Richard Megna/Fundamental Photographs, NYC

Parallel coherent light from a laser, which acts as nearly a point source, illuminates these shears. Instead of a clean shadow, there is a dramatic diffraction pattern, which is a strong confirmation of the wave theory of light. Diffraction patterns are washed out when typical extended sources of light are used, and hence are not seen, although a careful examination of shadows will reveal fuzziness. We will examine diffraction by a single slit, and how it affects the double-slit pattern. We also discuss diffraction gratings and diffraction of X-rays by crystals. We will see how diffraction affects the resolution of optical instruments, and that the ultimate resolution can never be greater than the wavelength of the radiation used. Finally we study the polarization of light.

# Diffraction and Polarization

CONTENTS



**CHAPTER-OPENING QUESTION — Guess now!**

[*Don't worry about getting the right answer now — the idea is to get your preconceived notions out on the table.*]

Because of diffraction, a light microscope has a maximum useful magnification of about

**(a)** 50×;
**(b)** 100×;
**(c)** 500×;
**(d)** 2000×;
**(e)** 5000×;

and the smallest objects it can resolve have a size of about

**(a)** 10 nm;
**(b)** 100 nm;
**(c)** 500 nm;
**(d)** 2500 nm;
**(e)** 5500 nm.

Young's double-slit experiment put the wave theory of light on a firm footing. But full acceptance came only with studies on diffraction more than a decade later, in the 1810s and 1820s.

Diffraction, with regard to water waves as well as for light, refers to the spreading or bending of waves around edges. Here we look at diffraction in more detail, including its important practical effects of limiting the amount of detail, or *resolution*, that can be obtained with any optical instrument such as telescopes, cameras, and the eye.

Note: Sections marked with an asterisk (*) may be considered optional by the instructor.


From Chapter 35 of *Physics for Scientists & Engineers with Modern Physics*, Fourth Edition, Douglas C. Giancoli.

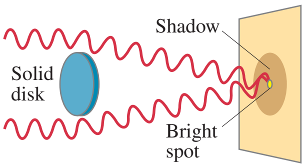
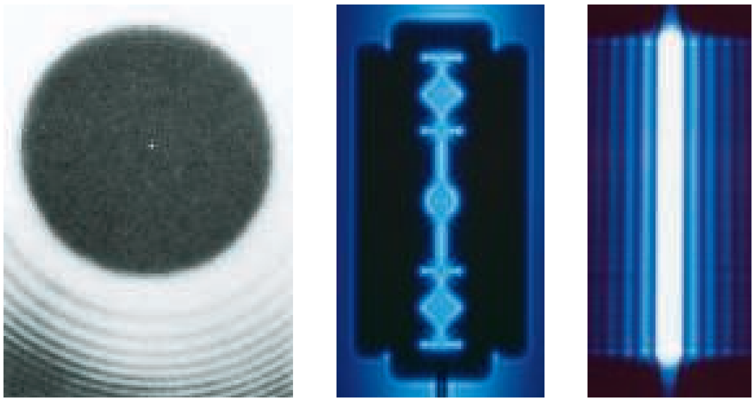



**FIGURE 1** If light is a wave, a bright spot will appear at the center of the shadow of a solid disk illuminated by a point source of monochromatic light.

In 1819, Augustin Fresnel (1788–1827) presented to the French Academy a wave theory of light that predicted and explained interference and diffraction effects. Almost immediately Siméon Poisson (1781–1840) pointed out a counter-intuitive inference: according to Fresnel's wave theory, if light from a point source were to fall on a solid disk, part of the incident light would be diffracted around the edges and would constructively interfere at the center of the shadow (Fig. 1). That prediction seemed very unlikely. But when the experiment was actually carried out by François Arago, the bright spot was seen at the very center of the shadow (Fig. 2a). This was strong evidence for the wave theory.

Figure 2a is a photograph of the shadow cast by a coin using a coherent point source of light, a laser in this case. The bright spot is clearly present at the center. Note also the bright and dark fringes beyond the shadow. These resemble the interference fringes of a double slit. Indeed, they are due to interference of waves diffracted around the disk, and the whole is referred to as a **diffraction pattern**. A diffraction pattern exists around any sharp-edged object illuminated by a point source, as shown in Fig. 2b and c. We are not always aware of diffraction because most sources of light in everyday life are not points, so light from different parts of the source washes out the pattern.



(a) (b) (c)

P. M. Rinard/American Journal of Physics

Ken Kay/Fundamental Photographs, NYC

Ken Kay/Fundamental Photographs, NYC

**FIGURE 2** Diffraction pattern of (a) a circular disk (a coin), (b) razorblade, (c) a single slit, each illuminated by a coherent point source of monochromatic light, such as a laser.

# 1 Diffraction by a Single Slit or Disk

To see how a diffraction pattern arises, we will analyze the important case of monochromatic light passing through a narrow slit. We will assume that parallel rays (plane waves) of light fall on the slit of width $D$, and pass through to a viewing screen very far away. If the viewing screen is not far away, lenses can be used to make the rays parallel.† As we know from studying water waves and from Huygens' principle, the waves passing through the slit spread out in all directions. We will now examine how the waves passing through different parts of the slit interfere with each other.

