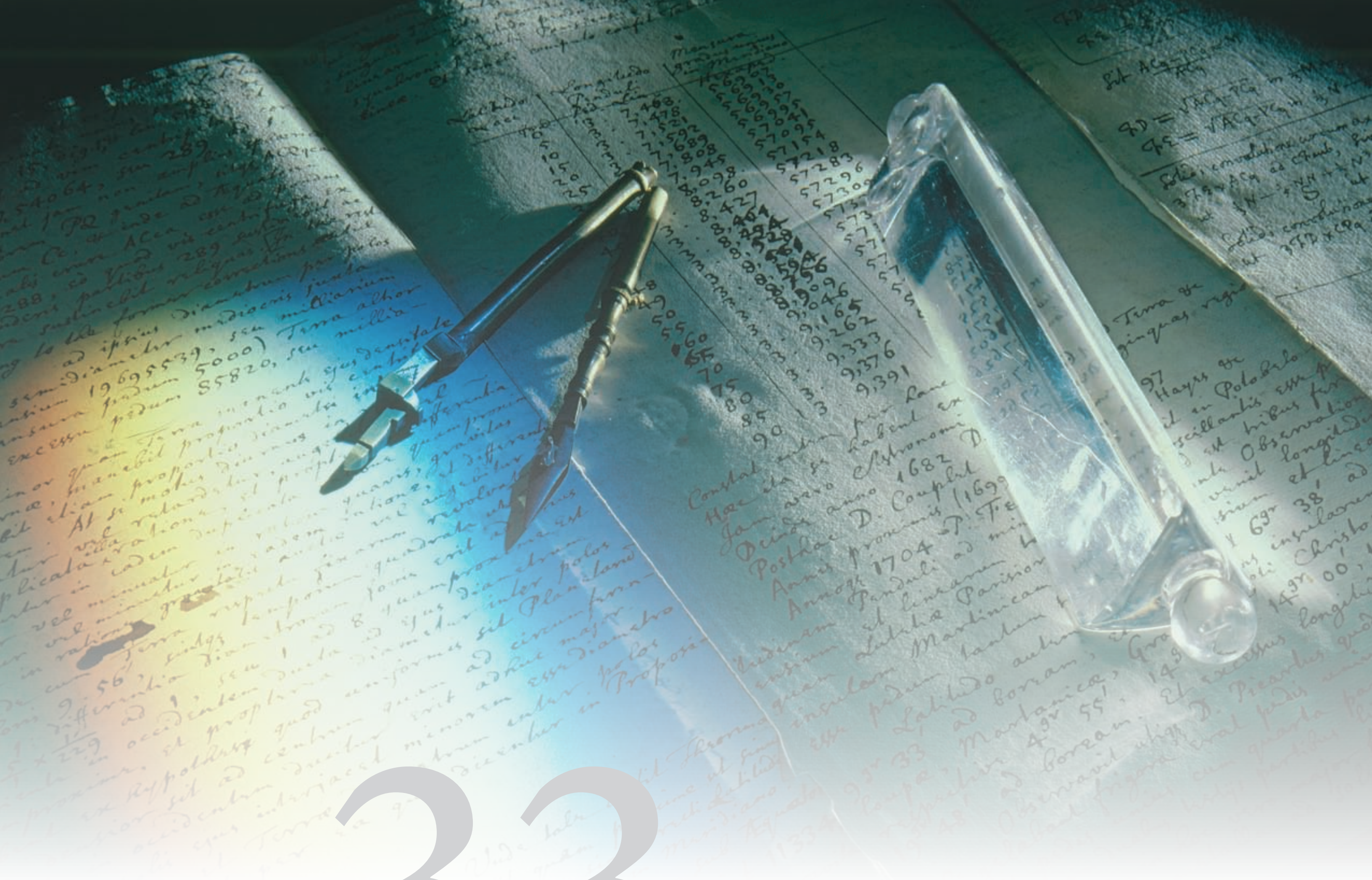


33 Ray Optics

CONCEPTS

QUANTITATIVE TOOLS

You can read these words because this page reflects light toward you; your eyes intercept some of the reflected light, and the lenses of your eyes redirect it, forming an image of the page on the retina. Where does the light reflected from the page come from? Our primary source of light during the day is the Sun, and our secondary source is the brightness of the sky. Indoors and at night, our light sources are flames in candles, white-hot filaments in light bulbs, and glowing gases in fluorescent bulbs. The light from all these sources comes from the accelerated motion of electrons as this motion produces electromagnetic waves.

In Chapter 30 we studied the propagating electric and magnetic fields that constitute electromagnetic waves, and we learned that a narrow frequency range of these waves corresponds to what we know as visible light. In this chapter we continue to study light, particularly its propagation and its interactions with materials. We shall not consider the electric and magnetic fields individually, but instead think of the behavior of rays of light. Such behavior, which is called *ray optics,* was understood long before it was known that light is an electromagnetic wave.

33.1 Rays

If you pierce a small hole in a piece of cardboard and then hold the cardboard between a lamp and a screen, the position where the light transmitted through the hole strikes the screen lies on a straight line connecting the lamp and the hole (**Figure 33.1**). This observation suggests that we can think of a light source as made up of many straight beams that spread out in three dimensions from the source. Each beam travels in a straight line until it interacts with an object. That interaction changes the beam's direction of travel.

We can represent the propagation of light by drawing **rays**:

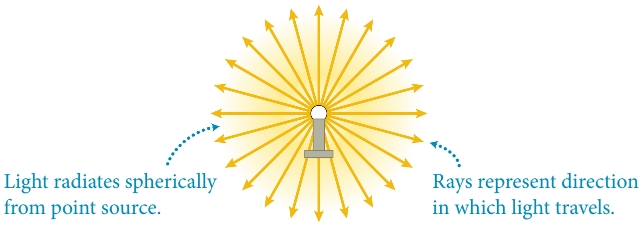
> **A ray is a line that represents the direction in which light travels. A beam of light with a very small cross-sectional area approximately corresponds to a ray.**

In order to see an object, our eyes form an image by collecting light that comes from the object. If the object is a light source, we see it by the light it emits. We can also see an object that is not a light source because such an object interacts with light that comes from a light source. The light is then redirected toward our eyes by means of this interaction.

When you stand outside on a sunny day, some of the rays from the Sun are blocked by your body while others travel in straight lines to the ground around you. You cast a shadow—a region on the ground that is darker than its surroundings because the Sun's rays that are blocked by your body do not strike this region. (The shadow region is not completely dark because it is still illuminated by light from the sky and by sunlight reflected from nearby objects.)

Figure 33.2 illustrates how rays can be used to represent the directions of light beams emanating from a light source. Just as with field line diagrams, we draw only a few rays to represent all the rays that could possibly be drawn; a ray could be drawn along any line radially outward from the source. Although most sources of light—the Sun, a flame, a light bulb—are extended, when the distance to the source is much greater than the extent of the source, we can treat that source as a *point source* (See Section 17.1). That is, we can treat the source as if all the light were emitted from a single point in space. In the first part of this chapter, we develop a feel for which rays to draw in a given situation.

Figure 33.2 Rays emanating from a source of light.



33.1 Suppose a second bulb is added to the left of the one in Figure 33.1, as illustrated in **Figure 33.3**. What happens to (*a*) the brightness of the spot created on the screen by the first bulb and (*b*) the brightness at locations close to the vertical edges of the original shadow on the screen (points P, Q, R, and S)?

Figure 33.1 A light beam that is not disturbed travels in a straight line.

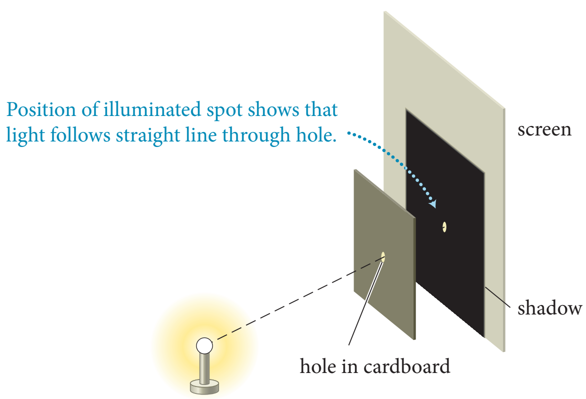
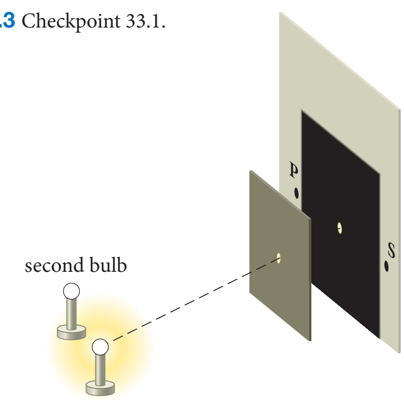


Figure 33.3 Checkpoint 33.1.

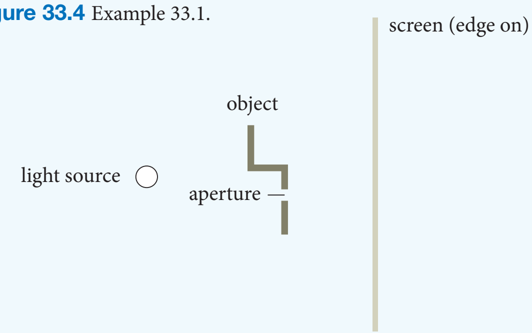


CONCEPTS

Example 33.1 Light and shadow

An object that has a small aperture is placed between a light source and a screen, as shown in **Figure 33.4**. Which parts of the screen are in the shadow?

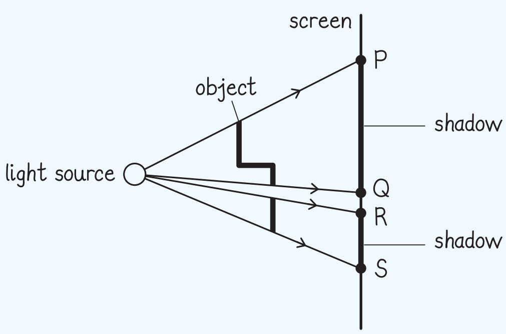
Figure 33.4 Example 33.1.



1 GETTING STARTED The rays emitted by the source radiate outward in all directions following straight paths. The shadow is cast because the object prevents some of the rays from reaching the screen (except for the rays that make it through the aperture).

2 DEVISE PLAN To locate the edges of the shadow, I draw straight lines from the source to the edges of the object (including the edges of the aperture) and extend these rays to the screen (**Figure 33.5**).

Figure 33.5



3 EXECUTE PLAN The top and bottom edges of the shadow correspond to the highest and lowest screen locations (P and S) reached by light rays that are not blocked by the object. The gap in the shadow between locations Q and R corresponds to light rays that pass through the aperture, which means that this region of the screen is not in shadow. ✔

4 EVALUATE RESULT A shadow that is taller than the object makes sense. Because the light rays from the source emanate in all directions, most of them reach the screen at angles other than 90°. This means that the distance from P to S must be greater than the object height. Indeed, I know from experience that the shadow cast by my hand gets larger as I move my hand closer to a lamp, increasing that angle.

33.2 Hold a piece of paper between your desk lamp (or any other source of light) and your desk or a wall. How does the sharpness of the edges of the shadow change as you move the paper closer to the bulb? Why does this happen?

33.2 Absorption, transmission, and reflection

Different materials interact differently with the light that strikes them, which is how you can visually distinguish wood from metal, fabric from skin, and a white piece of paper from a blue one. When light strikes an object, the light can be transmitted, absorbed, or reflected.

Transmitted light passes through a material. Objects that transmit light, such as a piece of glass, are said to be *transparent* (**Figure 33.6*a***). In *translucent* materials, such as frosted glass, light rays are *transmitted diffusely*—that is, they are redirected in random directions as they pass through, so that the transmitted light does not come from a definite direction (Figure 33.6*b*). Because translucent materials scatter light in this manner, we cannot see objects clearly through them.

Absorbed light enters a material but never exits again. Objects that absorb most of the light that strikes them, such as a piece of wood, are said to be *opaque*. When light strikes such materials, the energy carried by the light is converted to some other form (usually thermal energy) and the light propagation stops.

Reflected light is any light that is redirected away from the surface of the material (**Figure 33.7** on the next page). Smooth surfaces reflect light *specularly*—that is, each ray bounces off the surface in such a way that the angle between it and the normal to the surface doesn't change (Figure 33.7*a*). The angle between the incoming ray and the normal to the surface is called the **angle of incidence** θ_i; the angle between the outgoing ray and the normal is called the **angle of reflection** θ_r.

Figure 33.6 We see objects clearly through a sheet of clear glass but diffusely through frosted glass.

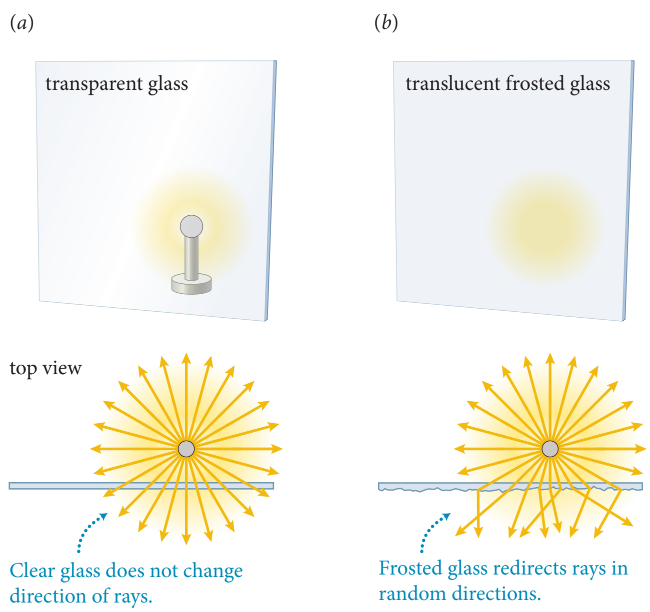
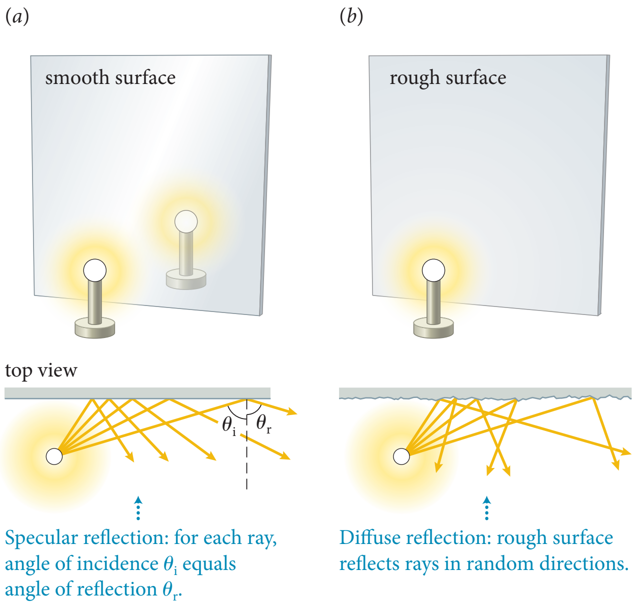


Figure 33.7 Light reflects specularly from a smooth surface, forming a mirror image. From a rough surface, it reflects diffusely (in random directions), so no image forms.



Empirically we find:

For a ray striking a smooth surface, the angle of reflection is equal to the angle of incidence, and both angles are in the same plane.

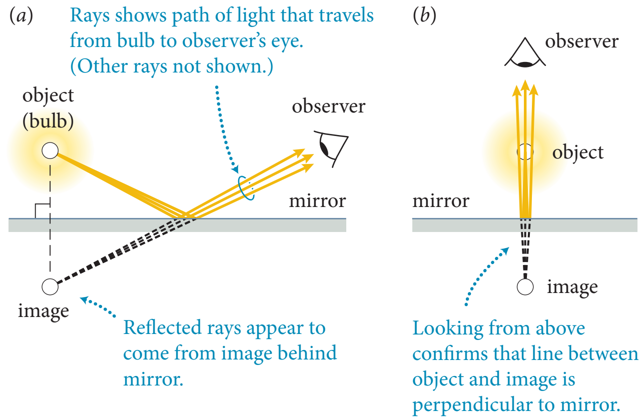
This **law of reflection** holds at smooth surfaces for any angle of incidence.

Surfaces that are not smooth reflect light in many directions (Figure 33.7*b*). For such *diffuse reflection,* each ray obeys the law of reflection, but the direction of the surface normal varies over the surface and so the angle of reflection also varies.

How smooth is smooth? If the height and separation of irregularities on the surface are small relative to the wavelength of the incident light, the surface acts like a smooth surface and most light is reflected specularly. For example, paper appears smooth to microwaves, which have wavelengths ranging from millimeters to meters, and therefore microwaves are reflected specularly from paper. Visible light, however, has wavelengths of hundreds of nanometers, and so paper reflects visible light diffusely.

Rays that come from an object and are reflected from a smooth surface form an **image,** an optically formed duplicate of the object (Figure 33.7*a*). **Figure 33.8*a*** shows the paths taken by light rays emitted by a light bulb placed in front of a mirror. A diagram like Figure 33.8*a* showing just a few selected rays is called a **ray diagram.** If we trace the reflected rays back to the point at which they appear to intersect, we see that point is behind the mirror. Consequently, the brain interprets the reflected rays as having come from that point, creating the illusion that the light bulb is behind the mirror. The directions of the rays that reach the eyes of the observer are the same as if they had come from an object located behind the mirror.

Figure 33.8 Diagrams showing the paths taken by light rays that are produced by a bulb and reflected by a mirror into an observer's eye. The reflected rays appear to come from behind the mirror, forming an image behind the mirror.



Note that the image is located on the line through the object and perpendicular to the mirror, because if we look along that line, the image lies behind the object (Figure 33.8*b*).

Rays that do not actually travel through the point from which they appear to come, like the rays in Figure 33.8*a,* are said to form a *virtual image*. A *real image* is formed when the rays actually do intersect at the location of the image. (Flat mirrors cannot form real images; we'll encounter real images when we discuss lenses in Section 33.4 and curved mirrors in Section 33.7.)

Example 33.2 How far behind the mirror?

If the light bulb in Figure 33.8*a* is 1.0 m in front of the mirror, how far behind the mirror is the image?

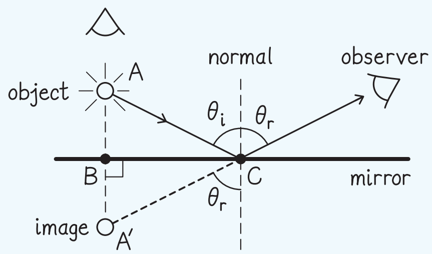
❶ GETTING STARTED The location of the image is the location from which the rays reflected by the mirror appear to come—that is, the point at which they intersect. From Figure 33.8 I know that because the rays intersect directly behind the bulb, a line that passes through the bulb and is normal to the mirror passes through the image.

❷ DEVISE PLAN I can obtain the distance of the image behind the mirror by considering one ray that travels from the bulb to the observer and then tracing that ray back through the mirror to its intersection with the line that is perpendicular to the mirror and passes through the bulb.

❸ EXECUTE PLAN I begin by drawing a ray that travels from the bulb to the mirror and is reflected to the location of the observer (**Figure 33.9**). In my drawing, A denotes the bulb location, B denotes the point where the line connecting the bulb and its image intersects the mirror, and C denotes the point at which the ray that is reflected to the observer hits the mirror. According to the law of reflection, $\theta_r = \theta_i$.

CONCEPTS

Figure 33.9



I now extend the reflected ray to behind the mirror (dashed line). I know that the image must lie somewhere along that dashed line and must also lie on the line that passes through the object and is perpendicular to the mirror. The image must therefore lie at the intersection of this line and the dashed ray extension; I denote that intersection point by A′.

To determine the distance BA′, which is how far behind the mirror the image is, I note that angle A′CB is equal to $90° - \theta_r$ and angle ACB is equal to $90° - \theta_i$. Because $\theta_i = \theta_r$, angles A′CB and ACB are equal. Therefore triangles ABC and A′BC are congruent, and AB = BA′. That is, the image appears at the same distance behind the mirror as the object is in front of it: 1.0 m behind the mirror. ✔

❹ **EVALUATE RESULT** My result makes sense because I know from experience that as I walk toward a mirror, my image also approaches it.

33.3 If the observer in Figure 33.8 moves to a different position, does the location of the image change?

The colors of visible light we see correspond to different frequencies of electromagnetic waves. Red corresponds to the lowest frequency of the visible spectrum. As the frequency increases, the color changes to orange, yellow, green, blue, indigo, and finally violet (the highest frequency of the visible spectrum). The range of visible frequencies is quite small relative to the range of the complete electromagnetic spectrum, as **Figure 33.10** shows. Frequencies

Figure 33.10 Visible light makes up only a small part of the electromagnetic spectrum.

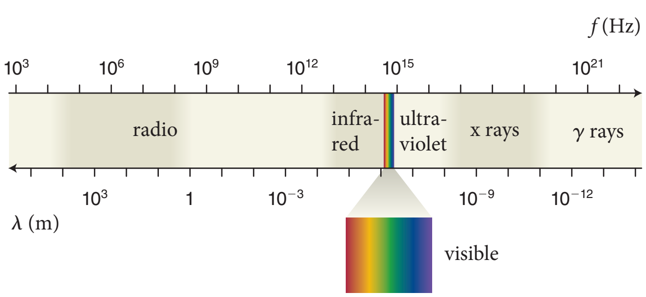
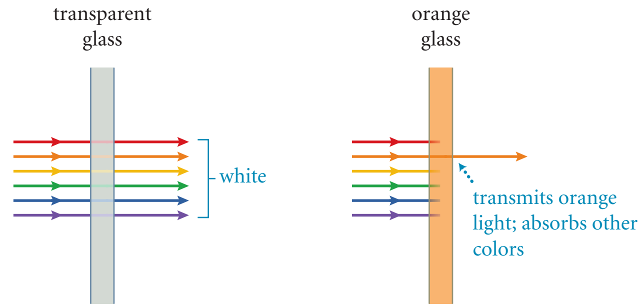


Figure 33.11 All colors of light pass through colorless glass (shown light blue for illustration purposes); orange glass transmits orange light and absorbs all colors of light except orange.

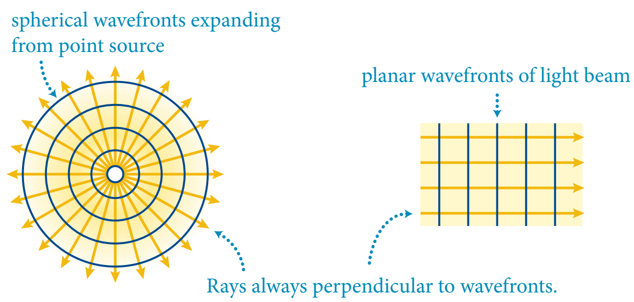


lower than the visible correspond to infrared radiation, and higher frequencies to ultraviolet. When a light source produces all the frequencies of the visible spectrum at roughly the same intensities, the emitted light appears white.

Different colors of light interact differently with different objects, affecting the color we perceive the object as being. Colorless materials, like a piece of ordinary window glass, transmit all colors of the visible spectrum. A piece of orange glass, on the other hand, transmits only the orange part of the visible spectrum. All other colors are absorbed in the glass (**Figure 33.11**). A red apple absorbs all colors of the visible spectrum except red, which is redirected to our eyes. Grass absorbs all colors except green, which is diffusively reflected at its surface.

Because light is a wave phenomenon, it is sometimes useful to represent the propagation of light with wavefronts, which we introduced in Section 17.1. Wavefronts are drawn perpendicular to the direction of propagation of the wave.* Because light rays point along the direction of propagation of the light, light wavefronts are perpendicular to light rays. **Figure 33.12*a*** shows the spherical

Figure 33.12 A point source of light produces spherical wavefronts; a beam of light contains planar wavefronts.



* For mechanical waves, the wavefronts are drawn at the locations of the wave crests, spaced by the wavelength of the wave. Because the wavelength of light is very short, the wavefronts of light cannot be represented to scale.

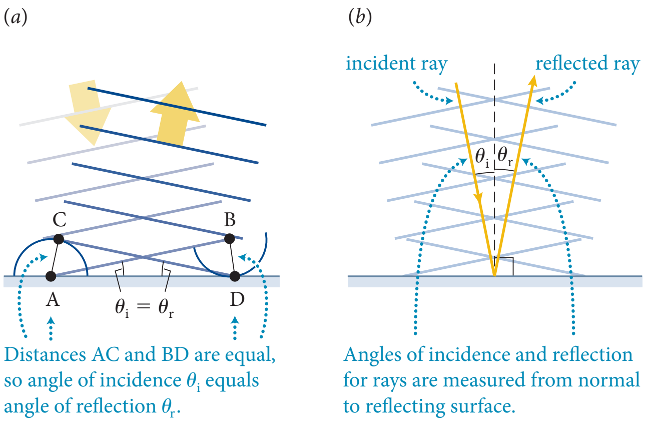


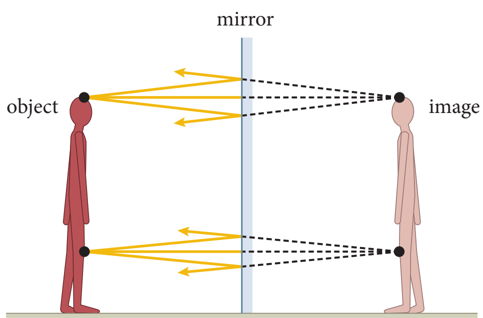
Figure 33.13 (*a*) The reflection of wavefronts from a smooth surface explains the law of reflection. (*b*) The corresponding rays and their angles of incidence θ_i and reflection θ_r.

wavefronts for light coming from a point source, and Figure 33.12*b* shows the straight-line rays and wavefronts corresponding to a planar electromagnetic wave. Note that a planar wave is represented with rays that are parallel to one another because all the wavefronts are parallel to one another.

By looking at how wavefronts behave, we can understand the law of reflection. When a light ray strikes a smooth surface at an incidence angle $\theta_i \neq 0$ (**Figure 33.13**), the left end of the first wavefront to reach the surface gets there, at A, before the right end does. In the time interval it takes the right end to reach the surface at D, the left end has traveled back from the surface to C. The distance traveled by the right end toward the surface, BD, is the same as that traveled by the left end away from the surface, AC, so the angles BAD and CDA must be equal. The angle of incidence θ_i equals angle BAD. Likewise, the angle of reflection θ_r equals angle CDA. So $\theta_i = \theta_r$.

So far we have treated the object (and consequently the image) as a single point. **Figure 33.14** shows how images are formed of extended objects. Each point on the object reflects (or emits) light rays, and the reflections of these rays

Figure 33.14 Paths taken by rays from more than one point on the object, showing how extended images form.



appear to come from a corresponding point on the image. A flat mirror thus produces behind the mirror an exact mirror image of the entire extended object.

33.4 In order for the person in Figure 33.14 to see a complete image of himself, does the mirror need to be as tall as he is?

33.3 Refraction and dispersion

As we found in Chapter 30, light propagates with speed $c_0 = 3 \times 10^8$ m/s in vacuum. In air, the speed of light is almost the same as that in vacuum. In a solid or liquid medium, however, light propagates at a speed c that is generally less than c_0.* In glass, for example, visible light propagates at two-thirds of the speed of light in vacuum (see Example 30.8).

How does this change in speed affect the propagation of an electromagnetic wave? Recall from Chapter 16 that harmonic waves are characterized by both a wavelength λ and a frequency f and that the product of the wavelength and frequency equals the wave's speed of propagation (Eq. 16.10). The frequency of the wave must remain the same because the oscillation frequency of the electromagnetic field that makes up the wave is determined by the acceleration of charged particles at the wave's source. The acceleration of the source does not alter when the wave travels from one medium to another, and thus the frequency of the traveling wave also cannot change.

33.5 In vacuum, a particular light wave has a wavelength of 400 nm. It then travels into a piece of glass, where its speed decreases to two-thirds of its vacuum speed. What is the distance between the wavefronts in the glass?

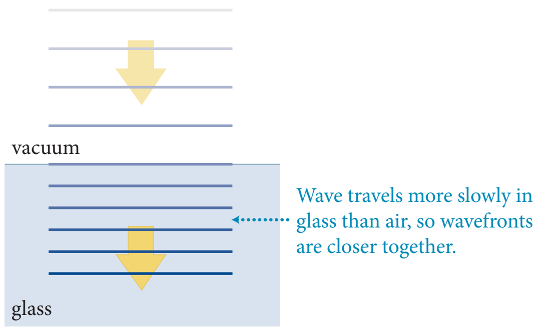
As we found in Checkpoint 33.5, when rays of light pass through the interface between vacuum and a transparent material, the wavefronts inside the material are more closely spaced than they are in vacuum, due to the lower speed of the wavefronts. **Figure 33.15** illustrates this effect for wavefronts incident normal to the surface of the material.

What if the wavefronts strike the transparent material at an angle? In such a case, one end of the wavefront arrives at the surface before the other (**Figure 33.16**). Once the end that reaches the surface first (this happens to be the left end in Figure 33.16) enters the material, it travels at the lower speed while the other end of the wavefront (the right end in our example) continues to travel at the

*For visible light, c is less than c_0. For x rays, c can be greater than c_0.

CONCEPTS

Figure 33.15 Wavefronts for a ray traveling from vacuum into transparent glass in a direction normal to the glass surface.



vacuum speed. This means the distance AC traveled by the left end is less than the distance BD traveled by the right end during the same time interval (Figure 33.16*a*). Consequently, the wavefront CD in the material is no longer parallel to the wavefront AB that has not yet entered the material.

The direction of the ray associated with these wavefronts therefore changes on entering the material. As shown in Figure 33.16*b*, the angle of incidence θ_1, between the ray in vacuum and the normal to the interface between the two materials, is greater than the angle θ_2, between the ray in the material and the normal to that interface. This bending of light as it moves from one material into another is referred to as **refraction,** and the angle θ_2 between the refracted ray and the normal to the interface between the materials is called the **angle of refraction.** Whenever light is refracted, the angle between the ray and the normal is always greater in the material in which the light travels faster, so:

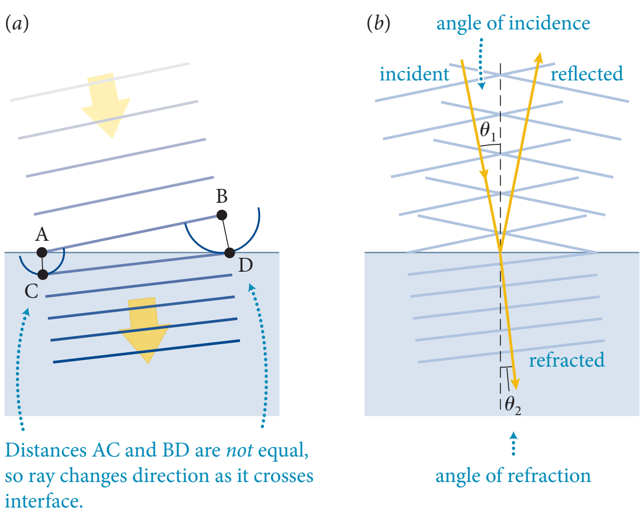
> **When a light ray travels from one material into a second material where light travels more slowly, the ray bends toward the normal to the interface between the materials.**

Generally, the speed of light decreases as the mass density of the material increases. Note also that, as shown in Figure 33.16*b*, both reflection and refraction take place at the interface between two media (or between vacuum and a medium).

The amount of bending depends on the angle of incidence and on the relative speeds in the two media. There is no bending for normal incidence (as we saw in Figure 33.15); the bending is less near normal incidence and becomes more pronounced as the angle of incidence increases. In Section 33.5, we'll work out a quantitative expression relating angles θ_1 and θ_2.

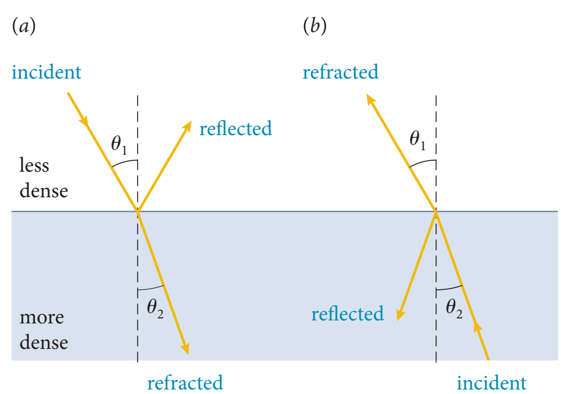
33.6 Suppose the ray in Figure 33.16 travels in the opposite direction—that is, from the denser medium to the less dense medium. If the angle of incidence is now θ_2, how does the angle of refraction compare with θ_1?

Figure 33.16 (*a*) Refraction is explained by the behavior of wavefronts that cross at an angle into a transparent medium in which they travel more slowly. (*b*) Incident, reflected, and refracted rays, showing the angles of incidence θ_1 and refraction θ_2 (measured from the normal to the surface).



Because the relationship between the angles of incidence and refraction is completely determined by the speed of the wavefronts in the two media, the angles do not depend on which is the incident ray and which is the refracted ray. As shown in **Figure 33.17**, θ_1 and θ_2 have the same values whether θ_1 is the angle of incidence (Figure 33.17*a*) or the angle of refraction (Figure 33.17*b*). Keep in mind, however, that the reflected ray is always on the same side of the interface as the incident ray, and so the angle of reflection is *not* the same in Figure 33.17*a* and Figure 33.17*b*.

Figure 33.17 Because refraction is caused by the relative speeds of the wavefronts in two media, the angles of incidence and refraction do not depend on which is the incident ray and which the refracted ray.



CONCEPTS

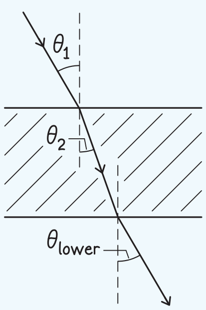
Example 33.3 Crossing a slab

Consider a light ray incident on a parallel-sided slab of glass surrounded by air, as shown in Figure 33.17*a*. The ray travels all the way through the slab and emerges into air on the other side. In what direction does the ray emerge?

❶ **GETTING STARTED** This problem involves two successive encounters of a light ray with interfaces between glass and air. At each interface, the ray is refracted. I need to determine the direction of the ray (its angle to the normal to the slab) after it crosses the lower interface of the slab represented in Figure 33.17*a*.

❷ **DEVISE PLAN** Because I want to know the direction of the emerging ray, I construct an appropriate ray diagram. Figure 33.17*a* shows the direction of the ray inside the slab. I extend the ray through the slab to the lower interface (**Figure 33.18**) and draw the emerging ray, labeling its angle to the normal θ_{lower}. To determine this angle, I need to consider the refraction that occurs at the lower interface.

Figure 33.18



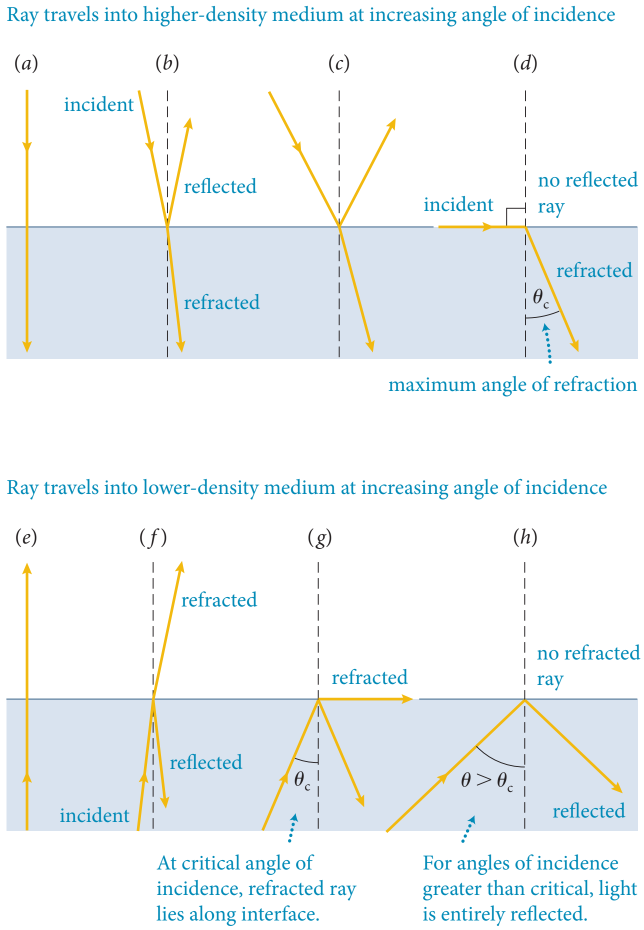
❸ **EXECUTE PLAN** Because the two interfaces are parallel, their normals are also parallel, and so the angle at which the ray is incident on the lower interface is equal to the angle θ_2 at which it is refracted at the upper interface. I saw in Figure 33.17*b* that if the angle between the ray and the normal in the slab is θ_2, it doesn't matter whether the ray in the slab is the incident ray or the refracted ray; either way, the angle between the ray in the air and the normal is θ_1. Therefore $\theta_{\text{lower}} = \theta_1$. ✔

❹ **EVALUATE RESULT** Crossing the lower interface from glass into air, the ray bends away from the normal, as it should because glass is denser than air. With $\theta_{\text{lower}} = \theta_1$, in fact, the ray emerges parallel to the original direction it had before entering the slab. This makes sense because the two air-glass interfaces are parallel. (Note, though, that the ray is shifted sideways by a small distance relative to its original path.)

