shocked. **Household electricity**, on the other hand, **is dangerous and should not be used in experiments**.

Pathways

The invention of batteries provided experimenters with a way of producing a continuous flow of charge: a **current**. Although the properties of interacting electric charges are unchanged from the preceding chapter, new effects are observed when charges move through conducting pathways in a continuous fashion. The simplest way to learn about these is to experiment with a simple flashlight battery, some wire, and a few flashlight bulbs. The basic properties of household electricity are then simple extensions of these ideas.

Imagine that you have a battery, a wire, and a bulb. From your previous experiences, can you predict an arrangement that will light the bulb? A common response is to connect the bulb to the battery, as shown in Figure 21-4. The bulb doesn't light. It doesn't matter which end of the battery is used or which part of the bulb is touched. The charges don't flow from one end of the battery to the bulb and cause it to light. Suppose you hold the wire to one end of the battery and the metal tip of the bulb to the other end. Touching the free end of the wire to various parts of the bulb will eventually yield success. Two of the four possible arrangements for lighting the bulb are shown in Figure 21-5.

Comparing the methods for lighting the bulb with the ways that don't light it indicates that we must use two parts of each object: the two ends of the wire, the two ends of the battery, and the two metal parts of the bulb. The two parts of the bulb are the metal tip and the metal around the base (this is threaded on many bulbs). Whenever all six ends are connected in pairs, forming a continuous loop (no matter how you do it), the bulb lights.

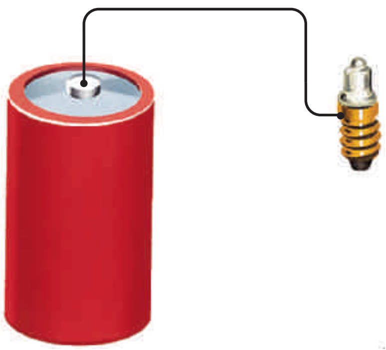


Figure 21-4 An unsuccessful attempt to light a bulb.

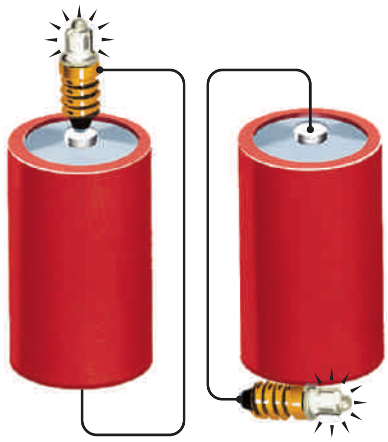


Figure 21-5 Successful ways to light a bulb.

Q: Would the bulb still light if you turned the battery in the opposite direction in the arrangements shown in Figure 21-5?

A: Either end of the battery works because reversing the battery does not change the conducting path.

Figure 21-6 Schematic drawing showing the continuous conducting path through a bulb.

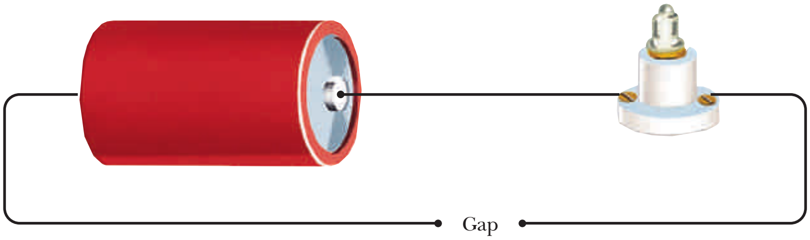


Figure 21-7 When a conducting material is placed between wires at the gap, the bulb lights.

The continuous conducting path (known as a **complete circuit**) allows the electric charges to flow from one end of the battery to the other end. If the bulb is to light, the conducting path must go through the bulb. You can verify this by examining a bulb with unfrosted glass, as illustrated in Figure 21-6. The part of the bulb that glows is a thin wire (called the *filament*) that is supported by two thicker wires coming out of the base. The bulb has to be taken apart to see that one of these support wires is connected to the metal tip and the other to the metal side of the bulb. The rest of the bulb is made from insulating materials so that there is a single conducting path through the filament within the bulb.

You can use the concept of a complete circuit to investigate the conductivity of different household materials. Connect a flashlight bulb to a flashlight battery, leaving a gap in the circuit, as shown in Figure 21-7. Test a variety of materials to see whether they are conductors by inserting them in the gap.

Combining the concept of a complete circuit with the idea of the conservation of charge leads to the conclusion that electricity flows out one end of the battery and back into the other end. The charge that leaves one end of the battery returns to the other end. Charge does not get lost or used up along the way.

Q: How does the information about pathways explain why the bulb in Figure 21-4 doesn't light?

A: Because there is no pathway for the charges to flow back to the other end of the battery, there is no complete circuit.

A Water Model

The flow of charge in a complete circuit is analogous to the flow of water in a closed system of pipes. The battery is analogous to a pump, the wires to the pipes, and the bulb (something that transforms energy in the flow to another form) to a paddle wheel that turns when the water is flowing (Figure 21-8). In a closed system of pipes, the water that leaves the pump returns to the intake side of the pump.

The water model is useful for clarifying the difference between current, charge, and voltage. The current is a measure of the amount of charge flowing past a given section of the circuit in a unit time. If we measure a charge ΔQ passing the point in an interval of time Δt, the current I in the circuit is

current = $\frac{\text{charge}}{\text{time}}$ ▶

$$I = \frac{\Delta Q}{\Delta t}$$

This expression is analogous to the rate at which water flows through a pipe. The flow of water is measured in units such as liters per second. Electric current is measured in **coulombs** per second, a unit known as the **ampere** (A). Flashlight batteries usually provide less than 1 ampere of current, typical household circuits are usually limited to a maximum of 20 amperes, and a car battery provides more than 100 amperes while starting the car. (The precise definition of the ampere will be described in the next chapter.)

The voltage between two points in a circuit is a measure of the change in electric potential between these two points. That is, voltage is the work done in moving one unit of electric charge between the two points. The work is equal to the distance the charge moves along the circuit multiplied by the force on the unit electric charge due to the electric field that exists in the circuit. You can also think of voltage as being analogous to pressure in our water model.

Although the water model is useful, it is important to realize its limitations. The electric circuit is always "filled" with charges; it does not need to be filled as you would a garden hose. A break in an electric circuit causes the electricity to stop flowing; it doesn't spill out the end of the wire as water would from a broken pipe. An electric circuit is not analogous to a sprinkler system used to water lawns. A valve blocks the flow of water, whereas a switch puts a gap in the conducting path to stop the flow of electricity.

Another difference is that one or both types of charge may be free to move in a conductor. In plasmas and ionized liquids and gases, both charges can move. Not until 1879 did experiments determine that only some of the negatively charged electrons are mobile in metals. For most macroscopic effects, positive charges moving in one direction are equivalent to negative charges moving in the opposite direction. For example, consider two neutral metal spheres. If you move 1 coulomb of positive charge from sphere A to sphere B, sphere

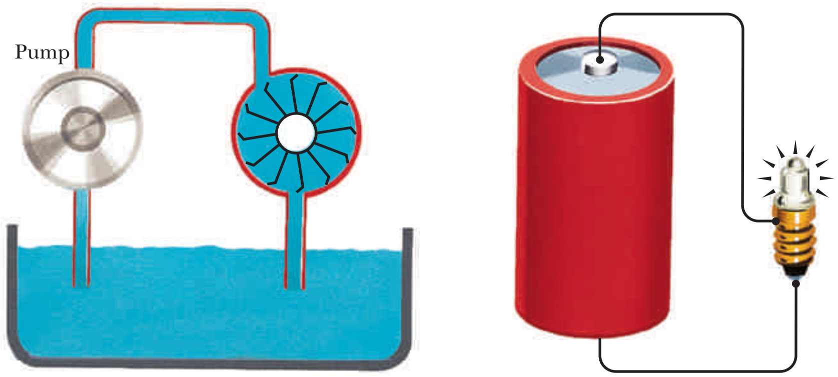


Figure 21-8 The flow of charge is analogous to the flow of water through a closed system.

B has an excess of 1 coulomb of positive charge. Sphere A has a deficiency of 1 coulomb of positive charge, or the equivalent, an excess of 1 coulomb of negative charge. The net charges on the two spheres would be identical if 1 coulomb of negative charge were moved from B to A. Because of this equivalence, as well as the historic origins of the subject, when discussing the macroscopic effect of current, we adopt the convention of assuming that the direction of the electric current is the same as the flow of the positive charges.

Resistance

Extended presentation available in the *Problem Solving* supplement

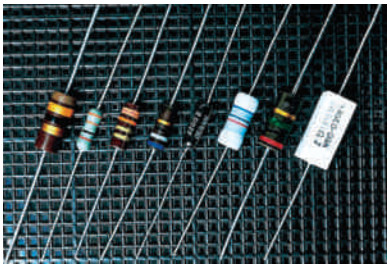
If a bulb is left connected to a battery, the battery runs down in less than a day. (The actual time depends on the particular bulb and battery used.) The bulb has the same brightness for most of this time, but near the end it dims and goes out. However, if you connect just a wire across the ends of the battery, the battery runs down in less than 1 hour, and the wire often gets too hot to touch. Because the battery is drained much faster and something is obviously happening in the wire, we infer that the current is larger through the wire than through the pathway with the bulb. The bulb offers more **resistance** to the flow of electricity. It is important to notice that the amount of current delivered by a battery depends on what is connected to the battery. Increasing the resistance in the circuit decreases the current through the battery, and decreasing the resistance in the circuit increases the current through the battery.

This notion of resistance also makes sense according to the water model. More water flows through wide pipes each minute than through narrow pipes. The bulb's filament is a thin section of wire and should offer more resistance. Our analogy also tells us that long pipes offer more resistance than short pipes. Therefore, we would expect the resistance of wires to increase with length and decrease with diameter. The other factor—not obvious from our water model—is that the resistance depends on the type of material used for the wires. These ideas are verified by experiments with wires of different sizes and materials. Table 21-1 compares the resistances of wires with the same length and diameter but made of different metals.

Table 21-1 *The Resistance of Various 100-Meter-Long Wires with a Diameter of 1 Millimeter*

Metal	Resistance (Ω)
Silver	2.02
Copper	2.16
Gold	3.11
Aluminum	3.59
Tungsten	7.13
Iron	12.7
Nichrome	191

Resistance is a result of the interaction of the pathway with the flow of charge. The electrons in a wire feel a force due to the repulsion of the negative terminal and the attraction of the positive terminal and are accelerated; the electrons experience forces due to the electric field that exists in the wire. However, the electrons don't go very far before they bump into atoms, which causes them to lose speed and to be deflected in random directions. Although the average speed of the electrons due to their thermal motion is quite high, all of these collisions keep them from moving very fast along the wire; a typical average speed along the wire is on the order of millimeters per second. This impedance to the flow of charge determines the resistance of the wire.



Resistors are used in electronic circuits.

As the electrons move through a wire, they undergo collisions that transfer kinetic energy to the atoms and cause the wire's temperature to increase. If enough energy is transferred, the wire gets hot enough to glow. Heating coils and elements in stoves, ovens, toasters, baseboard heaters, and lights do this. Resistors are used to control the voltages and currents in circuits used in such devices as radios and curling irons.

Resistance R is defined to be the voltage V across an object divided by the current I through the object:

$$R = \frac{V}{I}$$

$$\text{resistance} = \frac{\text{voltage}}{\text{current}}$$

Resistance is the number of volts across an object required to drive 1 ampere of current through the object, and is therefore measured in volts per ampere, a unit known as the **ohm** (Ω, uppercase Greek omega). This definition is always valid but is most useful when the resistance is constant or relatively constant. In this case this relationship is known as **Ohm's law**. The resistances of pieces of metal, carbon, and some other substances are approximately constant if they are maintained at a constant temperature. The resistance of the filament in a lightbulb increases as the filament heats up.

WORKING IT OUT *Ohm's Law*

Ohm's law provides us with a relationship between the resistance of, the potential difference (voltage) across, and the current through a circuit element. For example, suppose we have a 12-V battery and want to produce a current of 2 A in a particular circuit. What resistance should the circuit have?

$$R = \frac{V}{I} = \frac{12\ \mathrm{V}}{2\ \mathrm{A}} = 6\ \Omega$$

Are You On the Bus?

Q: If a lightbulb draws a current of 0.5 A when connected to a 120-V circuit, what is the resistance of its filament?

A: $R = V/R =$ (120 V)/(0.5 A) = 240 Ω.

Ohm's law can be rearranged to find any of the three quantities when the other two are given. Suppose a heating element has a resistance of 9.6 Ω when hot. What current will it draw when connected to 120 V?

$$I = \frac{V}{R} = \frac{120\ \mathrm{V}}{9.6\ \Omega} = 12.5\ \mathrm{A}$$

Are You On the Bus?

Q: If a 3-V flashlight bulb has a resistance of 9 Ω, how much current will the bulb draw?

A: $I = V/R =$ (3 V)/(9 Ω) = $\frac{1}{3}$ A.

The Danger of Electricity

Electricity is dangerous because our bodies are electric machines. Our muscles constrict when neurons "fire." This reaction is normally triggered by a complex chain of chemical reactions but can also occur when electric charges flow through the muscle. This process is the origin of twitches! This situation can become dangerous when the muscle is part of the heart or breathing system. In these cases the current can also affect the pacing signals, interrupting the normal processes. Currents as small as 50 milliamperes can interrupt breathing in some individuals.

Often people believe that the danger with electricity is the voltage. Although high voltages can be dangerous, it is the current through a body that is lethal. Ohm's law tells us that the current is equal to the voltage divided by the resistance of the path. In this case the body is part of the path. If the conditions are such that the total resistance is low, even a low voltage can cause a dangerously high current. The resistance of dry skin is high enough that you can place a finger across the terminals of a 9-volt battery without feeling anything. However, if you touch the terminals of the 9-volt battery to your tongue, you will feel it. The moisture of the tongue lowers the resistance.

One way of preventing high currents from passing through your body, especially through your heart, is to keep one hand in your pocket. Wearing rubber-soled shoes is also good.

A Model for Electric Current

We can use flashlight bulbs to develop a simple model of more complicated electric circuits. Each bulb serves as a visual indicator of the current through the bulb. Bulbs don't glow at all until the current exceeds a certain value; after that the brightness increases with increased current. Although the relationship between the current and the bulb's brightness is complicated, it is reasonable to assume that if one bulb is glowing more brightly than an identical bulb, it must have more current. We say that "more flow means more glow."

We begin by creating a "standard" to which we can refer; the brightness of a single bulb connected to a single battery will represent a standard current. We will also assume that all batteries and bulbs are identical.

Two bulbs can be connected to a battery so that there is a single path from the battery through one bulb, through the second bulb, and back to the other end of the battery. In this arrangement (Figure 21-9), the two bulbs are said to be in series with each other. We observe that the two bulbs have the same brightness and that they are dimmer than the single bulb in the standard circuit. This decrease in brightness indicates that there is less current in the series circuit than in our standard circuit. This conclusion is supported by the observation that the battery's lifetime in this series circuit is longer than the lifetime of the battery in the standard circuit. Therefore, the resistance of two bulbs in series is greater than that of a single bulb. (It is tempting to say that there is twice as much resistance in the series circuit because there are two bulbs. However, the resistance of the lightbulbs changes with brightness; consequently, the resistance does not double.)

Of special note is the observation that the two bulbs have the same brightness. This tells us that the current through each bulb is the same; whatever charge flows through the first bulb also flows through the second bulb. Electric charge is not being used up or lost along the pathway in agreement with the conservation of charge. The same amount of charge flows back into the battery as left it. It is the flow of charge through a bulb that causes the filament to heat up and glow. If the same amount of charge flows through each bulb each second, the two identical filaments will reach the same temperature and glow with the same brightness.

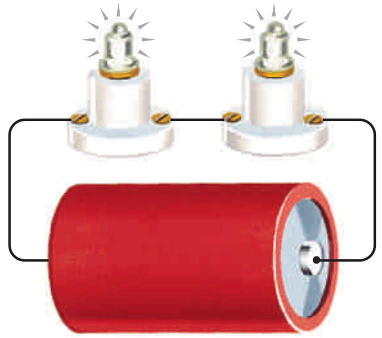


Figure 21-9 Two bulbs wired in series are equally bright but dimmer than the standard.

FLAWED REASONING

Alan connects a single bulb to a battery and notices that it is fairly bright. He then adds a second bulb in series with the first and finds that both bulbs are now dimmer and equally so. He makes the following conclusion: "The single bulb got all the current from the battery. The two bulbs must now *share* the current, each getting half. It makes sense that they should be dimmer."

Alan thinks he understands what is going on, but he has made a serious error in his reasoning. **Can you find it?**

ANSWER The word *share* has two meanings. We can share a book or we can share a pizza. If we share a book, you read it first, and then I read it. If we share a pizza, we first divide the pizza in two pieces and each take a piece. Alan believes that the current through the battery is always the same size and that the two bulbs are "sharing" this current as we would share a pizza. In fact, the bulbs share the current as we would share a book; first the current passes through one bulb (heating its filament), and then through the other (heating its filament). The two bulbs are dimmer than the single bulb because the increased resistance in the circuit reduces the current through the battery.

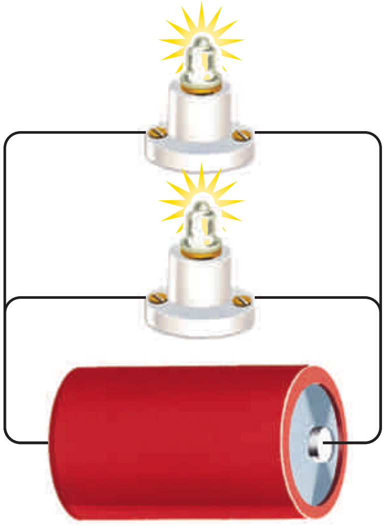


Figure 21-10 Two bulbs wired in parallel are equally bright and have the same brightness as the standard.

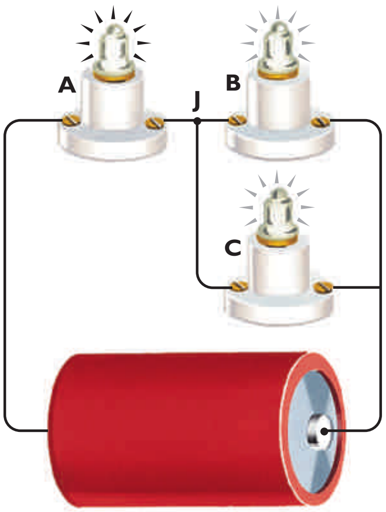


Figure 21-11 Bulbs B and C are in parallel, and the pair is in series with bulb A.

Another consequence of the equal brightness is that bulbs cannot be used to determine a direction for the flow of electricity. Everything can be explained equally well assuming a flow of negative charges in one direction (the actual situation in wires), a flow of positive charges in the other direction, or both charges flowing simultaneously in opposite directions (as happens in fluids). In fact, because only the motion of charges matters, one charge moving back and forth would also work. In household electricity the negative charges move back and forth with a frequency of 60 hertz.

Two bulbs can also be connected so that each bulb has its own path from one end of the battery to the other, as shown in Figure 21-10. In this arrangement the two bulbs are said to be wired in parallel. In contrast to the bulbs in series, the current in one bulb does not pass through the other, which can be seen by disconnecting either bulb and observing that the other is not affected. The two bulbs in parallel are equally bright, and each is as bright as our standard. Because each bulb has its own path, the battery in this circuit supplies twice as much current as the battery in the standard circuit. This can be verified experimentally; in this arrangement the battery runs down in one-half its normal lifetime.

Adding the extra bulb in parallel has increased the current through the battery, indicating that the resistance of the circuit must have decreased. Although a bulb can be correctly thought of as an obstacle to the flow of charge, the addition of one more bulb in a circuit can either increase or decrease the total resistance of the circuit, depending on how we add the bulb. When the new bulb is added in series (adding a new resistance on an existing line), the resistance of the circuit increases and the current through the battery decreases. When the new bulb is added in parallel (on a new path that did not exist before), the resistance of the circuit goes down and the current through the battery increases. Even though the new path contains resistance (the new bulb), it represents a new opportunity for flow that did not exist before.

Consider the following analogy. When a popular movie ends, people jostle to get out of the theater through the front door. The finite width of the door represents resistance to the flow of people. If the back door is also opened, the flow of people out of the building increases, even if the back door is narrow. It is a new opportunity for flow. Any path that is added in parallel lowers the total resistance of a circuit, regardless of how much resistance is contained on the new path.

Three or more bulbs can be connected in series or parallel or in combinations of series and parallel. The relative brightnesses of these bulbs can be predicted using the ideas that we have discussed. As an example, consider the combination of bulbs in Figure 21-11. Bulb A will be the brightest because all of the current must pass through it, but it will not be as bright as our standard. Bulbs B and C will be dimmer because the current splits—part going through bulb B and part through bulb C. The charge flowing into the junction J must be equal to that flowing out of the junction. This is just a consequence of the conservation of charge. Because the bulbs are identical, the resistances of the two paths are equal, and the current will split equally; bulbs B and C will be equally bright.

In general, when charge reaches the junction between two parallel branches of unequal resistance, more charge will flow through the easier branch. We say that "current favors the path of least resistance." This does not mean that all of the current takes the easier path; some current takes the more difficult path. If parallel path 1 has twice the resistance of path 2, path 1 will have one-half the current of path 2.

The German physicist Gustav Kirchhoff formalized two rules for analyzing the current in circuits. **Kirchhoff's junction rule** states:

Kirchhoff's junction rule ▶

The sum of the currents entering any junction in a circuit must equal the sum of the currents leaving that junction.

As we have seen, this is a consequence of the conservation of charge. We will discuss Kirchhoff's other rule in the next section.

If one of the paths in a circuit is a conducting wire without a lightbulb as shown in Figure 21-12, the path has very little resistance and virtually all of the current takes this path. This is known as a **short circuit**. When there is a short circuit, so little current flows through any bulbs in parallel to the short that the bulbs go out.

In general, parallel paths are not independent of each other. Making a change in one parallel path usually affects the current in the other path. The only exception to this rule is the case we started with in Figure 21-10, in which each path had direct connections to both sides of an ideal battery. These paths are said to be "parallel across the battery," and each path acts as if it has its own battery (for reasons that we will discuss later).

Figure 21-12 When the wire on the left is connected from one end of the battery to the other, producing a short circuit, the light goes out.

FLAWED REASONING

Three students are considering what will happen to the brightness of bulb B in Figure 21-11 when bulb C is removed from its socket.

Terry: "The current splits after passing through A. When C is removed, all the current must pass through B, so B will get brighter."

Judy: "Removing bulb C removes a pathway from the circuit, which will increase the resistance of the circuit. The current through the battery will go down, and B will get dimmer."

Kay: "You're both half right. The current through A will go down, but B will now get all the current instead of sharing it with C. It's like choosing between some of a large pizza or all of a small one. We need more information to make a decision."

Which student (if any) do you agree with?

ANSWER Kay correctly recognizes the conundrum. We know that the current through the battery (and hence through bulb A) must decrease when bulb C is removed, but we do not know by how much it will decrease. The model that we have developed so far does not allow us to make a prediction; we will need to add something to our model.

A Model for Voltage

If you leave your house and climb a nearby mountain, you gain a certain amount of gravitational potential energy. As you return to your house, you lose this same amount of gravitational potential energy, regardless of which path you choose to descend the mountain. It is the same with electric circuits. A 12-volt battery delivers 12 joules of electric potential energy to every coulomb of charge that passes through it. As that coulomb of charge travels through the circuit and returns to the battery, it must lose 12 joules of electric potential energy, regardless of which path it takes. This electric potential energy is delivered to the resistive elements in the circuit (such as bulbs) and converted to heat and light. This basic application of the conservation of energy is called **Kirchhoff's loop rule:**

Along any path from the positive terminal to the negative terminal of a battery, the voltage drops across the resistive elements encountered must add up to the battery voltage.

◀ Kirchhoff's loop rule

This means that a single bulb connected to a battery will have a voltage drop equal to the voltage of the battery. We also know from experience that a bulb will glow more brightly when connected to a 12-volt battery than when it is connected to a 6-volt battery. This suggests that the bulb's brightness can be used to indicate the voltage drop across the bulb. If one bulb is brighter than another, it must have a larger voltage drop across it (assuming again that both bulbs are identical). This important idea, coupled with Kirchhoff's loop rule, allows us to answer many questions that are left unanswered by our present model for electric current.

Let's return to the conundrum presented in the last Flawed Reasoning feature. Our model for electric current allows us to predict that bulb A in Figure 21-11 will get dimmer when bulb C is removed from its socket. As Judy explains, removing bulb C removes a pathway from the circuit, increasing the resistance of the circuit. The current through the battery must decrease, and because all of the current passes through bulb A, bulb A must get dimmer. The current model, however, cannot predict what happens to bulb B. The total current in the circuit has decreased, but bulb B's share of that current has increased (from 50% to 100%). Which effect wins out? Our model for electric voltage answers the question without ambiguity. The sum of the voltage across bulb A and the voltage across bulb B must be equal to the battery voltage by Kirchhoff's loop rule. If bulb A gets dimmer when bulb C is removed, the voltage across bulb A must decrease. Because the voltage drop across bulbs A and B must be equal to the battery voltage, the voltage across bulb B must increase by the same amount and bulb B will therefore get brighter.

We are now in a position to understand the independence of branches that are parallel across an ideal battery—why each parallel branch acts as if it has its own battery. The voltage drops across each branch must equal the battery voltage by Kirchhoff's loop rule, regardless of what changes are made to the other branches. An ideal 6-volt battery is one that can maintain a potential difference of 6 volts across its terminals regardless of how much current is drawn by the circuit. In practice, real batteries have trouble delivering large currents because of limitations of the chemical processes inside the battery. This limitation manifests itself in the form of a drop in the battery voltage as the current increases. The addition of a second bulb in parallel to our standard bulb will actually dim the first bulb slightly if you are using an alkaline battery. The effect will not be noticeable with a rechargeable nickel-cadmium (NiCad) battery. The NiCad battery is said to be more ideal, and many branches can be connected across a NiCad battery before the effect appears. This assumes that each of those branches contains resistance and draws a small amount of current. If a parallel branch contains nothing but a conducting wire, it demands more current than the battery can supply and the voltage of the battery drops dramatically, causing the bulbs on the other branches to go out. This is called "shorting out the battery" and is similar to what happens on a hot summer day when Los Angeles has a "brownout" caused by the huge demand for current to run air conditioners.

The circuits in houses are wired in parallel so that electric devices can be turned on and off without affecting one another. As each new appliance is turned on, the electric company supplies more current. Each parallel circuit, however, is wired in series with a circuit breaker, as shown in Figure 21-13, to deliberately put a "weak link" in the circuit. If too many devices are plugged into one circuit, they will draw more current than the wires can safely carry. For example, the wires may heat up at a weak spot and start a fire. The circuit breaker interrupts the circuit, shutting everything off.

Some modern Christmas tree lights are wired in parallel. If one bulb burns out, it is the only bulb that goes out. The older style was wired in series; when one bulb went out, they all went out. That was rather inconvenient, because when one bulb burned out, all of them had to be tested until the defective one was found. The latest style of Christmas tree lights has the bulbs wired in series, but when a bulb burns out, a short circuit is formed across the burned-out bulb that allows the rest of the bulbs to stay lit.

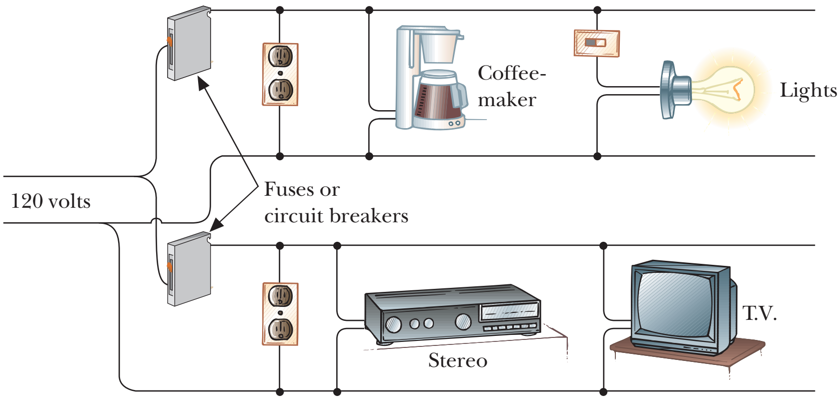
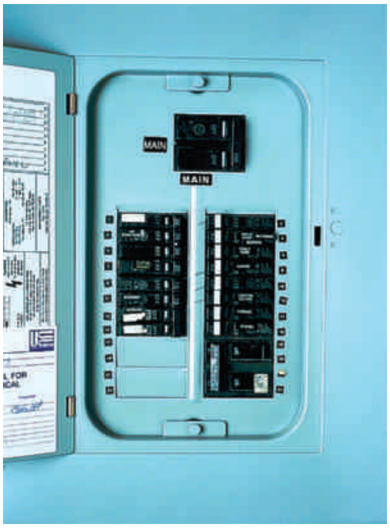


Figure 21-13 Appliances are connected to household circuits in parallel. The circuit is protected from drawing too much current by a circuit breaker wired in series.



A circuit breaker protects household circuits from drawing too much current and starting a fire.

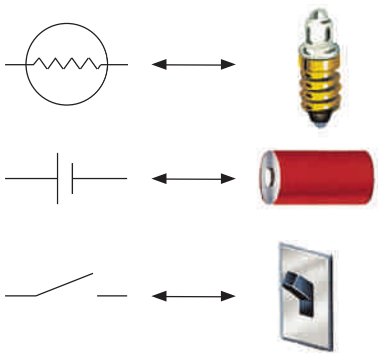


Figure 21-14 Common symbols used in diagrams of electric circuits. (*Top to bottom:* resistor, battery, switch.)

Over the years, a convention for drawing circuit elements has been developed. Like all symbols, the electrical symbols capture the functional essence of the element and omit nonessential characteristics. A lightbulb, for example, has no directional characteristic, but a battery does. Their symbols reflect this difference. Figure 21-14 gives the common symbols paired with a diagram like those we have been using.

Circuits are also drawn with sharper corners than exist in the actual circuits. This is simply a technique that has helped communication among experimenters.

Electric Power

When you buy electricity, it is energy that you are paying for. This electric energy is converted to heat, light, or motion. Why is it, then, that we talk about power? As we discussed in Chapter 7, **power** is the amount of energy used per unit time. But we need to be careful here; conservation of energy tells us that energy can be neither created nor destroyed. Formally, power is the amount of energy transformed from one form to another form divided by the elapsed time. In electricity, the energy the electrons gain from the electric field is converted to such forms as kinetic energy, heat, sound, and light. Most electric devices in our houses are rated by their power usage. Power is measured in **watts** (W). Many household lightbulbs are rated at 60, 75, or 100 watts. Electric heaters and hair dryers may use 1500 watts.

The electric meter connected between your house and the energy or power company records the energy you use, much as the odometer in your car records the miles you drive. Electric energy is usually billed at 8 to 25 cents per kilowatt-hour. A typical household with an electric range and an electric clothes dryer (but no electric heat) uses about 900 kilowatt-hours per month.

There are times when the energy losses in connecting wires become important. Sending electric energy long distances through wires from a power plant can result in significant energy losses. Two things can be done to minimize these losses. First, the wires should have as little resistance as possible, which means using large-diameter wires and low-resistance materials. Second, transformers (discussed in Chapter 22) allow the utility company to send the same energy through the wire by raising the voltage and simultaneously lowering the current. Lower current means less thermal loss in the transmission lines. Economic and engineering considerations dictate the extent to which each of these methods is used to minimize energy loss. Electricity is carried between energy companies and towns at very high voltages (up to 765,000 volts). The

This electric meter measures the energy used in a house or business. This meter reads 95,109 kilowatt-hours.

WORKING IT OUT *Electric Power*

We can obtain an expression for electric power by using the definition for electric energy we developed in Chapter 20 and the definition for current. Because energy is the charge multiplied by the voltage, we have

$$P = \frac{\Delta E}{\Delta t} = \frac{\Delta QV}{\Delta t} = \frac{\Delta Q}{\Delta t}V$$

This equation tells us that the power is equal to the current times the voltage:

power = current × voltage ▶

$$P = IV$$

Therefore, an electric appliance that draws a current of 10 A at a voltage of 120 V uses energy at a rate of

$$P = IV = (10\ \text{A})(120\ \text{V}) = 1200\ \text{W}$$

Q: **What power is required to operate a clock radio if it draws 0.05 A from the household circuit?**

A: $P = IV =$ **(0.05 A)(120 V) = 6 W.**

By rearranging the expression for power, you can determine the current through a 100-W bulb:

$$I = \frac{P}{V} = \frac{100\ \text{W}}{120\ \text{V}} = 0.83\ \text{A}$$

We can obtain an alternative expression for electric power by using the relationship that the voltage is equal to the current multiplied by the resistance:

power = current squared × resistance ▶

$$P = IV = I(IR) = I^2R$$

In this form it is more obvious why a bulb in a circuit glows, yet the connecting wires do not. The current through the circuit is the same everywhere, but the section that has the highest resistance receives the most thermal energy per unit time. Because the filament is made of a fine wire of high-resistance material, it has the highest resistance; its temperature goes up, and it glows. The connecting wires also heat up, but they just don't get hot enough to glow or even feel warm.

voltage is lowered by transformers for distribution around towns and then lowered once more to the safer 120 volts used in homes.

Summary

Volta used the electric potential difference between two different metals to build the first battery. Batteries make charges flow continuously as a current, producing heat and light and running motors and other appliances.

The voltage produced by an individual cell depends on the materials used but not on its size. The size determines the amount of chemicals contained in the cell and therefore the total amount of charge that can be transferred. The

WORKING IT OUT *Cost of Electric Energy*

A device running at constant power uses an amount of energy equal to the power multiplied by the time; the longer we use it, the more energy it consumes. For example, a 60-W bulb burning for 10 h uses

$$\Delta E = P\Delta t = (60\ \text{W})(10\ \text{h}) = 600\ \text{Wh} = 0.6\ \text{kWh}$$

of electric energy. It is important to recognize that a kilowatt-hour (kWh) is a unit of energy, not power. Because 1 kW is equal to 1000 J/s and 1 h is the same as 3600 s, 1 kWh is 3.6 million J.

We can now calculate the monthly cost of electric energy for this typical household using 900 kWh per month. Let's assume that electric energy costs 10¢ per kWh:

$$\text{cost} = (900\ \text{kWh})(\$0.10/\text{kWh}) = \$90$$

Are You On the Bus?

Q: At this price, what would it cost to run a 1500-W heater continuously during an 8-h night?

A: $E = Pt =$ **(1500 W)(8 h) = 12,000 Wh = 12 kWh. Cost = (12 kWh)(10¢/kWh) = $1.20.**

voltages of cells in series are additive. Cells placed in parallel do not increase the voltage but increase the effective size of the battery.

The electricity in homes is usually supplied at 120 volts and as 60-hertz alternating current. The electricity from a battery is direct current and usually at a much lower voltage.

Electric charges can flow continuously only when there is a complete pathway, or circuit, made of conducting materials. The current is determined by the voltage and the total resistance of the circuit, $V = IR$. The voltage between two points in a circuit is equal to the change in electric potential between these two points, which is the work done in moving a unit of positive charge through the circuit. The most charge flows through the path with the least resistance. Conservation of charge requires that electricity flow out one end of the battery and back into the other. Charge does not get lost or used up along the way.