33.7 When the ray reflected from the bottom surface in Figure 33.18 reemerges from the top surface, how does the angle it makes with the normal compare with θ_1?

What range of refraction angles is possible? To answer this question, let's first consider the case where the ray travels from a low-density medium into a denser medium (**Figure 33.19*a–d***). Because the angle of incidence is always greater than the angle of refraction in this situation, as the angle of incidence approaches 90°, the angle of refraction remains less than 90° (Figure 33.19*d*). The full 90° range of incidence angles gives a range of refraction angles that is less than 90°.

Figure 33.19 The range of possible refraction angles for a ray crossing into a medium of either higher or lower density.

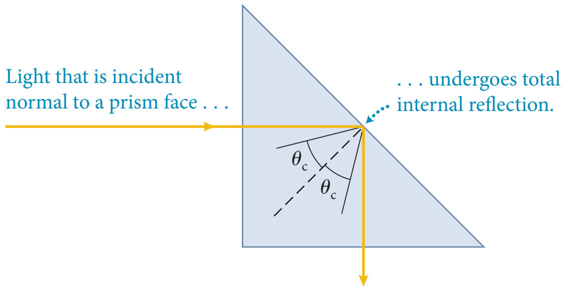


Next consider the case where the ray travels from a high-density medium into a lower-density medium (Figure 34.18*e–g*). The angle of incidence is now less than the angle of refraction. Consequently, as the angle of incidence increases, it reaches a value for which the refracted ray emerges along the interface (Figure 33.19*g*). This angle of incidence is called the **critical angle** θ_c and is equal to the angle of refraction shown in Figure 33.19*d*. For angles of incidence greater than θ_c, the angle of refraction would have to be greater than 90°, which is impossible. Therefore, no light is refracted. Instead, all the light is *reflected* back into the higher-density medium (Figure 33.19*h*), a phenomenon called **total internal reflection.**

Several optical devices make use of total internal reflection to direct light. The glass prism shown in **Figure 33.20**

CONCEPTS

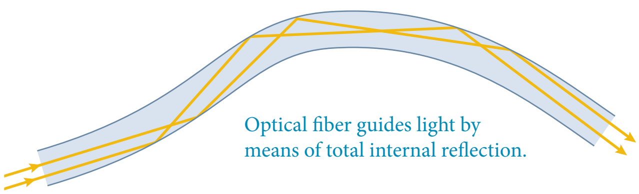
Figure 33.20 A prism can act as a perfect mirror by means of total internal reflection.



reflects light just as a mirror would. A light ray enters the prism's front surface at normal incidence. Because the back surface is slanted relative to the front surface, the angle at which the ray hits the back surface is less than 90°. This back-surface angle of incidence is greater than the critical angle for the glass, however, and so the light is totally reflected from the back surface. Such prisms are actually better mirrors than most regular mirrors; they reflect very close to 100% of the incident light, whereas mirrors are less reflective due to imperfections in the reflecting surface.

Optical fibers also guide light by means of total internal reflection. An optical fiber is a long, thin fiber made of a transparent material such as glass. If light shines into one end of the fiber at an angle greater than the critical angle, the light travels along the fiber through repeated total internal reflections, and essentially all of the light that entered the fiber emerges at the other end (**Figure 33.21**). Because very little light is lost as the light travels, only a faint glow comes from the rest of the fiber.

Figure 33.21 How optical fibers work.



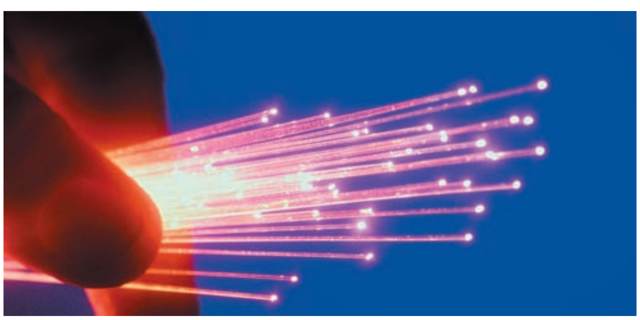
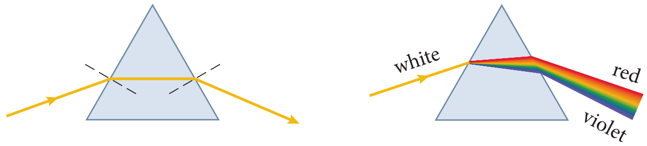


Figure 33.22 The phenomenon of dispersion, which results from the fact that the speed of light in a given medium (and hence the angle of refraction) depends slightly on the frequency of the light.

(*a*) Prism refracts light of single frequency

(*b*) Dispersion: different colors have different angles of refraction



(*c*) Rainbows result from dispersion of sunlight by raindrops

Because the speed of light in any given medium depends slightly on the frequency of the light, the angle of refraction also depends on frequency. This phenomenon is called **dispersion** because it causes rays of different colors to separate—to be *dispersed*—when refracted. Prisms like the one shown in **Figure 33.22** are designed to separate colors by the frequency dependence of the angle of refraction. In most media, high-frequency light travels more slowly than low-frequency light, and so high-frequency light bends more strongly toward the normal. The lowest frequency of visible light is red and the highest is violet, which means violet light bends the most, as the rainbow of Figure 33.22*c* shows.

Both rainbows and the brilliance of gems result from a combination of total internal reflection and dispersion. In a rainbow (Figure 33.22*c*), the combination of total internal reflection and dispersion means that we see different colors coming from water droplets at different viewing angles. Gems are cut with many internal surfaces from which total internal reflection takes place. Because the light is also dispersed, colorless gems such as diamonds shine with many distinct colors.

33.8 Because of dispersion, the critical angle for total internal reflection in a given medium varies with frequency. Is the critical angle for a violet ray greater or less than that for a red ray?

CONCEPTS

Fermat's principle

Figure 33.23 shows four ways in which a light ray can travel between two locations A and B: directly, reflected from a mirror, refracted through a glass slab, and refracted through a prism.* You could say that in each case the ray reaches B because it is aimed properly from A. However, an entirely different way of looking at the path followed by the light was suggested by the French mathematician Pierre de Fermat (1601–1665) in a formulation today known as **Fermat's principle:**

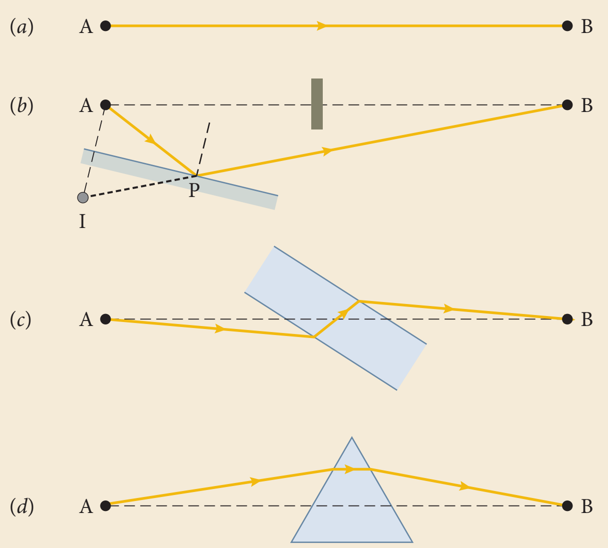
> **The path taken by a light ray between two locations is the path for which the time interval needed to travel between those locations is a minimum.**

This principle may seem to imply that light always travels in a straight line. However, the *quickest path* between two locations is not necessarily the *shortest distance* when the speed of light differs in different regions.

Let's consider the four paths in Figure 33.23 using Fermat's principle. In Figure 33.23*a*, the ray does follow a straight path because the medium in which the ray travels is uniform. As a result, the quickest path is indeed the shortest distance: a straight line from A to B.

In Figure 33.23*b*, the fact that the straight-line path from A to B is blocked means that the ray must reflect somewhere off the mirror in order to travel from A to B. The path shown, which satisfies the law of reflection, is the shortest distance from A to B involving reflection from the mirror. Because the distance from A to the reflection location P equals the distance between the image location I and P, the straight line IB is equal in length to the path traveled by the ray from A to B. Moving the reflection location to either side of P, so that the angle of incidence does not equal the angle of reflection, increases the length of the path. Thus, Fermat's principle implies the law of reflection.

Figure 33.23 Ray diagrams illustrating the *quickest path* for a light ray traveling from A to B for four situations.



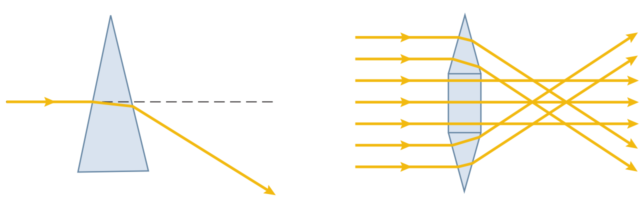
When the ray must travel through some air and some glass, as in Figure 33.22*c* and *d*, the quickest path is not a straight line because the ray's speed in the glass is only two-thirds of its speed in air. To minimize the time interval needed to travel from A to B, the ray bends on entering and exiting the glass. Such a bent path reduces the distance traveled through the glass without increasing the distance traveled in air so much that it offsets the amount of time saved. In Example 33.7, we shall see that calculating the bending angles with Fermat's principle gives the same result as with ray optics.

*Note that what distinguishes a glass slab from a glass prism is the way I use the terms: In a slab, the two opposite surfaces are parallel to each other; in a prism, they are not.

33.4 Forming images

As shown in **Figure 33.24**, by combining two prisms and a glass slab we can create a device that steers parallel light rays toward each other. The rays through the center of the device pass straight through, those through the top prism are refracted downward, and those through the bottom prism are refracted upward.

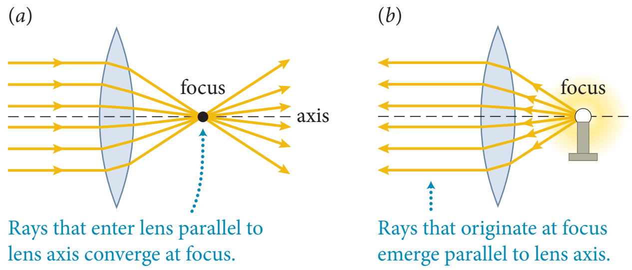
Figure 33.24 A device that redirects parallel light rays toward each other.



To bring all parallel incident rays to a single point, a structure called a **lens** is used. A lens is designed with curved surfaces so that the refraction of incident rays increases gradually as we move away from the center. To accomplish this, lenses are typically made with spherical surfaces, which are easy to manufacture.

Figure 33.25*a* shows a lens with *convex* spherical surfaces, where a *convex surface* is defined as one that curves like the outside of a sphere. Rays parallel to the lens *axis*—a line perpendicular to the lens through its center—converge through such a lens onto a single point called either the **focus** or the **focal point.** A lens with convex surfaces is therefore called a *converging lens*. The distance from the center of the lens to the focus is called the **focal length** f.

Figure 33.25 Converging lens with convex spherical surfaces.



What if we place a light source at the focus of a lens, as in Figure 33.25*b*? As we saw in the preceding section, the path followed by a light ray is unaffected by reversing the direction of propagation of the ray, as long as the ray is not absorbed by the medium. So, if we place a light source at the focus of a lens, a beam of parallel rays emerges on the other side of the lens.

> **33.9** Sketch the wavefronts corresponding to all the rays in Figure 33.25*a*, both the parallel ones on the left and the refracted ones on the right.

We can reverse the direction of the rays through the lens of Figure 33.25, so that parallel rays enter from the right (**Figure 33.26**). The rays converge again at a focus that is the same distance f from the center of the lens. Thus, every lens has two foci, one on either side of the lens at the same distance f from it.

If we tilt the parallel rays a bit relative to the lens axis (**Figure 33.27**), the rays still converge at a distance f from the center, but the focus is no longer on the axis. Provided the parallel rays make only a small angle with the lens axis, they all converge at a point on a plane—called the *focal plane*—that is perpendicular to the axis a distance f from the lens. Rays that run near the lens axis—either parallel to it or at a small angle—are said to be *paraxial*.

Now that we know how parallel rays and rays that emanate from the focus of a lens are refracted by the lens, we can determine where images are formed. The image of a point on an object is formed where all the light rays emanating from that point converge. (These light rays then diverge from the location of the image; when they enter the eye, the brain interprets them as having come from the location of the image.) An image of the entire object is made up of the images of all the individual points on the object.

Figure 33.26 A lens has two equivalent foci, one on each side, at equal distances from the center of the lens.

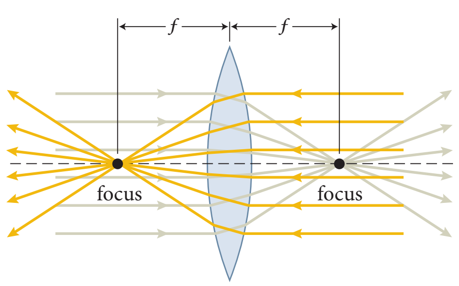
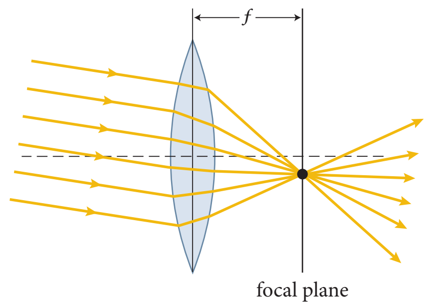


Figure 33.27 If the rays strike the lens at an angle, they no longer converge on the focus, but they still converge on a focal plane at the focal distance f.

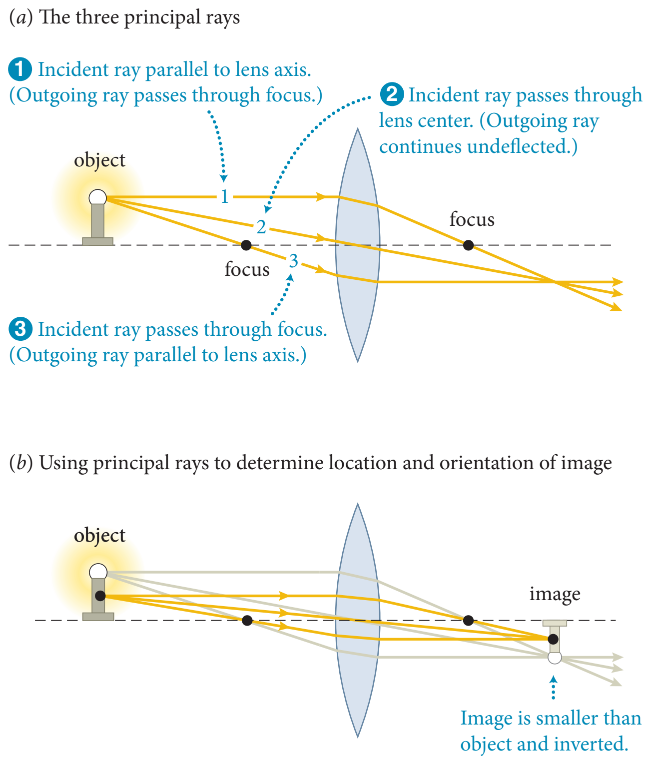


To determine where the rays emanating from a point on an object converge, we don't need to draw all the rays. Instead, we draw three special ones, called **principal rays,** and see where they converge:

1. a ray that travels parallel to the lens axis before entering the lens,
2. a ray that passes through the center of the lens, and
3. a ray that passes through the focus that is on the same side of the lens as the object.

These three principal rays are shown in **Figure 33.28*a*** for the case where the object lies beyond the focus of the

Figure 33.28 The three principal rays for a spherical lens can be used to determine the location, size, and orientation of the image for a given object.



Procedure: Simplified ray diagrams for lenses

To determine the location and orientation of an image formed by a lens, follow this procedure:

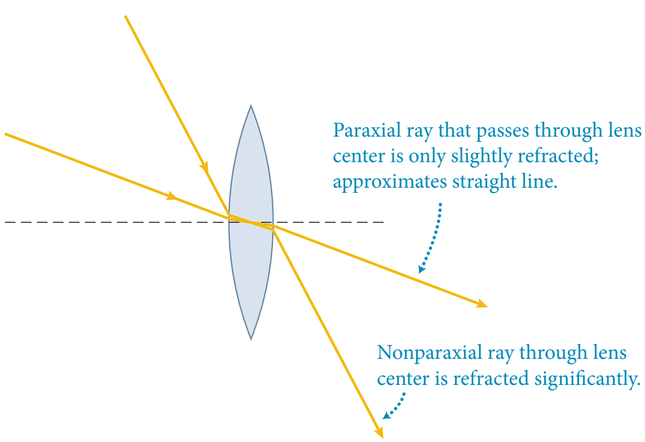
1. Draw a horizontal line representing the lens axis (the line perpendicular to the lens through its center). In the center of the diagram, draw a vertical line representing the lens. Put a + above the line to represent a converging lens or a − to represent a diverging lens.
2. Put two dots on the axis on either side of the lens to represent the foci of the lens. The dots should be equidistant from the lens.
3. Represent the object by drawing an upward-pointing arrow from the axis at the appropriate relative distance from the lens. For example, if the distance from the object to the lens is twice the focal length of the lens, put the arrow twice as far from the lens as the dot you drew in step 2. The top of the arrow should be at about half the height of the lens.
4. From the top of the arrow representing the object draw two or three of the three *principal rays* listed in the Procedure box "Principal rays for lenses" on page 888.
5. The top of the image is at the point where the rays *that exit the lens* intersect (if they diverge, trace them backward to determine the point of intersection). If the intersection is on the opposite side of the lens from the object, the image is real; if it is on the same side, the image is virtual. Draw an arrow pointing from the axis to the intersection to represent the image (use a dashed arrow for a virtual image).

In general it is sufficient to draw two principal rays, but depending on the situation, some rays may be easier to draw than others. You can also use a third ray to verify that it, too, goes through the intersection. (If it doesn't, you have made a mistake.)

lens. We already know how rays 1 and 3 travel. Ray 1 passes through the focus on the other side of the lens, and ray 3 emerges from the lens parallel to the axis. As for ray 2, as long as it is paraxial, it passes straight through with negligible refraction (**Figure 33.29**). (Nonparaxial rays are shifted significantly; our treatment of lenses in this chapter is restricted to images formed by paraxial rays.) Ray 2 can therefore be drawn as traveling in a straight line through the center of the lens.

To determine the location and orientation of the image of an extended object, we work out the locations of the images of several points on the object. Figure 33.28*b* shows where the rays converge for two points on the object, and the extended image that can be inferred from these points. The image is smaller than the object and inverted. See the Procedure box "Simplified ray diagrams for lenses" on this page for a general description of how to draw simplified ray diagrams for lenses.

Figure 33.29 Paraxial and nonparaxial rays. For a paraxial ray or a lens that is not too thick, the refraction is so slight that we can consider this ray to be one uninterrupted straight line.



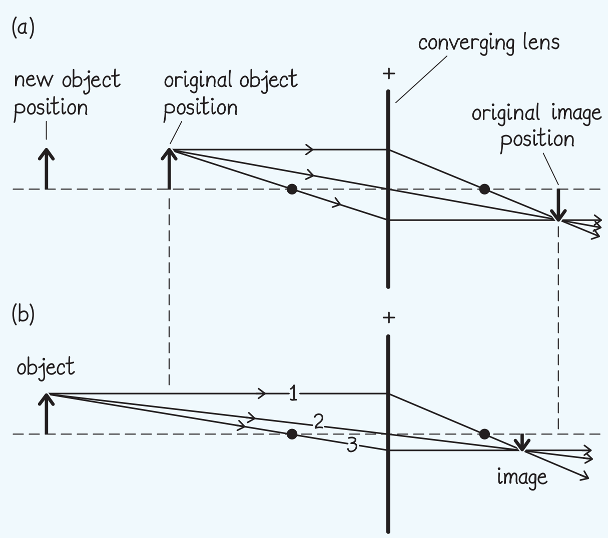
33.10 Do you need to draw all three principal rays to determine the location of an image?

Example 33.4 Where is the image?

Consider the light bulb that is the object in Figure 33.28. If you move the bulb to the left, does the image shift left, shift right, or stay in the same place?

❶ GETTING STARTED Using the procedure for drawing simplified ray diagrams, I represent the lens, object, and image of Figure 33.28, and draw the bulb at its new position (**Figure 33.30*a***).

Figure 33.30



❷ **DEVISE PLAN** I can determine which way the image shifts by drawing the principal rays for the light bulb in its new position.

❸ **EXECUTE PLAN** With the rays drawn (Figure 33.30*b*), I see that the image shifts left. ✔

❹ **EVALUATE RESULT** As I move the light bulb to the left, principal ray 1 remains the same, but principal ray 2 makes less of an angle with the lens axis than before. Consequently, the location at which these two rays intersect is closer to the lens than before, shifting the image to the left, as I concluded from the diagram.

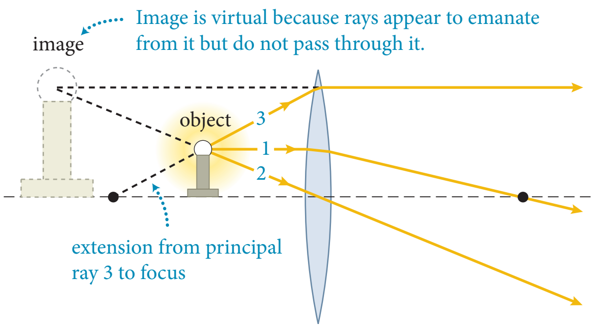
Notice in Figure 33.30 that the image gets smaller as the object moves farther away from the lens. Conversely, moving the object closer to the lens makes the image larger. Indeed, one of the most common uses of lenses is to enlarge images. If the object is placed at the focus, no image forms because the rays all emerge from the lens parallel to each other—in other words, the rays do not converge.

Consider placing an object *between* the lens and its focus, as in **Figure 33.31**. In this configuration, principal ray 3 does not pass through the focus. Instead, it lies on the line that joins the focus to the point of interest on the object.

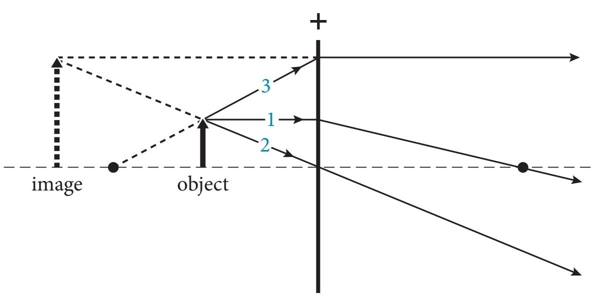
The image formed in the configuration of Figure 33.28 (object beyond focus) is real because the rays really do converge at the point where the image is formed. In contrast, the image in Figure 33.31 is virtual because the rays do not actually converge at the point where the image is formed. (The extensions of these rays do cross the image point, however, and so an observer interprets the rays emerging from the lens as having traveled along straight lines from the location of the image, as indicated in Figure 33.31*a*.)

Figure 33.31 (*a*) When an object is located between the lens and the focus, the image is virtual and enlarged. (*b*) In a simplified ray diagram, the object and image are replaced with solid and dashed arrows, respectively, and the lens is replaced by a vertical line. (The + indicates a converging lens.)

(*a*) Ray diagram for an object located between the focus and the lens



(*b*) Simplified version of ray diagram



An important difference between real and virtual images is that if a screen is placed at the location of a real image, the image can be seen on the screen. Placing a screen at the location of a virtual image does not display the image because the light rays do not actually pass through the image location.

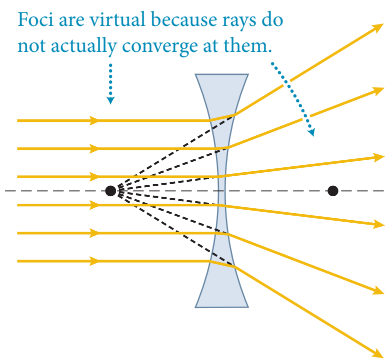
Figure 33.31*a* shows that, for this configuration of object and lens (object between focus and lens), the image is larger than the object and upright (unlike the image in Figure 33.28, which is inverted). A magnifying glass is designed to produce an enlarged, upright image of an object, which means that magnifying glasses are made with converging lenses and are held close to the object of interest (so that the object is between the lens and the focus).

33.11 As the object in Figure 33.31 is moved closer to the lens, does the size of the image increase, decrease, or stay the same?

Just as with electric circuits, it is convenient to use a simplified notation for ray diagrams. Figure 33.31*b* shows such a simplified version of the ray diagram of Figure 33.31*a*. Note that objects and real images are denoted by solid arrows and virtual images are denoted by dashed arrows.

Lenses can also be made with concave spherical surfaces rather than convex ones, where a *concave surface* is one that curves like the inside of a sphere. Such a lens is called a *diverging lens*, and **Figure 33.32** shows why: A series of parallel rays entering the lens are no longer parallel when they emerge. If we follow the path of the emerging rays back to the left side of the lens, we see that the diverging rays appear to all come from the same location on the left side. This location corresponds to the focus of a converging lens, but in a diverging lens it is a *virtual focus* rather than a real focus because the rays never actually travel through this location. (Just as for converging lenses, there is an equivalent focus on the other side of the lens.)

Figure 33.32 A diverging lens.



Procedure: Principal rays for lenses

The propagations of principal rays for converging and diverging lenses are very similar. The description below holds for rays that travel from left to right.

Converging lens

1. A ray that travels parallel to the lens axis before entering the lens goes through the right focus after exiting the lens.
2. A ray that passes through the center of the lens continues undeflected.
3. A ray that passes through the left focus travels parallel to the lens axis after exiting the lens. If the object is between the focus and the lens, this ray doesn't pass through the focus but lies on the line from the focus to the point where the ray originates.

Diverging lens

1. A ray that travels parallel to the lens axis before entering the lens continues along the line from the left focus to the point where the ray enters the lens.
2. A ray that passes through the center of the lens continues undeflected.
3. A ray that travels toward the right focus travels parallel to the lens axis after exiting the lens.

Figure 33.33 Ray diagram for an object outside the focus of a diverging lens.

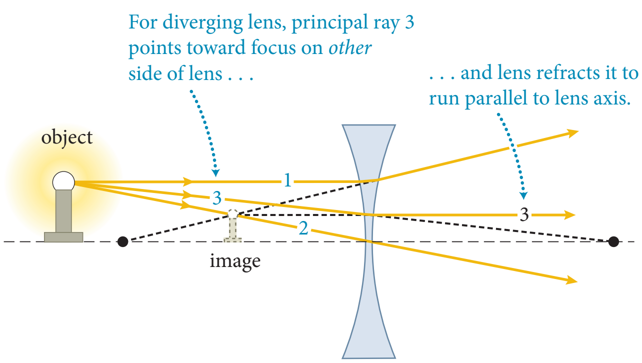
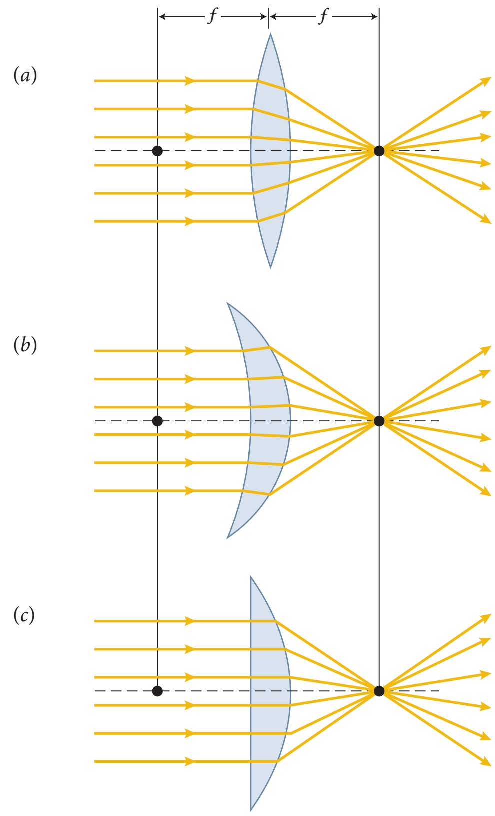


Figure 33.33 shows a ray diagram for a diverging lens. The same principal rays are drawn, but now ray 3 does not pass through the focus on the same side of the lens as the object. Instead, for diverging lenses ray 3 is drawn on the line that runs from the point where the ray originates to the focus on the other side of the lens; once refracted by the lens, this ray travels parallel to the lens axis. The general procedure for drawing ray diagrams for diverging lenses is still the same as the one for converging lenses (see the Procedure box "Simplified ray diagrams for lenses" on page 886), but the drawing of principal rays is a little bit different (see the Procedure box "Principal rays for lenses" above).

The lenses we have considered so far all have identical curved surfaces on each side. Many lenses have different surfaces, however. For example, it is possible to construct a converging lens with a certain focal length with two identical curved surfaces, two differently curved surfaces, or even a flat and a curved surface (**Figure 33.34**), as long as the lens is thicker at its center than at the edges. Regardless of the *radii of curvature* (that is, the radii of the spheres that best fit the surfaces), the lens has two foci, one on either side of the lens at the same distance f from it.

33.12 Is the image in Figure 33.33 real or virtual?

Figure 33.34 These lenses are all converging because each is thicker at the center than at the edges.

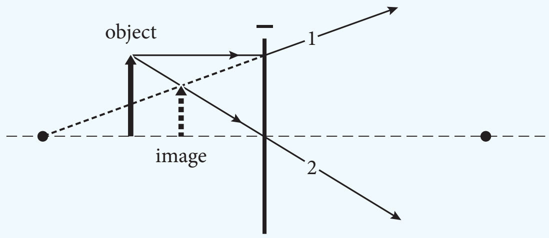


Example 33.5 Demagnifying glass

Suppose the object in Figure 33.33 is placed between the focus and the lens. (*a*) Is the image real or virtual? (*b*) Is it larger than, smaller than, or the same size as the object?

❶ GETTING STARTED To sketch the situation (**Figure 33.35**), I represent the diverging lens as a vertical line with a minus sign above it, and draw the horizontal lens axis. I add the focal points at equal distances from the lens along its axis. Finally, I add a solid arrow, representing the object, between the left focal point and the lens.

Figure 33.35



❷ **DEVISE PLAN** To locate the image and determine its size and whether it is real or virtual, I can draw the principal rays.

❸ **EXECUTE PLAN** (*a*) I add to my diagram the principal rays coming from the tip of my object arrow. All I need is rays 1 and 2; I do not need ray 3 because the intersection of ray 2 and the dashed extension of ray 1 unambiguously determines the location of the tip of the image arrow. The dashed extension I had to draw for ray 1 tells me that the rays do not actually intersect at this location; they only appear to intersect here. Therefore the image is virtual. ✔

(*b*) My diagram tells me that the image is smaller than the object. ✔

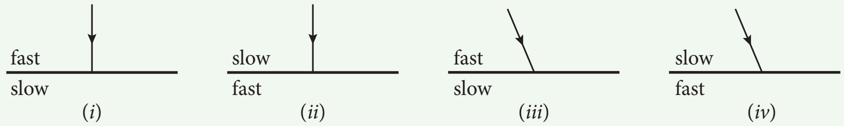
❹ **EVALUATE RESULT** Because a diverging lens spreads rays out rather than bringing them together, it makes sense that a virtual image will form. I also know from experience that in contrast to a converging lens, which magnifies images, diverging lenses create smaller images, as I found.

33.13 (*a*) Draw the third principal ray in **Figure 33.35**. Is there any position for the object in Figure 33.35 for which (*b*) the image is larger than the object and (*c*) the image is real?

Self-quiz

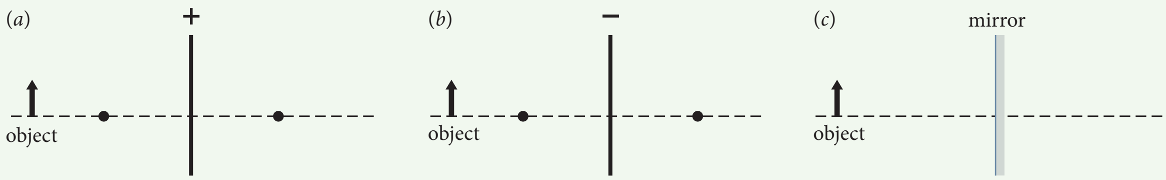
1. Why do you get a clear reflection from the surface of a lake on a calm day but little or no reflection from the surface on a windy day?
2. (*a*) As light travels from one medium into another, as shown in **Figure 33.36** ("fast" and "slow" refer to the wave speed in each medium), what happens to the wavelength of the light? (*b*) Draw the reflected and refracted rays at each surface.

Figure 33.36



3. What is the difference between a real image and a virtual image?
4. In each situation in **Figure 33.37**, draw the three rays emanating from the top of the object and reflecting or refracting from the optical element shown. Show the image, and state whether it is real or virtual.

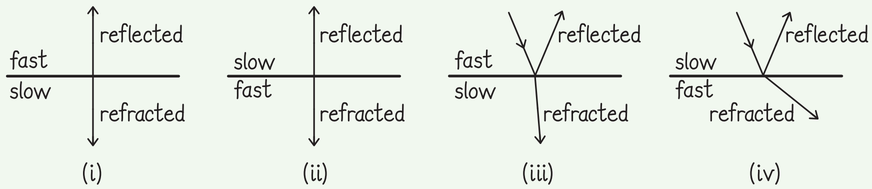
Figure 33.37



Answers:

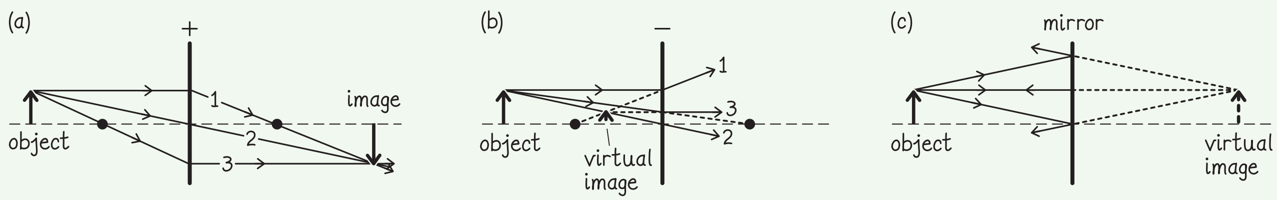
1. On a calm day, the lake surface is smooth, and specular reflection is like that of a mirror. On a windy day, the surface is rough, which makes the reflection diffuse and prevents the formation of an image.
2. (*a*) The wavelength decreases when the wave travels more slowly in the second medium (*i* and *iii*) and increases when the wave travels faster in the second medium (*ii* and *iv*). (*b*) See **Figure 33.38**.

Figure 33.38



3. Real image: All rays actually pass through the location of the image, and the image can be seen on a screen placed at the image location. Virtual image: All rays do not pass through the location of the image (only the extensions of the rays do), and the image cannot be seen on a screen placed at the image location.
4. See **Figure 33.39**. For the lenses, the three principal rays can be used to locate the image; for the mirror, any rays and the law of reflection can be used to locate the image. The images are (*a*) real, (*b*) virtual, (*c*) virtual.

Figure 33.39



33.5 Snel's law

In the first part of this chapter, we saw that light refracts when it travels from one medium into another because the speed of light depends on the medium. The speed of light in a medium is specified by the **index of refraction:**

$$n \equiv \frac{c_0}{c}, \tag{33.1}$$

where c is the speed of light in the medium and c_0 is the speed of light in vacuum. (By definition, $n_{\text{vacuum}} = 1$; in air $n_{\text{air}} \approx 1$.) If a light wave of frequency f travels from one medium into another, the frequency doesn't change because the source determines the frequency (see also Checkpoint 16.10). The wavelength, however, does change; it is greater in the medium in which wave speed is greater.

The wavelength λ of the light is related to the wave speed and frequency, in the same manner that these quantities are related for harmonic waves (Eq. 16.10). In vacuum, for example,

$$\lambda = \frac{c_0}{f} \text{ (vacuum)}. \tag{33.2}$$

In a medium in which a wave has speed c_1, the wavelength λ_1 is given by

$$\lambda_1 = \frac{c_1}{f} = \frac{c_0/n_1}{f} = \frac{1}{n_1}\lambda, \tag{33.3}$$

where λ is the wavelength of the wave in vacuum and n_1 is the index of refraction of the medium. Thus, the wavelength decreases as the index of refraction increases. As discussed in Section 33.3, the amount of refraction a light wave undergoes varies somewhat with wavelength (see Figure 33.22) because different wavelengths of light travel at different speeds. Therefore the index of refraction depends on the wavelength. Table 33.1 lists the indices of refraction for some common transparent materials at a wavelength of 589 nm.

Let us now work out the quantitative relationship between the angle of incidence and the angle of refraction. Figure 33.40 shows wavefronts and one ray for a beam of light incident on the interface between medium 1 and medium 2 at angle θ_1 from the normal. The angle of refraction is θ_2. Using right triangles ABD and ACD, we can express angles θ_1 and θ_2 in terms of the wavelengths λ_1 and λ_2:

$$\sin\theta_1 = \frac{\text{BD}}{\text{AD}} = \frac{\lambda_1}{\text{AD}} \tag{33.4}$$

and

$$\sin\theta_2 = \frac{\text{AC}}{\text{AD}} = \frac{\lambda_2}{\text{AD}}. \tag{33.5}$$

Figure 33.40 Relationship between angle of incidence θ_1 and angle of refraction θ_2.

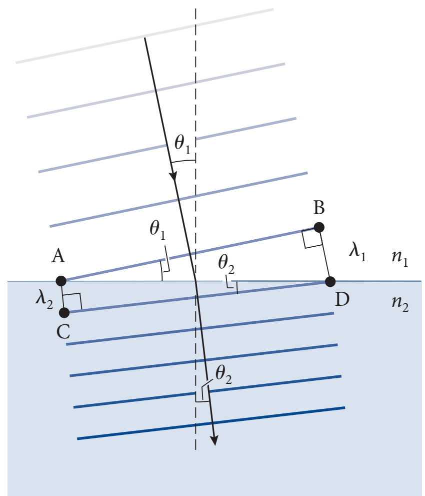


Table 33.1 Indices of refraction for common transparent materials

Material	n (for λ = 589 nm)
Air (at standard temperature and pressure)	1.00029
Liquid water	1.33
Sugar solution (30%)	1.38
Sugar solution (80%)	1.49
Microscope cover slip glass	1.52
Sodium chloride (table salt)	1.54
Flint glass	1.65
Diamond	2.42

Combining these equations to eliminate AD and substituting $\lambda_1 = \lambda/n_1$ (Eq. 33.3), we get

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{\lambda_1}{\lambda_2} = \frac{\lambda/n_1}{\lambda/n_2} = \frac{n_2}{n_1}, \tag{33.6}$$

which can be written as

$$n_1 \sin\theta_1 = n_2 \sin\theta_2. \tag{33.7}$$

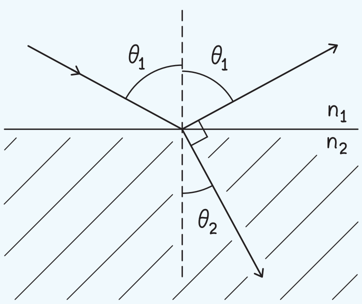
This relationship between the indices of refraction and the angles of incidence and refraction is called **Snel's law,** after the Dutch astronomer and mathematician Willebrord Snel van Royen (1580–1626).

Example 33.6 Bending 90°

A ray traveling through a medium for which the index of refraction is n_1 is incident on a medium for which the index of refraction is n_2. At what angle of incidence θ_1, expressed in terms of n_1 and n_2, must the ray strike the interface between the two media for the reflected and transmitted rays to be at right angles to each other?

❶ GETTING STARTED This problem involves both reflection and refraction at an interface between two media. To visualize the problem, I draw the incident, reflected, and refracted rays and indicate that the reflected and refracted rays are 90° apart (**Figure 33.41**).

Figure 33.41



❷ DEVISE PLAN Snel's law (Eq. 33.7), the law of reflection, and the indices of refraction determine the paths taken by the reflected and refracted rays. Therefore I need to use those relationships to obtain an expression that tells me the value of θ_1 that produces reflected and refracted rays oriented 90° to each other. To obtain θ_1 in terms of n_1 and n_2, I need to eliminate θ_2 from Eq. 33.7. To do so, I use the fact that the angles on the right side of the normal to the interface must add to 180°. Thus, with reflected and refracted rays forming a 90° angle, I can say $180° = \theta_1 + 90° + \theta_2$. Solving this expression for θ_2 gives $\theta_2 = 90° - \theta_1$, which I can substitute into Eq. 33.7.

❸ EXECUTE PLAN Substituting $\theta_2 = 90° - \theta_1$ into Eq. 33.7, I get

$$n_1 \sin\theta_1 = n_2 \sin(90° - \theta_1) = n_2 \cos\theta_1,$$

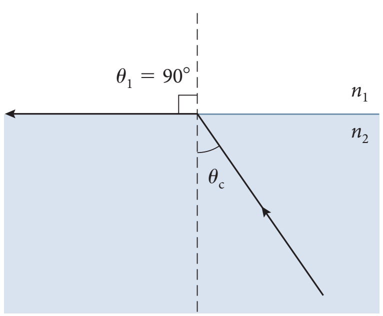
and isolating the terms that contain θ_1 gives

$$\frac{\sin\theta_1}{\cos\theta_1} = \tan\theta_1 = \frac{n_2}{n_1}$$

$$\theta_1 = \tan^{-1}\left(\frac{n_2}{n_1}\right). ✔$$

❹ EVALUATE RESULT My result says that θ_1 increases as n_2 increases. This makes sense because as n_2 increases, the refracted ray bends more, meaning that θ_2 becomes smaller. To keep the reflected and refracted rays perpendicular to each other, the angle of reflection must increase, and so θ_1 must also increase.

Figure 33.42 Critical angle for a ray traveling from a denser medium (n_2) to a less dense medium ($n_1 < n_2$).



Earlier in this chapter, we found that for rays traveling from a denser medium to a less dense medium, we can define a critical angle of incidence θ_c such that the angle of refraction θ_1 is equal to 90° (**Figure 33.42**); beyond this critical angle θ_c, total internal reflection occurs. We can calculate the critical angle θ_c for an interface between two media with indices of refraction n_1 and n_2 ($n_2 > n_1$) by applying Snel's law (Eq. 33.7) and setting $\theta_1 = 90°$:

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{1}{\sin\theta_c} = \frac{n_2}{n_1}. \tag{33.8}$$

Solving for θ_c gives

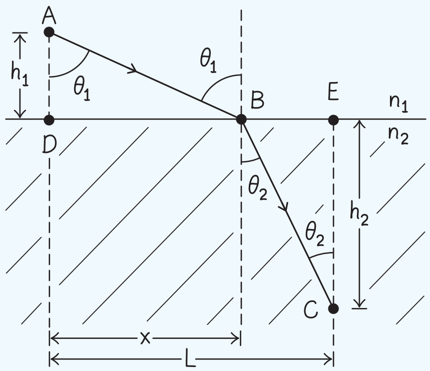
$$\theta_c = \sin^{-1}\left(\frac{n_1}{n_2}\right). \tag{33.9}$$

Example 33.7 Fermat's principle

For a light ray that crosses the interface between medium 1 having index of refraction n_1 and medium 2 having index of refraction n_2, what relationship between θ_1 and θ_2 follows from Fermat's principle (page 884)?

❶ **GETTING STARTED** I begin with a diagram that shows the two media and a ray traveling from an arbitrary point A in the n_1 medium to an arbitrary point C in the n_2 medium (**Figure 33.43**). Fermat's principle states that the path the ray takes from A to C is the path for which the time interval needed for the motion is a minimum. Therefore this ray must cross the interface at a point B that makes the time interval a minimum.

Figure 33.43



An alternative way to express this problem is: Given the locations of A and C, where must B lie so as to minimize the time interval needed to travel from A to C?

❷ **DEVISE PLAN** I add two more location labels to my drawing: D directly below A and lying on the interface, and E directly above C and lying on the interface. Doing so gives me two right-angle triangles that permit me to express the angles in terms of the distance traveled. I write h_1 for the distance AD and h_2 for the distance EC . I can think of the distance from D to B as unknown—I'll call it x—and the distance from D to E, which I'll call L, is fixed by the locations of A and C. My goal is to determine the value of x for which the travel time from A to C is minimized. Once I obtain x, I hope to obtain a relationship between θ_1 and θ_2.

❸ **EXECUTE PLAN** I begin by expressing the time interval Δt_{AC} the ray needs to travel from A to C in terms of the distances shown in Figure 33.43 and the speed of light in the two media:

$$\Delta t_{AC} = \Delta t_{AB} + \Delta t_{BC} = \frac{AB}{c_1} + \frac{BC}{c_2} = \frac{AB}{c_0/n_1} + \frac{BC}{c_0/n_2}. \quad (1)$$

Next I express AB and BC in terms of h_1, h_2, L, and x:

$$AB = \sqrt{h_1^2 + x^2}$$

$$BC = \sqrt{h_2^2 + (L - x)^2}.$$

Substituting these two expressions into Eq. 1 gives me

$$\Delta t_{AC} = \frac{\sqrt{h_1^2 + x^2}}{c_0/n_1} + \frac{\sqrt{h_2^2 + (L - x)^2}}{c_0/n_2}.$$

Except for x, all quantities in this expression are constants.

The path for which the time interval Δt_{AC} is a minimum—as it must be from Fermat's principle—is the path for which the derivative of Δt_{AC} with respect to x is zero:

$$\frac{d}{dx}(\Delta t_{AC}) = \frac{xn_1}{c_0\sqrt{h_1^2 + x^2}} - \frac{(L - x)n_2}{c_0\sqrt{h_2^2 + (L - x)^2}} = 0. \quad (2)$$

Solving this equation for x would tell me where the light ray crosses the interface, but I do not have values for L and x. However, the right triangles ADB and BEC in Figure 33.43 allow me to express these distances in terms of θ_1 and θ_2:

$$\sin\theta_1 = \frac{x}{\sqrt{h_1^2 + x^2}} \quad \text{and} \quad \sin\theta_2 = \frac{L - x}{\sqrt{h_2^2 + (L - x)^2}}. \quad (3)$$

I can now use these expressions to rewrite Eq. 2 in terms of θ_1 and θ_2. From Eq. 2 I obtain

$$\frac{xn_1}{c_0\sqrt{h_1^2 + x^2}} = \frac{(L - x)n_2}{c_0\sqrt{h_2^2 + (L - x)^2}}.$$

Canceling the c_0 factors that appear on both sides and substituting from Eq. 3, I get

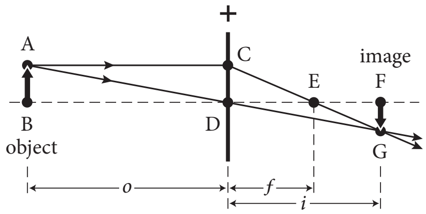
$$n_1 \sin\theta_1 = n_2 \sin\theta_2. \checkmark$$

❹ **EVALUATE RESULT** My result is identical to Snel's law—which I derived by considering the effect of changing speed on the propagation of wavefronts. So Fermat's principle yields the same result as Snel's law, which I know to be correct.

Fermat's principle applies to all of ray optics, not only to refraction. As discussed in the box "Fermat's principle" on page 884, the law of reflection also follows from this principle.

33.14 We found in Example 33.3 that a light ray is refracted twice when it passes completely through a slab of transparent material (see Figure 33.18). The result of these two refractions is that the exiting ray is shifted sideways relative to the entering ray. Let the slab be in air with an index of refraction $n_1 = 1$. (*a*) Derive an expression for the distance (perpendicular to the ray) over which the ray is shifted sideways for an angle of incidence θ_1, slab thickness d, and slab index of refraction n_2. (*b*) Calculate the value of the shift for $\theta_1 = 30°$, $n_2 = 1.5$, $d = 0.010$ m.

Figure 33.44 Simplified ray diagram for the formation of an image by a converging lens.



33.6 Thin lenses and optical instruments

In the first part of this chapter, we found that converging lenses form images of objects by focusing the light rays emanating from those objects. Let us now work out quantitatively the location and size of such images. We shall restrict our discussion to lenses that are thin enough that we can ignore the type of effects shown in Figure 33.29. Such lenses are called *thin lenses*.

A simplified ray diagram of the image formed by a converging lens is shown in **Figure 33.44**. The focal length f of the lens is DE, the distance o from the lens to the object (also called the *object distance*) is BD, and the distance i from the lens to the image (also called the *image distance*) is DF. The height of the object is AB, and the height of the image is FG. Let us denote the height of the object by h_o and the height of the image by h_i. We choose the values of h_o and h_i to be positive for upright objects and images and negative for inverted objects and images. We want to obtain a relationship between h_i and h_o, which will tell us how large the image is relative to the object. We also want a relationship among f, i, and o, which will tell us how the positions of the object and the image are related.

We begin by noting that triangles ABD and DFG are similar, which means

$$\frac{\mathrm{AB}}{\mathrm{DB}} = \frac{\mathrm{FG}}{\mathrm{DF}}. \tag{33.10}$$

Because the image is inverted, $h_i = -\mathrm{FG}$, we can rewrite Eq. 33.10 as

$$\frac{h_o}{o} = \frac{-h_i}{i}. \tag{33.11}$$

Rearranging this expression gives

$$-\frac{h_o}{h_i} = \frac{o}{i}. \tag{33.12}$$

In this case, the absolute value of the ratio of the object height to the image height equals the ratio of the object distance to the image distance.

Triangles CDE and EFG are also similar, which means

$$\frac{\mathrm{DE}}{\mathrm{CD}} = \frac{\mathrm{EF}}{\mathrm{FG}}, \tag{33.13}$$

which can be written as

$$\frac{f}{h_o} = \frac{i - f}{-h_i}. \tag{33.14}$$

Using Eq. 33.12 to rewrite Eq. 33.14 in terms of f, o, and i gives us

$$\frac{f}{o} = \frac{i - f}{i} = 1 - \frac{f}{i}. \tag{33.15}$$

Dividing by f and rearranging terms yield

$$\frac{1}{f} = \frac{1}{o} + \frac{1}{i}. \tag{33.16}$$

This result is known as the **lens equation.**

QUANTITATIVE TOOLS

It can be shown that Eq. 33.16 is generally true for either real or virtual images formed by either converging or diverging lenses, as long as we choose the signs of f, i, and o properly. For a converging lens, f is positive and o is positive if the object is in front of the lens. (This is always true for a single lens and for the first lens in a lens combination. For situations involving multiple lenses, however, it is possible that the object imaged by a secondary lens is on the opposite side of the lens from the side where the rays enter it—that is, the object is "behind the lens." In that case o is negative.) If the image is on the same side of the lens as the emerging light, the image is real and i is positive; if the image is on the opposite side of the lens from the emerging light, the image is virtual and i is negative.

For a diverging lens (a lens with concave surfaces), the focal length f is negative because the focus is virtual rather than real—that is, parallel rays appear to come from the same side of the lens where the light source is rather than converging on the other side of the lens. The same sign convention applies to o as for converging lenses. A single diverging lens always produces a virtual image (see Example 33.5), so i is always negative for such lenses.

The sign conventions for f, i, and o are similar for images formed by spherical mirrors, which are discussed in the next section. **Table 33.2** summarizes these sign conventions.

The **magnification** of the image is defined as the ratio of the signed image height to the object height. Using Eq. 33.12, we get

$$M \equiv \frac{h_\mathrm{i}}{h_\mathrm{o}} = -\frac{i}{o}. \qquad (33.17)$$

We define M this way so that the magnification of upright images is positive and that of inverted images is negative. Examining Figures 33.30, 33.31, and 33.33, we can see that for a single lens, when the image distance and object distance are both positive, as in Figure 33.30, we obtain an inverted image, whereas when the image distance is negative, as in Figures 33.31 and 33.33, we obtain an upright image.

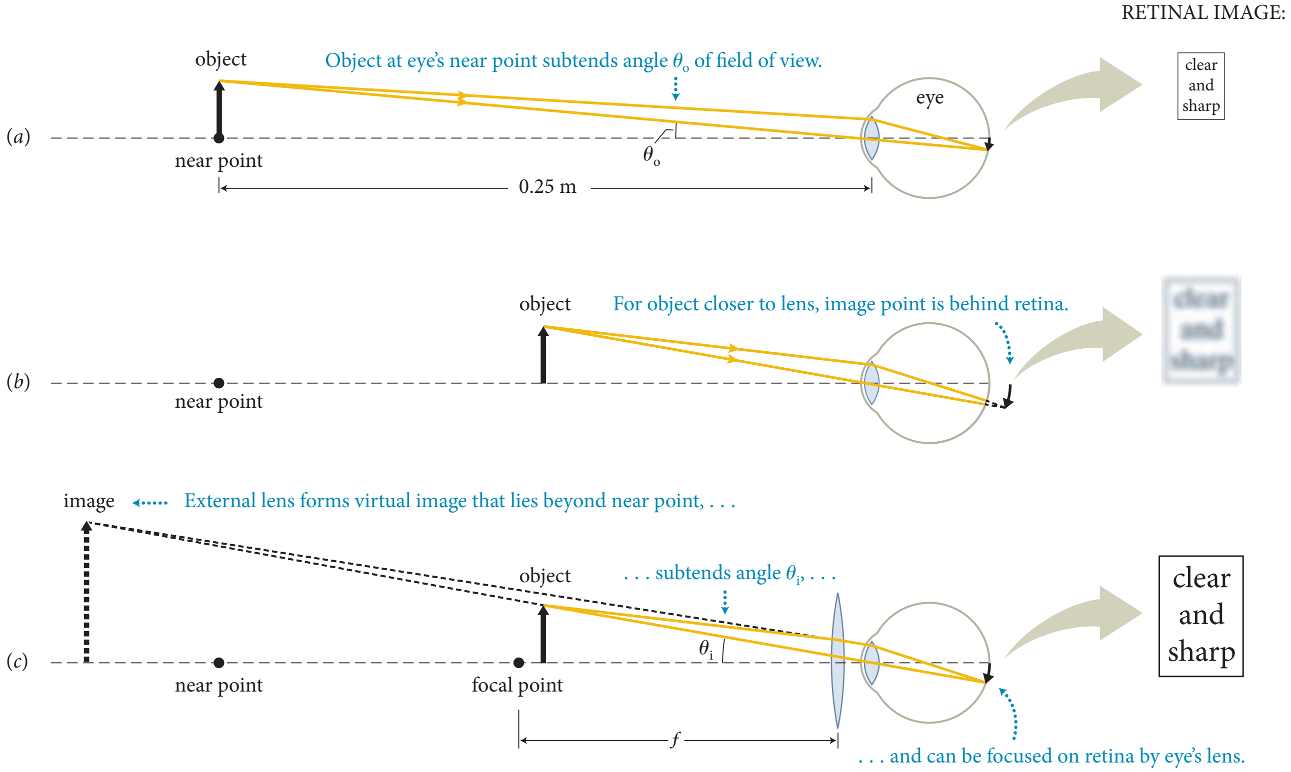
33.15 If the diverging lens in Figure 33.33 has a focal length of 80 mm and the object is located 100 mm from the lens, (a) what is the image distance and (b) how tall is the image relative to the object?

Table 33.2 Sign conventions for f, i, and o (positive = real; negative = virtual)

Sign	Lens	Mirror
$f > 0$	converging lens	converging mirror
$f < 0$	diverging lens	diverging mirror
$o > 0$	object in front[b] of lens	object in front of mirror
$o < 0$[a]	object behind lens	object behind mirror
$i > 0$	image behind lens	image in front of mirror
$i < 0$	image in front of lens	image behind mirror
$h_\mathrm{i} > 0$	image upright	image upright
$h_\mathrm{i} < 0$	image inverted	image inverted
$\lvert M \rvert > 1$	image larger than object	image larger than object
$\lvert M \rvert < 1$	image smaller than object	image smaller than object

[a] Encountered only with lens or mirror combinations.
[b] For both lenses and mirrors, "in front" means on the side where the rays originate; "behind" refers to the opposite side.

Figure 33.45 An eye cannot focus on an object that is closer than its near point (which represents the limit of the biological lens's ability to change curvature). However, an external converging lens (such as a magnifying lens) makes it possible to see objects that are closer than the near point. It also enlarges them.



The human eye focuses incoming light rays, forming an image on the retina of the eye (**Figure 33.45*a***). One part of the eye is its lens, but unlike the lenses we have examined so far, the focal length of the eye's lens is variable, which allows us to see objects clearly over a wide range of distances. When the muscle around the lens is fully relaxed, the lens flattens out and the retina lies in the focal plane of the lens. Thus, light rays from distant objects focus onto the retina.

To the unaided eye, the largest (and thus most detailed) image of an object is observed when we bring the object as close as possible to the eye. However, there is a limit to how much the eye's lens can adjust. The *near point* is the closest object distance at which the eye can focus on the object comfortably. Typically, for an adult, the near point is about 0.25 m from the eye. An object positioned at the near point appears clear and sharp to the observer, as shown in Figure 33.45*a*. With age the distance between the near point and the eye tends to increase, and when an object is brought closer than the near point, the plane where the image is formed lies behind the retina and the image "seen" by the retina is blurry (Figure 33.45*b*). This situation can be corrected by an external lens that works in combination with your eye's lens to focus the image on the retina (Figure 33.45*c*).

An external converging lens properly placed between the object and the eye, as in Figure 33.45*c*, magnifies the object. To maximize the size of the image, the object is held near the focus of the external lens, and the lens is held as close as possible to the eye. The image formed by the external lens then serves as the object for the eye's lens. The image formed by the external lens is virtual and subtends an angle θ_i that is greater than the angle θ_o subtended by the object in Figure 33.45*a*, permitting the viewer to see finer details. The image is also outside the near point; if the object is placed exactly at the focus of the external lens, the image is at infinity and can be viewed comfortably. We can define the *angular magnification* produced by the lens as

$$M_\theta \equiv \left|\frac{\theta_i}{\theta_o}\right|. \tag{33.18}$$

For small angles and an object placed close to the focus of the external lens, as in Figure 33.45*c*, the angle θ_i subtended by the image can be expressed in terms of the object height h_o and the focal length f of the lens:

$$\theta_i \approx \tan\theta_i \approx \frac{h_o}{f} \quad \text{(object close to focus, small } \theta_i\text{)}. \tag{33.19}$$

For small angles and an object placed at the eye's near point, as in Figure 33.45*a*, the angle subtended by the object is approximately

$$\theta_o \approx \tan\theta_o = \frac{h_o}{0.25\text{ m}} \quad \text{(object close to near point, small } \theta_o\text{)}. \tag{33.20}$$

Substituting Eqs. 33.19 and 33.20 into Eq. 33.18 gives an angular magnification of

$$M_\theta \approx \frac{0.25\text{ m}}{f}. \tag{33.21}$$

This expression gives what is called either the *small-angle approximation* or the *paraxial approximation* to the angular magnification because it is obtained with the small-angle approximations of Eqs. 33.19 and 33.20. These approximations are good to within 1% for angles of 10° or less.

Lenses placed near the eye (in the form of eyeglasses) are used to correct vision for far-sighted or near-sighted eyes. The strength of eyeglass lenses (and of magnifying lenses, too) is commonly symbolized by d and measured in *diopters*:

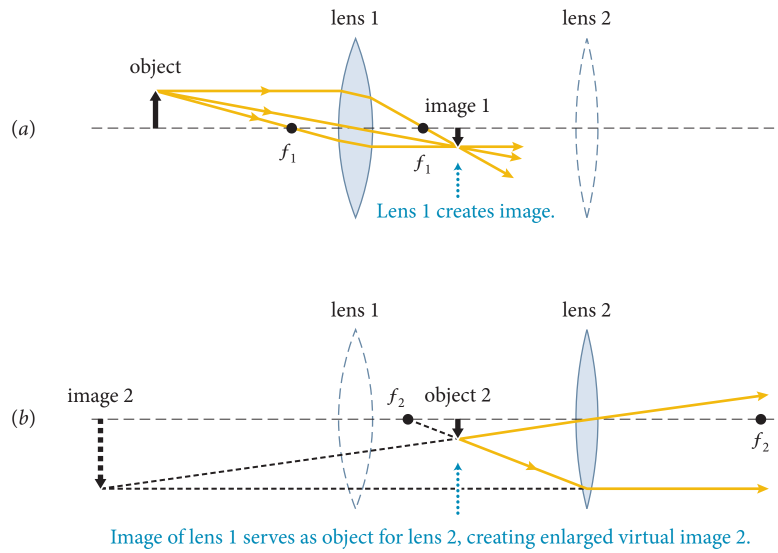
$$d \equiv \frac{1\text{ m}}{f}. \tag{33.22}$$

The *lens strength* d, like the lens focal length f, is positive for converging lenses and negative for diverging lenses. For example, a +4-diopter lens is a converging lens with a focal length of 0.25 m. Diverging lenses are typically used to correct nearsightedness, with lens strengths ranging from −0.5 to −4 diopters.

33.16 A single-lens magnifying glass used to examine photographic slides produces eightfold angular magnification. (*a*) What is the lens strength in diopters? (*b*) What is the focal length of the lens?

Many optical instruments combine two or more lenses to increase magnification. To trace rays through a combination of lenses, use the following procedure: The image formed by the first lens serves as the object for the second lens, the image formed by the second lens serves as the object for the third lens, and so on. **Figure 33.46** on the next page shows a ray diagram constructed in two steps for a combination of two lenses. Figure 33.46*a* shows the object, image, and rays for lens 1, and Figure 33.46*b* shows these elements for lens 2. Note that the rays from object 2 (which is the image formed by lens 1) are *not* the continuation of those used to locate image 1.

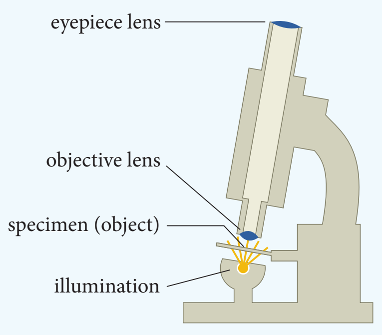
Figure 33.46 Two-step process for tracing rays through a combination of two lenses. When lenses are combined, the image of each lens serves as an object for the next lens.



Example 33.8 Compound microscope

A compound microscope consists of two converging lenses, the *objective lens* and the *eyepiece lens*, positioned on a common optical axis (**Figure 33.47**). The objective lens is positioned to form a real, highly magnified image 1 of the sample being examined, and the eyepiece lens is positioned to form a virtual, further magnified image 2 of image 1. It is image 2 that the user sees. A knob on the microscope allows the user to move the objective lens upward and downward to change both the sample-objective lens distance and the distance between the two lenses. (*a*) How must the sample and the two lenses be positioned relative to one another so that the user sees a highly magnified, virtual image of the sample? (*b*) What is the overall magnification produced by the microscope?

Figure 33.47 Example 33.8.

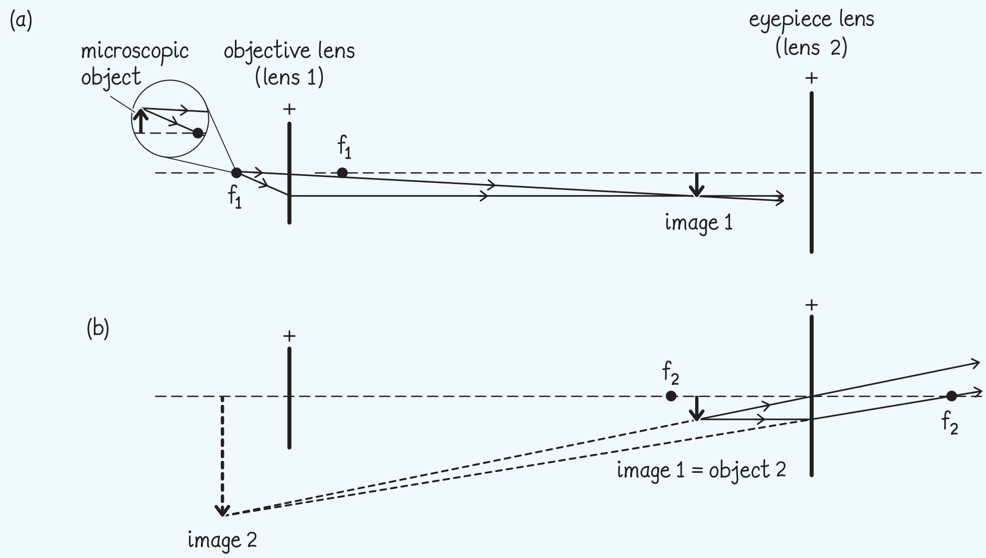


❶ GETTING STARTED I begin by examining Figure 33.46, which shows how, in a combination of two lenses 1 and 2, the image formed by lens 1 serves as the object for lens 2. The objective lens in a compound microscope corresponds to lens 1 in Figure 33.46, and the eyepiece lens corresponds to lens 2. Thus to keep things simple I refer to the objective lens as 1 and the eyepiece lens as 2.

❷ DEVISE PLAN To determine the relative positioning of the two lenses relative to each other, I must examine ray diagrams for various lens-sample distances and determine for which arrangement I get the greatest magnification. To determine the magnification M_1 of image 1, I can use the lens equation (Eq. 33.16) together with the relationship among magnification, image distance, and object distance (Eq. 33.17). The focal length of the lenses is fixed by their construction, and in operating a microscope, the observer can adjust the distance between the sample and the objective lens, so I can express this magnification in terms of f_1 and o_1. To determine the magnification of image 2, I recognize that lens 2 is used as a magnifying glass, so I can use Eq. 33.21, which gives the angular magnification $M_{\theta 2}$ produced by a simple magnifier. The overall magnification produced by the microscope is the product $M_1 M_{\theta 2}$.

❸ EXECUTE PLAN (*a*) In Figure 33.46, lens 1 produces an image that is smaller than the object. I am told that the image formed by lens 1 in a microscope is larger than the sample, and so I must choose a different sample position, one that yields an image 1 larger than the sample. I am also told that this image is real. Placing the sample just outside the focal point of lens 1 gives me an image 1 that is larger than the sample. My choices are to increase or decrease the sample-lens 1 distance. Drawing a ray diagram for each possibility, I see that moving the sample farther and farther from lens 1 makes the image smaller and smaller. Therefore I should position the sample closer to lens 1 than in Figure 33.46*a*. Should I choose a position inside or outside the lens focus? I know from the

Figure 33.48



problem statement that this image is real, and I know from Figure 33.31 that an object inside the focus of a converging lens produces a virtual image. Thus my best choice is to adjust the sample-lens 1 distance so that the sample is just outside the lens focus (**Figure 33.48*a***).

I am told that image 2 is virtual and larger than image 1. I know from Figure 33.31 that a converging lens produces a virtual, magnified image when the object is inside the lens focus. I again draw ray diagrams for various positions inside the focus and see that the greatest magnification is obtained when I adjust the distance from lens 1 to lens 2 to make image 1 fall just inside the focal point of lens 2, as shown in Figure 33.48*b*. ✔

(*b*) To determine M_1, I use the lens equation to write i_1 in terms of f_1 and o_1:

$$i_1 = \frac{1}{\left(\frac{1}{f_1} - \frac{1}{o_1}\right)}.$$

I substitute this expression into Eq. 33.17:

$$M_1 = -\frac{i_1}{o_1} = -\frac{1}{o_1} \times \frac{1}{\left(\frac{1}{f_1} - \frac{1}{o_1}\right)} = -\frac{1}{\frac{o_1}{f_1} - 1},$$

which tells me that the magnification M_1 produced by lens 1 is determined by the ratio o_1/f_1. Because I have made o_1 slightly larger than f_1 in order to produce a real image 1, the denominator is positive and therefore M_1 is negative.

The angular magnification produced by lens 2 is approximately

$$M_{\theta 2} = \frac{0.25\text{ m}}{f_2}.$$

The overall magnification produced by the microscope is thus

$$M = M_1 M_{\theta 2} = \frac{-0.25\text{ m}}{f_2\left(\frac{o_1}{f_1} - 1\right)}. \checkmark$$

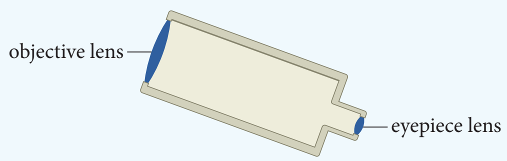
❹ **EVALUATE RESULT** Figure 33.48*a* indicates that image 1 is inverted, making M_1 negative and giving me confidence in my expression for M_1. Figure 33.48*b* tells me that image 2 is upright relative to its object, and so $M_{\theta 2}$ is positive, which agrees with my result. Because image 2 is inverted relative to the sample, the overall magnification is negative, as my result shows.

33.17 (*a*) Consider replacing the objective lens in Fig. 33.48*a* with one that has a greater focal length, and moving the sample in order to keep it just outside the focal point of the lens. Does the image formed by the objective lens move closer to the objective lens, stay in the same place, or move farther from the objective lens? (*b*) In practice it is desirable for a microscope to be fairly compact. To keep the microscope compact, should the focal length of the objective lens be chosen to be short or long, or does it matter?

Example 33.9 Refracting telescope

A refracting telescope, like a compound microscope, contains two converging lenses, the objective lens and the eyepiece lens, positioned on a common optical axis (**Figure 33.49**). However, a telescope is designed to view large, very distant objects, whereas a microscope is used to view very small objects that are placed very close to the objective lens. Consequently, the arrangement of lenses in a telescope is different from the arrangement in a microscope. The telescope's objective lens is positioned to form a real image of very distant objects, and the eyepiece lens is positioned to form a virtual image of the image produced by the objective lens, to be viewed by an observer. (*a*) How should the lenses be arranged to accomplish this? (*b*) What is the overall magnification produced by the telescope?

Figure 33.49 Example 33.9.



❶ **GETTING STARTED** I begin by examining Figure 33.46 and then construct a similar ray diagram with the object at a very great distance from the lenses. As in Example 33.8, I use lens 1 to refer to the objective lens and lens 2 to refer to the eyepiece lens. Because the object is very far away, light rays from it enter lens 1 as parallel rays. These rays form an image 1 in the focal plane of lens 1 (**Figure 33.50*a***). Because I know the location of image 1, I need to draw only one principal ray. As in the microscope, lens 2 is used as a simple magnifier to view image 1.

❷ **DEVISE PLAN** Because the original object is very distant and the final image is viewed by the observer's eye, I can calculate the angular magnification of this image. Although I could calculate the angular magnification produced by each lens and multiply them together, in this case it is simpler to determine the overall angular magnification because the angles θ_o and θ_i the object and image 2 subtend at the observer's eye are both very small. I can determine the overall angular magnification by taking the ratio of these angles while using the small-angle approximation.

❸ **EXECUTE PLAN** (*a*) In order for lens 2 to produce a magnified, virtual image of image 1, image 1 should be positioned just inside the focal plane of lens 2. If lens 2 is placed such that the image is at the focal plane of lens 2 (Figure 33.50*b*), lens 2 forms an infinitely distant, virtual image that can be viewed comfortably by the observer's relaxed eye. As my diagram shows, the lenses are then arranged such that their foci coincide. ✔

(*b*) **Figure 33.51*a*** shows the ray that passes through the foci of the lenses, labeled with the angles θ_o (subtended by the object) and θ_i (subtended by the image). Figure 33.51*b* shows the triangles I use to relate each of these angles to the height h_i of image 1 and the focal lengths of the lenses. The angular magnification is the ratio θ_o/θ_i. I can approximate these angles by

Figure 33.51

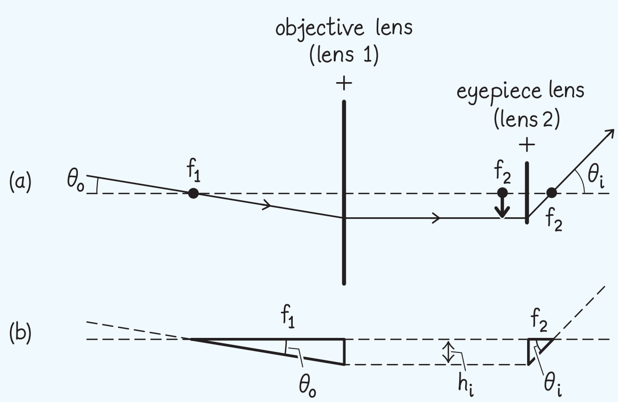
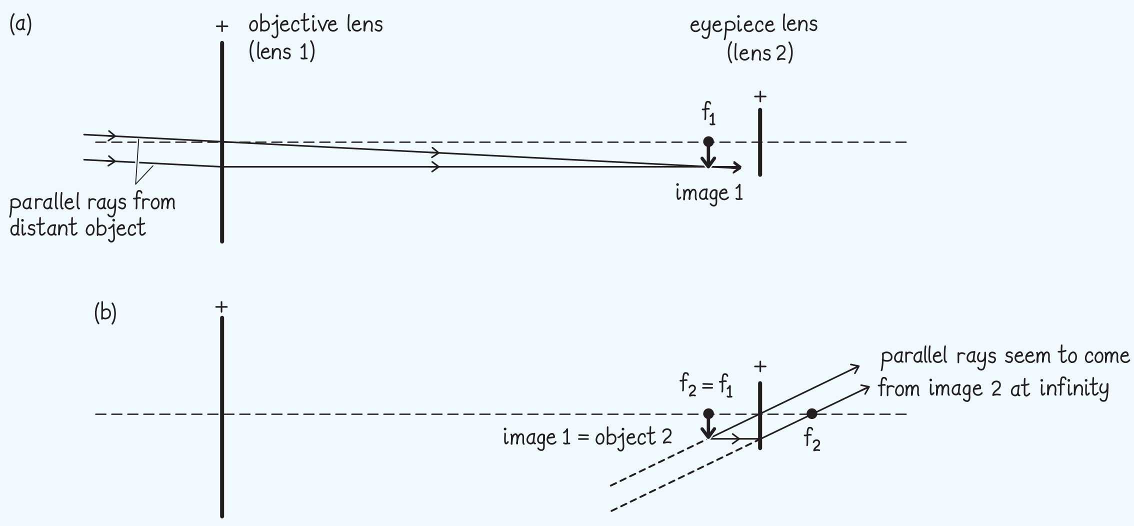


Figure 33.50



their tangents; substitute for the tangents of the angles in terms of h_i, f_1, and f_2; and simplify:

$$M_\theta = \left|\frac{\theta_i}{\theta_o}\right| \approx \left|\frac{\tan\theta_i}{\tan\theta_o}\right| = \frac{|h_i/f_2|}{|h_i/f_1|} = \left|\frac{f_1}{f_2}\right|. ✔$$

❹ **EVALUATE RESULT** Figure 33.50 indicates that image 1 is inverted and that image 2 is upright relative to image 1, which means that image 2 is inverted relative to the distant object. This makes sense because the incoming ray in Figure 33.51 is angled downward but the outgoing ray is angled upward.