Conservation of energy requires that any electric potential energy gained by a charge as it passes through the battery must be lost by the charge as it passes through the circuit. The voltage drops across the resistive elements along any path between the terminals of a battery must sum to the voltage of the battery.

One or both types of charge may be free to move in a conductor. In plasmas and ionized liquids and gases, both charges can move, whereas only some of the negatively charged electrons are mobile in metals. For most macroscopic effects, positive charges moving in one direction are equivalent to negative charges moving in the opposite direction.

The resistance of wires increases with increasing length, decreases with increasing cross-sectional area, and depends on the type of material. Resistance R is the ratio of the voltage V across an object to the current I through the object, $R = V/I$. Resistance is measured in volts per ampere, or ohms (Ω).

A power company sells energy, not power. Power is the rate of using energy and is used to rate most electrical devices in our houses. Power is measured in watts (W) or kilowatts (kW), whereas household electric energy usage is measured in kilowatt-hours (kWh).

Everyday Physics | *The Real Cost of Electricity*

The real cost of electricity varies a great deal, depending on the source. We are most familiar with purchasing electric energy from the local energy company for use in our homes. But what is the cost of this electric energy? Examining your electric bill shows that you are charged something between 8¢ and 25¢ per kilowatt-hour. Let's use a representative value of 10¢ per kilowatt-hour. Knowing that 1 kilowatt-hour is equal to 3.6 million joules, we can divide these two numbers to learn that we can use about 36 million joules for each dollar we spend.

We can also buy electric energy packaged in a variety of batteries. In this case we pay for properties of the batteries that appeal to the manufacturer of the electric appliances. This may include the batteries' size, voltage, lifetime, or their ability to maintain current as they wear out. To discover what we are paying for the electric energy, we need to know the prices and the total electric energy that the batteries provide during their lifetimes. As shown in the text, the energy can be obtained by multiplying the total charge by the voltage rating of the battery. The total charge is obtained by multiplying the current rating by the battery's lifetime.

A standard flashlight battery—usually a carbon–zinc D cell—costs about 75¢ and has a rating of 1.5 volts. Its lifetime depends on the way it is used. Flashlight D cells are normally used for short intervals over a much longer time. Under these conditions the battery could supply 375 milliamperes for about 400 minutes, yielding a total charge of 9000 coulombs and a total energy of 13,500 joules. Therefore, you are getting 18,000 joules per dollar, 2000 times as expensive as household electricity. Alkaline batteries were a major improvement in performance over their carbon–zinc cousins, primarily in applications requiring fairly large, continuous currents, such as portable stereo radios. Alkaline batteries typically last about 50–75% longer than a standard battery, but under some conditions can last much longer. However, their costs are also higher. Are alkaline batteries a good buy? Only if their increase in cost is less than their corresponding increase in lifetime.

The development of mercury, silver, and lithium batteries allowed manufacturers to pack more energy in smaller volumes. These devices have a much more stable current over their lifetimes, so they are especially useful in devices such as wristwatches and photocells in cameras, which require a particular value for the current. A watch battery delivers 1.3 volts and 10 millionths of an ampere for a year, giving a total energy of 410 joules. If these batteries sell for about $2, they have an energy–cost rating of 205 joules per dollar, or 175,000 times the cost of household electricity!

1. Your car battery is recharged by the alternator while the car is running. Is this electric energy cost-free?
2. A typical 6-volt lantern battery can supply a total charge of 10 ampere-hours, or about 36,000 coulombs, during its lifetime. If this battery cost $8.00, how much more expensive is this electric energy than household electricity?

CHAPTER 21 *Revisited*

When electrons move through a conductor, they collide with the atoms in the conductor, raising the average kinetic energy of the atoms and therefore the temperature of the conductor. The rise in temperature depends on the thermal properties of the material and its electrical resistance. Filaments in lightbulbs are made of thin wires of high-resistance materials to enhance the heating. If the rise in temperature is great enough, the wire glows, emitting light.

Key Terms

ampere The SI unit of electric current, 1 coulomb per second.

complete circuit A continuous conducting path from one end of a battery (or other source of electric potential) to the other end of the battery.

coulomb The SI unit of electric charge. The amount of charge passing a given point in a conductor carrying a current of 1 ampere.

current A flow of electric charge. Current is measured in amperes.

Kirchhoff's junction rule The sum of the currents entering any junction in a circuit must equal the sum of the currents leaving that junction.

Kirchhoff's loop rule Along any path from the positive terminal to the negative terminal of a battery, the voltage drops across the resistive elements encountered must add up to the battery voltage.

ohm The SI unit of electric resistance. A current of 1 ampere will flow through a resistance of 1 ohm under 1 volt of potential difference.

Ohm's law The resistance of an object is equal to the voltage across it divided by the current through it.

parallel Two circuit elements are wired in parallel when the current can flow through one or the other but not both. Elements that are wired parallel to each other are directly connected to each other at both terminals.

power The rate at which energy is converted from one form to another. In electric circuits the power is equal to the current times the voltage. Power is measured in joules per second, or watts.

resistance The impedance to the flow of electric current. The resistance is equal to the voltage across the object divided by the current through it. Resistance is measured in volts per ampere, or ohms.

series An arrangement of resistances (or batteries) on a single pathway so that the current flows through each element.

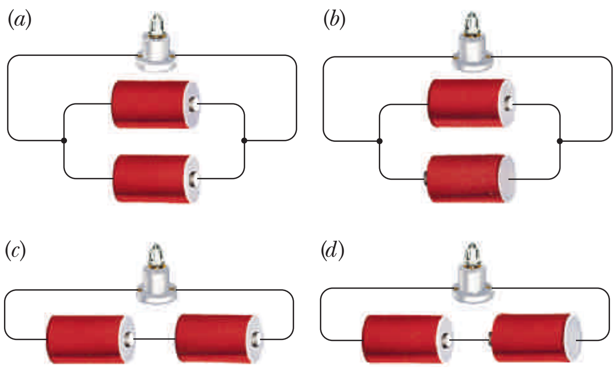
short circuit A path in an electric circuit that has very little resistance.

watt The SI unit of power, 1 joule per second.

Questions and exercises are paired so that most odd-numbered are followed by a similar even-numbered.
Blue-numbered questions and exercises are answered in Appendix B.
indicates more challenging questions and exercises.
WebAssign Many Conceptual Questions and Exercises for this chapter may be assigned online at WebAssign.

Conceptual Questions

1. Does the size of a cell affect its voltage, its lifetime, or the current it delivers?
2. When you stack three flashlight batteries in the same direction, you get a voltage of $3 \times 1\frac{1}{2}$ volts $= 4\frac{1}{2}$ volts. What voltage do you get if one of the batteries is turned to face in the opposite direction?
3. You have two 9-volt cells. What voltage would you get if you wired them in series?
4. If you have four 1.5-volt cells, what voltage would you get if you wired them in parallel?
5. What are the differences and similarities of the electricity supplied by a flashlight battery and a car battery?
6. What are the differences and similarities of the electricity supplied by a car battery and your local energy company?
7. Which bulbs in the following figure are the brightest? Which are the dimmest but still glow? Why?
8. Which bulbs in the following figure do not light? Explain.



Questions 7–10 Identical bulbs are connected to various combinations of identical batteries.

9. Which arrangements of batteries in the preceding figure will burn out most quickly? Which will last the longest? Why?
10. Which arrangements of batteries in the preceding figure will light the bulb for the longest time? Explain.

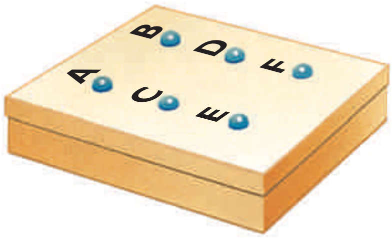
11. There are four different ways of connecting a battery, a bulb, and a wire so that the bulb lights. Two of these are shown in Figure 21-5. What are the other two?

12. Write a short statement describing how you would connect a battery, a bulb, and a single wire so that the bulb lights. Your statement should include all four ways without favoring any one way.

13. What is the difference, if any, in a circuit between a bulb burning out and removing the bulb from its socket?

14. Suppose that you ran speaker wires from the stereo in the family room to your bedroom, but you forgot to mark the wires. What could you do to tell which wires you should connect to which speaker?

15. You are given a mystery box with six bolt heads on top, as shown in the following figure. You connect a battery and a bulb across each pair of bolt heads and find that the bulb lights for the following pairs: AB, CF, CE, and EF, but not for any of the other combinations. Draw diagrams showing all of the ways the box could be wired.



16. A friend examines a box similar to the one in Question 15 and finds that the bulb lights for only the following combinations: AB, CD, and DE. Can your friend's observations be correct? Explain.

17. What is the difference between a volt and an ampere?

18. Car batteries are often rated in ampere-hours. What does this mean?

19. In the water model for electricity, what are the analogues of a battery, a switch, a wire, and a lightbulb?

20. In the water model for electricity, what are the analogues of charge, current, and voltage?

21. Which of the following affects the resistance of a wire: diameter, type of metal, length, or temperature? Explain.

22. Some power tools run poorly when connected to very long extension cords. Why? What could you do to improve the situation?

23. If the resistance connected to a battery is cut in half, what happens to the current through the battery?

24. A resistor has a current of 2 amperes when connected to a single battery. If a second identical battery is connected in series with the first, what is the current through the resistor? What is the current if the batteries are connected in parallel?

25. Why does a short circuit cause a battery to run down quickly?

26. Why is it dangerous to use a blow dryer while taking a bath?

27. If the only voltage you have is 120 volts, how could you light some 6-volt lightbulbs without burning them out?

28. In England the voltage supplied to homes is 240 volts. What would happen if you used an American lightbulb in England?

29. In a circuit with a single lightbulb, the current through the bulb is found to be 2 amperes. Is the current that flows through the battery greater than, equal to, or less than 2 amperes? Explain.

30. Two bulbs in series are connected to a battery. The current in the first bulb is found to be 1 ampere. What is the current through the battery?

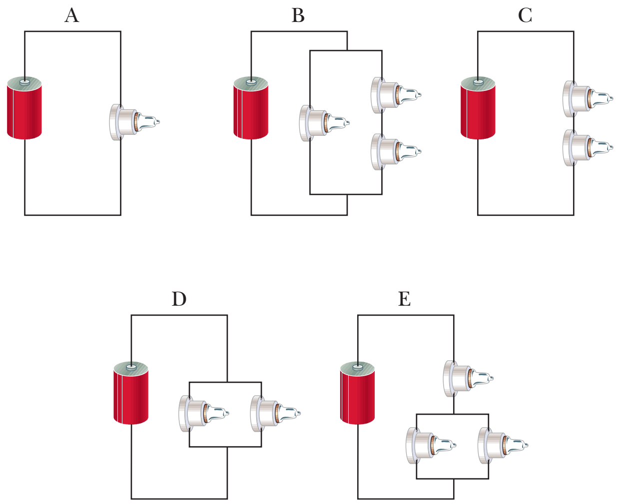
31. If a second bulb is added in series to a circuit with a single bulb, does the resistance of the circuit increase, decrease, or remain the same? Does the current through the battery increase, decrease, or remain the same? Why?

32. When a bulb is added in parallel to a circuit with a single bulb, does the resistance of the circuit increase, decrease, or remain the same? Does current increase, decrease, or remain the same? Why?

33. If a string of five bulbs in series is added in parallel to a single bulb connected to a battery, you find that the brightness of the original bulb does not change. Has the current through the battery increased, decreased, or stayed the same? Has the resistance of the circuit increased, decreased, or stayed the same? Explain.

34. Which circuit offers the greater resistance to the battery, two bulbs in series or two bulbs in parallel? Why?

35. For the circuits in the following figure, is the current through the battery in circuit A greater than, equal to, or less than the current through the battery in circuit E? Explain your reasoning.



Questions 35–38 Identical batteries are connected to various combinations of identical bulbs.

36. For the circuits in the preceding figure, is the current through the battery in circuit A greater than, equal to, or less than the current through the battery in circuit B? Explain your reasoning.

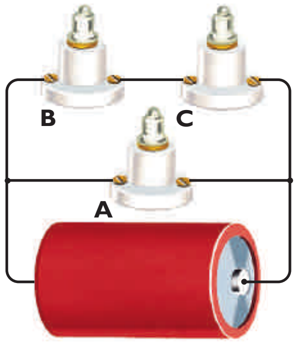
37. For the circuits in the preceding figure, is the current through the battery in circuit C greater than, equal to, or less than the current through the battery in circuit E? Explain your reasoning.

38. For the circuits in the preceding figure, is the current through the battery in circuit D greater than, equal to, or less than the current through the battery in circuit E? Explain your reasoning.

39. How would you connect two batteries and two bulbs to get the most light?

40. How would you wire two batteries and two bulbs to produce light for the longest time?

41. Which bulbs are the brightest and dimmest in the following figure? Explain.



Questions 41–46.

42. Would the resistance of the circuit in the preceding figure increase, decrease, or stay the same if bulb A were removed from its socket? What if bulb C were removed instead? Explain.

43. What happens to the brightness of bulbs B and C in the preceding figure when bulb A burns out? Explain.

44. What happens to the brightness of bulbs A and C in the preceding figure when bulb B is removed from its socket? Explain.

45. If you put a wire across the two terminals of bulb A, what happens to the brightness of each of the bulbs in the preceding figure? Why?

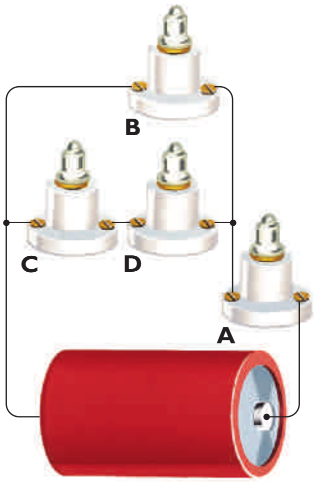
46. What happens to the brightness of each of the bulbs in the preceding figure if you connect a wire across the terminals of bulb B?

47. How does the brightness of bulb C compare to the brightness of bulb D in the following figure? Explain your reasoning.

48. How does the brightness of bulb D compare to the brightness of bulb B in the following figure? Explain your reasoning.

49. What happens to the brightness of bulbs B, C, and D in the following figure when bulb A burns out? Explain.

50. What happens to the brightness of bulbs A, B, and C in the following figure when bulb D is removed from its socket? Explain.



Questions 47–52.

51. If you put a wire across the two terminals of bulb A, what happens to the brightness of each of the bulbs in the preceding figure? Why?

52. What happens to the brightness of each of the bulbs in the preceding figure if you connect a wire across the terminals of bulb B?

53. A box has three identical bulbs mounted on its top with the wires hidden inside the box. Initially, bulb A is the brightest, and bulbs B and C are equally bright. If you unscrew A, bulbs B and C go out. If you unscrew B, A gets dimmer and C gets brighter so that A and C are equally bright. If you unscrew C, A gets dimmer and B gets brighter so that A and B are equally bright. If you unscrew B and C, A goes out. How are the bulbs wired?

54. A box has three identical bulbs mounted on its top with the wires hidden inside the box. Initially, bulb A is the brightest, and bulbs B and C are equally bright. If you unscrew A, B and C remain the same. If you unscrew B, A remains the same and C goes out. If you unscrew C, A remains the same and B goes out. If you unscrew B and C, A remains the same. How are the bulbs wired?

55. Are the headlights in cars and trucks wired in series or parallel? Explain.

56. Why should you not replace a 5-ampere fuse in your car with a 10-ampere fuse?

57. What happens to the power supplied by a battery if the resistance connected to the battery is doubled? Explain.

58. A single resistor dissipates 8 watts of power when connected directly across a battery. If a second identical resistor is added in series with the first, what power will the original resistor dissipate?

59. A resistor dissipates 2 watts of power when connected to a single battery. If a second identical battery were connected in parallel with the first, what power would the resistor dissipate?

60. A resistor dissipates 2 watts of power when connected to a single battery. If a second identical battery were connected in series with the first, what power would the resistor dissipate?

61. Two bulbs have ratings of 60 watts and 120 watts. Which bulb carries the higher current? Explain.

62. Two bulbs have ratings of 60 watts and 120 watts. Which bulb has the higher resistance? Explain.

Exercises

63. What is the resistance of a lightbulb that draws 0.5 A when it is plugged into a 120-V outlet?

64. You plug a vacuum cleaner into a 120-V outlet with nothing else on the circuit. If the 15-A circuit breaker is tripped, what is the maximum possible resistance of the vacuum's motor?

65. The coils of a heater have a resistance of 10 Ω when hot. What current does the heater draw when plugged into a 120-V outlet?

66. You plug your new 9-Ω hair dryer into a 120-V outlet and trip the 10-A circuit breaker. If you plug the hair dryer into a line with a 15-A breaker, will it trip the breaker? Assume that nothing else is plugged into either circuit.

67. If a lightbulb has a resistance of 8 Ω and a current of 0.5 A, at what voltage is it operating?

68. A lightbulb has a resistance of 250 Ω. What voltage is required for the bulb to draw a current of 0.5 A?

69. Two $1\frac{1}{2}$-V batteries are connected in series to a 3-Ω resistor. How much current flows through each battery?

70. Two $1\frac{1}{2}$-V batteries are connected in parallel to a 3-Ω resistor. How much current flows through each battery?

71. Two 3-Ω resistors are connected in series to a 12-V battery. What single resistor, if connected to the battery alone (called the *equivalent resistance*), would draw this same current? What is the current through the battery?

72. A 2-Ω and a 4-Ω resistor are connected in series to a 12-V battery. What is the current through each resistor? Use Ohm's law to show that the voltage drops across the individual resistors add up to 12 V.

73. Two 4-Ω resistors are connected in parallel to a 12-V battery. Use the fact that the voltage across each of the resistors is 12 V to find the total current through the battery. What single resistor, if connected to the battery alone (called the *equivalent resistance*), would draw this same current?

74. A 3-Ω resistor is connected in parallel with a 12-Ω resistor, and the combination is connected to a 12-V battery. How much current does the battery supply?

75. A heating element is rated at 1400 W. How much current does it draw when it is connected to a 120-V line?

76. What is the power used by a toaster that draws a current of 8 A when connected to a 120-V line?

77. A DVD player has a maximum power rating of 24 W when plugged into a 120-V outlet. What is the maximum current that it requires?

78. If a clock draws a maximum current of 5 mA from a 120-V line, what is its maximum power consumption?

79. A 4-Ω resistor is connected to a 12-V battery. What is the current through the battery? What is the power dissipated by the resistor?

80. A 3-Ω resistor draws a current of 2 A when hooked up to a battery of unknown voltage. What is the battery's voltage? What is the power dissipated by the resistor?

81. What is the power rating of a heating coil with a resistance of 12 Ω that draws a current of 20 A?

82. A coffeemaker has a resistance of 12 Ω and draws a current of 10 A. What power does it use?

83. What is the resistance of a 60-W lightbulb in a 120-V circuit?

84. What is the resistance of the coil in a 1500-W heater?

85. If a hair dryer is rated at 1200 W, how much energy is used to operate the hair dryer for 8 min?

86. If a 60-W bulb is left on continuously in a secluded hallway, how much energy is used each month?

87. If electricity costs 15¢/kWh, how much does it cost to burn a 100-W bulb for one day?

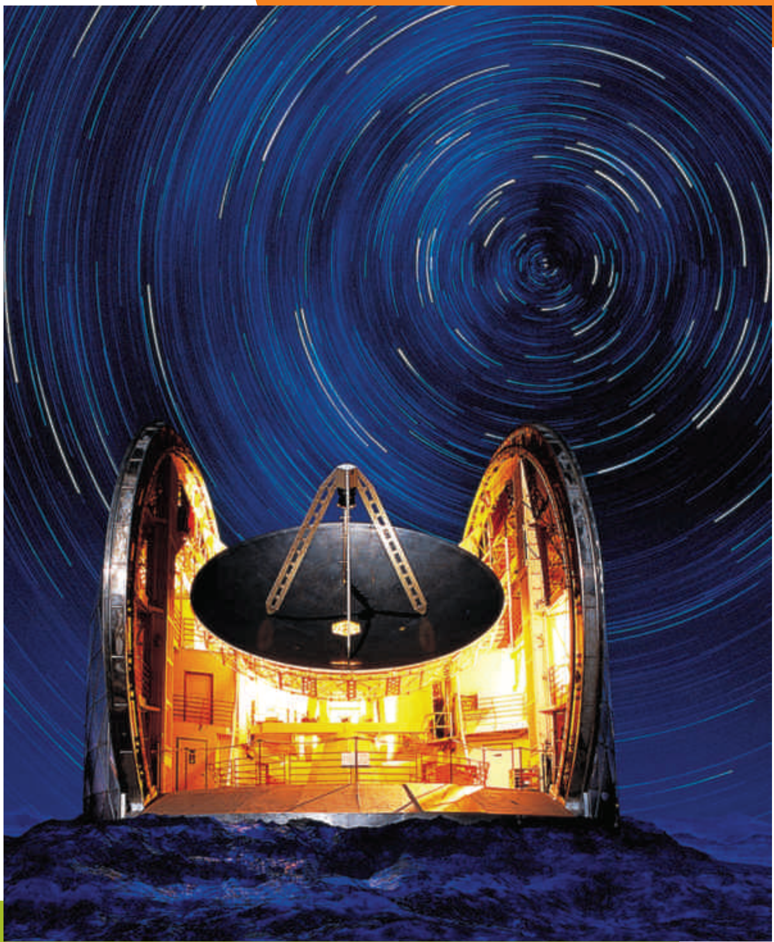
88. A 1500-W heater for a sauna requires 40 min to heat the sauna to 190°F. What does this cost if electricity sells for 12¢/kWh?

22

Electro-magnetism

▶ During the 19th century, scientists discovered that electricity and magnetism were not separate phenomena but two different aspects of something called electromagnetism. It is electromagnetism that enables us to pop corn in our microwave ovens and brings television to us while we enjoy our popcorn. What is this connection between electricity and magnetism?

(See page 488 for the answer to this question.)



A radio telescope scanning the sky over Hawaii. The stars appear to move in circles as Earth spins on its axis.

A collection of permanent magnets.

MOST of us have played with magnets. Through play we learn that these little pieces of material attract and repel each other and that they attract some objects but have no effect on others. For instance, magnets stick to refrigerator doors, but not to the walls of a room; to paper clips, but not to paper; and so on.

Although magnetic properties are interesting, they, like electricity, appear to be rather isolated phenomena. Unlike electricity, however, magnetism does seem more permanent. And although there are similarities between these two phenomena, they seem to be separate properties of matter: a magnet with no net electric charge does not deflect the foils of an electroscope. So magnetism appears to be a different phenomenon from electricity. Yet some experimenters in the 17th century were more fascinated by the similarities than the differences between magnetism and electricity, and they searched for a connection.

Magnets

It has been known since at least the sixth century BC that lodestone, a naturally occurring mineral, attracts iron. It was also known that iron can be made magnetic by rubbing it with lodestone. The magnetized piece of iron then attracts other pieces of iron. The only elements that occur naturally in the magnetized state are iron, nickel, and cobalt. Many permanent magnets are made from alloys of these metals.

Imagine having three unmarked, magnetized rods. If we arbitrarily mark one end of one of the rods with an X, we find that it attracts one end of each of the other two magnets (Figure 22-1), and it repels their other ends. Let's mark these ends A (for attract) and R (for repel), respectively. We then find that the two A ends repel each other, the two R ends repel each other, and the A and R ends attract each other. Because the two A ends were determined in the same way using the first magnet, they should be *alike*. Therefore, the ends of the magnets (called **magnetic poles**) behave like the electric charges we studied in Chapter 20. We can summarize the behavior of the magnetic poles with the following simple statement:

magnets ▶

Like poles repel; unlike poles attract.

Both electricity and magnetism seem to have two kinds of "charge." In electricity, however, it is possible to separate the two charges, or at least to put more of one kind of charge on one object than on the other. But in magnetism the poles always come in pairs that have the same strength.

Are You On the Bus?

Q: How might you search for a third kind of magnetic pole?

A: You could search for a magnetized material that attracts or repels both north poles *and* south poles. No such material has been discovered.

Suppose you try to separate the poles by breaking a magnet in half. You get two magnets, and each of these new magnets has two poles (Figure 22-2). If you break each of these in half, you get four magnets, and so on. Even if you continue dividing down to the atomic level, you always obtain magnets with two poles. It does not seem to be possible to isolate a single magnetic pole, a **magnetic monopole**. Many extensive searches for magnetic monopoles have been conducted. Although the existence of a magnetic monopole is not ruled out by the present, well-established theories, all searches have ended in failure and it does not seem likely that one will be discovered. The discovery of a magnetic monopole would most likely result in a Nobel Prize.

In naming the ends, or poles, of the magnets, we make use of the observation that a freely swinging magnet aligns itself along the north–south direction. The end that points north is called *north* and the other *south*. The compasses used for navigation are simply tiny magnets that are free to rotate. Magnets have been used as compasses since the 11th century, but not until 1600 did William Gilbert, an English physician and scientist, hypothesize that they work because Earth itself is a giant lodestone. Gilbert even made a spherically shaped piece of lodestone to show that a compass placed near it behaved as it did on Earth.

Magnets are surrounded by **magnetic fields** in the same way electric charges are surrounded by electric fields. These magnetic fields can be detected using a small compass. The direction of a magnetic field at a particular point is the direction in which the north pole of the compass points when placed at this point, and the torque that aligns the compass is a measure of the strength of the magnetic field. The photograph in Figure 22-3(a) was obtained by sprinkling iron filings on a piece of glass placed over a bar magnet. The iron filings line up along the direction of the magnetic field.

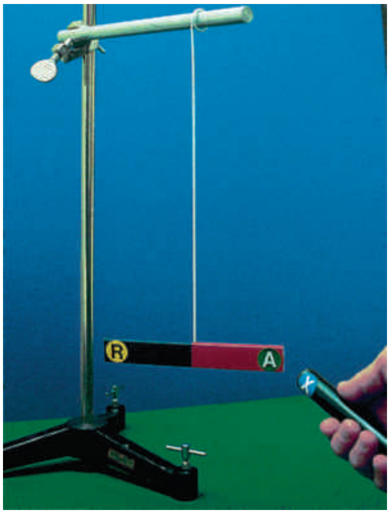
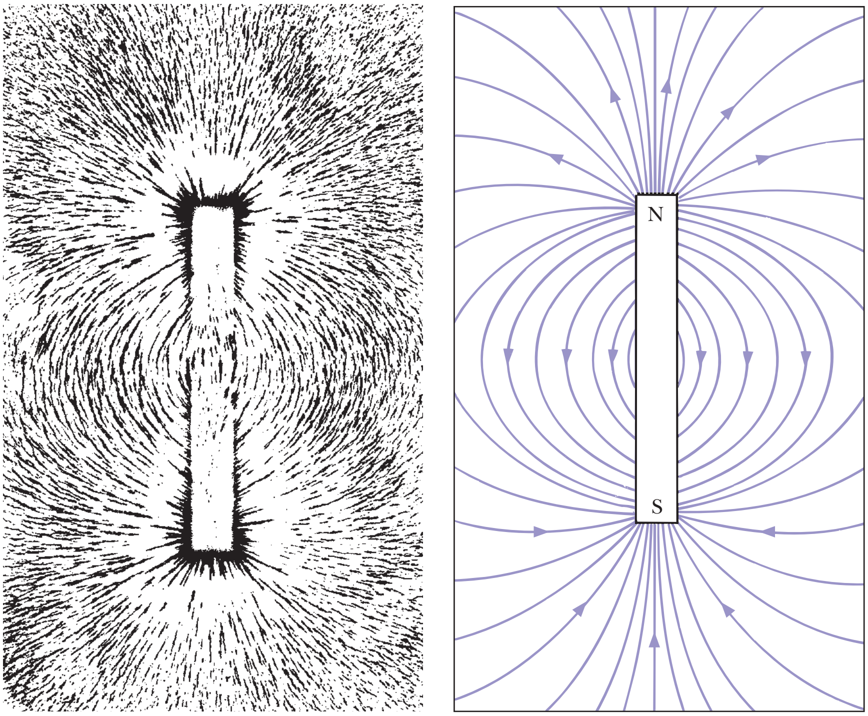


Figure 22-1 Magnetic poles can be identified by their attraction to and repulsion by known poles.

Are You On the Bus?

Q: How would the photograph in Figure 22-3 change if the poles of the magnet were reversed?

A: The photograph would not change because the iron filings do not tell us which way the field points.



(*a*) (*b*)

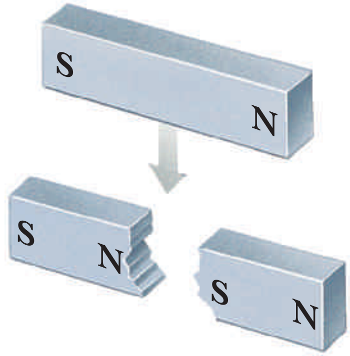
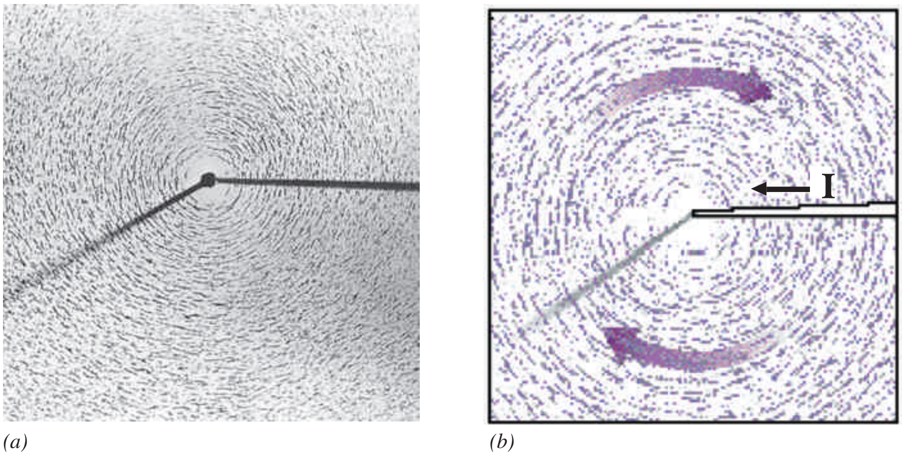


Figure 22-2 Dividing a magnet results in two smaller magnets, each with both poles.

Figure 22-3 (a) Iron filings line up along the magnetic field lines surrounding a bar magnet. (b) The drawing indicates the direction of the magnetic field outside the magnet.



PSSC PHYSICS, 2nd edition, 1965; D. C. Heath & Co. and Education Development Center, Inc., Newton, MA.

Figure 22-4 (a) Iron filings show that the magnetic field lines surrounding a long, straight, current-carrying wire are circles around the wire. The dark line at the left is the shadow of the wire. (b) The diagram shows the direction of the magnetic field for a current flowing away from you (into the page).

Electric Currents and Magnetism

Although electricity and magnetism were well known for centuries, any connection between the two phenomena eluded experimenters until the 19th century. In 1820 Hans Christian Oersted, a Danish scientist, discovered a connection while performing a lecture demonstration: a compass needle experiences a force when it is brought near a current-carrying wire. This means that the current-carrying wire produces a magnetic field in the surrounding space. This discovery was fascinating and yet confusing. It was known that stationary charges did not produce magnetic fields, so the motion of the charges must produce the field. The photograph in Figure 22-4(a) shows that the magnetic field lines form circles about the wire.

The direction of the magnetic field is given by the *right-hand rule*: if you grasp the wire with the thumb of your right hand pointing in the direction of the positive current, your fingers encircle the wire in the direction of the magnetic field, as shown in Figure 22-5.

Bending the wire produces different field patterns. For instance, the magnetic field of a circular loop of wire is shown in Figure 22-6. The contributions to the magnetic field from each segment of the wire add together inside the loop to produce a field pointing into the page. Adding more loops to form the cylindrical structure, called a *solenoid* (Figure 22-7), produces a magnetic

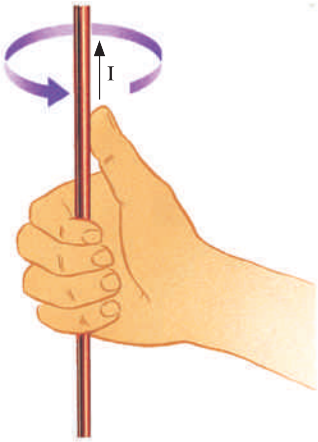


Figure 22-5 The direction of the magnetic field is given by the right-hand rule. If your thumb points in the direction of the current, your fingers encircle the wire in the direction of the field.

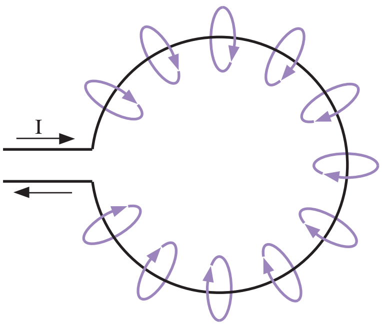


Figure 22-6 The magnetic field of this single loop of wire points into the page inside the loop and out of the page outside the loop.

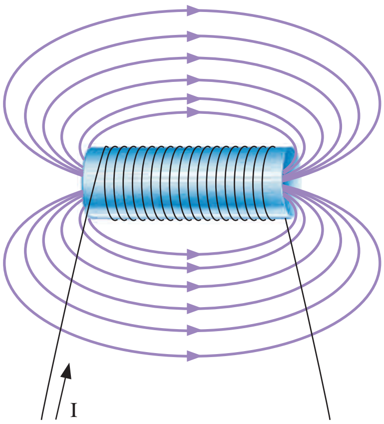


Figure 22-7 The magnetic field of a solenoid is like that of a bar magnet.

field like that of the bar magnet in Figure 22-3. The magnetic field inside the solenoid is relatively strong and rather uniform except near the ends.

Observations like these led French physicist André Ampère to suggest that *all* magnetic fields originate from current loops. We now believe that magnetism originates in current loops at the atomic level. In a simplified model of the atom, we visualize these currents arising from electrons orbiting atomic nuclei or spinning about their axes. The macroscopic magnetic properties of an object are determined by the superposition of these atomic magnetic fields. If their orientations are random, the object has no net magnetization; if they are aligned, the object is magnetized.

Making Magnets

Like electric charge, our universe is filled with magnetism. We do not normally detect the magnetism because of the random orientation of the atomic current loops. The atoms can be aligned, however, by the presence of a magnetic field. In most materials this alignment disappears when the magnetic field is removed. However, in some materials the atoms remain aligned and therefore retain their macroscopic magnetism.

A piece of iron can be magnetized by placing it in a strong magnetic field, by stroking it with a magnet, or by hitting it while it is in a magnetic field. Tapping an iron rod on the floor while holding it vertically will magnetize it because Earth's magnetic field has a vertical component. You can check for the resulting magnetization by holding a compass near the top and bottom of a metal object, such as a filing cabinet. The opening and closing of the file drawers provides the tapping; Earth's magnetic field does the rest. Figure 22-8 shows the orientation of a compass near the top and bottom of a file cabinet. On the other hand, an iron object that has been magnetized can lose its magnetism if it is dropped because the impact tends to randomize the alignment of the atoms. Heating the object can also randomize the atomic magnetic fields.