33.18 A telescope with a magnification of 22× has an eyepiece lens for which the focal length f_2 is 40.0 mm. (*a*) What is the focal length f_1 of the objective lens? (*b*) What is the length of the telescope?

33.7 Spherical mirrors

Just like lenses, spherical mirrors focus parallel rays (**Figure 33.52**). A concave mirror focuses rays to a point in front of the mirror, corresponding to a real focus; a diverging mirror makes rays diverge so that they appear to come from a point behind the mirror, corresponding to a virtual focus. Thus, just as with lenses, we can have both converging and diverging spherical mirrors. Unlike lenses, however, spherical mirrors have only a single focus.

To obtain the location of the focus of a converging mirror, we examine the reflection of the two rays shown in **Figure 33.53*a***. Ray 1 comes in parallel to the axis of the mirror, striking the mirror at A, and ray 2 comes in along the axis of the mirror. The focus of the mirror is at D, where the two reflected rays cross, and the center of the sphere on which the surface of the mirror lies is at C. Line CA is therefore a radius of the sphere and perpendicular to the mirror surface. We denote the length of the radius by R. This distance is called the *radius of curvature*.

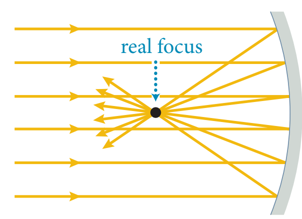
Ray 2 strikes the mirror at normal incidence and so is reflected back along the axis. Ray 1 is reflected through the focus at D. Consequently, the distances CD and AD in Figure 33.53*a* are equal. Dividing triangle ACD into two congruent right triangles by drawing a line perpendicular to the base from D, we see that $\text{CD} = (R/2)/\cos\theta_i$. For small θ, $\cos\theta \approx 1$, and so $\text{CD} = R/2$ and $\text{BD} = R - \text{CD} = R/2$, independent of θ. Therefore, the focus, which is located at D because that is where the two reflected rays cross, lies halfway between the mirror and the center of the sphere, making the focal length

$$f = \frac{R}{2}. \tag{33.23}$$

As Figure 33.53*b* shows, the geometry and hence the position of the focus are exactly the same for diverging mirrors except that here the focus lies behind the mirror.

Figure 33.52 Spherical mirrors focus parallel rays just like lenses do.

(*a*) Concave spherical mirror



(*b*) Convex spherical mirror

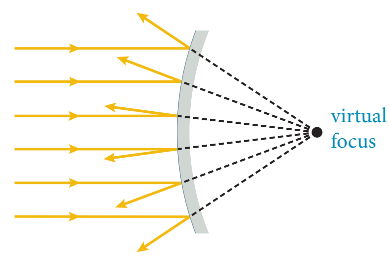
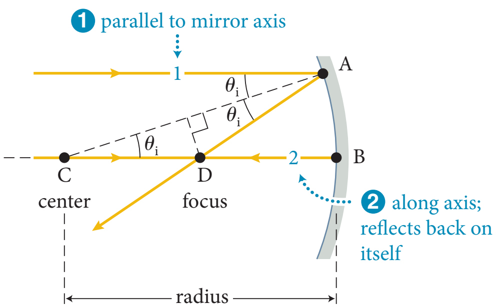
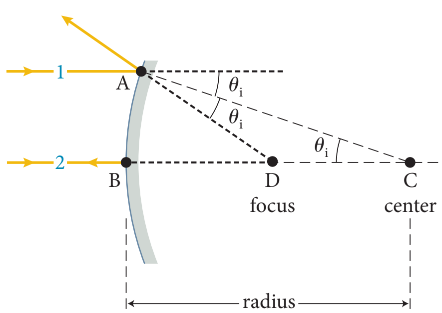


Figure 33.53 The two principal rays used to determine the focus of a spherical mirror.

(*a*) Concave spherical mirror



(*b*) Convex spherical mirror



Procedure: Ray diagrams for spherical mirrors

Ray diagrams for spherical mirrors are very similar to those for lenses. This procedure is for rays traveling from the left to the right.

1. Draw a horizontal line representing the mirror axis. In the center of the diagram, draw a circular arc representing the mirror. A converging mirror curves toward the left; a diverging mirror curves toward the right.
2. Put a dot on the axis at the center of the circular arc and label it C. Add another dot on the axis, halfway between C and the mirror. This point is the focus. Label it f.
3. Represent the object by drawing an upward-pointing arrow from the axis at the appropriate relative distance to the left of the mirror. For example, if the distance from the object to a converging mirror is one-third the radius of curvature of the mirror, put the arrow a bit to the right of the focus. The top of the arrow should be at about half the height of the mirror.
4. From the top of the arrow representing the object draw two or three of the following three so-called *principal rays* listed in the Procedure box "Principal rays for spherical mirrors."
5. The top of the image is at the point where the rays that are reflected by the mirror intersect. If the intersection is on the left side of the lens, the image is real; if it is on the right side, the image is virtual. Draw an arrow pointing from the axis to the intersection to represent the image (use a dashed arrow for a virtual image).

To determine the location of images formed by mirrors, we follow the same ray-tracing procedure we used for lenses. The three principal rays emanating from a given point on the object are analogous: ray 1 approaching the mirror parallel to the mirror axis, ray 2 passing through the center C of the mirror, and ray 3 passing through the focus on its way to the mirror. Now, however, "center" refers to *the center of the sphere on which the mirror surface lies* rather than the center of the lens. **Figure 33.54** shows a ray diagram for an image formed by a converging mirror. Ray 1 is reflected through the focus, ray 2 strikes the mirror at normal incidence and thus reflects back on itself, and ray 3 is reflected parallel to the mirror axis. As Figure 33.54 shows, there is a fourth ray that can easily be drawn: A ray that hits the mirror on the axis is reflected back symmetrically about the axis. The procedures for drawing ray diagrams and principal rays for spherical mirrors are given in the Procedure boxes on this page.

Object distance, image distance, and focal length are measured from the surface of the mirror, and the relationship among o, i, and f is the same as that for lenses:

$$\frac{1}{f} = \frac{1}{o} + \frac{1}{i}. \tag{33.24}$$

Procedure: Principal rays for spherical mirrors

This description holds for rays that travel from left to right.

Converging mirror

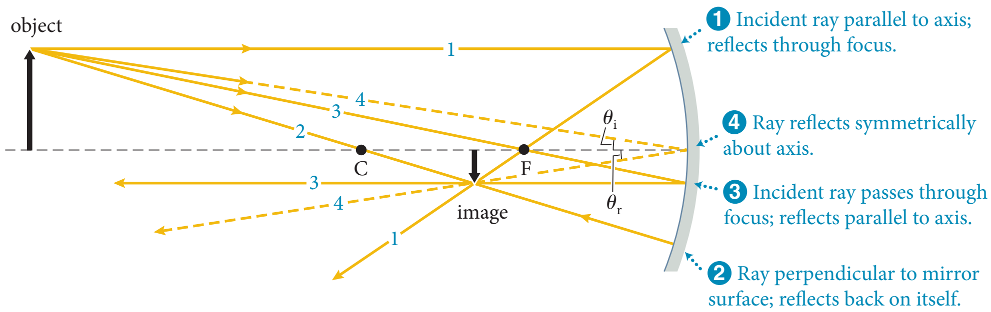
1. A ray that travels parallel to the mirror axis before reaching the mirror goes through the focus after being reflected.
2. A ray that passes through the center of the sphere on which the mirror surface lies is reflected back onto itself. If the object is between the center and the mirror, this ray doesn't pass through the center but lies on the line from the center to the point at which the ray originates.
3. A ray that passes through the focus is reflected parallel to the axis. If the object is between the focus and the mirror, this ray doesn't pass through the focus but lies on the line from the focus to the point at which the ray originates.

Diverging mirror

1. A ray that travels parallel to the mirror axis before reaching the mirror is reflected along the line that goes through the focus and the point where the ray strikes the surface.
2. A ray that passes through the center of the sphere on which the mirror surface lies is reflected back onto itself.
3. A ray whose extension passes through the focus is reflected parallel to the axis.

For both converging and diverging mirrors, a ray that hits the mirror on the axis is reflected back symmetrically about the axis.

Figure 33.54 Principal ray diagram for an object outside the focus of a concave spherical mirror. The image is real, inverted, and smaller than the object.



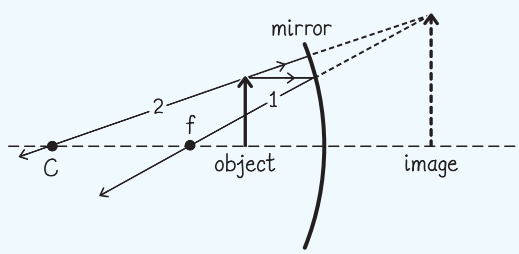
Just as for lenses, the focal length f for any spherical mirror is positive for a real focus and negative for a virtual focus, and o is positive when the object is in front of the mirror and negative when it is behind the mirror (that can happen only when the object is an image formed by another mirror or a lens). Similarly, i is positive for a real image and negative for a virtual image. With mirrors, however, a real image is located on the same side of the mirror as the object and a virtual image is located on the opposite side—the opposite of what happens with lenses. These sign conventions are summarized in Table 33.2. Finally, the relationship between object and image distances and heights for lenses (Eq. 33.12) also applies to mirrors, so that equation can be used to determine the size of the images formed by mirrors.

Example 33.10 Funny mirror

An object is placed 0.30 m in front of a converging mirror for which the radius of curvature is 1.0 m. (*a*) On which side of the mirror is the image? Is the image real or virtual? (*b*) If the object is 50 mm tall, what is the height of the image?

❶ **GETTING STARTED** To visualize the situation, I draw the mirror and its axis, and indicate its center of curvature C and its focal point f halfway between the mirror surface and C. I represent the object as a solid arrow (**Figure 33.55**).

Figure 33.55



❷ **DEVISE PLAN** To locate the image and identify whether it is real or virtual, I can draw the principal rays. I can use Eq. 33.12 to obtain the image height from the object height and the image and object distances. I can determine the image distance from the focal length and the object distance using Eq. 33.24.

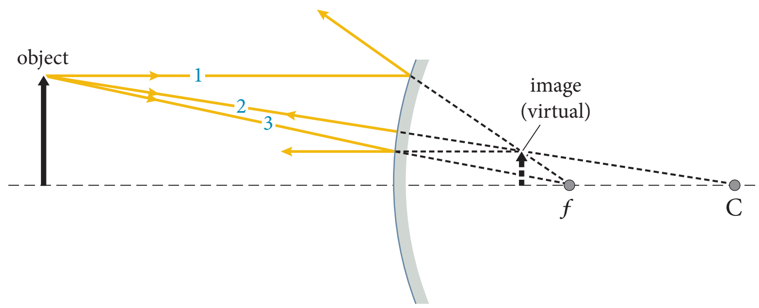
❸ **EXECUTE PLAN** (*a*) As Figure 33.55 shows, I need to draw only two principal rays because the intersection of two rays unambiguously determines the position of the image. I draw ray 1 parallel to the mirror axis and reflecting through the focus. Because the object is between the mirror and the center of curvature, I draw ray 2 along the line defined by C and the tip of the object. Ray 2 does not pass through C until after it is reflected from the mirror, however, and I indicate this by adding an arrowhead pointing toward the mirror on the part of the ray to the right of the object and an arrowhead pointing away from the mirror on the part to the left of the object. My diagram shows that the rays do not actually meet but appear to come from a point behind the mirror. Therefore the image is behind the mirror and virtual. ✔

(*b*) I begin by determining the image distance i. The focal length of the mirror is half the radius of curvature, $f = 0.50$ m. Substituting this value and $o = 0.30$ m into Eq. 33.24 and solving for i give me $i = -0.75$ m. The negative sign indicates that the image is virtual and therefore behind the mirror; this is consistent with my result from part *a*. I then solve Eq. 33.12 for h_i (the signed image height) and substitute the values from this problem:

$$h_i = \frac{-i\,h_o}{o} = \frac{-(-0.75\text{ m})(0.050\text{ m})}{(0.30\text{ m})} = 0.13\text{ m.}\ ✔$$

❹ **EVALUATE RESULT** My ray diagram (Figure 33.55) indicates that the image is enlarged and upright, so h_i should be positive and greater than h_o. This agrees with my result.

Figure 33.56 Ray diagram for an image formed by a convex spherical mirror (object distance greater than focal length). The image is virtual, upright, and smaller than the object.



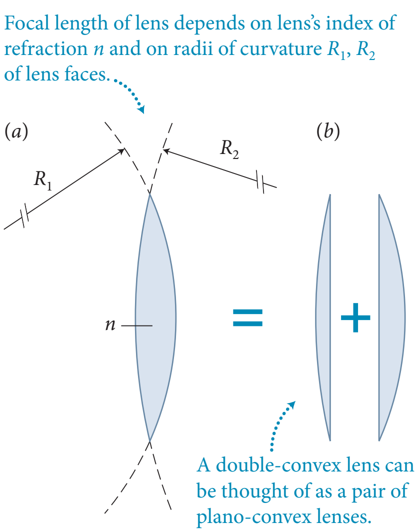
Example 33.10 shows that placing an object inside the focus of a converging mirror produces an upright, virtual image. Compare this with the situation in Figure 33.54, where a real, inverted image is formed for an object placed outside the mirror's focus. The same occurs with a converging lens: Placing the object inside the focus produces an upright, virtual image (Figure 33.31), while placing the object outside the focus produces a real, inverted image (Figure 33.30).

Diverging mirrors, like diverging lenses, always form virtual images when used alone because the light rays must diverge from the mirror surface. **Figure 33.56** shows a ray diagram for an image formed by a diverging mirror. The image is much smaller than the object, which allows a relatively large scene to be captured on a small mirror surface, and is upright. For these reasons, wide-angle rear-view mirrors on the passenger side of cars and trucks and wide-angle surveillance mirrors are typically convex. (A converging mirror also produces small images of distant objects, but the images are inverted, as Figure 33.54 shows.)

33.19 An object is placed 1.0 m in front of a diverging mirror for which the radius of curvature is 1.0 m. (*a*) Where is the image located relative to the mirror? Is the image real or virtual? (*b*) If the object is 0.30 m tall, what is the height of the image?

33.8 Lensmaker's formula

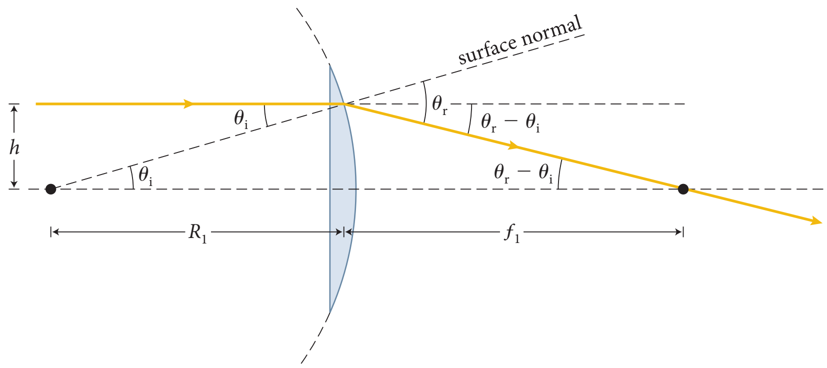
Figure 33.57 Analysis of a double-convex lens.



The focal length of a lens is determined by the refractive index n of the material of which the lens is made and by the radii of curvature R_1 and R_2 of its two surfaces (**Figure 33.57*a***). In this section we work out the relationship among f, R_1, R_2, and n. In this analysis, we can think of a double-convex lens as two plano-convex lenses placed with the two flat surfaces facing each other (Figure 33.57*b*). Remember that both foci of a thin lens are the same distance f from the center of the lens. Because this is true, we can interchange the two surfaces of a thin lens without changing its focal length.

We begin by determining the focal lengths of the two plano-convex lenses in Figure 33.57*b*. **Figure 33.58** shows a ray diagram for light that passes through the right lens only. To calculate the focal length f_1 of this lens, consider a ray incident from the left that comes in parallel to the axis at a distance h above the axis. Because the ray is normal to the planar surface of the lens, it is not refracted at that surface. After passing through the lens, it strikes the curved surface at an angle θ_i measured from the normal to the curved surface and is refracted as it leaves the lens. The refracted ray emerges at an angle θ_r measured from the normal to the curved surface and crosses the lens axis a distance f_1 from the lens. Therefore the angle that the emerging ray makes with the lens axis is $\theta_r - \theta_i$.

Figure 33.58 Ray diagram for light passing through the right-hand plano-convex lens of Figure 33.57*b*.



Applying Snel's law (Eq. 33.7) to this situation gives

$$n \sin \theta_i = \sin \theta_r. \tag{33.25}$$

(We do not need to show an n on the right because the medium is air and $n_{air} = 1$.) For paraxial rays, we can approximate the sines of the angles in Eq. 33.25 by the angles

$$n\theta_i = \theta_r \quad \text{(small angles).} \tag{33.26}$$

Using this relationship, we can express the angle between the emerging ray and the lens axis as

$$\theta_r - \theta_i = n\theta_i - \theta_i = (n - 1)\theta_i. \tag{33.27}$$

For small angles, the angles are approximately equal to their tangents, which means this relationship can be expressed as

$$\theta_r - \theta_i = (n - 1)\theta_i \approx \frac{h}{f_1} \tag{33.28}$$

and

$$\theta_i \approx \frac{h}{R_1}. \tag{33.29}$$

Substituting Eq. 33.29 into Eq. 33.28 and dividing both sides by h, we get

$$\frac{1}{f_1} = \frac{n - 1}{R_1}. \tag{33.30}$$

We can follow the same procedure for the left lens in Figure 33.57*b*, using R_2 as our radius of curvature and a ray that originates at the left focus of that lens. The result analogous to Eq. 33.30 is

$$\frac{1}{f_2} = \frac{n - 1}{R_2}. \tag{33.31}$$

Now let us determine the focal length of the lens combination by working out the lens equation (Eq. 33.16) for the combination, just as we did for the microscope and telescope in Section 33.6. Consider both the object and light source to be on the left side of the lens combination in Figure 33.57. First, the light from the object strikes the left lens from the left and forms an image someplace to the

QUANTITATIVE TOOLS

right of the right lens. The lens equation that relates the location of this object and image, in terms of the focal length f_2 calculated in Eq. 33.31, is

$$\frac{1}{o_2} + \frac{1}{i_2} = \frac{1}{f_2} = \frac{n-1}{R_2}. \tag{33.32}$$

The image formed by the left lens now serves as the (virtual) object for the right lens. Consequently, $o_1 = -i_2$. (The object for the right lens is virtual, and therefore the object distance o_1 is negative because the object is located on the right side of the lens and the illumination comes from the left side of the lens.) The lens equation for the right lens is thus

$$-\frac{1}{i_2} + \frac{1}{i_1} = \frac{1}{f_1} = \frac{n-1}{R_1}, \tag{33.33}$$

where the rightmost equality comes from Eq. 33.30.

Adding Eqs. 33.32 and 33.33 yields

$$\frac{1}{o_2} + \frac{1}{i_1} = (n-1)\left(\frac{1}{R_1} + \frac{1}{R_2}\right). \tag{33.34}$$

The lens equation for the lens as a whole is simply

$$\frac{1}{o} + \frac{1}{i} = \frac{1}{f}, \tag{33.35}$$

where f, o, and i are the focal length, object distance, and image distance of the lens combination, respectively. The object of the left lens is the actual object, and the image formed by the right lens is the final image, which means that in Eq. 33.35 $o = o_2$ and $i = i_1$. Comparing Eqs. 33.34 and 33.35 gives us the **lensmaker's formula** for the focal length of our lens combination:

$$\frac{1}{f} = (n-1)\left(\frac{1}{R_1} + \frac{1}{R_2}\right). \tag{33.36}$$

Our derivation was for a double-convex lens, but it can be shown that the lensmaker's formula applies to any thin lens, not just a double-convex lens. The radii of curvature are positive for convex surfaces, negative for concave surfaces, and infinity for planar surfaces. For a double-convex lens, f is positive. For a double-concave lens, f is negative (because $n > 1$ for any material used for lenses).

33.20 How should the lensmaker's formula be modified if a lens for which the index of refraction is n_1 is submerged in a medium for which the index of refraction is n_2?

Chapter Glossary

SI units of physical quantities are given in parentheses.

absorbed, reflected, and **transmitted light** Light that enters a material but never exits again, light that is redirected away from the surface of the material, and light that passes through a material, respectively.

angle of incidence θ_i The angle between a ray that is incident on a surface and the normal to that surface

angle of reflection θ_r The angle between a ray that is reflected from a surface and the normal to that surface.

angle of refraction θ The angle between a ray that is refracted after crossing the surface between one medium and another and the normal to that surface.

critical angle θ_c (unitless) The angle of incidence for which the angle of refraction equals 90° when a ray travels from a medium with an index of refraction n_2 to one with an index of refraction $n_1 < n_2$:

$$\theta_c = \sin^{-1}\left(\frac{n_1}{n_2}\right). \tag{33.9}$$

dispersion The spatial separation of waves of different wavelength caused by a frequency dependence of the wave speed.

Fermat's principle The path taken by a light ray between any two locations is the path for which the time interval needed to travel between those locations is a minimum.

focal length f (m) The distance f from the center of the lens or the surface of the mirror to the focus. The value of f is positive for a converging lens or mirror and negative for a diverging lens or mirror.

focus (also called **focal point**) The location where parallel rays come together. If the rays cross at the focus, the focus is *real*. If only the extensions of the rays cross at the focus, it is *virtual*.

image An optical likeness of an object produced by a lens or mirror. The image is at the point from which the rays emanating from the surface of the lens or mirror appear to originate. If the rays travel through the point from which they appear to come, the image is *real;* if they do not travel through that point, the image is *virtual*.

index of refraction n (unitless) The ratio of the speed of light in vacuum to the speed of light in a medium:

$$n \equiv \frac{c_0}{c}. \tag{33.1}$$

law of reflection The angle of reflection for a ray striking a smooth surface is equal to the angle of incidence, and both angles are in the same plane.

lens An optical element that redirects light in order to form images. A *converging lens* directs parallel incident rays to a single point on the other side of the lens. A *diverging lens* separates parallel incident rays in such a manner that they appear to all come from a single point on the side of the lens where the rays came from.

lens equation The equation that relates the object distance o, the image distance i, and the focal length f of a lens or mirror:

$$\frac{1}{f} = \frac{1}{o} + \frac{1}{i}. \tag{33.16}$$

lensmaker's formula The relationship among the focal length f of a lens, the refractive index n of the material of which the lens is made, and the radii of curvature R_1 and R_2 of its two surfaces:

$$\frac{1}{f} = (n - 1)\left(\frac{1}{R_1} + \frac{1}{R_2}\right). \tag{33.36}$$

magnification M (unitless) The ratio of the signed image height h_i ($h_i > 0$ for upright image, $h_i < 0$ for inverted image) to the object height h_o:

$$M \equiv \frac{h_i}{h_o} = -\frac{i}{o}. \tag{33.17}$$

The *angular magnification* is defined as the ratio of the angle θ_i subtended by the image and the angle θ_o subtended by the object:

$$M_\theta \equiv \left|\frac{\theta_i}{\theta_o}\right|. \tag{33.18}$$

Provided these angles are small and for an object that is placed close to both the focus of the lens and the eye's near point, the angular magnification is $M_\theta \approx (0.25\ \text{m}/f)$.

principal rays a set of rays that can be used in ray diagrams to determine the location, size, and orientation of images formed by lenses or spherical mirrors.

ray A line that represents the direction in which light travels. A beam of light with a very small cross-sectional area approximately corresponds to a ray.

ray diagram A diagram that shows just a few selected rays, typically the so-called *principal rays* (see the Procedure boxes on pages 888 and 902).

refraction The changing in direction of a ray when it travels from one medium to another.

Snel's law The relationship among the indices of refraction n_1 and n_2 of two materials and the angle of incidence θ_1 and angle of refraction θ_2 at the interface of the materials

$$n_1 \sin\theta_1 = n_2 \sin\theta_2. \tag{33.7}$$

total internal reflection Mirrorlike reflection that occurs when a ray traveling in a medium strikes the medium boundary at an angle greater than the critical angle. The ray is completely reflected back into the medium.

34 Wave and Particle Optics

CONCEPTS

QUANTITATIVE TOOLS

In Chapter 33, we considered the propagation of light along a straight path. The chapter title, "Ray Optics," reflects the fact that we considered propagating light only in the simplest way—as straight-line motion. You know from Chapter 30, however, that light is an electromagnetic wave. This means that it must undergo interference and diffraction, just like any mechanical wave. As you will learn in this chapter, light waves can interfere with one another and diffract when they pass through small openings.

Another fact about light you will learn in this chapter is that it has a dual nature: It is a wave, yes, but also has the properties of a particle!

34.1 Diffraction of light

As we saw in Chapter 17, when a water wave strikes a barrier that has a small opening, the wave diffracts (spreads out) after it passes through the opening. **Figure 34.1*a***, for example, shows surface water waves diffracting nearly circularly after they pass through an opening.

Given that light is a wave, as we discussed in Chapter 30, why don't we see light diffract in a similar fashion after it travels through, say, a window? As Figure 34.1*b* shows, after passing through a window, light continues to travel in a straight line, casting a sharp-edged shadow with no discernible diffraction.

The reason light does not diffract through a window is that the wavelength of the light is very much smaller than the size of the window. In Figure 34.1*a*, the wavelength of the water wave is about the same as the width of the opening, but the wavelength of the light in Figure 34.1*b* is about a million times smaller than the width of the window.

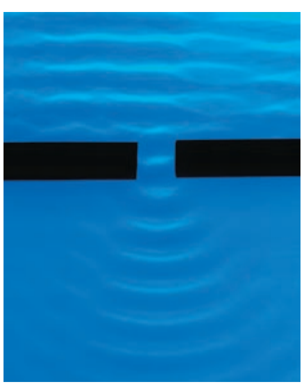
Diffraction is indeed observed with light waves but only when the width of the opening through which the light passes is not much greater in size than the wavelength of the light. Empirical evidence shows that diffraction occurs through openings approximately two orders of magnitude greater than the wavelength. Thus, visible light, with a wavelength on the order of 1 μm, diffracts through apertures up to hundreds of micrometers wide.

To understand diffraction, it is useful to consider the propagation of wavefronts. As discussed in Sections 17.1 and 33.2, a wavefront is a surface on which a wave spreading through space has constant phase. Wavefronts are everywhere perpendicular to the direction of propagation of the wave. By convention, wavefronts are drawn at the crests of the waves, which means the separation between adjacent wavefronts equals the wavelength (Figure 17.2). Although in principle wavefronts can take any shape, usually we consider only those that are either planar or spherical. Most sources of light, from light bulbs to stars, can be modeled as point sources—single points that produce concentric spherical wavefronts. As discussed in Section 17.1 (see especially Figure 17.7), far away from a point source, the radius of the spherical wavefronts is so great that the wavefronts are very nearly planar. As a result, distant point sources can be considered to be sources of planar waves. Lasers also produce planar waves, even very close to the source. For this reason, we can use a laser beam in seeing how electromagnetic waves behave. Keep in mind, however, that our analysis applies to any type of electromagnetic radiation, not just to laser beams.

Let us now determine under what conditions a planar wave spreads out as it propagates—in other words, under what conditions it undergoes diffraction. **Figure 34.2** shows planar wavefronts from a beam of electromagnetic waves propagating to the right, with point Q located at the upper end of the wavefronts and point P located outside the region reached by the wavefronts. Because a wavelet (see Section 17.4) centered on Q radiates toward P along the line QP, we expect the beam to spread out as it propagates. However, such spreading is not observed. The reason is that the wavelets centered on points below Q also radiate toward P,

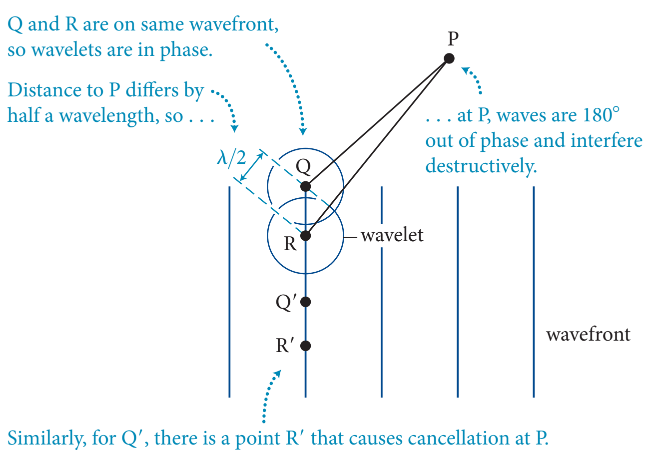
Figure 34.1 Water waves diffract when they pass through a gap, whereas light coming in through a window seems not to diffract—it forms a sharp-edged shadow. Notice that the gap in the breakwater is roughly as wide as the wavelength of the water waves.

(*a*)



(*b*)

Figure 34.2 The reason we don't usually see diffraction for light beams, provided the beam is very much wider than the wavelength of the light waves.



CONCEPTS

and we need to sum the contributions of all these wavelets. Consider, for example, the wavelet centered on R, for which the distance PR is exactly half a wavelength longer than the distance PQ. Because Q and R lie on the same wavefront, they produce coherent wavelets. (You should review Section 17.4 if you do not see why this is true.) This means that at P the electric field part of the wavelet from Q is 180° out of phase with the electric field part of the wavelet from R. Thus the two electric fields interfere destructively at P. Because the same is true for the magnetic field part of the wavelets, there is complete destructive interference (see Section 16.3) at P.

In the same manner, for points Q′ lying below Q on the wavefront, we can find on the same wavefront a point R′ for which the distance R′P is exactly half a wavelength longer than the distance Q′P. Thus the fields from the wavelet traveling along R′P cancel those from the wavelet traveling along Q′P. If the wavefronts extend far enough below Q, we can always find points that cancel the radiation from any other point. As a result, we conclude that the light does not spread outside the beam; in other words, there is no diffraction.

34.1 Consider the point P located ahead of the wavefronts shown in **Figure 34.3**. Following the line of reasoning used in the preceding discussion, what can you say about the intensity (power/area, as defined in Eq. 30.34 and Section 17.5) at P once the wave fronts reach it? (Hint: Consider separately points R above, below, and on the ray through P.)

Checkpoint 34 .1 demonstrates that the intensity of a planar wave is uniform as the wave propagates forward because the fields from individual wavelets cancel in any outward direction and reinforce only in the direction of propagation. Section 17.3 (especially Figure 17.20) shows that combining the waves from many adjacent point sources indeed produces a planar wave that propagates forward with very little spreading. This is true because the wavelets interfere destructively with one another in any direction other than the direction of propagation of the wavefronts.

The cancellation process we used in Figure 34.2 can be used only when the width of the laser beam, and hence the width of the wavefronts, is much greater than the wavelength of the light, so that all points on a wavefront can be paired with other points on that wavefront that are half a wavelength or more away from that point and the wavelets from these points cancel.

Figure 34.3 Checkpoint 34.1.

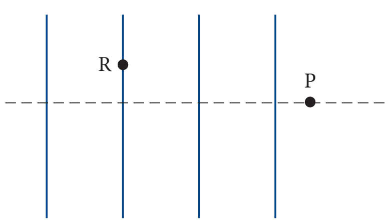
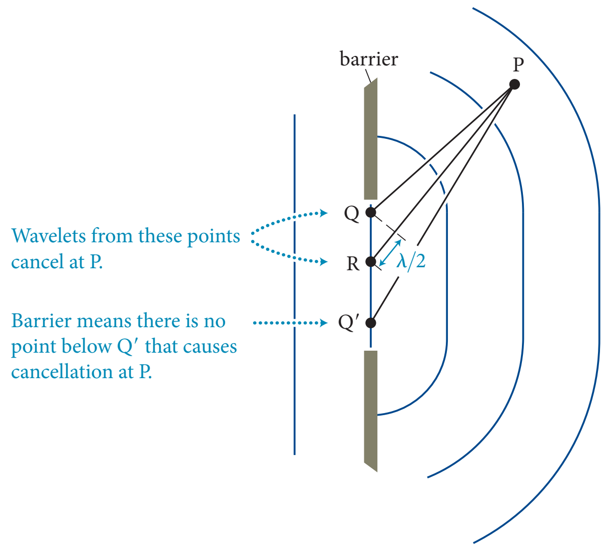


Figure 34.4 The reason we *do* see diffraction when a light beam is transmitted through a small aperture.



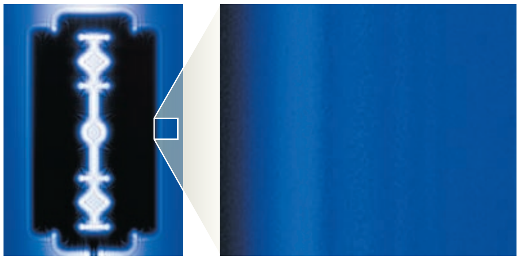
If the width of the beam is comparable to the wavelength of the light, not all points on the wavefronts can be paired in this manner. To see why this is so, let us place a barrier in front of our light source, as in **Figure 34.4**. You can think of this drawing as a bird's-eye view of a beam of light traveling to the right and running into a wall that has a gap in it. Light hitting the wall on either side of the gap cannot pass through. As each wavefront of the beam reaches the barrier, only the portion that hits the gap continues moving to the right. The width of each wavefront that passes through the gap is equal to the gap width.

Suppose the gap in Figure 34.4 has a width equal to 2λ. All radiation that reaches P from wavelets that originate above Q′ cancels, as indicated by the rays drawn from points Q and R in Figure 34.4; wavelets that originate at or below Q′ are not canceled at P because those wavelets lack corresponding wavelets at an appropriate distance below Q′. As a result, some of the light spreads out past the edges of the original path of the beam—the light is diffracted. Diffraction in this situation occurs for exactly the same reason as the diffraction of water waves pictured in Figure 34.1*a*. Diffraction of light occurs when a planar wave passes through an aperture that is only micrometers wide (much less than the width of a hair). As shown in Figure 34.1*a*, if the width of the aperture is equal to or less than the wavelength, the wavefronts coming from the aperture are spherical. If the aperture is a few wavelengths wide, the wavefronts are elongated right after they pass through the aperture, as shown in Figure 34.4.

An ordinary window, like that shown in Figure 34.1*b*, is effectively infinitely wide relative to the wavelength of the light, so only the light at the very edges of the window diffracts. In practice, not even diffraction from the edge of the window is observed because the edge is not perfectly smooth on a micrometer scale. However, it *is* possible to observe diffraction of light from the edge of a smooth razor blade, as shown in **Figure 34.5**.

CONCEPTS

Figure 34.5 Edge diffraction is not usually apparent for visible light because most edges are not smooth enough. However, it can be observed around the edge of a razor blade. The blade in this image is illuminated by a point source of monochromatic light.



Example 34.1 Spreading out

Do you expect to be able to observe the diffraction of light through (*a*) the front door to your house; (*b*) the holes in a button; (*c*) the gaps between threads of the fabric of an umbrella?

❶ **GETTING STARTED** I expect to see noticeable diffraction through openings up to roughly two orders of magnitude times the wavelength of the light. I therefore need to estimate the width of each opening and determine the width-to-wavelength ratio.

❷ **DEVISE PLAN** To estimate the width of the front door and of the holes in the button, I can draw on my experience; to estimate the widths of the gaps between the threads of the fabric of an umbrella, I shall use an upper limit. Then I shall take the ratios of these widths to the wavelength of light in the middle of the visible range, 500 nm.

❸ **EXECUTE PLAN** (*a*) A door is about 1 m wide, so the ratio of the door's width to 500 nm is 2×10^6, much too great to see diffraction. ✔

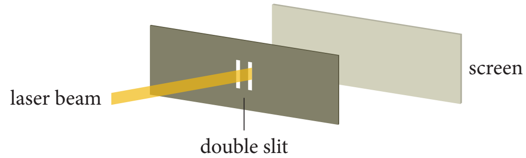
(*b*) The holes in a button are about 1 mm in diameter, so the ratio of this width to 500 nm is 2×10^3, still too great to see diffraction. ✔

(*c*) I know that a human hair, which is less than 100 μm in diameter, cannot easily be threaded through a piece of fabric. Therefore I estimate the gaps between the threads in the fabric to be one-tenth of a hair diameter, or 10 μm at most. The ratio of a gap width to 500 nm is therefore 20 or less, and I expect to see diffraction through the gaps in the fabric. ✔

❹ **EVALUATE RESULT** I know from experience that I do not see diffraction through a doorway. I can check my answers to parts *b* and *c* by looking through the holes in a button and through an open umbrella. When I do so, I see no diffraction through the button but I can see diffraction through the umbrella if the fabric is dark. (This diffraction is particularly noticeable when I look at a streetlight at night through the umbrella.)

34.2 In discussing how a planar wave propagates, we could turn our earlier argument around and say that for each point Q in Figure 34.2 there is a point S somewhere on the wavefront that radiates toward P along a path exactly one wavelength longer than that from Q, and therefore there should be a nonzero intensity at P. What is wrong with this argument?

Figure 34.6 When the planar electromagnetic waves of a laser beam pass through a pair of narrow slits, what do we see on the screen?



34.2 Diffraction gratings

What happens if instead of passing through a single small aperture, a planar electromagnetic wave strikes a barrier that contains two narrow slits at normal incidence, as shown in **Figure 34.6**? If the slit width is much less than the wavelength of the wave, the slits serve as two coherent point sources of electromagnetic waves of the same wavelength as the wave striking the barrier, and the waves from the two sources interfere with one another in the manner described in Section 17.3 for two adjacent sources of surface waves. The only difference is that electromagnetic waves are not confined to a planar surface but spread out in three dimensions. If the two slits are either round or square, the waves that emerge from them are spherical. If the slits are much taller than they are wide, as in Figure 34.6, the slits serve as lines of point sources and the waves that emerge from the slits have cylindrical wavefronts.

The crests of the waves from these two coherent point sources overlap in certain directions, as shown in **Figure 34.7** (and Figure 17.13). Between these directions where there is overlap are directions along which the waves from the two sources cancel. If we place a screen to the right of the slits, a pattern of alternating bright and dark bands appears on the screen, as shown in **Figure 34.8** on the next page. Such dark and bright bands are commonly called **interference fringes.** The bright fringes are labeled by a number called the **fringe order** m. The central bright fringe is the *zeroth-order bright fringe* ($m = 0$); around it are higher-order bright fringes ($m = 1, 2, \ldots$). Note that the pattern is symmetrical about this zeroth-order bright fringe.

Figure 34.7 Interference between the diffracted waves emerging from the two slits.

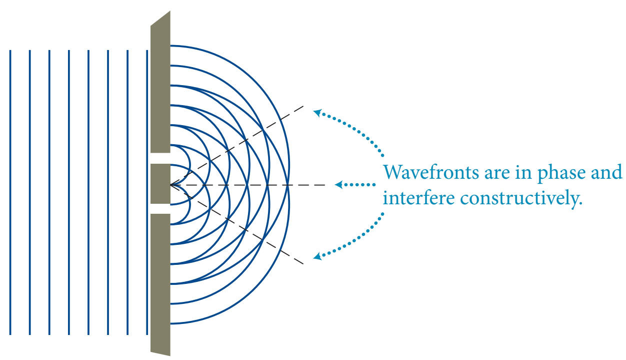
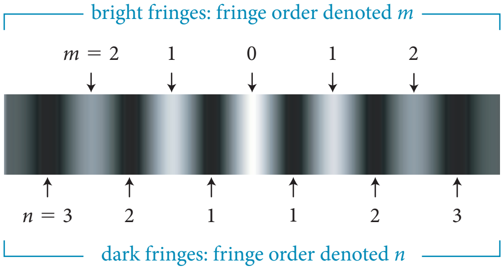


Figure 34.8 The interference pattern produced when the laser beam of Figure 34.6 passes through a pair of slits and strikes the screen.



How do we determine the locations of the bright fringes? The central (zeroth-order, $m = 0$) bright fringe is simplest to locate. As shown in Figure 34.7, waves from the two sources interfere constructively along a perpendicular line running through the midpoint between the two slits. Because the waves coming from the two slits travel the same distance to reach any point along this perpendicular line, the waves arrive in phase with each other.

To locate the other bright interference fringes, we need to work out the directions in which the difference in path length between waves coming from the upper source and waves coming from the lower source is an integer number of wavelengths. In these directions, constructive interference between waves from the two sources produces bright fringes.

Let's consider two rays, one from each slit, that meet at the screen to form a fringe, as shown in **Figure 34.9*a***. In general we shall take the distance from the sources to the screen to be much greater than the distance between the slits. Note from the figure how, even though the rays eventually meet at the screen, their paths are essentially parallel when they emerge from the slits (Figure 34.9*b*). If we denote the angle between the nearly parallel rays and the normal to the barrier as θ, we can say that the difference in path length for waves emitted at angle θ from the two sources is $d \sin \theta$, where d is the distance between the slits (Figure 34.9*b*). When this path-length difference is equal to an integer multiple of the wavelength, constructive interference occurs. The central bright fringe corresponds to $d \sin \theta = 0$ or $\theta = 0$. Subsequent bright fringes are located at angles given by $d \sin \theta_m = \pm m\lambda$, where $m = 1, 2, \ldots$ denotes the order of the bright fringe. The plus and minus signs give rise to bright fringes on either side of the central maximum.

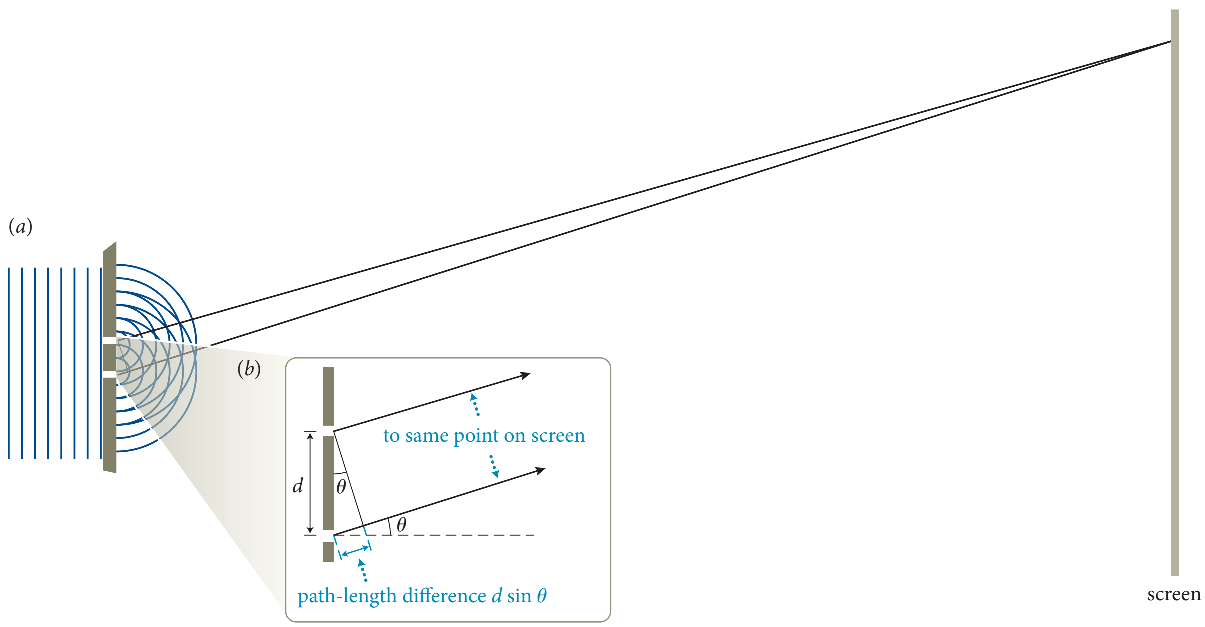
Likewise, in directions that correspond to path-length differences of an odd number of half-wavelengths, waves from the two sources interfere destructively, resulting in dark fringes. For these fringes, we use n to denote fringe order. The smallest angle at which destructive interference occurs corresponds to $d \sin \theta = \frac{1}{2}\lambda$. More generally, the angles at which dark fringes occur are given by $d \sin \theta_n = \pm(n - \frac{1}{2})\lambda$, where $n = 1, 2, \ldots$ denotes the order of the dark fringe. The dark fringes around the zeroth-order bright fringe are the first-order dark fringes.

Example 34.2 Two-slit diffraction grating

Coherent green light of wavelength 530 nm passes through two very narrow slits that are separated by 1.00 μm. (*a*) Where is the first-order bright fringe? (*b*) What is the angular separation between the $n = 1$ and $n = 2$ dark fringes?

❶ GETTING STARTED This problem involves interference between light rays passing through two closely spaced, very narrow slits, as shown in Figure 34.9.

Figure 34.9 Determining the path-length difference between two rays traveling to a point on a distant screen in a two-slit interference setup.



CONCEPTS

❷ **DEVISE PLAN** The angular locations of the centers of the bright fringes are given by the condition for constructive interference, $d \sin \theta_m = m\lambda$; the angular locations of the dark fringes are given by the condition for destructive interference, $d \sin \theta_n = (n - \frac{1}{2})\lambda$. (For both bright and dark fringes, I omit the $\pm$ signs because I'll only consider the fringes on one side of the central maximum.) I therefore need to use these relationships to calculate the angular locations of the fringes. To use the constructive and destructive interference conditions, I need to identify the appropriate values of m and n and then use them with the wavelength and the distance between slits to obtain the angular positions of the bright and dark fringes I am interested in.

❸ **EXECUTE PLAN** (*a*) The first-order bright fringe corresponds to $m = 1$, which means the center of the first-order bright fringe is located at the value θ_1 corresponding to $d \sin \theta_1 = \lambda$. Substituting for d and λ in this expression and solving for θ_1, I obtain

$$\theta_1 = \sin^{-1}\left(\frac{0.530\ \mu\text{m}}{1.00\ \mu\text{m}}\right) = 32.0°. ✔$$

(*b*) The two lowest-order dark fringes correspond to $n = 1$ and $n = 2$, and their centers occur at the angles corresponding to $d \sin \theta_1 = \lambda/2$ and $d \sin \theta_2 = 3\lambda/2$. Substituting and solving, I obtain

$$\theta_1 = \sin^{-1}\left(\frac{0.530\ \mu\text{m}}{2 \times 1.00\ \mu\text{m}}\right) = 15.4°$$

$$\theta_2 = \sin^{-1}\left(\frac{3 \times 0.530\ \mu\text{m}}{2 \times 1.00\ \mu\text{m}}\right) = 52.7°.$$

The angular separation is thus $52.7° - 15.4° = 37.3°$. ✔

❹ **EVALUATE RESULT** The center of the $m = 1$ bright fringe is roughly halfway between the centers of the $n = 1$ and $n = 2$ dark fringes, as I expect from Figure 34.8.

34.3 Does the spacing of the bright fringes in the two-slit arrangement in Figure 34.6 increase, decrease, or stay the same if we (*a*) increase the spacing d of the slits, or (*b*) increase the wavelength λ of the light incident on the arrangement?

Now consider the effect of many equally spaced narrow slits in a barrier on which planar waves are incident normally (**Figure 34.10**). Once again, we can determine the fringe pattern produced at a given location on a distant screen by combining all the waves (one from each slit) that travel to that location. The condition for constructive interference among all the waves is equivalent to that for two slits. As can be seen from Figure 34.10, the path-length difference between each pair of adjacent waves is the same. This means that if the waves from two of the slits interfere constructively, the waves from *all* of the slits do the same. Bright fringes therefore appear at angles corresponding to $d \sin \theta_m = \pm m\lambda$, with d the separation between adjacent slits. The location of bright fringes therefore does not depend on the number of slits as long as the separation d between adjacent slits is the same for all slits.

Figure 34.10 Path-length difference for planar waves striking a barrier with multiple slits.

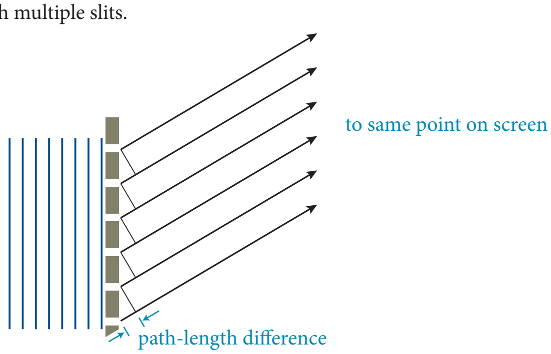
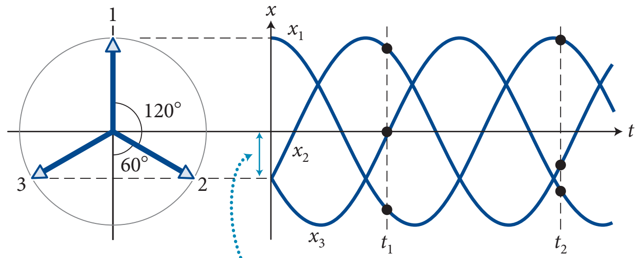


Figure 34.11 Three coherent waves interfere destructively when each is out of phase with the other two by one-third of a cycle.



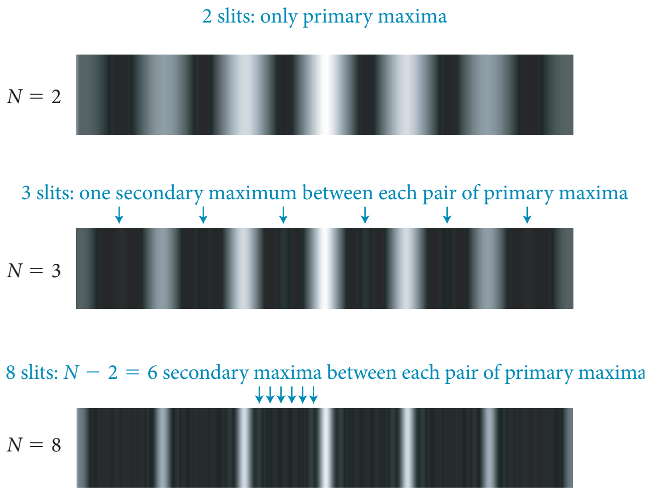
34.4 Suppose there are three slits in a barrier on which light is incident normally, with each slit separated from its neighbor by a distance d. Do the waves from all three slits cancel perfectly at angles given by $d \sin \theta = \pm(n - \frac{1}{2})\lambda$?

As Checkpoint 34.4 illustrates, the condition for complete destructive interference of all of the waves is not the same for three slits as for two. Instead, as **Figure 34.11** shows, three coherent waves cancel one another perfectly when each is out of phase with the other two by one-third of a cycle. The three waves also cancel one another perfectly when each is out of phase with the other two by two-thirds of a cycle. In that case, phasors 2 and 3 in Figure 34.11 are interchanged. So there are two dark fringes between each pair of bright fringes. In between these two dark fringes is a faint bright fringe (**Figure 34.12** on the next page). The brightest fringes are the *principal maxima* in the interference pattern. These correspond to constructive interference of the waves diffracted by all three slits. The fainter bright fringes are *secondary maxima*. At these locations the cancellation is not complete.

Four coherent waves cancel one another when adjacent waves are out of phase by one-fourth of a cycle. In the same manner, if there are N slits, the condition for complete destructive interference is that each wave must differ in phase by $1/N$ of a cycle from its immediate neighbors. Then the N waves are evenly distributed throughout one cycle of oscillation and add to zero. The condition for the path-length differences for the dark fringes is thus $d \sin \theta_k = \pm(k/N)\lambda$, where k is any integer that is *not* a whole-number multiple of N (because when $k/N = m$, we have constructive interference).

CONCEPTS

Figure 34.12 Interference pattern caused by the diffraction of a coherent beam of light through two, three, and eight narrow slits.



Although the bright fringes are in the same location regardless of the number of slits, there are now $N - 1$ dark fringes between the bright fringes and $N - 2$ secondary maxima between each pair of principal maxima. As a result, as N increases, the bright fringes become narrower and brighter, as shown in Figure 34.12. (The brightness of the pattern corresponds to the intensity of the light striking the screen.)

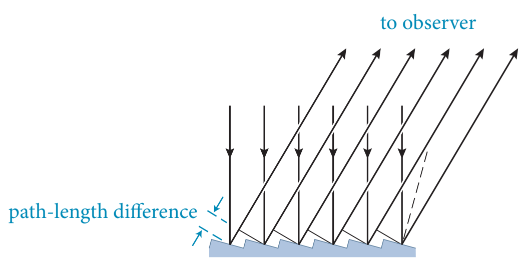
34.5 Why does the brightness of the fringes increase as the number of slits increases?

The interference of a planar electromagnetic wave as it passes through many closely spaced narrow slits is due to the diffraction that occurs at the slits. A barrier that contains a very large number of such slits is therefore called a **diffraction grating.** Diffraction gratings can be either transmissive (such as the one shown in Figure 34.10) or reflective. Reflective diffraction gratings are made by engraving grooves to reflect light from a surface, as shown in **Figure 34.13**. The grooves on a music compact disc (opening picture in this chapter) form a reflective diffraction grating.

Why is the light reflected from the compact disc surface so colorful? You found in Checkpoint 34.3 that the position of interference fringes produced by light of a single color depends on the wavelength. When white light, which contains many different colors and therefore many different wavelengths, falls on a diffraction grating, the fringes for each wavelength are displaced from each other, producing a series of rainbows.

34.6 Suppose the light striking the reflective diffraction grating in Figure 34.13 is white light—that is, light consisting of all the colors of the rainbow. Red light has the longest wavelength of the colors that make up white light; violet has the shortest. (*a*) Is the angle at which first-order constructive interference occurs for violet light less than, equal to, or greater than the angle for red light? (*b*) Why are multiple rainbows visible in the reflected light?

Figure 34.13 Reflective diffraction grating.



Very precisely manufactured diffraction gratings have many uses in scientific equipment. They are most widely used to disperse visible, ultraviolet, or infrared light into its constituent wavelengths because the resulting spectrum provides information about the object that emitted the light. For example, astronomers often use a diffraction grating attached to a telescope to identify the wavelengths present in the light from stars, in order to understand the chemical composition of the stars or their distance from Earth.

Certain common objects can also function as diffraction gratings. For example, if you look through a piece of dark, finely woven, taut fabric at a point source of light (say, a distant street light through the fabric of a dark umbrella), you will see fringes.

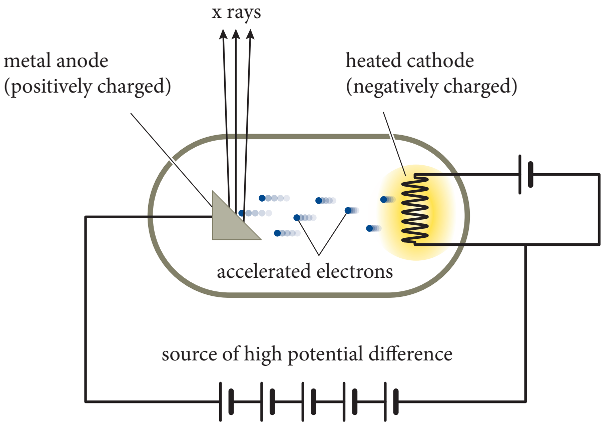
34.7 Diffraction gratings used in astronomical instruments must be able to separate wavelengths that are quite close together. (*a*) To increase the ability to do this, should the separation between slits be made less or greater? (*b*) Does the width of the slits affect the diffraction pattern?

34.3 X-ray diffraction

The interference of very-short-wavelength electromagnetic radiation is widely used to study the structure of materials. **X rays** are electromagnetic waves that have wavelengths ranging from 0.01 nm to 10 nm, more than 100 times less than the wavelengths of visible light.

Figure 34.14 shows one way to generate X rays. Electrons are ejected from a heated cathode on the right and accelerated by a potential difference of several thousand volts toward a metal anode on the left. The electrons crash into the atoms that make up the anode; this inelastic collision decelerates the electrons very rapidly and gives the atoms a great deal of internal energy. The atoms then re-emit this energy in the form of x rays. In addition, the rapidly decelerating electrons radiate x rays, typically at a 90° angle to the path of the accelerated electrons.

Figure 34.14 Schematic diagram for a cathode ray tube x-ray emitter.

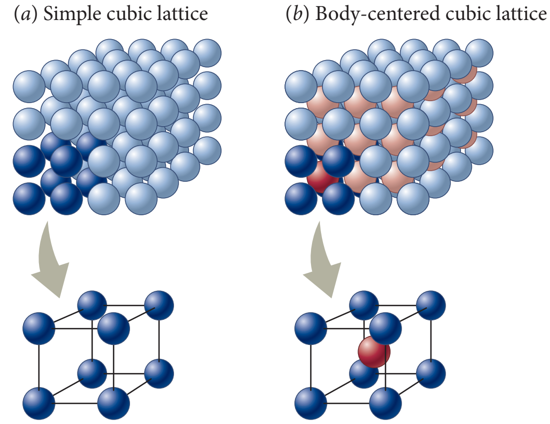


X rays can pass through many soft materials with low mass density that are opaque to visible light. For example, they pass through soft tissues in the human body but are strongly absorbed by bones and teeth. As a result, x rays are widely used to obtain photographic images of the skeleton (**Figure 34.15**). X-ray imaging of blood vessels or soft internal organs can be done by giving the patient a drug containing heavy atoms, such as iodine, because the heavy atoms absorb x rays. For example, one way cardiologists diagnose heart problems is to directly observe blood vessels in a patient's heart by injecting a heavy-atom drug into the patient's blood and taking x-ray movies of the beating heart as the blood is pumped through.

Because x-ray wavelengths are either shorter than or comparable to the typical distance between atoms in solid materials (0.1 nm to 1 nm), x-ray diffraction can be used to study atomic arrangements in solids. Many solids are *crystalline*, meaning that their atoms are arranged in a three-dimensional, regularly spaced grid called a *crystal lattice* (**Figure 34.16**). The lattice serves as a three-dimensional diffraction grating for x rays because the lattice spacing is comparable to the x-ray wavelength.

Figure 34.15 Bones and teeth absorb x rays, whereas soft tissues are nearly transparent to them.

Figure 34.16 Two examples of crystal lattices.



Consider what happens when x rays strike the top plane of atoms in a crystal lattice* at an angle θ (**Figure 34.17**). Each atom that is struck by the beam of x rays acts as the source of a wavelet emitting waves in all directions, much like the slits of the diffraction gratings we discussed in the preceding section. Waves emitted at $\theta' = \theta$ have the same path length and so they add constructively, yielding a strong reflected beam.

34.8 Considering only the top row of atoms in Figure 34.17, are there any other directions in which the x rays diffracted by the atoms interfere constructively?

As you saw in Checkpoint 34.8, x rays diffracted by the crystal in directions other than at angle $\theta' = \theta$ are much weaker than those diffracted at angle $\theta' = \theta$. We might therefore expect not to see any x-ray diffraction from crystals at other angles. However, a crystal consists of many planes of atoms, and some incident x rays penetrate into the

* The lattice shown is a so-called cubic lattice. Lattices can be much more complex than the cubic lattice, but the principles of diffraction are the same.

Figure 34.17 Diffraction of x rays by the atoms at the surface of a crystal lattice.

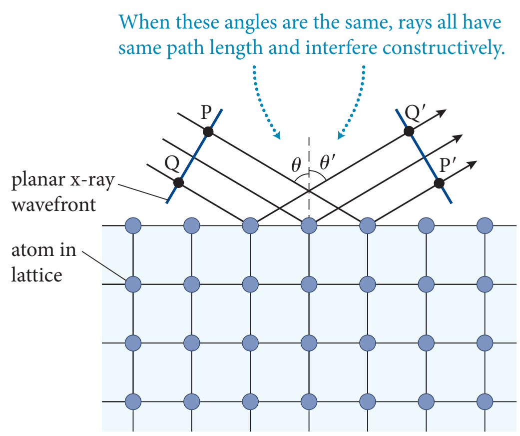


Figure 34.18 Interference of x rays diffracted by adjacent planes of a crystal.

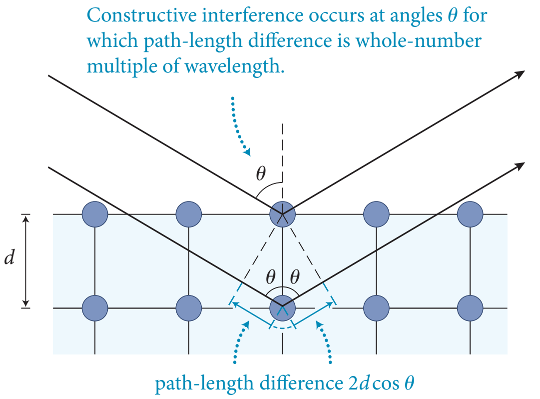
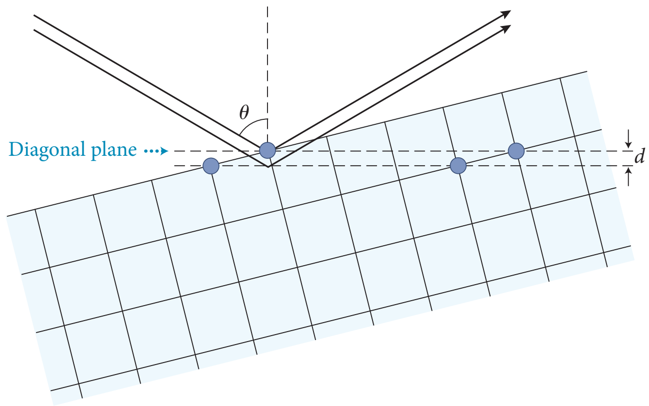


Figure 34.19 Constructive interference of x rays diffracted by two diagonal crystal planes.



crystal. We thus need to take into account the diffraction of the x rays by the atoms in multiple crystal planes to determine the diffraction of the crystal as a whole.

Figure 34.18 shows the x rays diffracted by atoms in two adjacent planes of a crystal. For most angles of incidence, the waves diffracted by atoms in different planes differ in phase and so interfere destructively. However, when the difference in path length between rays diffracted by atoms in different planes is a whole-number multiple of the x rays' wavelength, the rays are in phase and interfere constructively. The path-length difference equals $2d \cos \theta$, where d is the distance between adjacent planes. Therefore the condition for constructive interference is $2d \cos \theta = m\lambda$. This condition is called the **Bragg condition,** after the father-and-son team of physicists who formulated it. Because the atomic spacing and the x-ray wavelength are fixed, crystals reflect x rays only at those angles for which $2d \cos \theta$ is an integer multiple of the x-ray wavelength.

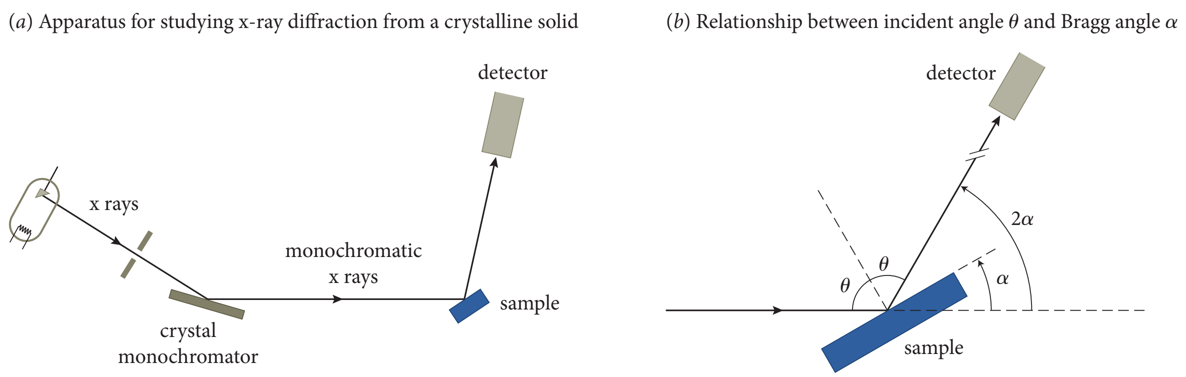
The atoms in the lattice of a crystal define many different lattice planes. **Figure 34.19** shows two sets of lattice planes in a cubic crystal, one indicated by dashed lines (planes parallel to the surface of the crystal) and the other indicated by solid lines (diagonal planes). If we tilt the crystal so that the angle the incident x rays make with its surface is different from the angle shown in Figure 34.18, a different set of planes with a different spacing can produce constructive interference of the diffracted waves.

By measuring the angles at which strong x-ray diffraction occurs, one can determine the arrangement of atoms in a crystalline solid. **Figure 34.20** shows how such a measurement is carried out. An x-ray source like that shown in Figure 34.14 is used to produce a beam of x rays of various wavelengths. The beam is then diffracted from a *crystal monochromator* (which is simply a crystal of known lattice spacing) positioned at an angle chosen so that the Bragg condition is satisfied for one desired wavelength of x rays. Because the other wavelengths in the original beam do not satisfy the Bragg condition, a *monochromatic* (single-wavelength) beam of x rays is diffracted from the monochromator to the sample.

The sample of crystalline material whose lattice is being studied is slowly rotated with respect to the monochromatic beam, and as this rotation takes place, the intensity of x rays diffracted from the sample is measured on a detector as a function of the angle α between the x rays and the sample surface (Figure 34.20*b*). This angle is often called the *Bragg angle* α. As this angle changes, the Bragg condition is

CONCEPTS

Figure 34.20 (*a*) Apparatus for studying x-ray diffraction from a crystalline solid. (*b*) Relationship between the incident angle θ and the Bragg angle α.



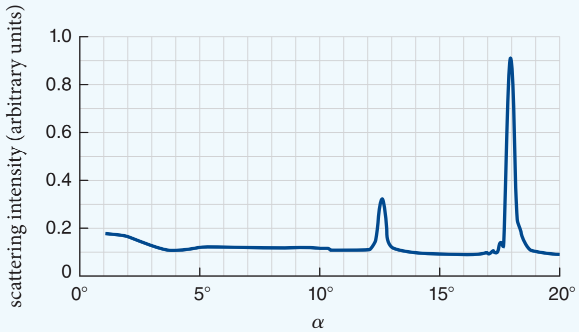
generally not satisfied. However, at specific Bragg angles, the various crystal planes in the sample satisfy the Bragg condition, producing a high intensity of x rays on the detector.

34.9 Express the Bragg condition $2d \cos\theta = m\lambda$ in terms of the Bragg angle α between the x rays and the surface of the sample rather than the angle θ between the x rays and the surface normal.

Example 34.3 X-ray diffraction

Figure 34.21 shows diffracted x-ray intensity as a function of the Bragg angle α, obtained using x rays having a wavelength of 0.11 nm. (*a*) Without calculating values for the lattice spacing d, identify which of the two peaks corresponds to a greater distance between adjacent planes in the sample being studied. (*b*) Calculate the distance between adjacent planes corresponding to each peak.

Figure 34.21 Example 34.3.



❶ **GETTING STARTED** The peaks in a graph of x ray intensity as a function of the Bragg angle α correspond to constructive interference and therefore values of α that satisfy the Bragg condition.

❷ **DEVISE PLAN** For part *a*, I can use the Bragg condition, together with the shape of the graph, to deduce which peak results from the greater d value. Because the graph gives intensity in terms of the Bragg angle, I shall need the Bragg condition expressed in terms of the Bragg angle.

For part *b*, I can solve this form of the Bragg condition for d and then insert my two given α values to determine the plane separation distance in each case. Looking at this form of the Bragg condition, I see that the Bragg angle α at which a peak occurs increases as m increases. Because this graph begins at $\theta = 0°$, the two peaks must correspond to $m = 1$ for the interference patterns when the crystal surface is oriented at the two Bragg angles I am working with.

❸ **EXECUTE PLAN** (*a*) From Checkpoint 34.9 I know that the Bragg condition in terms of the Bragg angle is $2d \sin\alpha = m\lambda$. In order for the product $2d \sin\alpha$ to remain constant, α must decrease as the distance d between adjacent planes increases. Therefore the peak at the smaller α value corresponds to the greater distance between planes. ✔

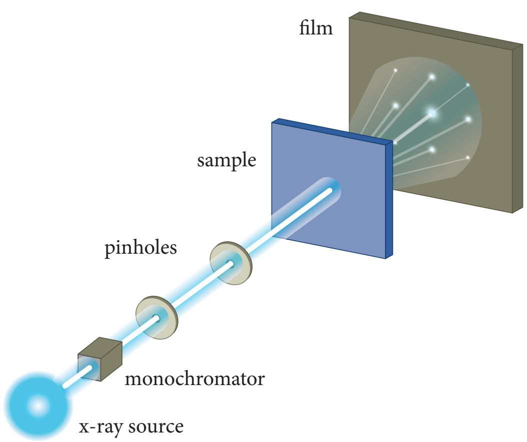
(*b*) To obtain d for each peak, I solve $2d \sin\alpha = m\lambda$ for d and then substitute $m = 1$ and the values for α and λ. For the short peak, $\alpha = 12.5°$, which gives $d = \lambda/(2 \sin\alpha) = (0.11\ \text{nm})/(2 \sin 12.5°) = 0.25$ nm. For the tall peak, $\alpha = 18°$, which gives $d = \lambda/2 \sin\alpha = (0.11\ \text{nm})/(2 \sin 18°) = 0.18$ nm. ✔

❹ **EVALUATE RESULT** The Bragg angle α at which constructive interference occurs decreases as the distance between planes increases. This is consistent with what I found previously for interference between two slits (Checkpoint 34.3), in which increasing the distance between slits also causes the angle between fringes to decrease. In general, the size of an interference pattern decreases as the distances between interfering sources increase. Finally, the smaller value of d multiplied by $\sqrt{2}$ gives the greater value of d, as it should for the distances between planes for the cubic lattice in Figure 34.19.

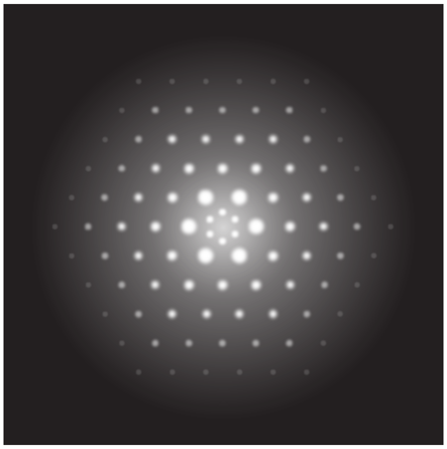
Many studies of crystal structure are done by passing x rays through the crystal rather than reflecting them from the various crystal planes. The crystal then acts like a three-dimensional transmissive diffraction grating for the x-ray beam. Instead of a single line of slits, the beam encounters many rows of slits. As a result, many rows of fringes usually called "spots" are formed. The experimental apparatus for such a measurement is shown in **Figure 34.22*a***. From the

Figure 34.22 X-ray crystallography.

(*a*) Schematic apparatus for x-ray crystallography



(*b*) X-ray diffraction pattern of diamond lattice



CONCEPTS

position and intensity of the spots in the resulting diffraction pattern (Figure 34.22*b*), we can deduce the arrangement of the atoms in the sample.

The earliest x-ray diffraction studies of crystals were done on simple crystals, such as metals that form cubic or other very simple lattices. However, x-ray diffraction has also been used to study the atomic structure of much more complicated molecules, by crystallizing solutions of such molecules. The double helical structure of DNA was determined using x-ray diffraction, also called *x-ray crystallography,* and crystallography is widely used today to determine the atomic structure of even more complicated biological molecules.

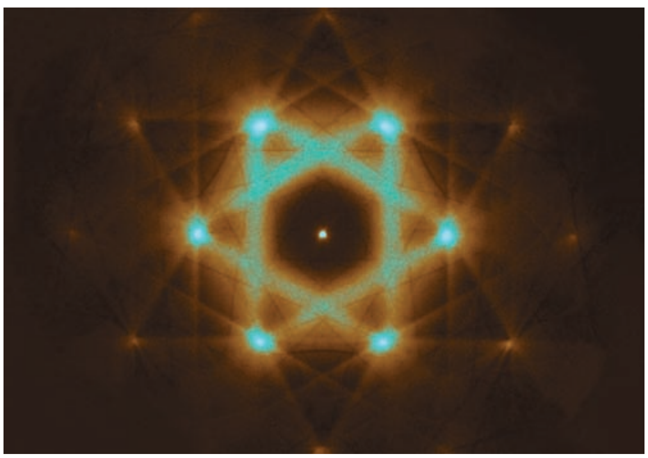
34.10 While comparing the x-ray diffraction patterns from two crystals, you determine that crystal A produces a pattern with more widely spaced diffraction spots than crystal B. Which crystal has the greater atomic spacing?

34.4 Matter waves

Patterns remarkably similar to x-ray diffraction patterns can be obtained by aiming a narrow beam of electrons at a crystal. **Figure 34.23** shows the pattern obtained by sending a beam of high-speed electrons through a solid crystalline sample in an instrument called an electron microscope. The shape of the pattern is similar to that obtained with x rays (Figure 34.22*b*), which tells us that electrons are also diffracted by crystals. This discovery, in turn, suggests that the electrons behave like waves because interference and diffraction are wave phenomena. Indeed, electrons have been found to exhibit interference in many other experiments. For example, a beam of electrons aimed at two very narrow slits produces an interference pattern similar to that of a beam of light aimed at two narrow slits (**Figure 34.24**).

Varying the speed of the electrons changes the spacing of the interference pattern, which indicates that the electron wavelength depends on speed.

Figure 34.23 Electron diffraction pattern for a diamond lattice. Notice the similarity to the x-ray pattern in Figure 34.22*b*.



CONCEPTS

34.11 Spots in an electron diffraction pattern, such as the one shown in Figure 34.23, move closer together as the speed of the electrons is increased. Does this mean the wavelength of the electrons increases or decreases with increasing speed?

Up to now we have considered electrons as being nearly pointlike particles, so it is surprising—to say the least—to discover that they also behave like waves. In fact, electrons can exhibit both particle behavior and wave behavior in a single experiment! This dual behavior was vividly demonstrated in an experiment done in 1989 in Japan using an apparatus similar to that shown in Figure 34.24. The number of electrons emitted by the source was kept very low, so as to ensure that at any given instant at most one electron was traveling from the source to the screen. The screen shows the place where each electron arrives as a bright dot, and at first these dots appear at seemingly random locations, as shown in **Figure 34.25*a*** and *b*. However, as more and more electrons reach the screen one after another, it becomes clear that the electron impacts are not randomly located (Figure 34.25*c* and *d*). Rather, they are arranged exactly in the two-slit interference pattern observed for a higher-intensity electron beam. Covering one slit and forcing the electrons to pass through the other makes the interference pattern disappear, just as it would for a light wave or a water wave.