A versatile magnet can be constructed by wrapping wire around an iron core and connecting the ends of the wire to a battery, as illustrated in Figure 22-9. When current passes through the coil of wire, it generates a magnetic

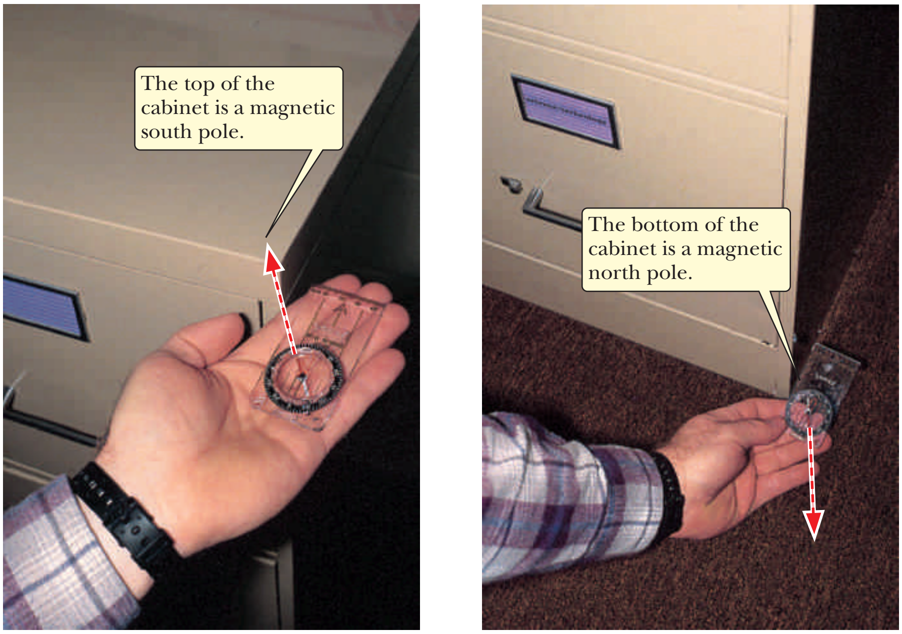


Figure 22-8 The magnetic fields near the top and bottom of a metal file cabinet point in opposite directions.

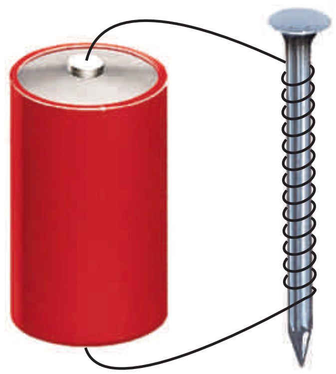
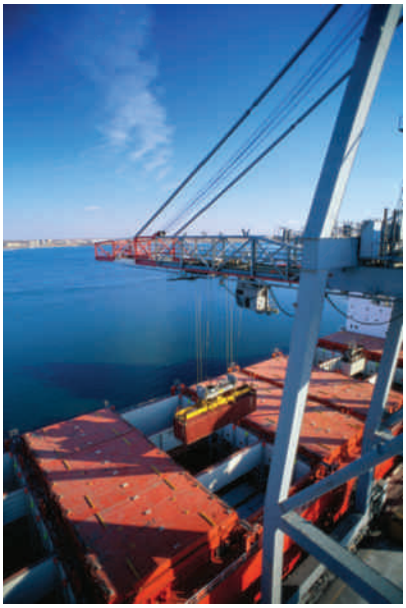


Figure 22-9 A simple electromagnet can be constructed from a nail and some wire.

field along the axis of the coil. The magnetic field that is induced in the iron adds to that of the solenoid, increasing the strength of the magnetic field. These **electromagnets** are useful because they can be turned on and off, the strength of the field can be varied by varying the current, and they can produce large magnetic fields. Electromagnets can be used for such tasks as moving cars or sorting scrap iron from nonmagnetic metals.



A large electromagnet is used to load cargo containers onto a ship.

The Ampere

A current-carrying wire exerts forces on compass needles (Figure 22-10). Therefore, by Newton's third law, a magnet should exert a force on the wire. This effect was quickly verified by experiment. A wire between the jaws of a large horseshoe magnet jumps out of the gap when the current is turned on (Figure 22-11). This force is biggest when the wire is perpendicular to the magnetic field. The force on the wire is always perpendicular to the wire and to the magnetic field.

These experiences indicate that two current-carrying wires should exert forces on each other. They do. The magnetic field produced by each current exerts a force on the other. If the currents are in the same direction, the wires attract each other (Figure 22-12); if the currents are in opposite directions, they repel.

This effect is used to define the unit of current—the basic electrical unit in the metric system. Consider two long parallel wires separated by 1 meter and carrying the same current. If the force between these wires is 2×10^{-7} newton on each meter of wire, the current is defined as 1 **ampere** (A). The *coulomb* (C) is then defined as the amount of charge passing a given point in one of these wires during 1 second. This is the same as the charge on 6.25×10^{18} protons.

The source of the force on each wire is the magnetic field produced by the other wire. We can also use the interacting wires to define a field strength for magnetism. The strength of the magnetic field at a distance of 1 meter from a long straight wire carrying a current of 1 ampere is 2×10^{-7} **tesla** (T). Another unit for magnetic field strength, the **gauss** (G), is often used instead of tesla, where 1 tesla = 10,000 gauss.

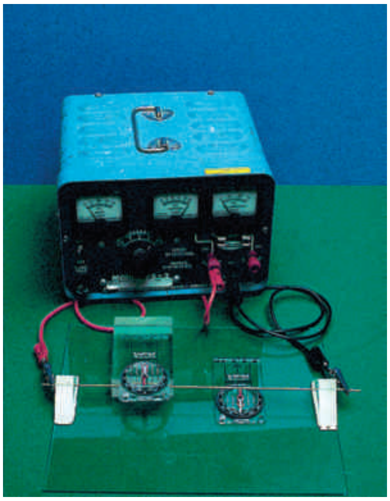


Figure 22-10 A current-carrying wire produces a magnetic field surrounding the wire. Notice that the compasses above and below the wire point in opposite directions.

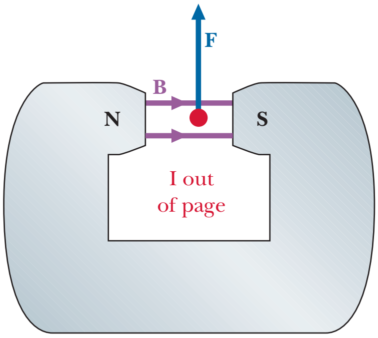
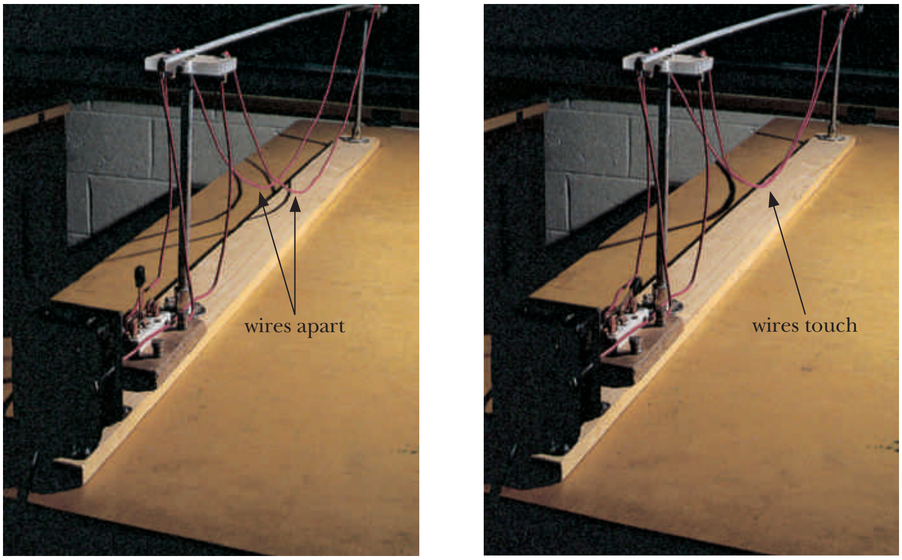


Figure 22-11 The magnetic field of the horseshoe magnet exerts a force on the current-carrying wire, causing it to jump out of the jaws of the magnet.



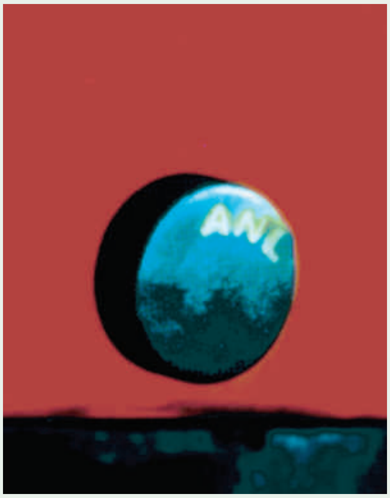
(a) (b)

Figure 22-12 (a) When the switch is open, there is no attractive force between the wires. (b) When the switch is closed, the currents are in the same direction, and the wires are attracted to each other.

Everyday Physics | *Superconductivity*

Under normal conditions all conductors exhibit resistance to the passage of electric current. This makes sense if you visualize the conduction electrons bumping their way through a metal consisting of an array of vibrating atoms. The electrons are accelerated along the wire by the electric potential difference provided by a battery and then lose their acquired speeds in collisions. The surprise is that the electrical resistance of some materials goes to zero at very low temperatures; they become *superconductors*.

This knowledge is not a new breakthrough; superconductivity was first discovered by Dutch physicist Heike Kamerlingh Onnes in 1911. Three years earlier Onnes had developed a process for liquefying helium and was therefore able to study the properties of materials at very low temperatures. The first superconductor was solid mercury, which lost its electrical resistance at a critical temperature of 4.15 K. Experimenters searched for metals or alloys that exhibited superconductivity, and new superconductors were found with critical temperatures as high as 23 K.



Argonne National Laboratory and the U.S. Department of Energy

A permanent magnet levitates above a superconductor because its magnetic field cannot penetrate the superconducting material.

In 1972 three American physicists, John Bardeen, Leon Cooper, and J. Robert Schrieffer, received the Nobel Prize for explaining superconductivity in 1957. Their theory—called the BCS theory—showed how electrons pair up to create a resonance effect that allows the electron pairs to travel effortlessly through the material. The electrical resistance does not simply drop to a very low value; it drops to zero! Once a current has been established in a superconducting material, it will persist without any applied voltage. Such supercurrents have been observed to last for years, which can have great practical applications in the generation of magnetic fields. Large magnetic fields require very large currents. In ordinary materials this means generating a lot of thermal energy at high cost and diverting the heat to avoid melting the magnets. In contrast, once the large currents have been established in superconducting magnets, they can be disconnected from the power supply; no further electric energy is needed.

Theoretical models for superconductivity predicted an upper limit for the critical temperature of about 30 K. Had this upper limit been true, superconductivity would have remained in the domain of very low temperatures. The only gases with boiling points this low are helium (which is expensive) and hydrogen (which is explosive). In 1986, nearly 80 years after the discovery of the first superconductor, a major new class of superconductors was discovered with much higher critical temperatures. These new superconductors are *ceramics* made from copper oxides mixed with such rare-earth elements as lanthanum and yttrium and have critical temperatures as high as 125 K. This new limit is significant because this temperature is higher than the boiling point of nitrogen, an abundant gas that is relatively inexpensive to liquefy and safe to use, allowing liquid nitrogen to be used to keep a material superconducting. A tremendous increase in the use of superconductors will occur if a superconducting material can be made with a critical temperature above room temperature (or at least above those obtained by ordinary refrigeration).

One important potential application of superconductivity is in the transportation of electric energy. With conventional transmission lines, much of the electricity produced at a distant power plant is lost to resistive heating effects in the wires that carry this energy to our homes. A superconducting transmission line would eliminate these losses. The problem, of course, is that we need to cool these conduits in order to make them superconducting. However, if we can use ordinary refrigeration, the cooling costs will be much less than the costs of resistive losses.

Superconductors have a second property that is important. They expel magnetic fields when they become superconducting. This also means that magnetic fields cannot penetrate the superconductors, and they will therefore repel magnets. This is dramatically shown by floating a permanent magnet above a superconductor, as shown in the photo above. This effect may find applications in improving the magnetic levitation of trains for high-speed transportation. A prototype train has been constructed in Japan using helium-cooled superconducting magnets to levitate the train as well as to propel it at speeds up to 300 mph.

Many other applications have been proposed. The switching properties of superconducting materials could have a large impact on the field of computer electronics. It may also be possible to construct superconducting generators and motors. These new superconducting materials may also find applications in the field of medical imaging. However, many of these applications will require major technological breakthroughs in fabrication (for instance, ceramics are brittle and therefore difficult to form into wires) and in finding materials that will carry larger currents.

1. The boiling point of nitrogen is 77.36 K. Why is it important that superconductors have a critical temperature higher than this value?
2. List three possible applications of high-temperature superconducting materials.

Figure 22-13 Earth is a giant magnet with its south magnetic pole in the Northern Hemisphere. This early view of Earth's magnetism has been replaced by one in which the magnetic field is generated by electric currents in Earth's core.

The magnets that are used to hold notes or pictures on refrigerator doors produce fields on the order of 0.3 tesla, whereas large laboratory magnets produce fields of 2.5 teslas. The theoretical limit for a permanent magnetic field is 5 teslas. Electromagnets made with ordinary wires have produced steady fields of 34 teslas, whereas those made with superconducting wires have not yet exceeded 22 teslas. Combinations of these two types of magnets have reached 45 teslas. At the upper extreme, pulsed magnets have produced fields of 70 teslas. Even higher fields approaching 1000 teslas have been generated for extremely short times, but the forces are so large that the magnets destroy themselves.

The Magnetic Earth

As mentioned earlier, a simple compass shows that Earth acts as if it had a huge magnet in its core, as illustrated by Figure 22-13. The strength of Earth's magnetic field at Earth's surface is typically 5×10^{-5} tesla (0.5 gauss). In the United States its horizontal component points generally northward and an even larger component points downward.

Measurements of the magnetic field show that one of Earth's magnetic poles is located in northeastern Canada just north of Hudson Bay, about 1300 kilometers from the geographic North Pole. Earth's other magnetic pole is almost directly on the other side of Earth. A line through Earth connecting these two poles is tilted about 12 degrees from Earth's rotational axis, which passes through Earth's geographic North and South Poles.

Q: Is the magnetic pole located in northern Canada a north or a south magnetic pole?

A: Because this pole attracts the north pole of a compass and because opposite poles attract, this magnetic pole must actually be a south magnetic pole. However, the geographical location of the pole is still known as the magnetic North Pole.

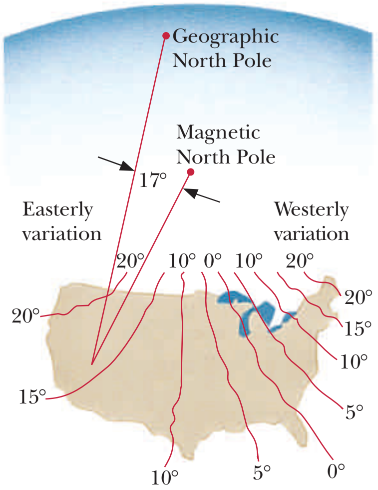


Figure 22-14 The magnetic variation between the directions to the geographic and magnetic North Poles changes across the United States.

The tilt of Earth's magnetic axis from its rotational axis makes finding true north with a compass a bit complicated. The only time your compass will point to true north is when the magnetic pole is between you and the geographic pole. Because of local variations in the magnetic field, this situation occurs along an irregular line running from Florida to the Great Lakes, as shown in Figure 22-14. If you are east of this line, your compass will point to the west of true north. If you are west of this line, your compass points to the east of true north. The difference in the directions to the magnetic and geographic poles is known as the *magnetic variation*. As an example, the magnetic variation in Bozeman, Montana, is currently 15.5 degrees east.

Airplane pilots measure magnetic headings and refer to their aeronautical maps to find the true headings. The large numbers painted on the ends of runways correspond to the magnetic directions of the runways divided by 10 and rounded.

At one time scientists imagined that Earth's magnetic field was caused by magnetized solid iron in Earth's interior, but they now believe this cannot be true. Earth's interior is known to be hot enough that the iron and nickel are in a liquid state. In a liquid state, the atomic magnetic fields do not remain aligned but rather take on random orientations, thus eliminating a macroscopic magnetic effect.

The best model is that large electric currents circulating in the molten interior cause Earth's magnetic field. Such currents could easily produce the field that we observe on Earth, as well as the general features of the magnetic fields of the other planets. However, there are some difficulties with this theory. No one understands the details of the mechanisms for producing and maintaining the currents.

The most puzzling aspect of Earth's magnetic field is its reversals, when the North and South Poles switch locations. There is strong evidence that Earth's magnetic field has reversed directions 171 times in the last 17 million years. This evidence comes from the rocks on both sides of the mid-Atlantic rift. As molten rock emerges from the rift, it cools and solidifies, and at that moment the direction of Earth's magnetization is locked into the rocks and preserved. Samples from the ocean's floor show that the directions of the magnetic fields in the rocks alternate as we approach the rift. Although we know the reversals have occurred, no one has been able to propose a satisfactory mechanism for these reversals or account for the source of the vast amount of energy required to reverse the field.

Charged Particles in Magnetic Fields

Extended presentation available in the *Problem Solving* supplement

Recall that a magnet has no effect on a charged object other than the normal electrostatic attraction between neutral and charged objects. On the other hand, a current-carrying wire in a magnetic field does experience a force. Because an electric current is a stream of charged particles, the motion of the charges must be important. As bizarre as it seems, the magnetic force on a charged particle is zero unless the charge is moving!

In addition, the direction of the magnetic force is probably not the direction you would predict. Recall that the current (the motion of charged particles) is along the wire in Figure 22-11 but that the wire experiences a force in a direction perpendicular to the wire and perpendicular to the magnetic field. Thus, we expect a charged particle moving in a magnetic field to experience a force that is at right angles to its velocity and to the magnetic field.

The strength of the force depends on the charge q, the strength of the magnetic field B, the speed v, and the angle between the field and the velocity. The magnetic force is maximum when the field and velocity are perpendicular:

$$F_{max} = qvB$$

maximum magnetic force = charge × speed × magnetic field

When the velocity and the magnetic field are parallel, the force is zero.