Experiments show that after many electrons are allowed to pass one at a time through the apparatus, the number of electrons arriving on each small region of the screen per unit time is proportional to the intensity that would be observed at that region if light of the appropriate wavelength were shone on the slits. In other words, the *probability* of any single electron arriving at that small region in a fixed time interval corresponds to the intensity of the two-slit interference pattern at that location.

The observation of individual electrons arriving at the screen one after another indicates that the electrons are individual particles. However, if they are particles in the sense we think of for material objects, they must pass through either the right slit or the left slit, and so how can they produce an interference pattern?

The very counterintuitive conclusion is that each electron somehow travels through *both* slits simultaneously! Such a statement is not surprising for a wavefront hitting the two slits, but it goes completely against our intuitive notion of what we call a "particle."

We cannot explain the results of this experiment by concluding that electrons are *waves* because a classical wave could not produce the individual pinpoint images on the screen

Figure 34.24 Apparatus for observing two-slit interference with an electron beam.

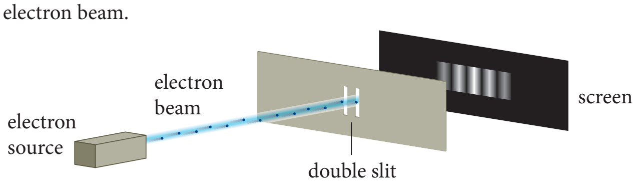
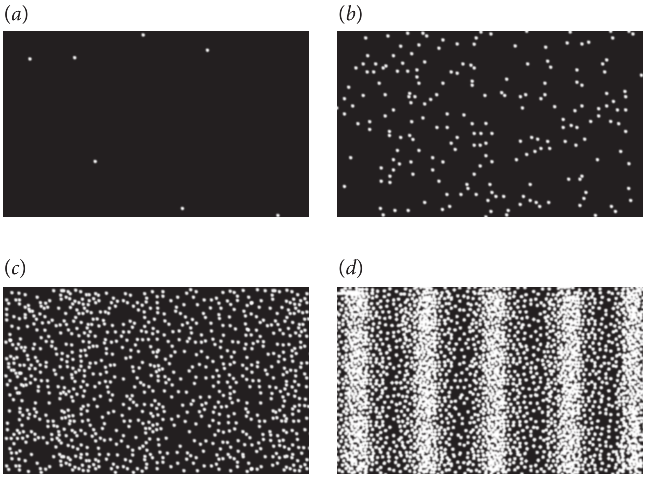


Figure 34.25 When we perform two-slit diffraction with a very weak electron beam, we can see the pattern build up over time. At first (*a*, *b*), the dots that mark electron impacts seem to be scattered randomly, but as more accumulate (*c*, *d*), the diffraction pattern becomes evident.



shown in, say, Figure 34.25*a*. If electrons were waves, the full diffraction pattern would be visible from the start (although it would be very faint). Scientists have concluded that electron behavior can be explained only if an electron has both particle properties and wave properties. This **wave-particle duality**—the possession of both wave properties and particle properties—has been observed not only for electrons but also for all other subatomic particles, for individual atoms, and, as we shall see in the next section, for light.

An expression for the wavelength of a particle, called the **de Broglie wavelength,** was proposed in 1924 by the French physicist Louis de Broglie* and confirmed experimentally a few years later. The de Broglie wavelength is inversely proportional to the momentum of the particle, $\lambda = h/p$, and the proportionality constant, called **Planck's constant,** is $h = 6.626 \times 10^{-34}\ \mathrm{J \cdot s}$. That this constant is a very small number indicates that the wavelength of macroscopic objects is extremely small. Because waves exhibit diffraction and interference only on length scales comparable to their wavelength, the wave nature of matter has been observed only with subatomic particles, atoms, and molecules. Exercise 34.4 illustrates this point.

Exercise 34.4 Electron versus baseball

Calculate the de Broglie wavelength associated with (*a*) a 0.14-kg baseball thrown at 20 m/s and (*b*) an electron of mass 9.1×10^{-31} kg moving at 5.0×10^6 m/s.

SOLUTION I use the expression $\lambda = h/mv$ for the de Broglie wavelength.

(*a*)
$$\lambda_{\text{de Broglie, baseball}} = \frac{6.626 \times 10^{-34}\ \mathrm{J \cdot s}}{(0.14\ \mathrm{kg})(20\ \mathrm{m/s})} = 2.4 \times 10^{-34}\ \mathrm{m}. ✔$$

This is 24 orders of magnitude less than the diameter of an atom!

(*b*)
$$\lambda_{\text{de Broglie, electron}} = \frac{6.626 \times 10^{-34}\ \mathrm{J \cdot s}}{(9.1 \times 10^{-31}\ \mathrm{kg})(5.0 \times 10^6\ \mathrm{m/s})} = 1.5 \times 10^{-10}\ \mathrm{m} = 0.15\ \mathrm{nm}. ✔$$

Electron diffraction takes place on distances comparable to the spacing between atoms, whereas a baseball would diffract only through apertures more than 10^{24} times smaller than the typical spacing between atoms. (But the diameter of a baseball is about 0.10 m, so such an experiment is not possible.)

34.12 How would the electron diffraction pattern in Figure 34.23 change if the electrons were traveling more slowly?

34.5 Photons

We have now found that particles have wave properties, but these properties are observed only when the particle wavelength is comparable to the size of the objects the particles interact with. Up until now, you have most probably thought of light as being solely a wave. However, the dual wave-particle nature of particles may lead you to wonder whether light, too, is not just a wave but also a particle.

34.13 Compare the two-slit interference pattern obtained with electrons (Figure 34.25) with the two-slit interference pattern obtained with light (Figure 34.8). If light and electrons exhibit a similar wave-particle duality, how might you modify the two-slit experiment with light shown in Figure 34.6 to observe light behaving like a particle?

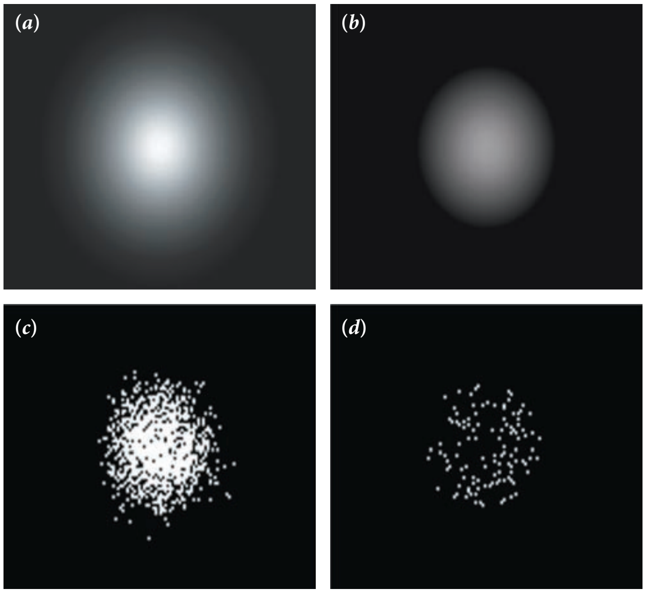
Figure 34.26*a* on the next page shows the image obtained by shining a beam of light onto the sensor chip of a digital camera. Pixels in the center of the beam register more light, and therefore the image of these pixels is brighter. The pixels at the edge of the beam are dimmer than those at the center, and outside the beam the pixels are black. If we place in front of the sensor chip a filter that lets only 50% of the light through, the image darkens because the brightness measured by each camera pixel is cut in half (Figure 34.26*b*).

Suppose we keep adding such filters, cutting the beam intensity in half with each addition. Does the image keep getting proportionally darker? If you carry out the experiment, you will discover that it does not, for one of two reasons. The first reason is mechanical: If you use the sensor chip of a digital camera, it stops detecting below a certain level. As you decrease the intensity of the beam by adding more and more filters, the image first gets grainy and then turns black.

The second reason has to do with the fundamental nature of light. Even if you use an extremely sensitive detector, you will still see that once the beam becomes very weak, adding another filter does not simply halve the image intensity. Instead, as shown in Figure 34.26*c* and *d*, the image of the beam breaks up into individual point-like flashes of equal

* "de Broglie" is pronounced "duh-Br-uh-y," the "y" sounding as in "yikes."

Figure 34.26 Images formed by using the sensor of a digital camera to record increasingly faint beams of light for the same exposure period. The separate dots recorded for the faintest beams reveal the particle-like behavior of light.



intensity, resembling the impacts of individual particles. The impacts may at first appear to be randomly distributed within the profile of the beam, but if you accumulate many of these impacts, you discover that the probability of observing a flash in a given location follows the intensity profile of the beam—the impacts are more likely to occur near the center of the beam. In fact, the probability of observing an impact in a particular location is proportional to the intensity of the beam at that location.

As the beam intensity is reduced, the individual impacts become separated in time. No two impacts occur at the same instant. Nor are any "half impacts" ever recorded. From these observations, we conclude that light indeed has particle properties as well as wave properties. The 'particles' of light are called **photons.** As we discuss further in Section 34.10, a photon represents the basic unit of a light wave and carries a certain amount of light energy. For a photon of frequency f, this energy equals hf, where h is again Planck's constant. Photons thus represent the quantum of electromagnetic energy—they cannot be subdivided.

Example 34.5 Photons from a light bulb

A 50-W incandescent light bulb emits about 5.0 W of visible light. (The rest is converted to thermal energy.) If a circular aperture 5.0 mm in diameter is placed 1.0 km away from the light bulb, approximately how many photons reach the aperture each second?

❶ **GETTING STARTED** This problem asks me to relate the power of light emitted by a light bulb to the rate at which photons pass through a certain area at a particular distance from the bulb. I can approximate the light bulb as a point source of light, meaning that its light is radiated uniformly in all directions. I can also simplify the problem by assuming that all of the light has the same wavelength, and I choose 500 nm for that wavelength. (This is a significant simplification because real light bulbs emit all visible wavelengths of light as well as infrared.)

❷ **DEVISE PLAN** I begin by determining the intensity—the power per unit area—of light produced by the light bulb over a sphere of radius 1 km, which tells me the intensity at the aperture. I can then multiply this intensity by the area of the aperture to obtain the power passing through the aperture, and I multiply that by 1 s to calculate the energy passing through the aperture in 1 s. Finally I shall use the relationship between photon energy and wavelength to determine the number of photons corresponding to that amount of energy.

❸ **EXECUTE PLAN** The intensity at the aperture is given by the power emitted by the bulb divided by the surface area of a sphere of radius 1.0 km:

$$\frac{(5.0\ \text{W})}{(4\pi)(1.0 \times 10^3\ \text{m})^2} = 4.0 \times 10^{-7}\ \text{J/m}^2 \cdot \text{s},$$

and therefore the amount of energy passing through a circular hole of radius 2.5 mm in 1.0 s is

$$(4.0 \times 10^{-7}\ \text{J/m}^2 \cdot \text{s})[\pi(2.5 \times 10^{-3}\ \text{m})^2] = 7.8 \times 10^{-12}\ \text{J}.$$

The energy of a single photon of wavelength 500 nm is

$$E = hf = \frac{hc_0}{\lambda}$$
$$= \frac{(6.626 \times 10^{-34}\ \text{J} \cdot \text{s})(3.00 \times 10^8\ \text{m/s})}{(500 \times 10^{-9}\ \text{m})}$$
$$= 3.98 \times 10^{-19}\ \text{J},$$

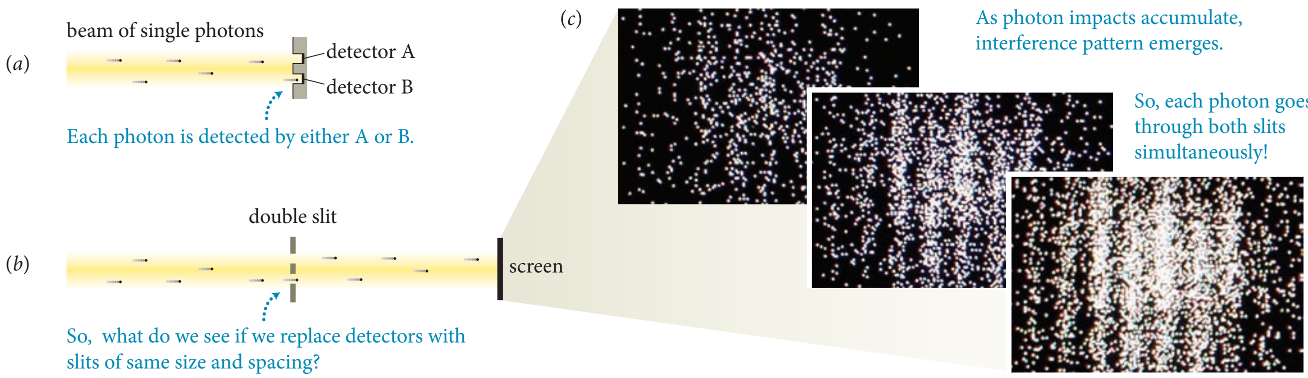
and so the number of photons that corresponds to the amount of energy passing through the hole is

$$\frac{7.8 \times 10^{-12}\ \text{J}}{3.98 \times 10^{-19}\ \text{J}} = 2.0 \times 10^7\ \text{photons.}\ ✔$$

❹ **EVALUATE RESULT** The number of photons I obtain is great even though I know from experience that the amount of light entering my eye 1.0 km from a 50-W bulb is exceedingly small. However, photons contain a vanishingly small amount of energy (about 10^{-19} J per photon), and so my answer is not implausible.

I assumed that all the photons have the same 500-nm wavelength. In reality the light bulb emits both longer- and shorter-wavelength photons. Suppose various wavelengths are equally distributed on both sides of 500 nm in the spectrum of the light emitted by the bulb. Photons with longer wavelengths have less energy and therefore there are more of them, but photons with shorter wavelengths have greater energy, and so there are fewer of them. The overall result is therefore about the same number of photons as the number I calculated using 500 nm as the only wavelength. To a first approximation, therefore, my result should be correct.

Figure 34.27 When we record a low-intensity beam of light with individual detectors, the beam acts like a stream of particles. Passing it through a double slit, however, causes an interference pattern to emerge.



What are photons? How can light be a wave *and* consist of photons that behave like particles? I cannot answer this question because no one really knows the answer. I can, however, describe how photons behave. As you will see, their behavior defies common sense—it is unlike anything we ever experience.

If we reduce the intensity of a light beam (or a beam of any other type of electromagnetic radiation) so greatly that the flashes from the impacts of individual photons are well separated in time and then aim this beam at two very small, very closely spaced detectors (**Figure 34.27*a***), we see that each photon is detected by either one detector or the other. A simultaneous impact on both detectors is never recorded. (The detectors can be made as small as 1 μm wide and spaced by just a fraction of a micrometer.) This observation suggests that each photon takes a definite path—toward either detector A or detector B.

Now imagine replacing the two detectors in Figure 34.27*a* by two narrow slits of the same size as the detectors and placing a screen some distance back from the two slits (Figure 34.27*b*). The pattern on the screen initially looks like random impacts from photons that make it through either one slit or the other. As we accumulate the impacts of many photons, however, an interference pattern emerges (Figure 34.27*c*) that is identical to the one obtained by shining an intense beam of light on the two slits.

Notice that replacing the two detectors with two slits doesn't change anything about the beam or the photons contained in it. However, the experiment using the two detectors suggests that photons are particles detected by one or the other of the detectors; the experiment using the two slits indicates that each photon is a wave traveling through both slits simultaneously! So which is it? Do the photons travel through one slit only or through both? If we physically cover one of the slits (and therefore *force* the photons to go through the other slit), the interference pattern disappears. The pattern that emerges after accumulating many photons behind the slit corresponds to the diffraction pattern of light behind a single slit.

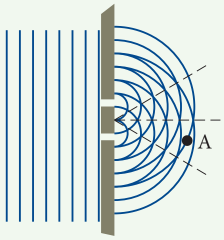
What this means is that photons behave as discrete particles when they are being detected. In transit, however, the wave nature of photons dictates their behavior. In other words, light behaves *both* like a wave and like a particle. It is impossible to explain the results of the above set of experiments by treating light as only a wave or only a particle; it must have qualities of both.

34.14 Figure 34.8 shows the interference pattern obtained by shining a strong laser beam on a pair of slits. If instead a very weak beam is shone on the same slits, so that the photons pass through the slits one photon at a time, at what angles is the probability of observing a particular photon the greatest?

Self-quiz

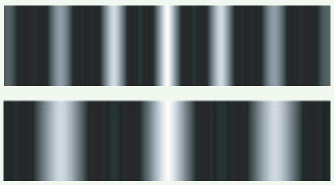
1. At point A in Figure 34.28, do the waves from the two slits add or cancel?

Figure 34.28



2. If the two sets of fringes shown in Figure 34.29 were produced by the same diffraction grating, which set is the product of the longer-wavelength radiation?

Figure 34.29



3. Coherent light of wavelength λ is normally incident on two slits separated by a distance d. What is the greatest possible fringe order?

4. Consider a proton and an electron moving at the same speed. Which has the longer wavelength?

5. Given the relationship between the energy E of a photon and its frequency f and the de Broglie expression relating momentum $p = mv$ and wavelength λ, determine the ratio E/p for a photon.

Answers:

1. At A, the crest of one wave overlaps the trough of the other wave, which means the waves cancel.
2. Because the wavelength is proportional to the sine of the angle the rays make with the normal to the diffraction grating, the fringes with the greater spacing were produced by the longer-wavelength radiation.
3. The fringe order is given by $d \sin\theta = \pm m\lambda$. Because the maximum value of $\sin\theta$ is 1, $d = m_{\text{max}}\lambda$ and so the maximum fringe order is $m_{\text{max}} = d/\lambda$. (Because m_{max} is an integer, you must truncate the value you obtain by dividing d by λ. For example, if $d/\lambda = 2.8$, then the greatest fringe order is 2.)
4. Because a proton has greater mass than an electron and because the de Broglie wavelength is inversely proportional to mass, $\lambda = h/mv$, the proton has a shorter wavelength than the electron.
5. The relationship between the energy of a photon and its frequency is $E = hf$ (see Section 34.5). The de Broglie wavelength is given by $\lambda = h/mv = h/p$, so $p = h/\lambda$. Therefore $E/p = hf/(h/\lambda) = f\lambda = c$, where c is the speed of light.

34.6 Multiple-slit interference

Let us now calculate the interference pattern produced by an electromagnetic wave normally incident on a barrier pierced by multiple closely spaced, very narrow slits. The width of each slit is much less than the wavelength of the radiation, so each slit serves as a point source of radiation. We begin by determining the pattern created when the barrier has just two very narrow slits.

As discussed earlier, the point sources corresponding to the two slits are in phase. Therefore the electric fields of the two waves that reach the screen differ in phase only due to any difference in the distance each wave travels from its slit to the screen. We shall refer to this difference as the *path-length difference* δs (**Figure 34.30*a***). Taking the screen to be at position $x = 0$, we can write the electric fields of the waves traveling in the directions shown in Figure 34.30*a* as

$$E_1 = E_0 \sin \omega t \tag{34.1}$$

and

$$E_2 = E_0 \sin(\omega t + \phi), \tag{34.2}$$

where E_0 is the amplitude of the electric field and ϕ is a phase constant that is equal to the phase difference that results from the path-length difference between the two waves. The phase difference divided by 2π is the fraction of a cycle by which the two waves differ. This equals the path-length difference δs divided by the wavelength, or $\phi/2\pi = \delta s/\lambda$, and so

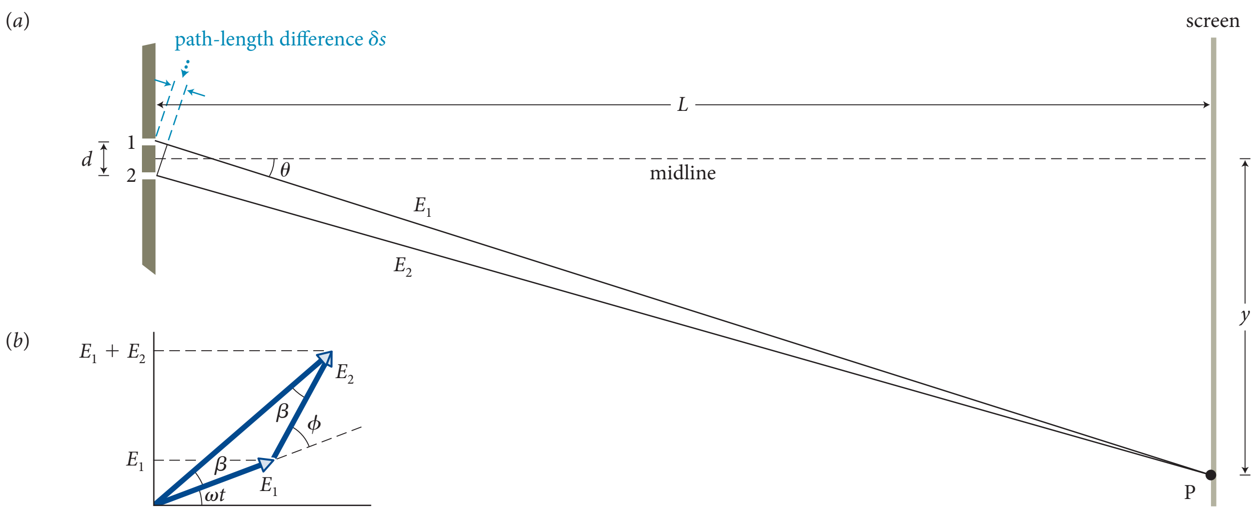
$$\phi = \frac{2\pi}{\lambda} d \sin\theta, \tag{34.3}$$

where d is the distance between the slits and θ is the angular position where the two rays meet on the screen.

We observe a bright fringe—a maximum in the intensity—when the two rays interfere constructively. This is the case when the phase difference ϕ equals an integer number times 2π:

$$\phi_m = \pm m(2\pi), \quad \text{for } m = 0, 1, 2, 3, \ldots. \tag{34.4}$$

Figure 34.30 (*a*) Interference of light diffracted by two very narrow slits. (*b*) We sum the phasors associated with the electric fields of the coherent light sources at slits 1 and 2.



QUANTITATIVE TOOLS

Combining Eqs. 34.3 and 34.4 and solving for sin θ determine the angles θ_m for which bright fringes occur:

$$\sin\theta_m = \pm\frac{m\lambda}{d}, \quad \text{for } m = 0, 1, 2, 3, \ldots \quad \text{(bright interference fringes)}, \tag{34.5}$$

as we found in Section 34.2.

A dark fringe—a minimum in the intensity—occurs when the two rays interfere destructively. This is the case when the phase difference equals a odd number times π:

$$\phi_n = \pm(2n - 1)\pi, \quad \text{for } n = 1, 2, 3, \ldots. \tag{34.6}$$

Substituting ϕ into Eq. 34.3 and solving for $\sin\theta$ determine the angles θ_n for which dark fringes occur:

$$\sin\theta_n = \pm\frac{(n - \frac{1}{2})}{d}\lambda, \quad \text{for } n = 1, 2, 3, \ldots \quad \text{(dark interference fringes)}. \tag{34.7}$$

To calculate the light intensity as a function of θ, we start by determining the sum of the electric fields E_1 and E_2 (Eqs. 34.1 and 34.2) using phasors. The amplitude of the sum of two sinusoidal quantities is equal to the vector sum of the phasors representing those two quantities (see Section 32.6). The phasors that represent E_1 and E_2 and their sum phasor are shown in Figure 34.30*b*. For the isosceles triangle formed by the E_1 and E_2 phasors, the exterior angle ϕ is equal to the sum of the two opposite interior angles β, and so $\beta = \frac{1}{2}\phi$. From this, we calculate the amplitude E_{12} of the sum of the two electric fields:

$$E_{12} = 2(E_0 \cos\beta) = 2(E_0 \cos\tfrac{1}{2}\phi). \tag{34.8}$$

Because this combined electric field oscillates at the same frequency as the incident wave, the time-dependent electric field at the screen is $E = E_{12}\sin\omega t$. At the central bright fringe, $\theta = 0$ and the two beams are in phase, so $\phi = 0$. The amplitude of E_{12} at the central bright fringe is therefore twice the amplitude of the incident wave, as we would expect.

We found in Chapter 30 (Eq. 30.36) that the intensity S of an electromagnetic wave is proportional to the product of the magnitudes of the electric and magnetic fields E and B: $S = EB/\mu_0$. Because B is proportional to E (Eq. 30.24, $E = Bc_0$), the intensity of light is commonly written in terms of E^2:

$$S = \frac{1}{\mu_0}EB = \frac{E(E/c_0)}{\mu_0}. \tag{34.9}$$

Substituting Eqs. 34.8 and 34.3 into Eq. 34.9 yields

$$S = \frac{4E_0^2}{\mu_0 c_0}\cos^2\left(\frac{\pi d\sin\theta}{\lambda}\right)\sin^2\omega t. \tag{34.10}$$

Visible electromagnetic waves oscillate at such high frequencies (10^{14} Hz to 10^{15} Hz) that we ordinarily measure the time-averaged intensity. The time average of $\sin^2\omega t$ is $\frac{1}{2}$. We can then write the time-averaged intensity of the interference pattern in terms of the time-averaged intensity of the incident wave:

$S_{0,\text{av}} = E_0^2/(2\mu_0 c_0)$. Thus substituting $\frac{1}{2}$ for $\sin^2 \omega t$ and $S_{0,\text{av}}(2\mu_0 c_0)$ for E_0^2 in Eq. 34.10 gives us

$$S_{\text{av}} = 4S_{0,\text{av}} \cos^2\left(\frac{\pi d \sin\theta}{\lambda}\right). \tag{34.11}$$

The maximum intensity of the interference pattern is *not* just the sum of the intensities of the two interfering waves—it is twice the sum! Because intensity is proportional to the square of the electric field, we must add the electric fields and then calculate the intensity from the square of the combined electric field.

34.15 How can the energy of the closed system made up of the two interfering waves remain constant (as the energy law states it must) if the maximum time-averaged intensity is four times the individual time-averaged intensities of the two waves?

Now let us work out how this pattern looks on the screen. If we consider only small angles θ, we can approximate $\sin\theta \approx \tan\theta = y/L$, where y is the position on the screen corresponding to the angle θ measured from the midline between the slits and L is the distance between the screen and the barrier that contains the slits, as shown in Figure 34.30*a*. (Positive y corresponds to positions above the midline; negative y to positions below the midline.) We can then write the time-averaged intensity as

$$S_{\text{av}} = 4S_{0,\text{av}} \cos^2(\phi/2) \approx 4S_{0,\text{av}} \cos^2\left(\pi d \frac{y}{L\lambda}\right). \tag{34.12}$$

The intensity varies periodically with the phase difference ϕ. Near the central maximum, where θ is small, the intensity also varies periodically with y (**Figure 34.31**).

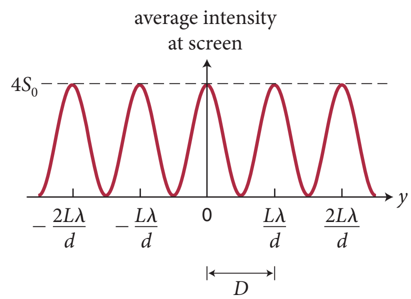


Figure 34.31 Average intensity produced on a screen by two-slit interference.

What is the distance D between adjacent intensity maxima of this pattern? Substituting y/L for $\sin\theta$ in Eq. 34.5, we obtain, for the positions of the maxima corresponding to any two values m and $m + 1$ of our order integer m,

$$m\lambda = d\frac{y_m}{L} \quad \text{and} \quad (m+1)\lambda = d\frac{y_{m+1}}{L}. \tag{34.13}$$

Subtracting the first of these equation from the second yields

$$\lambda = \frac{d}{L}(y_{m+1} - y_m). \tag{34.14}$$

The distance between adjacent maxima is $y_{m+1} - y_m = D$, so

$$D = \frac{L}{d}\lambda. \tag{34.15}$$

This expression tells us that the distance between the maxima is proportional to the wavelength of the light. The interference pattern "magnifies" the wavelength by the factor L/d.

If there are more than two slits in the barrier, the analysis proceeds in the same fashion as before. **Figure 34.32** on the next page shows a wave diffracting through six slits, each separated from its immediate neighbors by distance d. The condition for a maximum in the intensity pattern is the same regardless of the number of slits because if the difference in path length for waves from slits 1 and 2 is λ, then the difference in path length for any pair of adjacent slits is also λ. The principal maxima therefore appear at those locations where

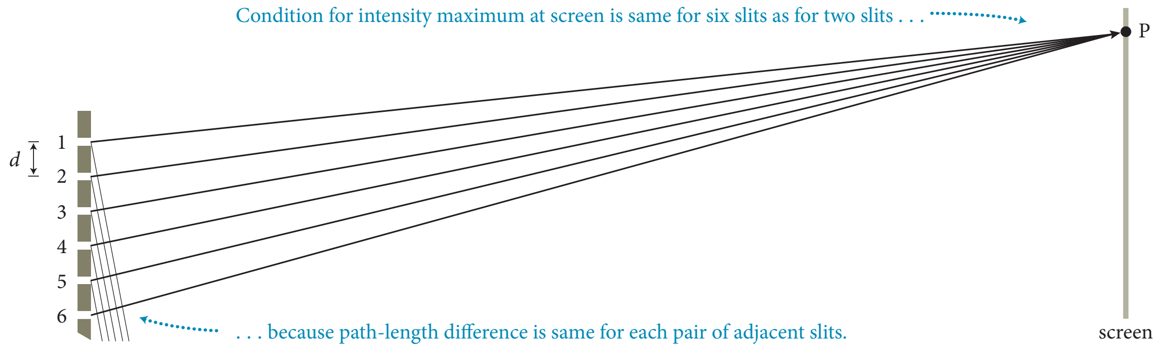
$$d \sin\theta_m = \pm m\lambda, \quad \text{for } m = 0, 1, 2, 3, \ldots \quad \text{(principal maxima)}, \tag{34.16}$$

just as we found for a two-slit barrier (Eq. 34.5).

The principal maxima occur at the same angles regardless of the number of slits. However, as discussed in Checkpoint 34.5, the *intensity* at the maxima increases with the number of slits. As we found in Section 34.2, as the number of

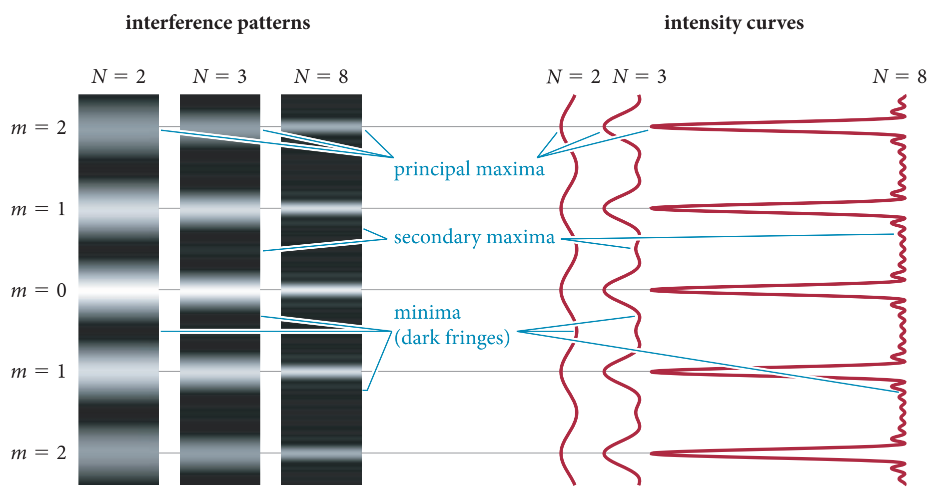
QUANTITATIVE TOOLS

Figure 34.32 Interference of light diffracted by six narrow slits.



slits increases, the interference fringes also become narrower (**Figure 34.33**). This is so because the minima closest to the m^{th} principal maximum occur for the ratio k/N that is as close as possible to m—namely, $(mN \pm 1)/N$. As the number of slits N becomes very great, the minima lie very close to m and so the interference pattern has extremely sharp maxima separated by broad dark regions with very faint secondary maxima.

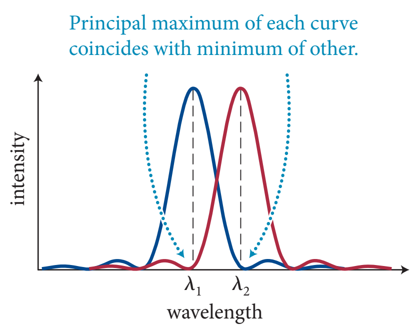
Figure 34.33 Interference pattern produced by gratings with two, three, and eight slits.



As discussed in Section 34.2, an important use of diffraction gratings is to disperse (separate) light into its constituent wavelengths in order to better understand the light source. The amount of information that can be obtained from a spectrum depends on whether the wavelengths of interest can be distinguished from one another in the spectrum (such wavelengths are said to be *resolved*).

Two wavelengths can be distinguished from each other if the principal maximum of one falls in the first dark region of the other, as shown in **Figure 34.34**. We found in Section 34.2 that, in general, a minimum occurs for interference through N slits when

Figure 34.34 Two clearly separated bright fringes for light of different wavelength.



$$d \sin \theta_{\text{min}} = \pm \frac{k}{N}\lambda, \quad \text{for integer } k \text{ that is not an integer multiple of } N \text{ (dark interference fringes).} \tag{34.17}$$

The angular position of the principal maxima increases with wavelength. Therefore, two wavelengths λ_1 and λ_2 can be distinguished from each other if the principal maximum for the longer wavelength falls at an angle greater than or equal to the angle for the $n = 1$ minimum for the shorter wavelength. As we first explored in Checkpoint 34.7, the separation between the slits is critical in determining the smallest wavelength difference that can be distinguished (this wavelength difference is called the *resolution*).

Example 34.6 Resolution of wavelengths in a diffraction grating

An astronomer wishes to determine the relative heights of the intensity peaks for the bright fringes produced by two wavelengths of radiation emitted by sodium atoms. The wavelengths are 589.0 nm and 589.6 nm, and she uses a diffraction grating with 500.0 slits/mm to disperse the light collected by her telescope. (*a*) In which order are the intensity maxima for these two wavelengths farthest apart from each other: $m = 0$, $m = 1$, or $m = 2$? (*b*) If the part of the diffraction grating covered by the light is 4.000 mm wide, are the second-order principal maxima produced by these two spectral lines distinguishable from each other?

❶ **GETTING STARTED** This problem is about the overlapping diffraction patterns produced by two very similar wavelengths of light as the light passes through a diffraction grating. To answer the questions, I need to calculate the angular positions of the principal intensity maxima for each wavelength in three orders of diffraction. I also need to determine the angular position of the minimum adjacent to the second-order principal maximum for each wavelength, in order to determine whether the second-order principal maxima of the two wavelengths are distinguishable.

❷ **DEVISE PLAN** For part *a*, I can use Eq. 34.16 to locate the $m = 0, 1,$ and 2 principal maxima for both wavelengths. For part *b*, I can locate the minimum adjacent to the second-order principal maximum for $\lambda = 589.0$ nm by applying the discussion following Eq. 34.17 and then compare that location with the location of the second-order principal maximum for $\lambda = 589.6$ nm.

❸ **EXECUTE PLAN** (*a*) I start by solving Eq. 34.16 for θ:

$$\theta_m = \pm\sin^{-1}\left(\frac{m\lambda}{d}\right). \qquad (1)$$

This result indicates that the principal maxima are farthest apart for $m = 2$.

To calculate the positions of the principal maxima, I substitute the appropriate values of m, λ, and d in Eq. 1. The separation d between the slits is $1/(500\ \text{slits/mm}) = 2.000 \times 10^{-3}\ \text{mm} = 2000$ nm. The $m = 0$ principal maximum occurs at $\theta = 0$ for any wavelength. The first-order principal maxima are located at

$$\theta_{589.0} = \pm\sin^{-1}\left(\frac{589.0\ \text{nm}}{2000\ \text{nm}}\right) = \pm 17.13°$$

$$\theta_{589.6} = \pm\sin^{-1}\left(\frac{589.6\ \text{nm}}{2000\ \text{nm}}\right) = \pm 17.15°,$$

and the second-order principal maxima are located at

$$\theta_{589.0} = \pm\sin^{-1}\left(\frac{2 \times 589.0\ \text{nm}}{2000\ \text{nm}}\right) = \pm 36.09°$$

$$\theta_{589.6} = \pm\sin^{-1}\left(\frac{2 \times 589.6\ \text{nm}}{2000\ \text{nm}}\right) = \pm 36.13°.$$

For these two wavelengths, the second-order principal maxima are the ones farthest apart from each other. ✔

(*b*) Using the information given in the discussion following Eq. 34.17, I can write that the condition for the minimum adjacent to the second-order principal maximum for $\lambda = 589.0$ nm is

$$d\sin\theta_{\text{min},589.0} = \pm\lambda\frac{mN+1}{N},$$

where I have replaced k in Eq. 34.17 by $mN + 1$ as explained in the discussion preceding that equation. Solving this expression for $\theta_{\text{min},589.0}$ gives

$$\theta_{\text{min},589.0} = \pm\sin^{-1}\left(\frac{\lambda}{d}\frac{mN+1}{N}\right).$$

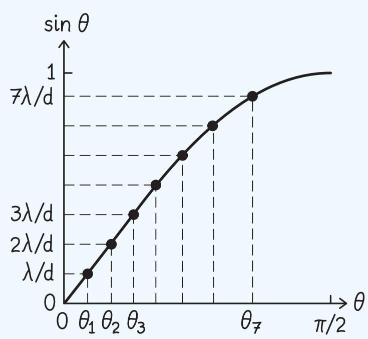
A region of the grating 4.000 mm wide contains 2000 slits. Substituting in the preceding expression, I get that, for $\lambda = 589.0$ nm, the minimum adjacent to the second-order principal maximum is at

$$\theta_{\text{min},589.0} = \pm\sin^{-1}\left(\frac{589.0\ \text{nm}}{2000\ \text{nm}} \times \frac{4001}{2000}\right) = \pm 36.10°. ✔$$

This minimum lies between 36.09°, the second-order principal maximum for $\lambda = 589.0$ nm, and 36.13°, the second-order principal maximum for $\lambda = 589.6$ nm, telling me that the second-order principal maxima for these two wavelengths are distinguishable from each other.

❹ **EVALUATE RESULT** Equation 34.16 tells me that the angles at which principal maxima occur are equal to the inverse sine of integer multiples of λ/d. For small angles $\sin\theta_m \approx \theta_m$, and so the angles θ_m are approximately equally spaced (**Figure 34.35**). As the curve for $\sin\theta$ bends toward the horizontal, however, the distance between adjacent θ_m values increases, in agreement with the result I obtained.

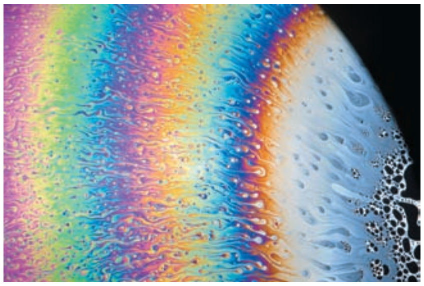
Figure 34.35



34.16 As the above discussion indicates, the separation distance between the principal maxima for different wavelengths increases as the fringe order m increases. Can you obtain an arbitrarily great separation distance by going to extremely high orders?

34.7 Thin-film interference

Figure 34.36 Soap bubble.



A familiar manifestation of the interference of light is the rainbow of colors reflected from thin films such as soap bubbles (**Figure 34.36**). This type of interference occurs in transparent materials whose thickness is comparable to the wavelengths of visible light. When such a thin material (we refer to it as a *film*) is either suspended in air, as in a soap bubble, or supported on a much thicker material with a different index of refraction, as in an oil slick on a puddle of water, white light reflecting from both the front and back surfaces of the film interferes, in the same manner as x rays reflecting from adjacent layers of atoms in a crystal. This is shown schematically in **Figure 34.37**. The film thickness t and index of refraction n_b, together with the angle of incidence, determine the path-length difference and corresponding phase difference between the reflected beams, and therefore determine which colors undergo constructive interference and which undergo destructive interference in any given direction.

To identify the conditions for constructive and destructive interference, we begin by expressing the electric fields of the waves reflected from the two surfaces, just as we did for waves passing through two slits in the preceding section:

$$E_1 = E_0 \sin(\omega t) \tag{34.18}$$

$$E_2 = E_0 \sin(\omega t + \phi). \tag{34.19}$$

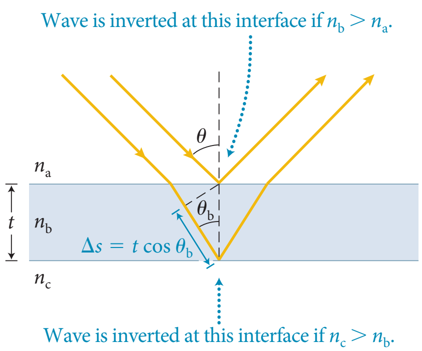
The phase difference ϕ is due to the path-length difference $2\Delta s$ and the effect of the reflections on the phases of each wave. The path-length difference gives rise to a phase difference

$$\phi_{\text{path}} = \frac{2\pi \Delta s}{\lambda_b} = \frac{2\pi(2t\cos\theta_b)}{\lambda/n_b} = \frac{4\pi n_b t \cos\theta_b}{\lambda}, \tag{34.20}$$

where the path-length difference Δs and the angle θ_b of the ray relative to the surface normal inside the film are shown in Figure 34.37, and $\lambda_b = \lambda/n_b$ is the wavelength of light inside the film (see Eq. 33.3). Because of refraction, expressing Δs in terms of the angle of incidence θ, instead of the angle in the film θ_b, involves Snel's law and produces a rather complicated result. However, for normal incidence $\cos\theta_b = 1$, and so the expression is greatly simplified.

To determine the phase difference due to the reflections from the two film surfaces, recall the discussion of mechanical waves at boundaries from Section 16.4. We saw there that when a wave pulse is launched in a first medium and travels into a second medium, the pulse is partially reflected and partially transmitted at the boundary. The reflected pulse is inverted relative to the incident pulse if the wave speed c_1 in the first medium is greater than the speed c_2 in the second medium. When $c_1 < c_2$, the incident pulse is not inverted upon reflection.

Figure 34.37 Thin-film interference.



Let us now extend these ideas to sinusoidal electromagnetic waves. When the incident wave in medium 1 reflects at the boundary with medium 2 of greater index of refraction ($n_1 < n_2$), the speed is less in medium 2 ($c_2 < c_1$), and so the wave is inverted upon reflection. This inverting of the wave is equivalent to the wave undergoing a phase shift of π upon reflection. If instead $n_1 > n_2$, the wave is not inverted and there is no phase shift upon reflection.

A phase shift of π can occur at each of the two interfaces, depending on the refractive indices n_a, n_b, and n_c of the three media, as indicated in Figure 34.37. Thus the phase difference associated just with the reflections for the two reflected waves is

$$\phi_r = \phi_{r2} - \phi_{r1}, \tag{34.21}$$

where ϕ_{r1} and ϕ_{r2} are either π or 0, depending on the values of the indices of refraction. Therefore the total phase difference between the two reflected waves is

$$\phi = \frac{4\pi n_b t \cos\theta_b}{\lambda} + \phi_{r2} - \phi_{r1}. \tag{34.22}$$

Because ϕ_{r1} and ϕ_{r2} must each be π or 0, the effect of the two reflections on the phase cancels out when $\phi_{r1} = \phi_{r2}$, and the phase difference is due entirely to ϕ_{path}. If ϕ_{r1} is not equal to ϕ_{r2}, then ϕ differs from ϕ_{path} by π, causing the constructive and destructive interference conditions to be switched from what they would be due to just ϕ_{path}.

In our discussion of x-ray diffraction, the light source consisted of monochromatic x rays, and we found that constructive interference between adjacent planes takes place at only certain angles. In the current context—light reflections from thin films—we usually deal with light across the visible spectrum. Therefore ϕ depends on both wavelength and angle of incidence. For normal incidence, Eq. 34.22 simplifies to

$$\phi = \frac{4\pi n_b t}{\lambda} + \phi_{r2} - \phi_{r2} \quad \text{(normal incidence).} \tag{34.23}$$

Then as before, constructive interference occurs when the phase difference corresponds to an integer number times 2π, and destructive when it corresponds to an odd number times π.

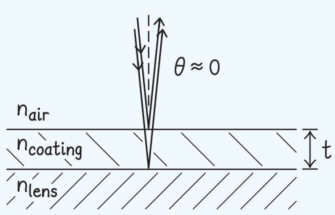
Example 34.7 Antireflective coating

Eyeglass lenses made of crown glass ($n = 1.52$) are given a thin coating of magnesium fluoride ($n = 1.38$) to minimize reflection of light from the lens surface. What is the minimum coating thickness for which reflection from the lens surface is minimized?

❶ **GETTING STARTED** Reflection is minimized at all values of the coating thickness that cause the waves reflected from the front coating surface to interfere destructively with those reflected from the back coating surface. My task therefore is to obtain the minimum thickness needed for destructive interference.

❷ **DEVISE PLAN** I see from Eq. 34.22 that the condition for destructive interference depends not only on the coating thickness t and index of refraction n_b but also on two variables I have no values for: the angle of incidence and the wavelength of the light. Thus I must make some simplifying assumptions to solve this problem. As in the text discussion, I shall consider only light normally incident on the coated lens and shall assume the lens and coating surfaces are flat, as shown in my sketch (**Figure 34.38**). I choose a representative visible wavelength, 500 nm—in the middle of the visible spectral range.

Figure 34.38



To obtain the condition for destructive interference, I can use Eq. 34.23 to determine the phase difference ϕ at normal incidence and equate it to an odd number times π radians. To work out the phase shifts ϕ_{r1} and ϕ_{r2} that occur at each reflection, I note that at the air-coating interface the light reflects from a medium for which the index of refraction is greater than that of the medium through which the incident wave travels, and so this wave undergoes a π phase shift upon reflection: $\phi_{r1} = \pi$. The wave that reflects from the coating-lens interface likewise reflects from a medium for which the index of refraction is greater than that of the medium through which the incident wave travels and thus also undergoes a π phase shift ($\phi_{r2} = \pi$). Consequently, ϕ_{r1} and ϕ_{r2} in Eq. 34.23 cancel, leaving only the phase difference due to the path-length difference. The minimum thickness produces a phase difference corresponding to the smallest number of cycles that gives destructive interference—namely, half a cycle.

❸ **EXECUTE PLAN** To determine the thickness that produces destructive interference for 500 nm light, I express the phase difference using Eq. 34.23 and equate it to a half-cycle phase difference:

$$\frac{4\pi n_b t}{\lambda} + \phi_{r2} - \phi_{r2} = \pi.$$

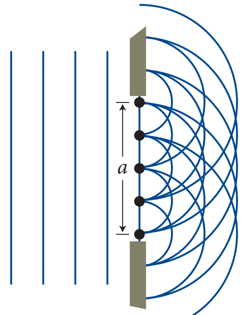
Substituting $\phi_{r1} = \phi_{r2} = \pi$, solving for the thickness t, and substituting values give

$$t = \frac{\lambda}{4n_b} = \frac{500 \text{ nm}}{4(1.38)} = 90.6 \text{ nm.} ✔$$

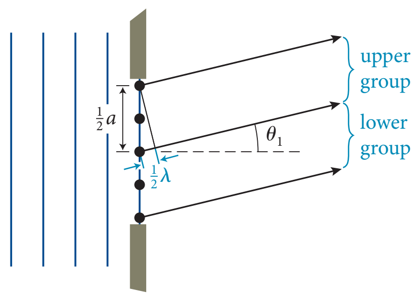
❹ **EVALUATE RESULT** The coating is thinner than the wavelength of the light traveling through the coating. That makes sense because, to create destructive interference, the wave that travels through the coating and reflects from the coating-lens interface must travel only half a wavelength farther than the wave that reflects from the air-coating interface.

Figure 34.39 Diffraction through a narrow slit.

(*a*) Wavefront in slit treated as a series of point sources



(*b*) Condition for first-order dark fringe: path lengths differ by $(\frac{1}{2}a)\sin\theta_1 = \lambda/2$



(*c*) Condition for second-order dark fringe: path lengths differ by $(\frac{1}{4}a)\sin\theta_2 = \lambda/2$

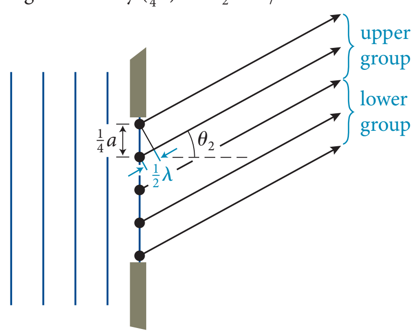
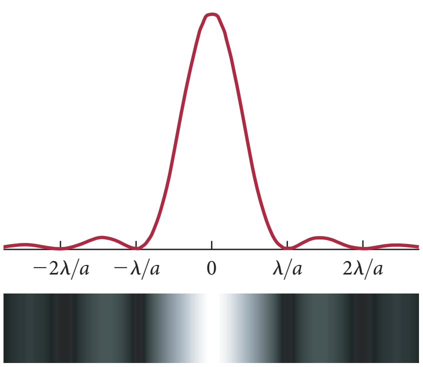


Figure 34.40 Intensity plot for a single-slit diffraction pattern.



34.17 When oil spreads on water, bands of different colors are visible. What causes the different colors?

34.8 Diffraction at a single-slit barrier

Let us now describe quantitatively the diffraction by a single slit. Huygens' principle (see Section 17.4) allows us to describe the wavefront that reaches the opening as a series of point sources that emit spherical wavelets (**Figure 34.39*a***). The wave beyond the slit is the superposition of all these spherical wavelets. In the original direction of propagation, all of the waves add in phase, and as a result the amplitude of the transmitted wave is the maximum possible: the sum of all the individual wave amplitudes. If we divide the rays representing the wavelets into two equal groups, as shown in Figure 34.39*b*, we can pair each ray in the upper half with a corresponding ray in the lower half. In Figure 34.39*b*, for example, we can pair the top ray in the upper group with the top ray in the bottom group. As we discussed in Section 34.2, these rays are essentially parallel when they emerge from the slits; as they travel and intersect at some location on a screen placed far to the right of the slits, the rays must travel different distances to reach that location. For rays traveling at an angle θ to the original propagation direction, the difference in path length from two such corresponding rays is $(a/2)\sin\theta$, where a is the width of the aperture. When the path lengths differ by half a wavelength, $(a/2)\sin\theta_1 = \lambda/2$, the rays interfere destructively, which means that the direction for the first-order dark fringe in a single-slit diffraction pattern is given by

$$\sin\theta_1 = \frac{\lambda}{a} \quad \text{(first-order dark diffraction fringe).} \tag{34.24}$$

Dividing the rays into four groups (Figure 34.39*c*) leads to a dark fringe (that is, a minimum in transmitted intensity) when $(a/4)\sin\theta = \frac{1}{2}\lambda$. Thus the direction for the second-order dark fringe is given by

$$\sin\theta_2 = 2\frac{\lambda}{a} \quad \text{(second-order dark diffraction fringe).} \tag{34.25}$$

The general condition for a dark fringe is thus

$$\sin\theta_n = \pm n\frac{\lambda}{a}, \quad n = 1, 2, 3, \ldots \quad \text{(dark diffraction fringes).} \tag{34.26}$$

Positive values of $\sin\theta_n$ correspond to dark fringes above the midline, while negative values correspond to dark fringes below the midline.

Calculation of the detailed intensity pattern is beyond the scope of this text, but we can examine some of the details of a representative pattern (**Figure 34.40**). In this single-slit diffraction pattern, the intensity of the first-order ($m = 1$) bright fringe is less than 5% of the intensity of the central ($m = 0$) bright fringe; the intensity of the second-order ($m = 2$) bright fringe is less than 2% of the central bright fringe intensity. In other words, most of the transmitted energy falls within the central peak. In general, in a single-slit diffraction pattern, the intensity is greatest at the central bright fringe and decreases rapidly with distance from the center of the pattern.

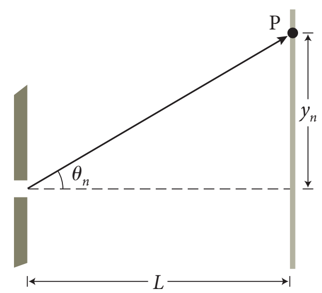
What if we want to calculate the linear positions of the dark diffraction fringes on a screen located a distance L away from a barrier containing a single slit (**Figure 34.41**) rather than the angular positions? For dark fringes located at small angles θ_n from the original direction of wave propagation, we can approximate $\tan\theta_n$ as $\sin\theta_n$:

$$y_n = L\tan\theta_n \approx L\sin\theta_n, \qquad (34.27)$$

and so, from Eq. 34.26,

$$y_n = \pm n\frac{\lambda L}{a} \quad \text{(dark diffraction fringes).} \qquad (34.28)$$

Figure 34.41 Calculating the positions of the dark fringes of a single-slit diffraction pattern.



Example 34.8 Spreading light

Consider the diffraction pattern shown actual size in Figure 34.40. If the pattern was formed by light from a 623-nm (red) laser passing through a single narrow slit and the screen on which the pattern was cast was 1.0 m away from the slit, what is the slit width?

❶ **GETTING STARTED** This problem asks me to relate the diffraction pattern in Figure 34.40, which was produced by a setup such as the one shown in Figure 34.41, to the width of the slit that produced it. Thus I need to relate the slit width to fringes whose position I can calculate from the parameters of the problem and can also measure on the image. Because the only variable I know how to calculate is the positions of minima in the diffraction pattern, I can measure the distance between the two first-order minima and relate that distance to the slit width and the geometry of the setup.

❷ **DEVISE PLAN** The positions of the two first-order minima, in terms of wavelength λ, slit-to-screen distance L, and slit width a, are given by Eq. 34.28 with $n = 1$. Subtracting the two values I obtain from each other gives me an expression for the distance between these two minima in terms of λ, L, and a. I am given the values of λ and L, and my task is to determine a. Thus if I know the distance between the n_1 minima, I can calculate a. Because the image in Figure 34.40 is actual size, I can measure this distance directly. Then I can solve Eq. 34.28 for a and insert my known values.

❸ **EXECUTE PLAN** Substituting $n = 1$ into Eq. 34.28 gives the linear positions of the two first-order minima:

$$y_1 = \pm\frac{\lambda L}{a},$$

so the distance w between the two minima is

$$w = \frac{2\lambda L}{a}. \qquad (1)$$

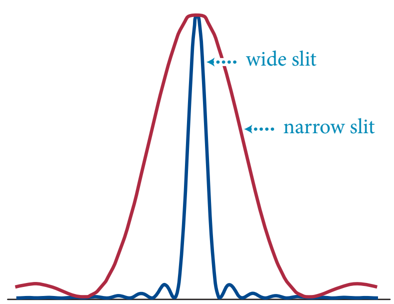
I measure the distance between the centers of the two dark fringes on either side of the central bright fringe in Figure 34.40 to be 23 mm. Solving Eq. 1 for the slit width thus gives me

$$a = \frac{2\lambda L}{w} = \frac{2(623\times10^{-9}\ \text{m})(1.0\ \text{m})}{23\times10^{-3}\ \text{m}}$$
$$= 5.4\times10^{-5}\ \text{m} = 0.054\ \text{mm} \quad ✔$$

❹ **EVALUATE RESULT** The slit width is about a factor of 100 greater than the wavelength of the light, and that ratio is consistent with the general range of slit sizes that produce noticeable diffraction of visible light.

As long as λ is small relative to the slit width a, so that the small-angle approximation for θ_n is valid, most of the diffracted light intensity falls within this region defined by Eq. 1 in Example 34.8. Note that w increases with decreasing a: The narrower the slit, the more the wave spreads out after passing through the slit (**Figure 34.42**). If the slit width is equal to or less than the wavelength of the light, there are no dark fringes, as we found in Checkpoint 34.16. The wave simply spreads out in all directions behind the slit, which means the slit behaves as a point source.

Figure 34.42 The width of the central maximum in a diffraction pattern decreases as the slit is widened.

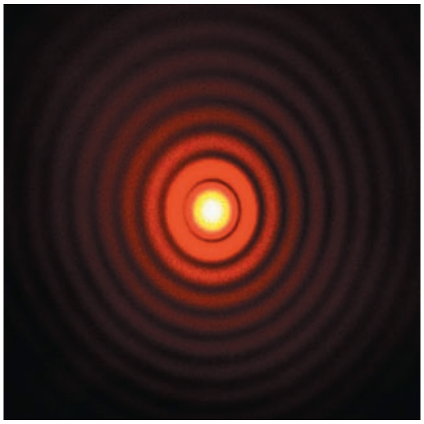


34.18 Using Eq. 34.24, calculate the angle at which the first dark fringe occurs when (*a*) $a < \lambda$ and (*b*) $a \gg \lambda$. Interpret your results.

34.9 Circular apertures and limits of resolution

When light passes through a circular aperture, the symmetry of the aperture causes the resulting diffraction pattern also to be circular (**Figure 34.43** on the next page). A circular central bright fringe is surrounded by circular dark diffraction fringes and additional diffraction bright fringes. The central bright

Figure 34.43 Diffraction pattern of light passing through a circular aperture.



fringe is called the *Airy disk*, after the British astronomer and mathematician Sir George Airy, who developed the first detailed description of diffraction in 1835. Calculation of the circular diffraction pattern is rather involved, so all I shall do here is state the location of the first dark fringe. When light is diffracted by a circular aperture of diameter d, the first dark fringe occurs at angle θ_1 given by

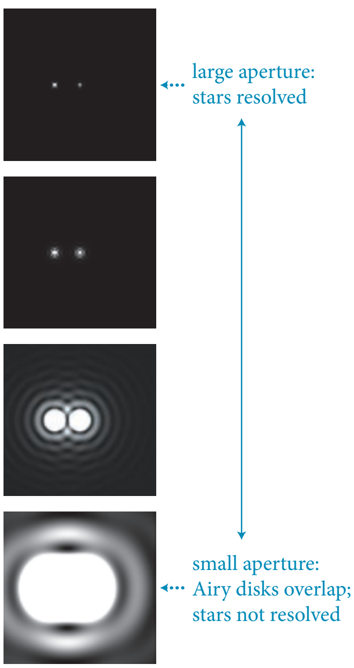
$$\sin\theta_1 = 1.22\frac{\lambda}{d}. \tag{34.29}$$

The result is similar to the one obtained in Eq. 34.24 for a slit of width a, except that now the sine of the angle is increased by a factor of 1.22. The increase in the angle for the same wavelength and aperture size can qualitatively be understood as follows: Equation 34.24 is obtained by considering the interference between wavelets coming from a slit whose width is a everywhere. A circular aperture has width d only across a diameter; the rest of the aperture is narrower. As the aperture gets narrower, diffraction through it becomes more pronounced. The factor of 1.22 quantitatively accounts for the varying horizontal width of the circular aperture.

The angular size of the Airy disk given in Eq. 34.29 determines the minimum angular separation of two point sources that can be distinguished by observing them with a (circular) lens—*regardless* of the magnification of the lens! To understand what this means, imagine imaging two distant, closely spaced point sources with a lens. These sources could be anything, from stars to organelles in a biological cell. The sources are not coherent, and so we can consider the Airy disks formed by the light from each source separately without considering interference between sources.

If there is overlap in the Airy disks of the images observed through the lens, it is difficult to tell whether there are two point sources or just one. Two objects being observed through a lens are just barely distinguishable when the center of one diffraction pattern is located at the first minimum of the other diffraction pattern. This happens when the angular separation between the two objects is at least the angle given in Eq. 34.29. If this is the case, we say that the two objects are *resolved*. This condition for distinguishability is called **Rayleigh's criterion.**

Figure 34.44 The resolution of these two stars improves as the aperture is made larger.



Because the diameter d of the lens is always much greater than the wavelength of the light, the angle in Eq. 34.29 is always small. Thus the minimum angular separation θ_r for which two sources can be resolved is approximately equal to the sine of the angle

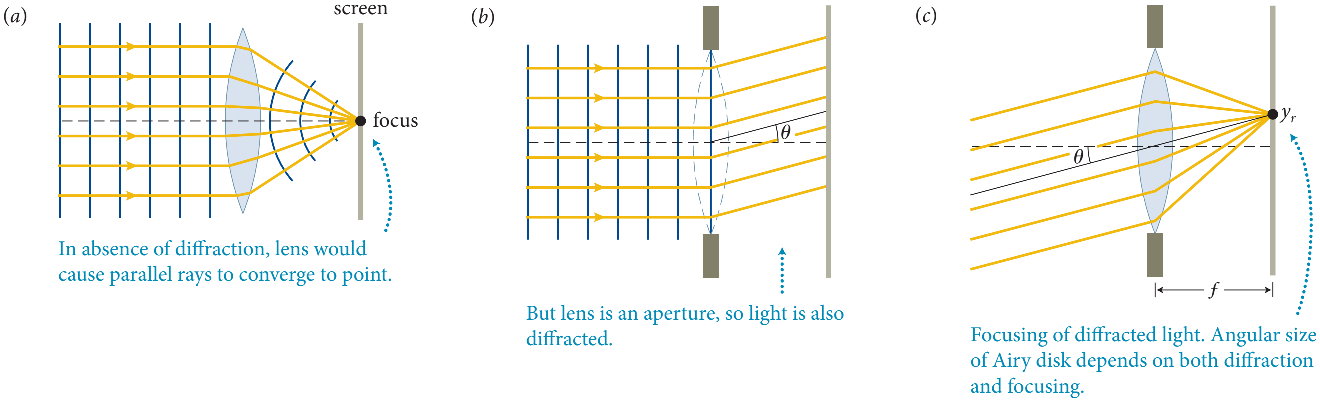
$$\theta_r \approx \sin\theta_r = 1.22\frac{\lambda}{d}. \tag{34.30}$$

Two objects that are separated by an angle equal to or greater than θ_r satisfy Rayleigh's criterion. For this reason, the closest two objects can be to each other and still be distinguished with an optical instrument such as a microscope or telescope depends not on the magnification but on the wavelength of the light and the size of the smallest aperture in the instrument.

Figure 34.44 shows the images of two stars obtained with a telescope. An aperture placed in front of the lens shows the effects of diffraction. When the opening of the aperture is small (bottom image in **Figure 34.44**), the images of the two stars are merged—the two stars cannot be resolved. As the aperture is opened, θ_r decreases and the two images separate cleanly.

Diffraction also determines the linear size of the images of point sources. In Chapter 33, we stated that the image of a point source formed on a screen by a lens is a point. **Figure 34.45*a*** shows parallel rays from a distant point source focused by a lens onto a screen placed in the focal plane of the lens. Without diffraction, the image formed by these rays would be an infinitesimally small point. In fact, because the aperture through which the light passes—the lens—has a

Figure 34.45 Analyzing the diffraction limit of a lens.



finite diameter, such rays do not focus to an infinitely small point, and the image formed is a diffraction pattern just like that shown in Figure 34.43. The angular size of the central bright fringe of this diffraction pattern—in other words, the Airy disk—is given by Eq. 34.29.

To calculate the radius of the Airy disk formed in the focal plane of the lens,* we must account for both diffraction and focusing. Figure 34.45*b* shows the diffraction of light through an aperture of the same diameter as the lens. Light that originally traveled parallel to the lens axis is diffracted by an angle θ. Now consider how the lens focuses the diffracted light. Parallel rays that make an angle θ with the lens axis are focused at a point located a distance y above the axis (Figure 34.45*c*, see also Figure 33.27). This distance y is given by $y = f\tan\theta$, where f is the focal length of the lens.

Our next step in determining the Airy disk radius is to calculate the distance $y_{\rm r}$ from the center of the disk to the first dark fringe in the diffraction pattern, which is found at the angle given by Eq. 34.29. In the small-angle limit, $\sin\theta \approx \tan\theta$ and so we can substitute $\sin\theta_{\rm r} = y_{\rm r}/f$ into Eq. 34.29, giving

$$y_{\rm r} = 1.22\,\frac{\lambda f}{d}. \tag{34.31}$$

This expression gives the radius $y_{\rm r}$ of the Airy disk and the minimum size of the area to which light can be focused with light of wavelength λ by a lens of focal length f and diameter d. The best ratio of f/d that can be achieved with a lens is approximately unity, and so the smallest diameter to which light can be focused is about 2.5λ. This means that the smallest diameter of "points" in the resulting image is also 2.5λ.

This diffraction-determined minimum size of the features in an image is commonly called the *diffraction limit*. An industry in which the diffraction limit poses a serious problem is the manufacture of integrated circuits, such as computer chips. The transistors and logic gates described in Section 32.4 are produced by a series of processes known as *photolithography*, in which the semiconductor substrate is coated with a polymer that is sensitive to ultraviolet light. The polymer is then exposed to ultraviolet light in the pattern of the desired metal electrodes. This pattern of light is produced by illuminating a metal mask with holes in the shape of the electrodes and then imaging the resulting pattern of light onto the surface. Finally, the exposed polymer is dissolved with a chemical rinse, leaving the semiconductor surface exposed where metal is desired. The metal is then deposited in a subsequent step.

*The radius of the Airy disk is smallest when the screen is in the focal plane of the lens, and it increases as the screen is moved closer to or farther from the lens.

The smallest electrode that can be made by this process is therefore determined by the diffraction limit for the ultraviolet light ($\lambda \leqslant 150$ nm) used to produce the pattern. With ordinary optical technology, this requires electrodes to be at least 300 nm wide. Many researchers are searching for other ways to produce electrodes that are not limited by diffraction.

Exercise 34.9 A point is not a point

A magnifying glass that has a focal length of 0.25 m and a diameter of 0.10 m is used to focus light of wavelength 623 nm. (*a*) What is the radius of the smallest Airy disk that can be produced by focusing light with this lens? (*b*) How large is the Airy disk formed when this lens focuses a laser beam that has a 2.0-mm diameter?

SOLUTION (*a*) The radius of the smallest Airy disk is given by Eq. 34.31 with d equal to the diameter of the lens:

$$y = 1.22\frac{(623 \times 10^{-9}\text{ m})(0.25\text{ m})}{0.10\text{ m}} = 1.9 \times 10^{-6}\text{ m} = 1.9\ \mu\text{m}. ✔$$

The minimum spot size is about three times greater than the wavelength of the light.

(*b*) In this case the radius is given by Eq. 34.31 with d equal to the diameter of the laser beam. I must use the beam diameter because the beam does not make use of most of the area of the lens but effectively defines its own aperture. I therefore have

$$y = 1.22\frac{(623 \times 10^{-9}\text{ m})(0.25\text{ m})}{0.0020\text{ m}} = 9.5 \times 10^{-5}\text{ m} = 95\ \mu\text{m}. ✔$$

This Airy disk is much bigger than the one found in part *a*. This Airy disk radius means the central bright fringe has a diameter of 190 μm = 0.19 mm, which is smaller than the original beam diameter due to the focusing of the lens. Because the laser beam is much smaller than the diameter of the lens, it cannot be effectively focused by this lens. The diffraction partially cancels the focusing.

Example 34.10 Blurry images

The widths of one pixel in the sensor for a digital camera is about 2.0 μm. If the camera lens has a diameter of 40 mm and a focal length of 30 mm, is the resolution of the resulting image limited by the lens or the sensor?

❶ **GETTING STARTED** This problem involves comparing, for the image formed by a digital camera, the resolution limit due to diffraction and the limit due to the size of the pixels in the sensor. The limit due to diffraction is the size of the image of a point source; the limit due to the sensor is the width of a single pixel. The greater limit determines the image resolution. The problem does not specify the wavelength of the light involved, but because the problem is concerned with forming images with visible light, I choose $\lambda = 500$ nm, near the center of the visible spectrum.

❷ **DEVISE PLAN** The diffraction-limited image of a point source is the Airy disk at the center of the diffraction pattern, so I can use Eq. 34.31 to obtain the radius of the Airy disk formed by the camera. I can then compare the diameter (not the radius) of that disk with the width of a pixel. Whichever is greater limits the resolution possible for the image.

❸ **EXECUTE PLAN** Substituting the values given into Eq. 34.31, I obtain for the Airy disk radius

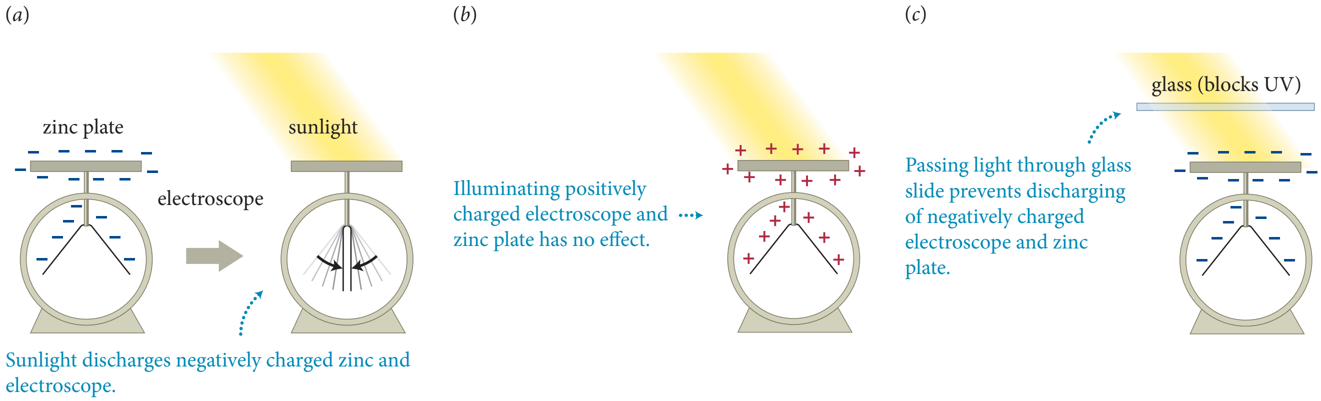
$$y = 1.22\frac{\lambda f}{d} = 1.22\frac{(0.500 \times 10^{-6}\text{ m})(30 \times 10^{-3}\text{ m})}{40 \times 10^{-3}\text{ m}} = 0.46 \times 10^{-6}\text{ m} = 0.46\ \mu\text{m}.$$

The diameter of the Airy disk is thus $2y = 0.92\ \mu$m. This is significantly less than the pixel width, which means the resolution of the image is limited by pixel width rather than by diffraction. ✔

❹ **EVALUATE RESULT** My result for the radius of the Airy disk is reasonable because typically the diffraction-limited width of the image of a point source is comparable to the wavelength of light emitted by the source. (Note that if I had chosen $\lambda = 700$ nm, the longest visible wavelength, the Airy disk radius would increase only by a factor of $700/500 = 1.4$ and thus its diameter, 1.3 μm, would still not exceed the pixel width. If I had chosen a wavelength shorter than 500 nm, the Airy disk diameter would be less than my calculated value. So my conclusion is the same for any visible wavelength: The resolution is limited by the pixel width.).

34.19 Which of these three lenses offers (*a*) the highest resolution and (*b*) the lowest resolution: (*i*) $f = 10$ mm, $d = 8$ mm; (*ii*) $f = 15$ mm, $d = 10$ mm, (*iii*) $f = 20$ mm, $d = 18$ mm?

Figure 34.46 The photoelectric effect.



34.10 Photon energy and momentum

In Chapter 30 we described light as an electromagnetic wave that has a wavelength λ and a frequency f and moves in vacuum at speed c_0, such that

$$c_0 = \lambda f. \tag{34.32}$$

We also found that the energy density in the electromagnetic wave is proportional to the square of the amplitude of the electric field oscillation. In addition, the experiment described in Section 34.4 suggests that light has particle properties. If light always propagates at speed c_0 in vacuum, what determines its energy? And if light is a particle that has energy, shouldn't it also have momentum?

The answer to the first question is provided by the **photoelectric effect,** a surprising phenomenon that cannot be explained by thinking of light as a wave (**Figure 34.46**). Place a piece of metal, such as zinc, on an electroscope (see Section 22.3) that is negatively charged, as shown in Figure 34.46*a*. Some of the charge immediately moves to the zinc so that it, too, is negatively charged. If you then shine sunlight on the metal, the light discharges the zinc and the electroscope. If the electroscope is positively charged, however, as in Figure 34.46*b*, nothing happens when light shines on it. If we place a piece of glass in the beam of light (Figure 34.46*c*), nothing happens even if the zinc plate is negatively charged and the light is very intense.

What is going on? In the situations of Figure 34.46*a* and *b*, the light knocks electrons out of the zinc plate. When the plate is initially negatively charged, like-charge repulsion causes the ejected electrons to accelerate away from the plate. When the plate is initially positively charged, opposite-charge attraction causes the ejected electrons to be attracted back to the plate, so that the charge on the plate does not change. Ultraviolet radiation cannot pass through ordinary glass, and thus we conclude from the situation in Figure 34.46*c* that ultraviolet light is essential in order for electrons to be ejected.

The apparatus illustrated in **Figure 34.47** is used to study the photoelectric effect. It allows us to measure the energy of the ejected electrons while separately controlling either the wavelength or the intensity of the light. A zinc target T is placed in an evacuated quartz bulb (quartz is transparent to ultraviolet light), along with another metal electrode called the collector (C). A power supply is used to maintain a constant potential difference V_{CT} between the target and the collector. The current from the target to the collector is measured with an ammeter. If the target is kept at a negative potential relative to the collector, so that V_{CT} is negative, any electrons ejected from the target by the light are accelerated by the electric field and move to the collector. With negative V_{CT}, the current measured is proportional to the intensity of the light source, suggesting that ejecting each electron requires a certain amount of light energy.

If the potential difference V_{CT} is made slightly positive (so that the target is positive relative to the collector), there is a small current detected, but the electric field

Figure 34.47 Apparatus to study the photoelectric effect. The potential difference V_{CT} is positive when $V_T > V_C$.