The magnetic force produces some interesting phenomena. First, because the force is always perpendicular to the particle's velocity, it never changes the particle's speed, only its direction. If the particle's velocity is parallel to the magnetic field, the particle does not experience a force, and it moves in a straight line. On the other hand, if the velocity is perpendicular to the field, the particle experiences a centripetal force that causes it to move in a circular path [Figure 22-15(a)]. If the velocity has components parallel and perpendicular to the magnetic field, these two motions will superimpose, and the particle will follow a helical path [Figure 22-15(b)] along the magnetic field.

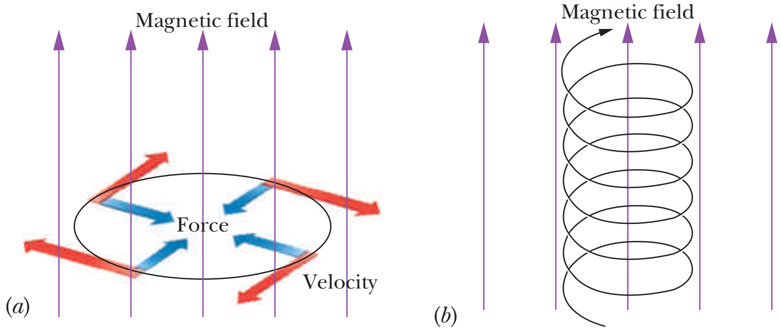


Figure 22-15 (a) The charged particle follows a circular path if its velocity is perpendicular to the magnetic field. (b) The path is helical if its velocity has components both perpendicular and parallel to the field.

©Kevin Schafer/Photographer's Choice/Getty

The aurora borealis is caused by cosmic rays following magnetic field lines toward the magnetic North Pole.

The magnetic force on charged particles is the cause of dramatic effects known as the *aurora borealis* (northern lights) and the *aurora australis* (southern lights). As charged particles, mostly from the Sun but also from outer space, approach Earth, they interact with Earth's magnetic field. This interaction causes the charged particles to follow helical paths along Earth's magnetic field so that they strike Earth's atmosphere in the regions of the magnetic poles. The collisions of these cosmic rays with the oxygen and nitrogen molecules in the atmosphere produce the spectacular evening light shows high over the north and south magnetic poles.

The magnetic forces on charged particles are important in many scientific and technologic devices as varied as containment vessels used in developing the future technology of fusion reactors, particle accelerators used in research, and television sets in our homes. We will discuss many of these in the remaining chapters.

WORKING IT OUT *Magnetic Force*

If a proton is moving perpendicular to a uniform magnetic field, it will experience a magnetic force that will pull the proton out of its straight-line path and cause it to travel in a circle of radius r_c, known as the "cyclotron radius." Find the cyclotron radius for a proton moving at 3×10^6 m/s perpendicular to a magnetic field of 1.5 T.

The maximum magnetic force on the proton (for motion perpendicular to the field) is given by

$$F = qvB = (1.6 \times 10^{-19}\ \text{kg})(3 \times 10^6\ \text{m/s})(1.5\ \text{T}) = 7.2 \times 10^{-13}\ \text{N}$$

This very small force causes a large acceleration because it is acting on a small mass (the proton).

$$a = \frac{F}{m} = \frac{7.2 \times 10^{-13}\ \text{N}}{1.67 \times 10^{-27}\ \text{kg}} = 4.3 \times 10^{14}\ \text{m/s}^2$$

This acceleration is centripetal, so we can now solve for the cyclotron radius:

$$a = \frac{v^2}{r_c}$$

$$r_c = \frac{v^2}{a} = \frac{(3 \times 10^6\ \text{m/s})^2}{4.3 \times 10^{14}\ \text{m/s}^2} = 0.021\ \text{m} = 2.1\ \text{cm}$$

Magnetism and Electric Currents

In the evolution of the physics world view, a small number of basic themes have emerged that reflect our biases but also fuel our searches. One such theme is symmetry. A number of times, progress has been made while searching for symmetric effects. Early in the development of our understanding of electricity and magnetism, the questions of symmetry haunted some experimenters. In particular, they asked: given that an electric current produces a magnetic field, does a magnetic field produce an electric current? To investigate this, you might try wrapping some wire around a magnet and connecting the wire to an *ammeter*, an instrument for measuring current. Although similar experiments were performed with large magnets and sensitive instruments, no effect was found.

British scientist Michael Faraday discovered the connection in 1831. He found that motion is the key to producing an electric current with a magnet. If

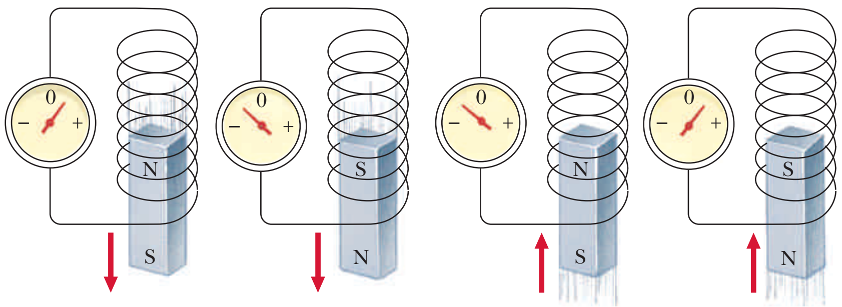


Figure 22-16 A magnet inserted into or removed from a coil of wire produces a current. (When the needle is centered, there is no current.)

a wire is moved through a magnetic field (but not parallel to the field), a current is produced in the wire. This current is due to the motion of the wire in the magnetic field; there are no batteries. The current is largest if the motion is perpendicular to the field and increases with the speed of the wire. Having the benefit of hindsight, we see that this is understandable because of the magnetic force on the charges in the wire.

Because the principle of relativity (Chapter 9) must be valid for electricity and magnetism, we know that it is the *relative* motion of the wire and the magnet that is important. It doesn't matter which is moving relative to the laboratory and which is stationary. Therefore, we should be able to generate a current in a stationary wire by moving the magnet. This principle can be demonstrated with a simple experiment. If a bar magnet is placed in a coil of wire and quickly withdrawn, a current is generated that is easily detected. Quickly inserting the magnet into the coil generates a current in the opposite direction. Reversing the direction of the bar magnet also reverses the direction of the current. The four possibilities are shown in Figure 22-16.

Faraday also discovered that motion is not the only way of producing a current with a magnetic field. A current is produced if the strength of the magnetic field varies with time, even when there is no relative motion of the wire and the magnet. An increasing field produces a current in one direction; a decreasing field produces a current in the opposite direction.

After a long series of experiments conducted over several years, Faraday was able to generalize his results in terms of magnetic field lines. These lines are visualized as pointing along the direction of the magnetic field at all points in space. The number of lines in a given region of space represents the strength of the magnetic field. The lines are closer together where the magnetic field is stronger.

Faraday showed that if the number of magnetic field lines passing through a loop of wire changed *for any reason,* a current was produced in the loop (Figure 22-17). The voltage (and hence the current) generated in the loop depends

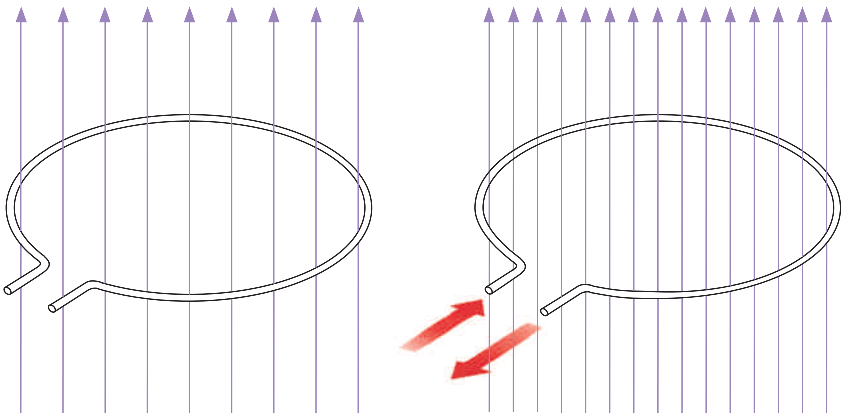


Figure 22-17 Changing the number of magnetic field lines passing through a loop of wire produces a current. The direction of the current is given by Lenz's law.

on how fast the number of field lines passing through the loop changes—the faster the change, the larger the voltage. This explanation of these phenomena is now known as *Faraday's law*.

Faraday's law ▶

Lenz's law ▶

The direction of the current induced in the coil by the changing number of magnetic field lines is given by *Lenz's law*, which states that the current always produces a magnetic field to oppose the change that created it. For example, in Figure 22-17 the number of lines in the upward direction is increasing. Therefore, the induced current will produce a magnetic field that points downward to try to cancel out the increase. Notice that the current shown in Figure 22-17 does this. If the number of lines had been decreasing, the current would be induced in the opposite direction to produce a magnetic field in the upward direction to try to maintain the original field.

Q: **What would you see on the ammeter if you dropped a magnet through the coil in Figure 22-16?**

A: **As the magnet enters the coil, the needle will swing to one side because of the increase in the number of field lines passing through the coil. As the magnet exits, the needle will swing to the other side because the number of lines decreases.**

FLAWED REASONING

The following question appears on the midterm: "A copper loop lies flat on a table. If the north pole of a bar magnet is lowered quickly toward the loop, as shown in Figure 22-18, a current is induced in the loop in the direction given by Lenz's law. Find this direction and explain how you used Lenz's law."

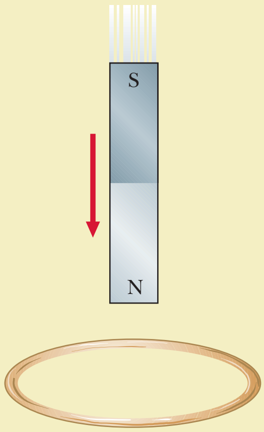


Figure 22-18

Jose gives the following answer: "Magnetic field lines leave the north pole of a bar magnet, so the field lines through the loop are directed downward. Lenz's law states that an induced magnetic field will be generated in the opposite direction—that is, upward. By the right-hand rule, the induced current must be counterclockwise to produce this field."

The induced current is counterclockwise, as Jose claims, but not for the reason given. **Explain Lenz's law to Jose.**

ANSWER There are two magnetic fields in the loop: the *external* field of the bar magnet and the *induced* field due to the current. Jose believes that the induced field opposes the external field, but this is not what Lenz's law asserts. The induced field opposes any *change* in the external field, not the field itself. Jose gets the correct direction by luck. If the bar magnet is moving upward, Jose's reasoning yields the wrong answer. The external field in the loop is downward and growing weaker. Jose predicts that the induced field is upward to oppose the downward external field. Lenz tells us that the induced field is downward to oppose the weakening of the external field.

Transformers

These discoveries about magnetism—especially its connections with electricity—have many practical uses. For example, they are fundamental to the operation of transformers used to change the voltage of alternating-current electricity. The same amount of electric power can be delivered through wires at low voltage and high current as at high voltage and low current (Chapter 21). For instance, 10 amperes at 12 volts supplies the same power as 1 ampere at 120 volts. The particular choice of what amperage and voltage are used depends on the circumstances. The large cylindrical containers on power poles are transformers for reducing the voltage to 120 volts before it enters our homes and businesses.

The schematic diagram in Figure 22-19 shows the essential features of a transformer. One coil, called the *primary coil*, is connected to a source of alternating current. The alternating current in the primary coil produces an alternating magnetic field that is transmitted to the *secondary coil* by the iron core. This alternating magnetic field produces an alternating voltage in the secondary coil.

The size of the voltage produced in the secondary coil depends on the ratio of the number of loops in the two coils. A transformer designed to reduce the voltage by a factor of 2 would have one-half as many loops in the secondary as in the primary coil.

Neighborhood transformers reduce the voltage for household use.

Q: Why won't a transformer work with direct-current electricity?

A: Direct current will not produce the varying magnetic field necessary to induce a current in the secondary coil.

Generators and Motors

The electric generator is another application of Faraday's discovery. If we rotate the loop of wire in Figure 22-17 by 90 degrees, the number of magnetic field lines passing through the loop drops to zero. The change in the number of field lines passing through the loop produces a voltage around the loop, which produces a current. The loop is usually rotated by a steam turbine or falling water in a hydroelectric facility. The steam can be produced by many means: by burning wood, coal, petroleum, or natural gas or by using heat from the Sun or nuclear reactors.

Figure 22-20 contains a series of drawings of the magnetic field and a wire loop in an electric generator that illustrates how the rotating loop produces electric voltages (and currents). Figure 22-20(a) shows the plane of the loop parallel to the magnetic field lines. The number of lines passing through the loop in this orientation is zero. As the loop rotates at a constant speed, the number of lines initially increases rapidly, producing a large voltage. As it continues to rotate (b), the number of lines passing through the loop continues to increase but at a slower rate. The voltage drops to zero when the plane of the loop is perpendicular to the field lines (c) and the number of lines through the loop is a maximum. The number of lines through the loop now decreases (d), and the voltage increases in the opposite direction. It increases to a maximum (e) and then returns to zero when the plane of the loop is perpendicular to the field lines once again. The voltage increases to a maximum in the opposite direction, and the entire cycle repeats. This generator produces the alternating voltage shown in Figure 22-21(a).

A simple change in the way the voltage is carried to the external circuit converts the generator to produce pulsating direct current [Figure 22-21(b)].

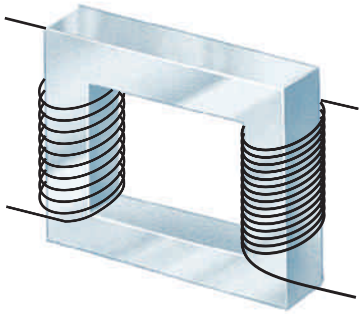


Figure 22-19 The essential features of a transformer.

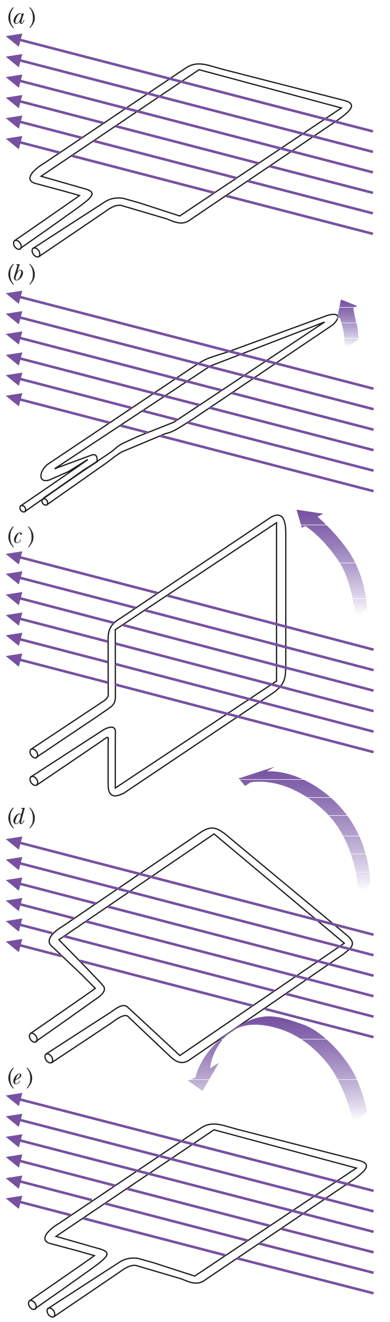


Figure 22-20 The number of magnetic field lines passing through a loop of wire changes as it rotates. This produces the alternating voltage shown in Figure 22-21.

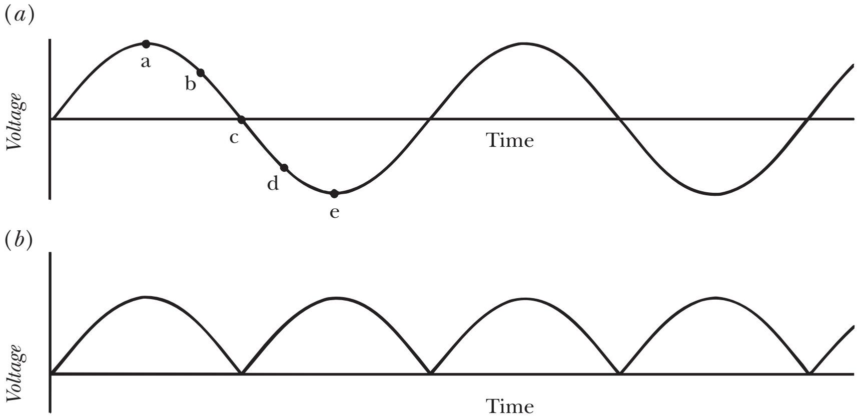


Figure 22-21 (a) A loop rotating in a magnetic field produces an alternating voltage. The letters show the voltage produced when the loop has the positions shown in Figure 22-20. (b) The pulsating direct current produced by using a commutator.

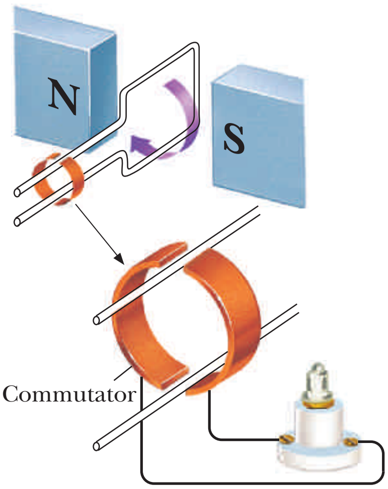


Figure 22-22 Schematic of a direct-current generator.

A connector (called a *commutator*), shown in Figure 22-22, reverses the connections from the loop to the outside circuit each half turn. This pulsating current can then be electronically smoothed to produce a constant direct current like that from a battery.

A direct-current motor is basically a direct-current generator run backward. In fact, the first such motor was produced during an exhibition in 1873 when a technician setting up a demonstration of generators hooked one up the wrong way and "discovered" a motor! In a generator we rotate the loop in a magnetic field, which produces a voltage that moves electric charges. In a motor the sequence is reversed; we apply a voltage, causing the charges to move. This current in a magnetic field produces a force that rotates the loop.