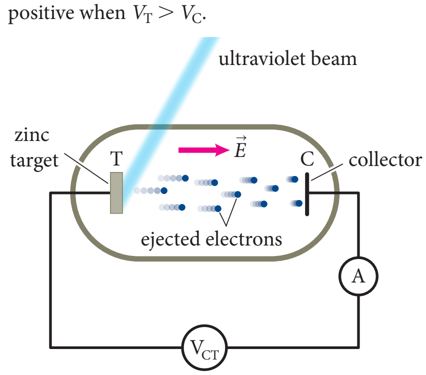
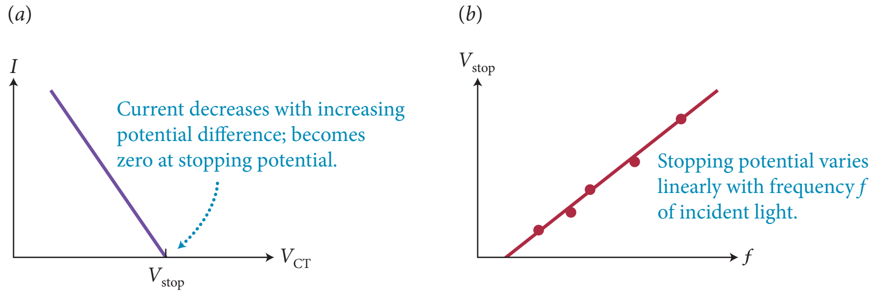


Figure 34.48 For the circuit in Figure 34.47, the current as a function of potential difference and the stopping potential as a function of the frequency of the incident light.



between T and C now *slows down* any ejected electrons that initially move toward the collector. As V_{CT} increases, there is a certain value of V_{CT} at which the flow of electron stops completely, as shown in the graph of current I versus V_{CT} in **Figure 34.48*a***. At this potential difference, called the *stopping potential difference,* the current is zero regardless of the intensity of the incident light. No matter how bright the light, there is no current between the target and the collector. This finding implies that the maximum kinetic energy with which the electrons leave the target does not depend on the intensity (and thus the incident power) of the light.

If we measure the kinetic energy of the ejected electrons (for $V_{CT} < V_{stop}$) directly, we discover that not all of them have the same kinetic energy. This happens because although the amount of energy absorbed from the photons is the same for all electrons, the energy required for the electron to make its way from its initial location in the target to the surface depends on the depth from which the electron is liberated. As a result, electrons released from the surface of the target have the maximum possible kinetic energy, which equals the amount of energy transferred to each electron by the light minus the energy required to liberate the electron from the metal.

For a given potential difference between the target and the collector, the electric field does work $-eV_{CT}$ on an electron as the electron moves from the target to the collector (see Eq. 25.17). The change in the electron's kinetic energy is thus

$$\Delta K = -eV_{CT}. \qquad (34.33)$$

Given that the electrons just barely reach the collector at the stopping potential difference, we know that their final kinetic energy is zero, and so for these electrons $\Delta K = K_f - K_i = -K_i$. The maximum kinetic energy with which the electrons leave the target is thus

$$K_{max} = K_i = eV_{stop}. \qquad (34.34)$$

Another clue to understanding the experiment in Figure 34.47 emerges when we change the frequency of the incident light and again measure the target-to-collector current as a function of V_{CT}. We observe that the stopping potential difference depends on the frequency; plotting this stopping potential difference as a function of the frequency of the light yields the results shown in Figure 34.48*b*.

34.20 (*a*) What does Figure 34.48*b* tell you about the relationship between the frequency of the incident light and the maximum kinetic energy of the ejected electrons? (*b*) What does the intercept of the line through the data points and the horizontal axis represent?

As Checkpoint 34.20, part *a* shows, the maximum kinetic energy of the ejected electrons depends on the frequency of the incident light, not its intensity. Electrons ejected by ultraviolet light, which has a higher frequency than visible light, have more kinetic energy than electrons ejected by visible light. (This is why putting glass in the beam of light in the experiment shown in Figure 34.46 essentially eliminates the effect. Although some electrons are ejected by the visible light, those electrons are liberated with less kinetic energy and are more likely to return to the target.) Furthermore, as you discovered in answering Checkpoint 34.20, part *b*, there is a certain minimum frequency of light below which electrons are not ejected at all, regardless of the intensity of the light.

Why does the photoelectric effect require us to think of light as a particle rather than as a wave? The stopping potential difference gives us the maximum kinetic energy with which electrons are released; the light must supply at least this much energy to the electrons in order to eject them. If light could be understood solely as a wave, the intensity of the wave, not its frequency, would determine the maximum amount of energy it could deliver to the electrons. Because the stopping potential difference depends not on light intensity but on frequency, we infer that light carries its energy in energy quanta and that the energy in each quantum is proportional to the frequency.

The photons described in Section 34.5 are these quanta. When an electron absorbs a photon, the electron acquires the photon's entire energy—the electron cannot absorb just part of a photon. The photon's energy frees the electron from the material and gives it additional kinetic energy. If we denote the minimum energy required to free the electron by E_0, we have

$$E_{\text{photon}} = hf = K_{\max} + E_0, \tag{34.35}$$

where $K_{\max}$ is the maximum kinetic energy of the electron as it is ejected. The energy E_0, called the **work function** of the target metal, is a property of the metal that measures how tightly electrons are bound to the metal.

The value of Planck's constant h can be determined by using the relationship between V_{stop} and f given in Figure 34.48*b*. Substituting Eq. 34.34 into Eq. 34.35 and solving the result for V_{stop}, we get

$$V_{\text{stop}} = \frac{h}{e}f - \frac{E_0}{e}. \tag{34.36}$$

This result shows that V_{stop} depends linearly on f and that the slope of the line in Figure 34.48*b* is h/e. By measuring the slope in Figure 34.48*b* and dividing that slope by the charge e of the electron, one obtains $h = 6.626 \times 10^{-34}\ \text{J} \cdot \text{s}$, the value given in Section 34.4.

As discussed in Section 34.5 and expressed in Eq. 34.35, Planck's constant relates the energy and frequency of a photon, $E_{\text{photon}} = hf_{\text{photon}}$, and in Section 34.4 we learned the relationship between the momentum and the wavelength of an electron (or anything else that is ordinarily thought of as a particle): $\lambda_{\text{electron}} = h/p_{\text{electron}}$. Because of the wave-particle duality, we can apply this expression for wavelength to photons as well as electrons: $\lambda_{\text{photon}} = h/p_{\text{photon}}$. If we calculate the momentum of a photon from its wavelength using this expression and then substitute $\lambda = c_0/f$, we see that the momentum of a photon is proportional to its energy:

$$p_{\text{photon}} = \frac{h}{\lambda_{\text{photon}}} = \frac{hf_{\text{photon}}}{c_0} = \frac{E_{\text{photon}}}{c_0}. \tag{34.37}$$

If we substitute this result into the equation relating energy and momentum derived in Chapter 14 (Eq. 14.57)

$$E^2 - (c_0 p)^2 = (mc_0^2)^2, \tag{34.38}$$

the left side of this equation becomes zero, and so we see that photons have zero mass ($m_{\text{photon}} = 0$). We derived Eq. 34.38 for particles that have nonzero mass. Now we see that we can treat photons as massless "particles of light." While these particles have no mass, they do have both momentum and energy:

$$E_{\text{photon}} = hf_{\text{photon}} \tag{34.39}$$

$$p_{\text{photon}} = \frac{hf_{\text{photon}}}{c_0}, \tag{34.40}$$

and, unlike ordinary particles, they always move at the speed of light c_0. Remember also that the mass of a particle is associated with the internal energy of that particle (Eq. 14.54), so $m_{\text{photon}} = 0$ means that photons have no internal energy and therefore no internal structure.

Example 34.11 Photoelectric effect

Light of wavelength 380 nm strikes the metal target in Figure 34.47. As long as the potential difference V_{CT} between the target and the collector is no greater than $+1.2$ V, there is a current in the circuit. Determine the longest wavelength of light that can eject electrons from this metal.

❶ **GETTING STARTED** To solve this problem, I recognize that the longest wavelength of light that can eject electrons corresponds to the lowest-energy photon that can eject an electron; this energy is equal to the work function. I therefore need to use the idea of stopping potential difference to determine the work function from the information given in the problem.

❷ **DEVISE PLAN** The fact that the current is zero when $V_{\text{CT}} > +1.2$ V tells me that the stopping potential difference is 1.2 V. Equation 34.36 gives the relationship between photon frequency and stopping potential difference. I can use the relationship between photon frequency and wavelength to rewrite Eq. 34.36 in terms of wavelength. Finally, Eq. 34.35 shows that the lowest photon energy comes when $K_{\max} = 0$ so that the lower energy equals the work function E_0. Therefore I must determine the wavelength of a photon that has an energy equal to the work function, and for this I can use the expression I developed for the relationship between photon energy and wavelength.

❸ **EXECUTE PLAN** Solving Eq. 34.36 for E_0, then substituting c_0/λ for f and inserting numerical values, I obtain

$$\begin{aligned} E_0 &= \frac{hc_0}{\lambda} - eV_{\text{stop}} \\ &= \frac{(6.626 \times 10^{-34}\ \text{J}\cdot\text{s})(2.998 \times 10^{8}\ \text{m/s})}{380 \times 10^{-9}\ \text{m}} \\ &\quad - (1.602 \times 10^{-19}\ \text{C})(1.2\ \text{V}) \\ &= 3.3 \times 10^{-19}\ \text{J}, \end{aligned}$$

where I have used the equality $1\ \text{V} \equiv 1\ \text{J/C}$ (Eq. 25.16). Because the longest wavelength that can eject electrons has energy equal to the work function, I solve $E_0 = hc_0/\lambda$ for λ and substitute the value of E_0 I just calculated to obtain this maximum wavelength:

$$\lambda = \frac{hc_0}{E_0} = \frac{(6.626 \times 10^{-34}\ \text{J}\cdot\text{s})(2.998 \times 10^{8}\ \text{m/s})}{3.3 \times 10^{-19}\ \text{J}} = 0.60\ \mu\text{m}. ✔$$

❹ **EVALUATE RESULT** This value for the longest wavelength that can eject electrons is greater than 380 nm, the wavelength corresponding to the stopping potential difference of 1.2 V, as it should be.

34.21 A photon enters a piece of glass for which the index of refraction is about 1.5. What happens to the photon's (*a*) speed, (*b*) frequency, (*c*) wavelength, and (*d*) energy?

QUANTITATIVE TOOLS

Chapter Glossary

SI units of physical quantities are given in parentheses.

Bragg condition The condition under which x rays diffracted by planes of atoms in a crystal lattice interfere constructively. For x rays of wavelength λ, diffracting from a crystal lattice with spacing d between adjacent planes of atoms at an angle θ between the incident rays and the normal to the scattering planes, the condition states that $2d \cos\theta = m\lambda$.

de Broglie wavelength λ (m) The wavelength associated with the wave behavior of a particle, $\lambda = h/p$.

diffraction grating An optical component with a periodic structure of equally spaced slits or grooves that diffracts and splits light into several beams that travel in different directions. When a diffraction grating is made up of slits, light passes through it and it is called a *transmission diffraction grating*; when a diffraction grating is made up of grooves, light reflects from it and it is called a *reflection grating*. The so-called *principal maxima* in the intensity pattern created by a diffraction grating occur at angles given by

$$d \sin\theta_m = \pm m\lambda, \quad \text{for } m = 0, 1, 2, 3, \ldots \tag{34.16}$$

and minima occur at angles give by

$$d \sin\theta_{\min} = \pm \frac{k}{N}\lambda \tag{34.17}$$

for an integer k that is not an integer multiple of N.

fringe order m, or n (unitless) A number indexing interference fringes; the central bright fringe is called zeroth order ($m = 0$), and the index increases with distance from the central bright fringe. The dark fringes flanking the central bright fringe are first order ($n = 1$).

interference fringes A pattern of alternating bright and dark bands cast on a screen produced by coherent light passing through very small, closely spaced slits, apertures, or edges.

photoelectric effect The emission of electrons from matter as a consequence of their absorption of energy from electromagnetic radiation with photon energy greater than the work function.

photon The indivisible, discrete basic unit, or quantum, of light. A photon of frequency f_{photon} has energy

$$E_{\text{photon}} = hf_{\text{photon}} \tag{34.39}$$

and momentum

$$p_{\text{photon}} = \frac{hf_{\text{photon}}}{c_0}. \tag{34.40}$$

Planck's constant h (J · s) The fundamental constant that relates the energy of a photon to its frequency and also the de Broglie wavelength and momentum of a particle: $h = 6.626 \times 10^{-34}$ J · s.

Rayleigh's criterion Two features in the image formed by a lens can be visually separated (and are then said to be *resolved*) if they satisfy Rayleigh's criterion. For a lens of diameter d and light of wavelength λ, the minimum angular separation θ_r for which two sources can be resolved is

$$\theta_r \approx 1.22 \frac{\lambda}{d}. \tag{34.30}$$

wave-particle duality The possession of both wave properties and particle properties, observed both for all atomic-scale material particles and for photons.

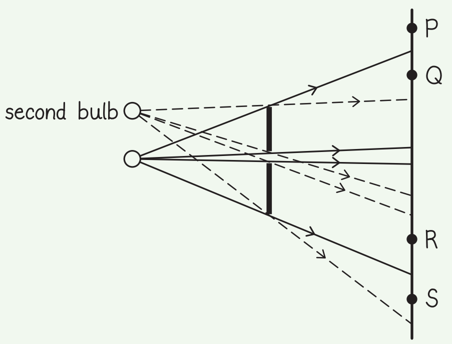
work function E_0 (J) The minimum energy required to free an electron from the surface of a metal. This energy measures how tightly the electron is bound to the metal.

x rays Electromagnetic waves that have wavelengths ranging from 0.01 nm to 10 nm.

Chapter 33

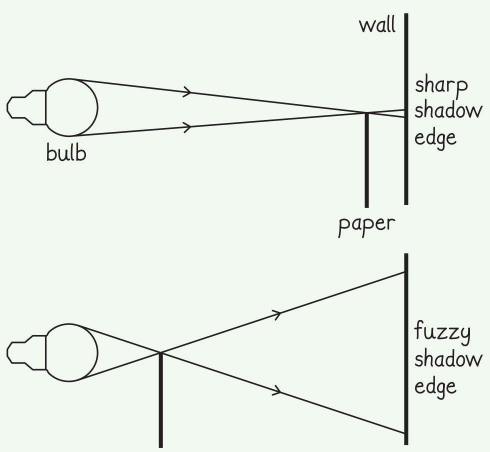
33.1 (*a*) See Figure S33.1. In considering whether the brightness of any location on the screen has changed, note that the distribution of light from the first bulb has not changed, which means the brightness of a particular location changes only if additional light is cast on that location by the second bulb. If the second bulb does not cast light on a particular location, the brightness does not change. (The fact that a given location would be shadowed if it were illuminated by only the second bulb does not decrease the brightness of the light from the first bulb.) The brightness of the spot created by the first bulb does not change because no light from the second bulb strikes this spot. (*b*) Locations P and Q are now brighter than before because some light from the second bulb strikes them. Locations R and S are unaffected because no light from the second bulb reaches them.

Figure S33.1



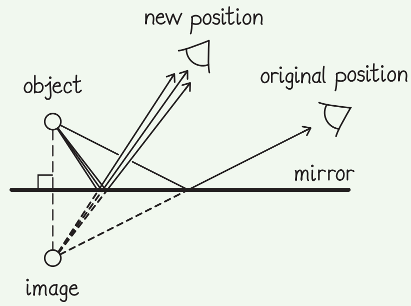
33.2 See Figure S33.2. The shadow edges become sharper as you move the paper farther from the bulb and fuzzier as you move the paper closer to the bulb. This happens because the bulb is not a point source of light. Rays from different parts of the bulb's surface pass by the edge of the paper at slightly different angles. When the paper is farther from the bulb, the difference in these angles is less and so the edge of the shadow is sharper. When the paper is close to the bulb, the angular size of the filament (as seen from the edge) is greater, which makes the edge of the shadow fuzzier.

Figure S33.2



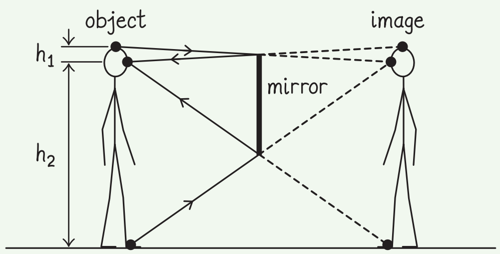
33.3 No. Figure S33.3 shows the reflected rays that reach the observer in the two locations. Because the angle of incidence always equals the angle of reflection when the reflecting surface is smooth, any ray from the object that reflects anywhere on the mirror can be traced back through the mirror to a location directly behind the object. Moving the observer changes only which subset of the reflected rays the observer sees and not where the rays appear to come from.

Figure S33.3



33.4 No. Figure S33.4 shows the rays that reflect into the person's eye from the highest and lowest points on his body. These rays strike the mirror at a height midway between where they originate on the person and the person's eye. So the mirror needs to extend only from halfway between the person's eye and the highest point on the person's body to halfway between the person's eye and the lowest point on the person's body. The ray from the highest point strikes the mirror at a height midway between the height h_1 in Figure S33.4 at which the ray originates and the height at which it strikes the eye, and the same is true for the ray from the lowest point. So the top of the mirror can be at a height that is half the height h_1 above the height of the eye, and the bottom of the mirror can be at a height that is half the height h_2 below the height of the eye. Because $h_1 + h_2$ is the person's height, this means the mirror needs to be only half this height. (However, the mirror must be positioned at the right height.)

Figure S33.4

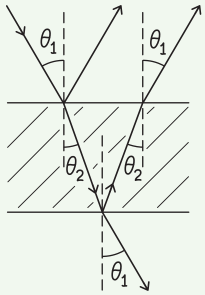


33.5 The wavefronts arrive at a given location in the glass at the same frequency that they arrive at any location in vacuum. However, they travel only two-thirds as fast in the glass as in vacuum. Sequential wavefronts arrive at the glass surface at instants separated by the period T of the wave ($T = 1/f$). If the wave traveled at the same speed in glass as it does in vacuum, then during one period, one wavefront would travel a distance into the glass equal to the vacuum wavelength (400 nm) before the next wavefront arrived at the surface of the glass, and the wavefronts would be separated by 400 nm. However, because the wavefronts travel at only two-thirds the speed of light in vacuum, the spacing between wavefronts in the glass must be two-thirds of the vacuum wavelength, or 267 nm.

33.6 The propagation of the wavefronts is exactly like that shown in Figure 33.16*a* with the direction of propagation reversed. Consequently, because the angle of incidence equals θ_2, the angle of refraction must equal θ_1.

33.7 The angle between the ray reflected from the bottom surface and the normal is θ_2, as shown in Figure S33.7 on the next page. This ray is now incident at the top surface and refracted as it emerges into the air. This is equivalent to the situation shown in Figure 33.17*b*, in which the ray originates in the glass and is refracted in the air. Therefore, as discussed in Example 33.3, in air the angle this refracted ray makes with the normal is θ_1. Note in Figure S33.7 that this refracted ray is shifted sideways relative to the ray *reflected* from the top surface.

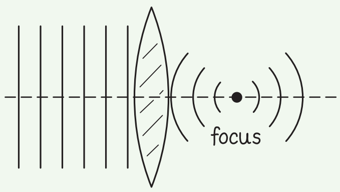
Figure S33.7



33.8 The critical angle corresponds to the angle of incidence for which the angle of refraction is 90°. When entering the medium, to obtain a given angle of refraction, the angle of incidence of a violet ray must be greater than that of a red ray; therefore the critical angle is smaller for a violet ray than for a red ray.

33.9 See Figure S33.9. The wavefronts are perpendicular to the light rays everywhere.

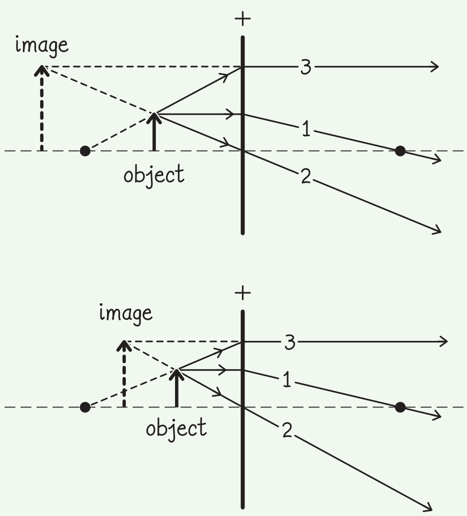
Figure S33.9



33.10 No. The image location can be found with any two of the three principal rays. Because all three intersect at one point, any two of them indicate the point of intersection and therefore the image location. It is, however, useful to draw all three rays in order to check that you have made no mistakes.

33.11 Decrease. See Figure S33.11. As the object moves closer to the lens, ray 3 makes a smaller angle with the lens axis and thus intercepts the lens closer to the axis. Ray 1 remains the same, and ray 2 makes a greater angle with the axis in order to pass through the center of the lens. As a result, the image point—the virtual intersection of the rays—is closer to the axis (as well as the lens), making the image smaller.

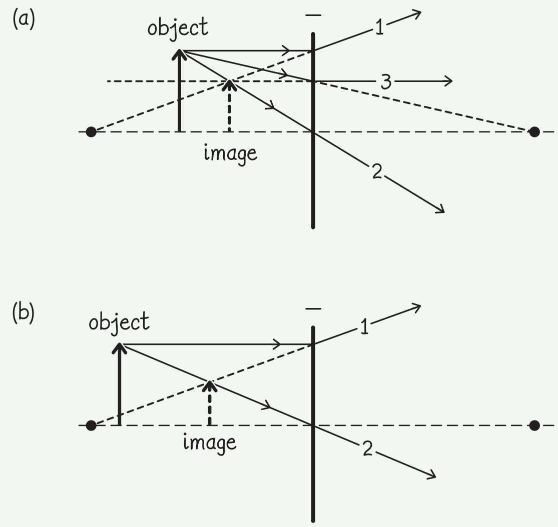
Figure S33.11



33.12 Virtual, because the rays do not actually intersect anywhere on the image. They merely appear to originate from a common point on the virtual image.

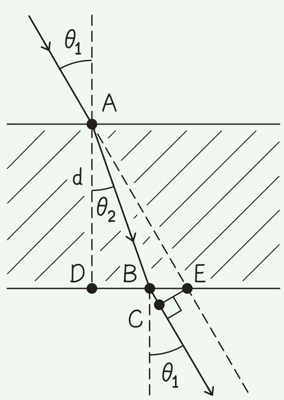
33.13 (*a*) See Figure S33.13*a*. You draw ray 3 by first drawing a line connecting the point of interest on the object with the focus on the side of the lens opposite the object. Run ray 3 along this line from the object to the lens surface; then change the ray direction so that it emerges parallel to the lens axis. (*b*) There is no position for which the image is larger than the object. To see why, all you need to consider is principal rays 1 and 2. When the object is moved, the direction of ray 1 does not change—only ray 2 changes direction. When the object is moved to the left, as in Figure S33.13*b*, ray 2 now intersects ray 1 closer to the lens axis and the image is smaller. Moving the object right, toward the lens, makes the image larger, but it cannot become larger than the object because on the left side of the lens where the image forms, ray 1 is always below the tip of the object. As shown in Figure 33.33, placing the object beyond the object-side focus also produces an image that is smaller than the object. (*c*) It is not possible to form a real image with a diverging lens because rays diverge when they go through it.

Figure S33.13



33.14 See Figure S33.14. (*a*) You need to determine the distance CE in terms of the angle of incidence θ_1, slab thickness d = AD, and slab index of refraction n_2. Angle CBE = 90° − θ_1, and so angle BEC = θ_1. Therefore CE = BE cos θ_1. To determine the distance BE, you can say DB + BE = DE, DE = $d \tan \theta_1$, and DB = $d \tan \theta_2$.

Figure S33.14



Substituting the second and third expressions in the first, you obtain $d \tan \theta_2$ + BE = $d \tan \theta_1$, which can be solved for BE:

$$\mathrm{BE} = d(\tan\theta_1 - \tan\theta_2) = d\left(\frac{\sin\theta_1}{\cos\theta_1} - \frac{\sin\theta_2}{\cos\theta_2}\right). \quad (1)$$

Applying Snel's law (Eq. 33.7) to this case and solving Eq. 33.7 for $\sin\theta_2$ give you $\sin\theta_2 = n_1\sin\theta_1/n_2$. Next you can use the identity $\sin^2\theta + \cos^2\theta = 1$ to get an expression for $\cos\theta_2$:

$$\cos\theta_2 = \sqrt{1-\sin^2\theta_2} = \sqrt{1-\left(\frac{n_1\sin\theta_1}{n_2}\right)^2}.$$

You can substitute these expressions for $\sin\theta_2$ and $\cos\theta_2$ into Eq. 1:

$$\mathrm{BE} = d\left[\frac{\sin\theta_1}{\cos\theta_1} - \frac{n_1\sin\theta_1}{n_2\sqrt{1-\left(\frac{n_1\sin\theta_1}{n_2}\right)^2}}\right],$$

which simplifies to

$$\mathrm{BE} = d\sin\theta_1\left(\frac{1}{\cos\theta_1} - \frac{n_1}{\sqrt{n_2^2 - n_1^2\sin^2\theta_1}}\right).$$

From this you can determine the distance CE:

$$\mathrm{CE} = \mathrm{BE}\cos\theta_1 = d\sin\theta_1\left(1 - \frac{n_1\cos\theta_1}{\sqrt{n_2^2 - n_1^2\sin^2\theta_1}}\right).$$

(*b*) $$\mathrm{CE} = (0.010\ \mathrm{m})(0.50)\left(1 - \frac{(1)(0.87)}{\sqrt{(1.5)^2 - (1)^2(0.5)^2}}\right)$$
$$= 0.0019\ \mathrm{m}.$$

Thus, at an angle of incidence of 30°, in passing through a glass slab that is 10 mm thick (about three times the thickness of a typical windowpane), a light ray is shifted sideways by 2 mm (20% of the thickness of the slab).

33.15 (*a*) f is negative (lens is diverging), so $f = -80$ mm; o is positive (object is on same side as illumination), so $o = 100$ mm. Substituting these values into Eq. 33.16 gives

$$\frac{1}{i} = \frac{1}{-80\ \mathrm{mm}} - \frac{1}{100\ \mathrm{mm}} = -0.0225\ \mathrm{mm}^{-1},$$

and solving for i gives $i = -44$ mm. The value of i is negative, as it should be for a virtual image, and the absolute value of i is less than o, which is consistent with Figure 33.33.

(*b*) Equation 33.17 gives

$$M = \frac{h_\mathrm{i}}{h_\mathrm{o}} = \frac{-i}{o} = \frac{-(-44\ \mathrm{mm})}{100\ \mathrm{mm}} = 0.44.$$

The image height is 44% of the object height.

33.16 (*a*) Solving Eq. 33.21 for f gives $f = (0.25\ \mathrm{m})/M_\theta$, and substituting this result into Eq. 33.22 gives $d = 4M_\theta$. For $M_\theta = 8$, you have $d = +32$ diopters.

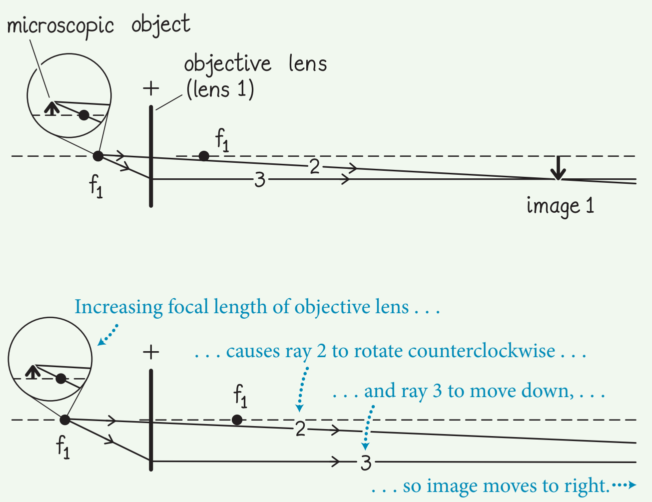
(*b*) $f = (0.25\ \mathrm{m})/M_\theta = (0.25\ \mathrm{m})/8 = 0.031\ \mathrm{m} = 31$ mm.

33.17 (*a*) The image moves farther from the objective lens. This can be seen either by constructing a ray diagram (Figure S33.17) or by considering the relationship among object distance, image distance, and focal length (Eq. 33.16). Solving Eq. 33.16 for i_1 gives

$$\frac{1}{i_1} = \frac{1}{f_1} - \frac{1}{o_1} = \frac{o_1 - f_1}{o_1 f_1} \quad \text{or} \quad i_1 = \frac{o_1 f_1}{o_1 - f_1}.$$

Increasing both o_1 and f_1 will increase the numerator of the expression for i_1; keeping the sample "just outside the focal point" of the lens implies that the distance between the sample and the lens, $o_1 - f_1$, is kept at least roughly the same, so increasing the numerator while keeping the denominator constant increases i_1. (Practical limitations on the construction of lenses actually allow the distance $o_1 - f_1$ to be smaller for shorter-focal-length lenses, so in fact, the denominator also increases as the focal length increases. However, this last effect is not a consequence of the simple thin-lens treatment but of ways that lenses are not ideal that are beyond the scope of this text.)

Figure S33.17



(*b*) Shorter focal lengths give a more compact microscope, because the first image must fall close to the focal point of the eyepiece lens, and thus the distance between the two lenses is roughly $i_1 + f_2$. From part *a* we know that i_1 depends on f_1, and so decreasing f_1 decreases i_1.

33.18 (*a*) Using the result of Example 33.9 gives you $f_1 = M_\theta f_2 = (22)\,(0.0400\ \mathrm{m}) = 0.88$ m.

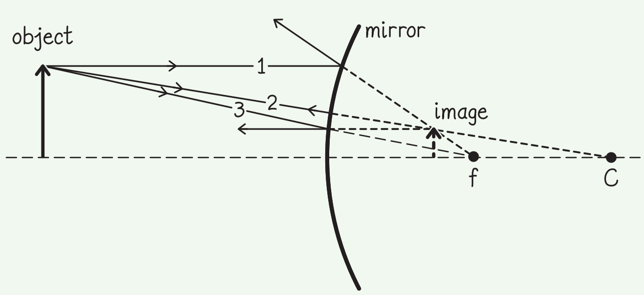
(*b*) The length of the telescope is roughly the sum of the focal lengths of the two lenses (see Figure 33.50): 0.88 m + 0.04 m = 0.92 m. Because the eyepiece lens is by design a short-focal-length lens, the length of the telescope is determined primarily by the focal length of the objective lens.

33.19 (*a*) The ray diagram in Figure S33.19 shows that the image is behind the mirror and virtual. You calculate the image distance with Eq. 33.24. The focal length is half the radius of curvature, C = 1.0 m, and is negative because it is a virtual focus:

$$i = \left(\frac{1}{f} - \frac{1}{o}\right)^{-1} = \left(\frac{1}{-0.50} - \frac{1}{1.0}\right)^{-1} = -0.33\ \mathrm{m}.$$

The negative value for i tells you that the image is located 0.33 m behind the mirror.

Figure S33.19



(*b*) $$h_\mathrm{i} = -\frac{h_\mathrm{o} i}{o} = -\frac{(0.30\ \mathrm{m})(-0.33\ \mathrm{m})}{(1.0\ \mathrm{m})} = 0.10\ \mathrm{m}.$$

The image is smaller than the object, and the positive value for h tells you that the image is upright.

33.20 The only change to the physics behind the derivation of the lensmaker's formula is that Eq. 33.25 becomes $n_1 \sin \theta_i = n_2 \sin \theta_r$. In the paraxial approximation, this result can be written as $(n_1/n_2)\theta_i = \theta_r$. Comparing this with Eq. 33.26, you see that the effect of submerging the lens is to substitute (n_1/n_2) for n. Once you make this change in the remainder of the derivation, the lensmaker's formula becomes

$$\frac{1}{f} = \left(\frac{n_1}{n_2} - 1\right)\left(\frac{1}{R_1} + \frac{1}{R_2}\right).$$

Chapter 34

34.1. Once the fronts reach P, the intensity at P is the same as the intensity at any point on any of the wavefronts. For any point R above or below the ray through P, you can locate a corresponding point R′ for which the distance R′P is exactly half a wavelength greater than the distance RP. Those points therefore do not contribute to the intensity at P. Only the point exactly on the ray (dashed line) contributes.

34.2. If you identify pairs of points that radiate in phase at P, then for each pair, there exists another pair that is exactly 180° out of phase with the first pair and thus cancels the radiation from the first pair.

34.3. (*a*) The spacing between fringes decreases as the separation between the slits increases because $\sin\theta$ is proportional to $1/d$. (*b*) The spacing between fringes increases as the wavelength increases because $\sin\theta$ is proportional to λ.

34.4. No. Two of the waves cancel each other perfectly, but the third wave is not canceled. The intensity observed on the screen, at angles given by $d \sin\theta = \pm(n - \frac{1}{2})\lambda$, is the intensity of one of the three waves.

34.5. Fringe brightness is determined by the intensity (power/area) that strikes the screen. Increasing the number of slits increases the amount of power that travels through the slits to the screen. In addition, increasing the number of slits causes the fringes to sharpen, with the result that the power that reaches the screen is concentrated into narrower fringes, decreasing the area and thus increasing the brightness further.

34.6. (*a*) As you found in Checkpoint 34.3, the angle of the fringes increases with wavelength, and so the angle is less for violet light than for red. White light is thus dispersed into a rainbow, with the violet end of the spectrum at smaller angles than the red. (*b*) Each rainbow corresponds to the fringes of a particular order for all colors, or said another way, each order produces its own rainbow.

34.7. (*a*) The spacing between fringes in a two-slit interference pattern increases with increasing wavelength and with decreasing distance between slits (see Checkpoint 34.3). The same is true for diffraction gratings: The separation distance between adjacent fringes is increased by decreasing the separation distance between adjacent slits. Therefore the distance between adjacent slits should be decreased if your aim is to increase the diffraction grating's ability to separate close wavelengths. (*b*) No, as long as the slits are very narrow. The fringe spacing is determined entirely by the separation distance between adjacent slits. (Slit width does, however, affect the brightness of the pattern by determining how much light gets through the diffraction grating.)

34.8. Yes. The diffracted x rays interfere constructively at all angles for which the difference in path length is either zero (as it is for $\theta' = \theta$) or an integer multiple of λ. The diffracted beam at $\theta' = \theta$ corresponds to the zeroth-order maximum; other beams correspond to higher orders and are much weaker.

34.9. From Figure 34.20*b* you can see that $2\theta + 2\alpha = \pi$, so $\theta = -\alpha + \frac{\pi}{2}$. Using trigonometry, I get $\cos\theta = \cos(-\alpha + \frac{\pi}{2}) = -\sin(-\alpha) = \sin\alpha$. Substituting this into the Bragg condition yields $2d \sin\alpha = m\lambda$.

34.10. Crystal B. Spots closer together correspond to greater distances between crystal planes and hence greater atomic spacing.

34.11. It decreases, because the spacing between spots is proportional to the wavelength.

34.12. If the electrons travel more slowly, their wavelength increases, and the diffraction spots spread farther apart.

34.13. You could perform the experiment shown in Figure 34.6 with a light source so weak that the light "particles" pass through the pair of slits only one at a time. If light indeed has particle properties, a source this weak would give you a pattern that initially is like the one in Figure 34.25*a*, showing where individual particles of light hit the screen.

34.14. The maximum probability is at the locations of greatest intensity in the interference pattern; these locations are at angles θ for which $\sin\theta$ is a multiple of λ/d.

34.15. Although the maximum time-averaged intensity is greater than the sum of the intensities of the original two beams, averaging this time-averaged intensity over the entire area filled by the bright and dark bands of the interference pattern gives just the sum of the intensities of the original two beams, $2S_{0,av}$ (because the average of the cosine squared is one-half).

34.16. No, because the higher-order bright fringes exist only as long as $m\lambda/d < 1$. For values of m such that $m\lambda/d > 1$, there are no bright fringes because these correspond to $\sin\theta_m > 1$, for which there is no angle θ_m.

34.17. The oil forms a thin film that causes thin-film interference. The thickness of the oil layer determines for which wavelength constructive interference occurs. If the thickness of the film varies spatially, different wavelengths interfere constructively at different locations.

34.18. (*a*) If $a < \lambda$, there are no dark fringes because that would require $\sin\theta > 1$. This indicates that no minimum exists, and the light coming through an aperture with a diameter less than the wavelength acts as a true point source as assumed in our original discussion of multiple-slit interference. When $a < \lambda$, the slit acts as a point source with the wavelets spreading out spherically from the slit. (*b*) For $a \gg \lambda$, the first dark fringe occurs very close to $\theta = 0$. This means there is essentially no diffraction, and the beam just propagates straight ahead. This illustrates that single-slit diffraction is observed primarily with slits from a few wavelengths wide to tens of wavelengths wide.

34.19. Resolution is determined by wavelength and the ratio f/d: $y_r = 1.22\,\lambda f/d$. For a given wavelength, the greater f/d is, the larger the Airy disk. For the three lenses, f/d is (*i*) 1.3, (*ii*) 1.5, and (*iii*) 1.1. (*a*) The highest resolution is obtained with the smallest Airy disk, produced by lens *iii*. (*b*) The lowest resolution is obtained with the largest Airy disk, produced by lens *ii*. (The wavelength determines the exact size of the Airy disk; the comparisons made here assume the same wavelength for all three lenses.)

34.20. (*a*) The stopping potential difference V_{stop} is proportional to the frequency (Figure 34.48*b*) of the incident photons, and the maximum kinetic energy K_{max} of the ejected electrons is proportional to V_{stop} (Eq. 34.34). Therefore, K_{max} must be proportional to the frequency of the incident photons. (*b*) The minimum frequency the light can have and be able to eject electrons. Lower-frequency light does not eject electrons.

34.21. (*a*) The photon slows down because the speed of light is less in a medium that has an index of refraction of 1.5 than in air (index of refraction 1). (*b*) The frequency remains unchanged, as discussed in Chapter 33. (*c*) The wavelength decreases because the wavefronts travel less far in a given time interval. (*d*) The photon's energy does not change because the medium does not take away any energy. This is why photon energy must be expressed in terms of frequency rather than wavelength, because neither frequency nor energy depends on the medium, whereas wavelength does.