Figure 22-20 can be used to illustrate the operation of a motor. When a voltage is applied to the loop when it is in the position shown in Figure 22-20(a), the magnetic field exerts a torque on the current-carrying loop. The forces on the long sides of the loop are in opposite directions because the currents are in opposite directions. The torque decreases as the loop rotates and becomes zero when the plane of the loop is vertical. If nothing changes as the loop coasts through this position, it will oscillate and eventually come to a stop. However, the commutator reverses the direction of the current so that the torque continues to act in the same direction.

The similarity of a motor and a generator is important in some electric-powered vehicles. While the vehicle is speeding up or cruising, the engine acts as an electric motor. However, when the brakes are applied, the connections

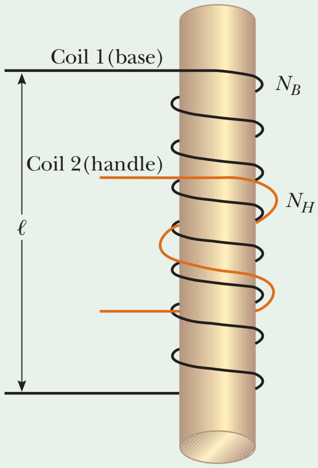
Everyday Physics | *"Wireless" Battery Charger*

An electric toothbrush has a base designed to hold the toothbrush handle when not in use. As shown in the following figure, the handle has a plastic cylindrical hole that fits loosely over a matching plastic cylinder on the base. How does the toothbrush get charged when there are no metal contacts? When the handle is placed on the base, a changing current in a solenoid inside the base cylinder (the primary coil) induces a current in a coil inside the handle (the secondary coil). This induced current charges the battery in the handle. The size of the induced current is amplified by inserting a core of ferromagnetic material, such as iron, into the primary coil. The primary coil then magnetizes the iron core such that its poles reverse 120 times each second (for 60-cycle current). This causes a greater change in the magnetic field inside the secondary coil.

Wireless charging is used in a number of other "cordless" devices. One significant example is the inductive charging that is used by some electric car manufacturers that avoids direct metal-to-metal contact between the car and the charger.

1. Why is N_H, the number of coils in the handle, less than N_B, the number of coils in the base unit?
2. Find the ratio N_B/N_H if the handle contains a 9-volt rechargeable battery.


Gilette



This electric toothbrush uses magnetic induction to charge its battery. The primary coil (coil 1) of N_B turns is in the base unit and the secondary coil (coil 2) of N_H turns is in the handle.

to the engine are changed so that it acts as a generator to recharge its batteries. The current exerts a torque on the engine, which in turn exerts a torque on the wheels to stop the vehicle. This is not a perpetual-motion machine because only part of the energy is recovered. However, regenerative braking can extend the range of the vehicle.

A Question of Symmetry

The connection between electricity and magnetism appears to be complete—a *changing* magnetic field produces an electric current, and an electric current produces a magnetic field. Notice, however, that the situation is not quite symmetric. It requires a changing magnetic field to produce the electric current, but the electric current produces a steady magnetic field.

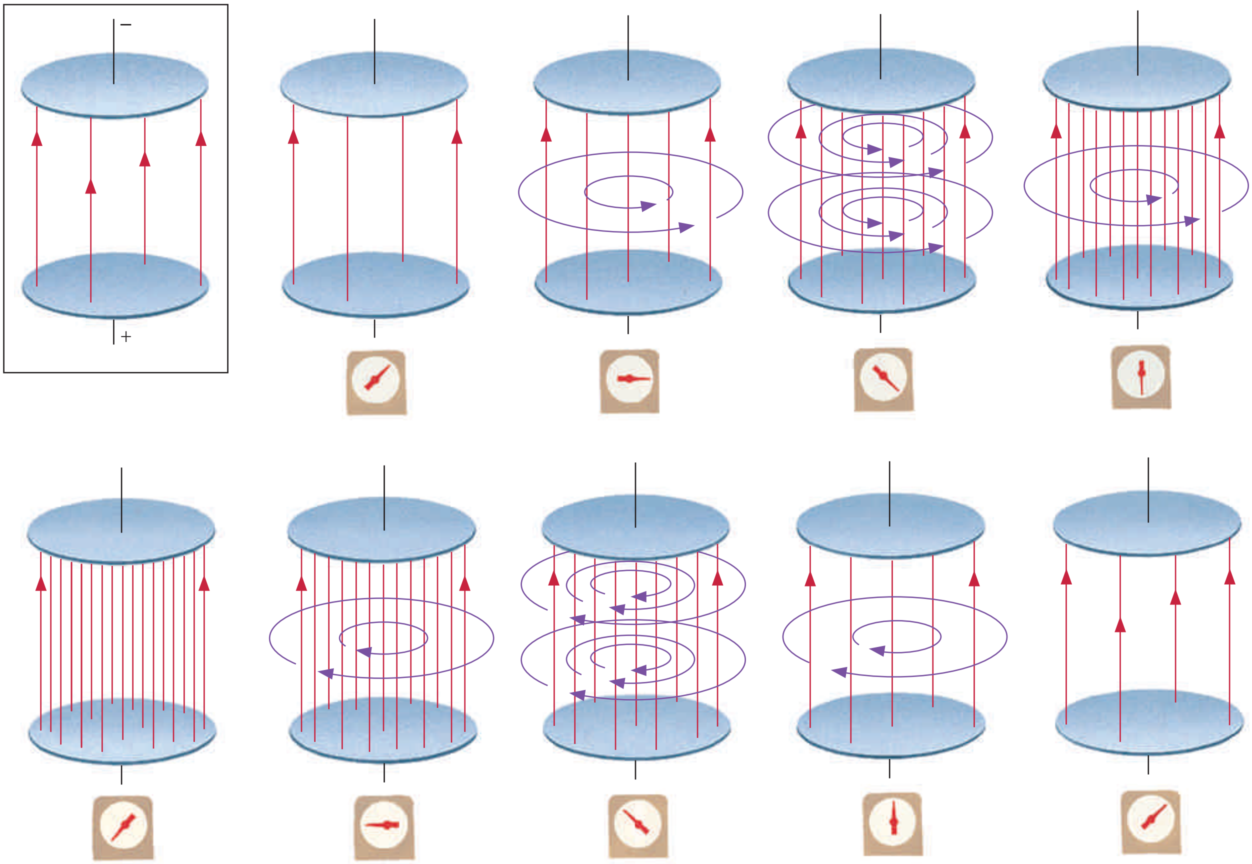


Figure 22-23 A steady electric field between the parallel plates shown in the boxed diagram at the left does not produce a magnetic field. An increasing electric field produces a magnetic field in one direction, whereas a decreasing field produces a magnetic field in the opposite direction.

Actually, there is more to this connection that becomes apparent only when everything is expressed in terms of the electric and magnetic fields. Consider two parallel plates connected to a battery, as shown in Figure 22-23. As current from the battery builds up charges on the plates, an electric field is generated in the region between the plates. Even though no charges are flowing between the plates, a magnetic field is produced in the region surrounding the plates that matches the magnetic field surrounding the wires. Therefore, the origin of the new field cannot be the charges. Furthermore, the magnetic field disappears when the current stops. The electric field between the plates continually increases as long as charge is flowing to the plates and remains constant when the charge stops flowing. Therefore, the magnetic field is produced by a *changing* electric field. As the charges leave the plates, the electric field decreases and a magnetic field in the opposite direction is produced.

We can make the analogy closer by looking at the region between the poles of an electromagnet. As the magnetic field between the poles increases, an electric field is produced in the surrounding region (Figure 22-24). When the magnetic field reaches its maximum strength and is no longer changing, the electric field disappears. As the magnetic field decreases, an electric field in the opposite direction is produced.

It is important to realize that a changing magnetic field produces an electric field in *empty* space. There is no need for a wire. If a wire is present, the electric field exerts forces on the charges within the wire and produces a current. But the important point is that even in the absence of the wire, the electric field is present.

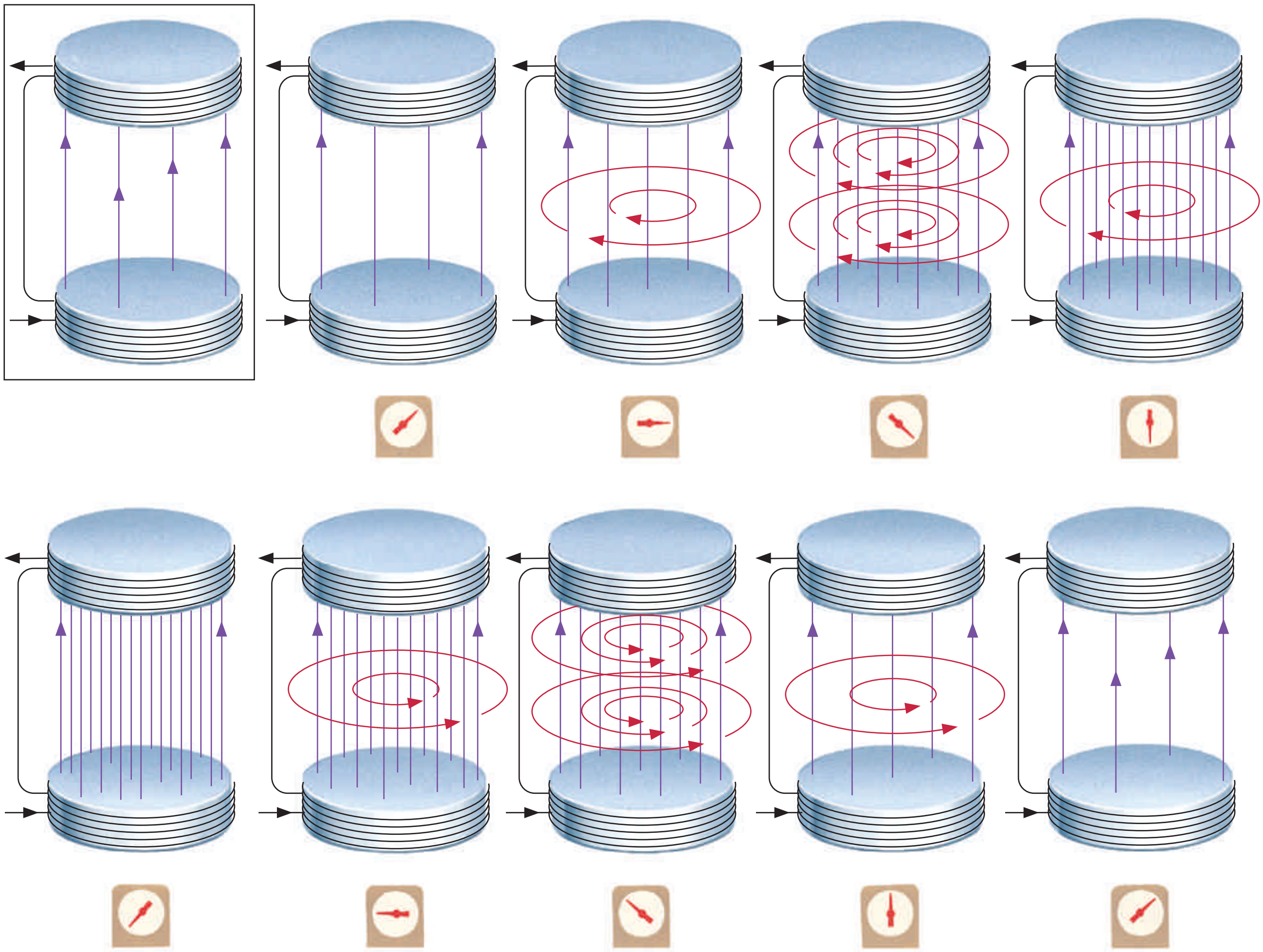


Figure 22-24 A steady magnetic field between the poles of the electromagnet shown in the boxed diagram at the left does not produce an electric field. An increasing magnetic field produces an electric field in one direction, whereas a decreasing magnetic field produces an electric field in the opposite direction.

If we use these results and focus on the fields rather than the currents, the situation is completely symmetric: a *changing* magnetic field generates an electric field, and a *changing* electric field generates a magnetic field. There is an intimate relationship between electricity and magnetism.

Electromagnetic Waves

If the magnetic field changes at a constant rate (that is, if it changes by the same amount each second), the electric field that is produced is constant. A rapidly changing magnetic field produces a large electric field, and a slowly changing magnetic field produces a smaller one. However, if the magnetic field starts out changing slowly and then increases its rate of change, the electric field starts out small and grows larger. Thus, it is possible for a changing magnetic field to produce a changing electric field.

The parallel plates in Figure 22-23 show the symmetric effect; a changing electric field produces a magnetic field. The rate of change of the electric field determines the size of the magnetic field. Therefore, a changing electric field can produce a changing magnetic field.

We have discussed a lot of "rates of change," and this can be quite confusing, but the payoff for understanding the process is worth the effort. We have discovered a sequential chain of field productions: one changing field

Q: How would you use an electric field to produce a magnetic field that increases in size?

A: The electric field must change slowly at the beginning and continually increase its rate of change.

produces another changing field, and then this new changing field produces the first kind of field again, and so on. The two fields generate each other in empty space.

We have only argued that this is possible, but the process was rigorously deduced in the 1860s by James Clerk Maxwell. Maxwell showed that this was a consequence of a set of four equations that he and others had developed to describe electricity and magnetism and the many connections between them. The equations, called *Maxwell's equations* in honor of his contributions, summarize all of electricity and magnetism.

Maxwell combined these equations into a single equation that had the same form as the equations that describe periodic waves, whether they are waves on a rope, water waves, or sound waves. Oscillating electric and magnetic fields can combine to produce waves that travel through space (Figure 22-25). As calculated by Maxwell, these **electromagnetic waves** take the form of oscillating electric and magnetic fields that travel with a speed equal to that of light. In 1887 German physicist Heinrich Hertz was able to produce electromagnetic waves on one side of a room and detect them on the other. The existence of electromagnetic waves affirmed Faraday's belief that these fields had their own identities.

Maxwell | *A Man for All Seasons*

James Clerk Maxwell was born on June 13, 1831, into a prominent Edinburgh family. *Clerk* (pronounced "Clark") was his original family name before the Clerks and the Maxwells intermarried. His "hyphenated" name was taken as a result of legal manipulations that prevented the family from holding extensive property in one unit.

Maxwell was brought up in a thoroughly Scottish environment, with kilts, Gaelic, and a burr. He was a curious and gifted child. His father, who had serious scientific interests, encouraged this bent. He attended Edinburgh Academy and in 1847 entered Edinburgh University, where a distinguished faculty welcomed him. One of his professors was famous for his textbooks in physics—some of which were used in the United States until after World War II. In 1850 Maxwell went to Cambridge University, first to Peterhouse College and then to Newton's old haunt, Trinity College. He was a brilliant student of mathematics and began a lifelong study of mathematical models of Saturn's rings. He won the most prestigious mathematical award in British school mathematics and began what would be a wonderfully productive review of the pioneering electromagnetic experiments of Michael Faraday.

Maxwell returned as a professor to Aberdeen, Scotland, and then to King's College in London. He married but had no children. He retired to his Galloway estate to write the monumental *Treatise on Electricity and Magnetism*. His work in this field generated important advances in the electrical industry, in technology, and in every branch of physical science.

North Wind Picture Archives

James Clerk Maxwell

Although retired and in ill health, Maxwell later designed and founded the Cavendish Laboratory and was its first director. This was the first major experimental laboratory explicitly designed with shielding from outside radiation and vibration.

He also developed an equally powerful kinetic theory of gases, which led in turn to further breakthrough work in thermodynamics. He continued the Saturn work and contributed to astrophysics, color vision, optics, photoelasticity, servomechanics, and, in the long run, color photography. He left a remarkable legacy. In tribute, Albert Einstein always displayed Maxwell's portrait alongside that of Isaac Newton. He endures as one of the greatest scientists in history.

—Pierce C. Mullen, historian and author

Source: C. W. Francis Everitt, *James Clerk Maxwell: Physicist and Natural Philosopher* (New York: Scribner, 1975).

FLAWED REASONING

All of us struggle with misconceptions as we build our world views. Even Isaac Newton made mistakes in his reasoning. He believed that "luminiferous ether" must fill all space to provide a medium for the propagation of light waves. If we jiggle gelatin on one side of a bowl, the jiggle moves across the bowl. Newton felt it was impossible for the jiggle to move across an empty bowl. **Explain how an electromagnetic wave can propagate through empty space.**

ANSWER The electromagnetic wave propagates through empty space by re-creating itself. A changing electric field creates a changing magnetic field—it waves the magnetic field. This changing magnetic field in turn creates a changing electric field—it waves the electric field. And then the process starts over again. It is somewhat like the popular preschool story of *Harold and the Purple Crayon* by Crockett Johnson. In this story, Harold gets to where he wants to go by drawing himself a set of stairs with his purple crayon and using his imagination.

Image not available due to copyright restrictions

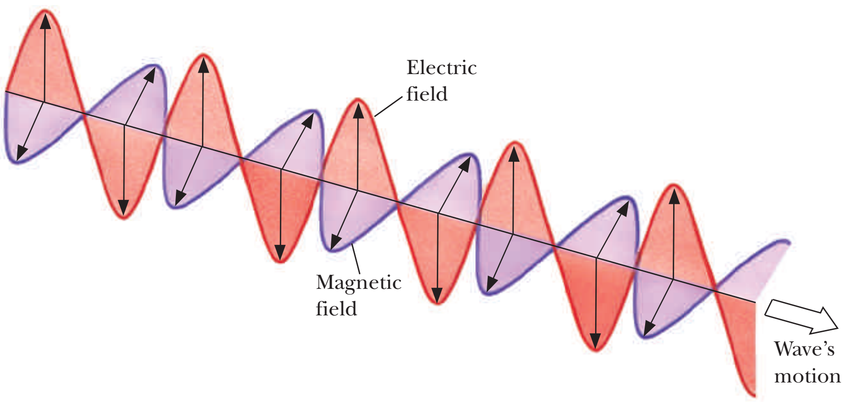


Figure 22-25 An electromagnetic wave propagating through space. The electric and magnetic fields are perpendicular to each other and to the direction the wave is moving.

Electromagnetic waves are produced whenever electric charges are accelerated. If the charges have a periodic oscillatory motion, the wave will have a fixed frequency and therefore a fixed wavelength according to the relationship we developed in Chapter 15 between speed, wavelength, and frequency:

$$v = \lambda f$$

◀ speed = wavelength × frequency

Maxwell's equations require that the speed be that of light, but place no restrictions on the frequency, as shown in the diagram of the electromagnetic

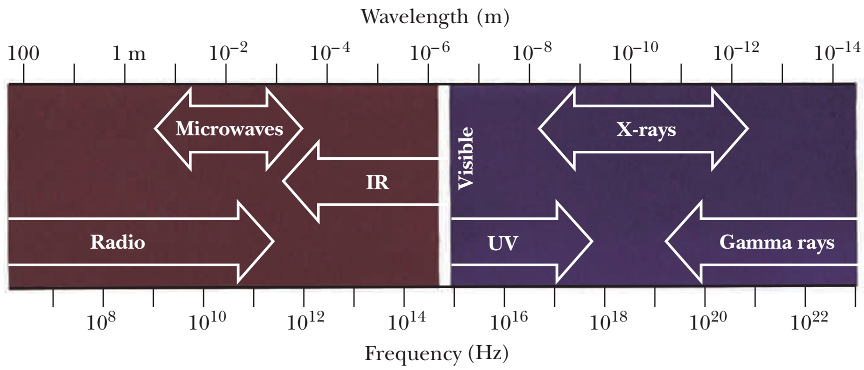
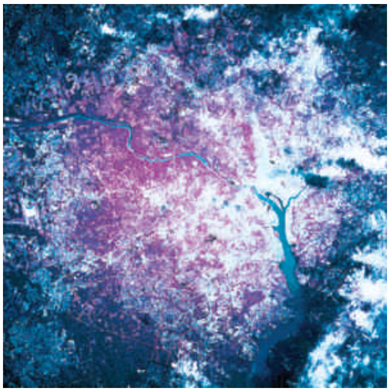


Figure 22-26 The electromagnetic spectrum.



NASA

An infrared photo of the Washington, D.C., area looking north, showing the Potomac River in blue and vegetation in red. This view from a space shuttle covers 68 square miles.

spectrum in Figure 22-26. Although these waves all have the same basic nature, the different ranges in frequency are produced (and detected) by different means. The boundaries between the various named regions are not distinct; they overlap quite a bit.

The lowest frequencies and the longest wavelengths belong to the radio waves. These are produced by large devices such as broadcast radio antennas. Microwaves are also produced electronically, but the devices are smaller, ranging in size from a few millimeters to a few meters. These devices are used in microwave ovens, radar, and long-distance transmission of telephone calls.

The frequencies of visible light extend from 4.0 to 7.5×10^{14} hertz. Although visible light occupies only a small region of the complete spectrum, it is obviously very important to us. This region is bounded on the low side by infrared (IR) radiation, waves whose wavelengths are too long (beyond the red) to be seen by the human eye. Infrared radiation is most noticeable when given off by hot objects, especially those that are red hot. This is the radiation you feel across the room from a fire or a heating element. The radiation that lies beyond violet is known as ultraviolet (UV) waves. This is the component of sunlight that causes suntans (and sunburns if it is excessive). Visible light and its neighbors are produced at the atomic level, and their properties are valuable clues about the structure of matter at the atomic level.

We are all familiar with X rays from visits to the doctor or dentist. X rays have high frequencies and are very penetrating. They are produced by the rapid acceleration of electrons in X-ray machines and are emitted by atoms. Gamma rays are an even higher-frequency radiation that originate in the nuclei of atoms. We will study these in more detail in later sections of this textbook.

Photo courtesy of DIRECTV, Inc.

A receiver dish for satellite television signals.

Radio and TV

Radio is a means of coding electromagnetic waves with the information in sound waves so that they can be transmitted through space, intercepted, and converted back into sound. Television is the same sort of process with the addition of the video information. Sound waves are changed to an electric signal by a microphone. In one version the sound waves cause a coil to vibrate in a magnetic field. This produces a current in the coil that can then be amplified. Audio frequencies are in the range of 20 hertz to 20 kilohertz.

If audio frequencies were broadcast directly, there could only be one radio station in any geographic region. Instead, the audio signal is combined with a broadcast signal. A station broadcasting at "1450 on your dial" sends out waves with a frequency of 1450 kilohertz (often called *kilocycles*). This is the *carrier frequency*. The audio signal is used to vary, or *modulate*, the carrier signal. Two modulation methods are *amplitude modulation* (AM) and *frequency modulation*

(FM). The frequency of the sound determines the frequency of the modulation, and the loudness of the sound determines the amplitude of the modulation. In AM radio the audio signal causes the amplitude of the carrier signal to vary; in FM radio it causes the frequency of the carrier signal to vary (Figure 22-27).

Q: Although modern police and fire sirens produce sound waves rather than radio waves, they are modulated. Is this modulation AM or FM?

A: Because these sirens have oscillating frequencies, they must be FM.

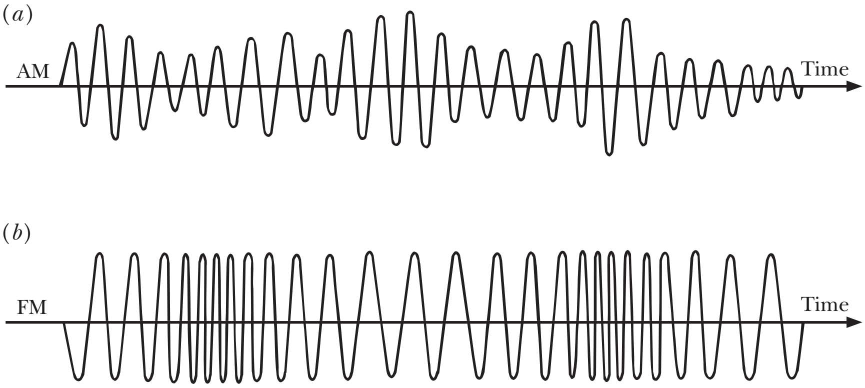


Figure 22-27 (a) AM and (b) FM radio waves.

In either case, the final signal is then amplified and sent to the antenna, where it causes electrons to move up and down the antenna wire. The accelerations of these electrons produce electromagnetic waves that are broadcast. As these waves hit the antenna in your radio, they cause the electrons in the antenna to move back and forth, producing oscillating currents. Although all nearby stations are received simultaneously, the radio is tuned so that it resonates with only one of them at a time. Your radio amplifies this station's signal.

The radio contains electric circuits to filter out the carrier frequency and retain only the electrical version of the sound information. These signals are then sent to a speaker for reconversion to sound waves. In one version of a speaker, the electric signal generates a magnetic field that interacts with a magnet to move a diaphragm. This diaphragm then moves air to generate sound.

The allocation of the possible ranges of carrier frequencies to various types of broadcast is a governmental responsibility that is complicated by the history of the development of such advances in broadcasting techniques as FM stereo and TV. AM radio stations broadcast between 550 and 1500 kilohertz, FM radio between 88 and 108 megahertz, and TV in three regions between 54 and 890 megahertz. Other regions are assigned to citizens-band receivers, ships, airplanes, police, amateur radio operators, and satellite communication.

Summary

Magnetic poles behave similarly to electric charges: like poles repel and unlike poles attract. However, magnetic poles always occur in pairs. Magnets attract some objects but have no effect on others.

Magnets have no effect on stationary charges and do not deflect the foils of an electroscope. On the other hand, a current-carrying wire produces a magnetic field and is attracted or repelled by other magnets or current-carrying wires. The force on the wire is always perpendicular to the wire and to the

WORKING IT OUT *FM Wavelength*

You are driving through Bozeman, Montana, listening to FM radio station 103.5 (KZMY). What is the wavelength of the carrier signal for this station?

The reading on the radio dial is the carrier frequency, expressed in megahertz. The carrier frequency is 103.5×10^6 Hz, or 1.035×10^8 Hz. Radio waves, like all other waves, obey the wave equation, $v = \lambda f$. Because radio waves are electromagnetic waves, they travel at the speed of light, $v = c = 3 \times 10^8$ m/s. We can easily find the wavelength of the KZMY carrier signal:

$$\lambda = \frac{c}{f} = \frac{3 \times 10^8 \text{ m/s}}{1.035 \times 10^8 \text{ Hz}} = 2.899 \text{ m}$$

or, about the length of your car.

magnetic field. Two current-carrying wires attract each other if the currents are in the same direction; they repel if the currents are in opposite directions. The field strength for magnetism is defined as 2×10^{-7} tesla at a distance of 1 meter from a wire carrying a current of 1 ampere.

All magnetic fields originate from current loops. Naturally occurring magnetism originates in current loops at the atomic level. Earth's magnetism has a strength at the surface of about 5×10^{-5} tesla (0.5 gauss) and is caused by large electric currents circulating in Earth's molten interior.

A charged particle moving in a magnetic field experiences a force at right angles to its velocity and to the magnetic field. The strength of the force depends on the angle between the field and the particle's motion. It is maximum when they are perpendicular and zero when they are parallel.

If a wire and magnetic field move relative to each other, a current is produced in the wire, provided the motion is not parallel to either the wire or the field. This current is largest if the motion is perpendicular to the field and increases with the relative speed. A current also occurs in a loop of wire if the magnetic field inside the loop varies with time. An increasing field produces a current in one direction; a decreasing field produces a current in the opposite direction.

Field lines can be used to represent magnetic fields. The magnetic field is strongest in regions where the field lines are close together. If the number of magnetic field lines passing through a loop of wire changes *for any reason*, a current is produced in the loop. The voltage (and hence the current) generated in the loop depends on the rate of change—the faster the change, the larger the voltage.

There is a close connection between changing electric and magnetic fields. A changing magnetic field can generate a changing electric field, and a changing electric field can generate a changing magnetic field, creating electromagnetic waves that travel through empty space. These waves are produced whenever electric charges are accelerated. When the charges have a periodic oscillatory motion, the wave has a fixed frequency and wavelength. The spectrum of these waves ranges from low-frequency radio waves, to visible light, to high-frequency X rays and gamma rays.

CHAPTER 22 *Revisited*

The connection between your favorite TV show and your microwave oven lies in the electromagnetic waves that are created by combining electric and magnetic effects. These waves can resonate with molecules in a microwave oven to raise the temperature of the popcorn and can be beamed through space to bring us television (and radio) programming.

Key Terms

ampere The SI unit of electric current. The current in each of two parallel wires when the magnetic force per unit length between them is 2×10^{-7} newton per meter.

electromagnet A magnet constructed by wrapping wire around an iron core. An electromagnet can be turned on and off by turning the current in the wire on and off.

electromagnetic wave A wave consisting of oscillating electric and magnetic fields. In a vacuum, electromagnetic waves travel at the speed of light.

gauss A unit of magnetic field strength; 10^{-4} tesla.

magnetic field The space surrounding a magnetic object, where each location is assigned a value determined by the torque on a compass placed at that location. The direction of the field is in the direction of the north pole of the compass.

magnetic monopole A hypothetical, isolated magnetic pole.

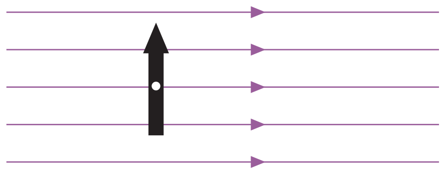
magnetic pole One end of a magnet; analogous to an electric charge.

tesla The SI unit of magnetic field.

Questions and exercises are paired so that most odd-numbered are followed by a similar even-numbered.
Blue-numbered questions and exercises are answered in Appendix B.
indicates more challenging questions and exercises.
WebAssign **Many Conceptual Questions and Exercises for this chapter may be assigned online at WebAssign.**

Conceptual Questions

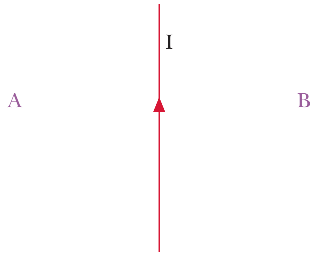
1. If you are given three iron rods, how could you use them to find the one that is not magnetized?
2. If the labels on a magnet are missing, how would you determine which pole is the north pole?
3. You have three iron bars, each of which may or may not be a permanent magnet. Each rod is painted green on one end and yellow on the other. You perform three experiments and find the following: the green end of bar A attracts the green end of bar B; the yellow end of bar A repels the green end of bar B; and the green end of bar B repels the yellow end of bar C. Which of these three results indicates that bar A must be a permanent magnet? Explain.
4. Consider the experiments described in Question 3. Would the green end of bar A attract, repel, or have no interaction with the green end of bar C? Explain.
5. You have three iron bars, each of which may or may not be a permanent magnet. Each rod is painted black on one end and white on the other. You perform three experiments and find the following: the black end of bar A repels the white end of bar B; the white end of bar A attracts the white end of bar C; and the white end of bar B attracts the white end of bar C. Is bar C a permanent magnet? Explain.
6. Consider the experiments described in Question 5. Would the black end of bar A attract, repel, or have no interaction with the white end of bar C? Explain.
7. If a bar magnet is broken into two pieces, how many magnetic poles are there?
8. We label magnetic poles such that like poles repel and unlike poles attract. Would it have been possible to use a labeling convention in which like poles attract and unlike poles repel? Why or why not?
9. How is the direction of the magnetic field defined at each point in space?
10. Your classmate claims: "To find the direction of the magnetic field at any point, you simply hold a compass in your hand and see what direction the needle points. Because the needle always points parallel to the ground, the magnetic field must also be parallel to the ground." This conclusion is inconsistent with the statement in the text that Earth's field points strongly downward in the United States. What is the flaw in your classmate's reasoning?
11. Magnetic field lines in a region are shown in the following figure. A freely rotating compass needle is released in the orientation shown. Will the needle rotate clockwise, counterclockwise, or not at all? Explain.



12. Oersted found that the magnetic field lines surrounding a current-carrying wire circle around and close on themselves. Because the field lines for a bar magnet also close on themselves, what does this imply about the direction of the field lines inside the magnet? Do the field lines point from north to south or from south to north?
13. When you look down the axis of a solenoid so that the current circulates in the counterclockwise direction, are you looking in the direction of the magnetic field or opposite of it? Explain.
14. How would the photograph in Figure 22-4 change if the current were in the opposite direction?

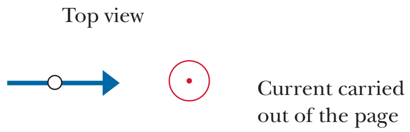
15. In the following figure, you are looking at the end of a long straight wire. The current in the wire is directed out of the page. At the point labeled A, is the direction of the magnetic field left, right, up, down, into the page, or out of the page? What is the direction at point B? Explain.

16. A long straight wire carries a current in the direction shown in the following figure. At the point labeled A, is the direction of the magnetic field left, right, up, down, into the page, or out of the page? What is the direction of the magnetic field at the point labeled B? Explain.



17. The following figure shows the top view of a compass next to a long straight wire. An arrow indicates the north pole of the compass. When a switch is closed, a current is established in the wire, and the compass needle initially rotates as indicated. Is the direction of the current in the wire into the page or out of the page? Explain.

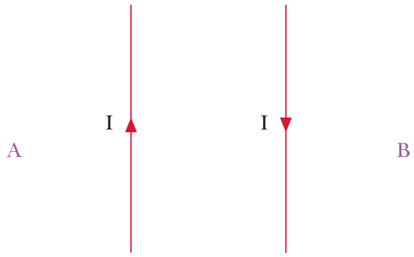
18. The following figure shows the top view of a compass next to a long straight wire. The north pole of the compass is indicated by an arrow. When a switch is closed, the current is directed out of the page as indicated. Will the compass needle initially rotate clockwise or counterclockwise? Explain.



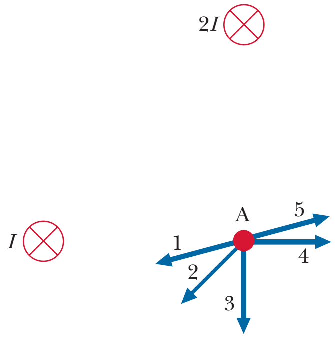
19. In electric circuits and telephone lines, two wires carrying currents in opposite directions are twisted together. How does this reduce the magnetic fields surrounding the wires?

20. Can you use a compass to locate the electric wires located inside the walls of your house? Explain.

21. Two long straight wires carry identical currents in opposite directions, as shown in the following figure. At the points labeled A and B, is the direction of the magnetic field left, right, up, down, into the page, or out of the page? Explain.



22. Two wires carry current into the page, as shown in the following figure. One wire carries a current I, and the other carries a current $2I$. Which of the arrows best represents the direction of the magnetic field at point A, which is the same distance from both wires? Explain.



23. In Melville's *Moby Dick*, Captain Ahab regained the confidence of his crew when he fixed the compass that had been damaged by an electrical storm so that it pointed north. How might he have done this?

24. Lodestone (magnetite) is an igneous rock, one that forms from molten material. How do you suppose it became magnetized?

25. What do you expect will happen to the strength of the magnetism of a bar magnet that is dropped on a hard floor?

26. Would you expect the head of a steel hammer to be magnetized? Explain.

27. Two long straight wires carry currents in opposite directions. The current in wire 1 is twice the current in wire 2. If the net magnetic force per unit length on wire 1 is 4 newtons per meter, what is the net magnetic force per unit length on wire 2?

28. Three long straight wires each carry identical current in the directions shown in the following figure. What is the direction of the net force on the wire at the right?