Parallel rays of monochromatic light pass through the narrow slit as shown in Fig. 3a. The slit width $D$ is on the order of the wavelength $\lambda$ of the light, but the slit's length (in and out of page) is large compared to $\lambda$. The light falls on a screen which is assumed to be very far away, so the rays heading for any point are very nearly parallel before they meet at the screen.

†Such a diffraction pattern, involving parallel rays, is called *Fraunhofer diffraction*. If the screen is close and no lenses are used, it is called *Fresnel diffraction*. The analysis in the latter case is rather involved, so we consider only the limiting case of Fraunhofer diffraction.

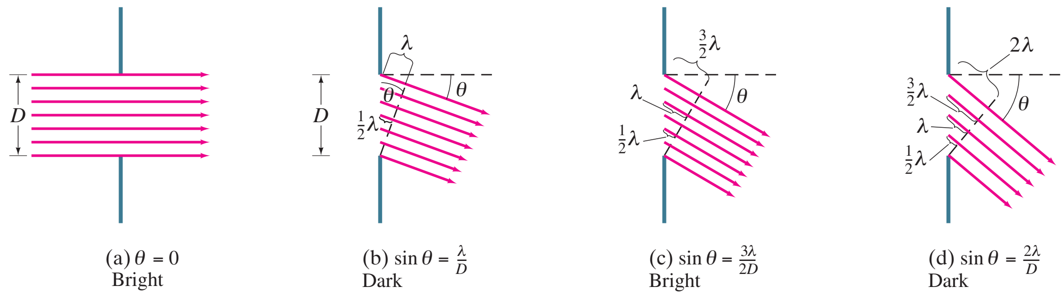



**FIGURE 3** Analysis of diffraction pattern formed by light passing through a narrow slit of width $D$.

First we consider rays that pass straight through as in Fig. 3a. They are all in phase, so there will be a central bright spot on the screen (see Fig. 2c). In Fig. 3b, we consider rays moving at an angle $\theta$ such that the ray from the top of the slit travels exactly one wavelength farther than the ray from the bottom edge to reach the screen. The ray passing through the very center of the slit will travel one-half wavelength farther than the ray at the bottom of the slit. These two rays will be exactly out of phase with one another and so will destructively interfere when they overlap at the screen. Similarly, a ray slightly above the bottom one will cancel a ray that is the same distance above the central one. Indeed, each ray passing through the lower half of the slit will cancel with a corresponding ray passing through the upper half. Thus, all the rays destructively interfere in pairs, and so the light intensity will be zero on the viewing screen at this angle. The angle $\theta$ at which this takes place can be seen from Fig. 3b to occur when $\lambda = D\sin\theta$, so

$$\sin\theta = \frac{\lambda}{D}. \qquad \text{[first minimum]} \quad \textbf{(1)}$$

The light intensity is a maximum at $\theta = 0°$ and decreases to a minimum (intensity = zero) at the angle $\theta$ given by Eq. 1.

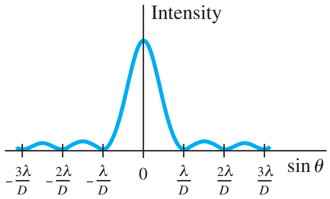
Now consider a larger angle $\theta$ such that the top ray travels $\frac{3}{2}\lambda$ farther than the bottom ray, as in Fig. 3c. In this case, the rays from the bottom third of the slit will cancel in pairs with those in the middle third because they will be $\lambda/2$ out of phase. However, light from the top third of the slit will still reach the screen, so there will be a bright spot centered near $\sin\theta \approx 3\lambda/2D$, but it will not be nearly as bright as the central spot at $\theta = 0°$. For an even larger angle $\theta$ such that the top ray travels $2\lambda$ farther than the bottom ray, Fig. 3d, rays from the bottom quarter of the slit will cancel with those in the quarter just above it because the path lengths differ by $\lambda/2$. And the rays through the quarter of the slit just above center will cancel with those through the top quarter. At this angle there will again be a minimum of zero intensity in the diffraction pattern. A plot of the intensity as a function of angle is shown in Fig. 4. This corresponds well with the photo of Fig. 2c. Notice that minima (zero intensity) occur on both sides of center at

$$D\sin\theta = m\lambda, \qquad m = \pm1, \pm2, \pm3, \cdots, \qquad \text{[minima]} \quad \textbf{(2)}$$

but *not* at $m = 0$ where there is the strongest maximum. Between the minima, smaller intensity maxima occur at approximately (not exactly) $m \approx \frac{3}{2}, \frac{5}{2}, \cdots$.

Note that the *minima* for a diffraction pattern, Eq. 2, satisfy a criterion that looks very similar to that for the *maxima* (bright fringes) for double-slit interference. Also note that $D$ is a single slit width, whereas $d$ in interference Equations is the distance between two slits.

**FIGURE 4** Intensity in the diffraction pattern of a single slit as a function of $\sin\theta$. Note that the central maximum is not only much higher than the maxima to each side, but it is also twice as wide ($2\lambda/D$ wide) as any of the others (only $\lambda/D$ wide each).




**CAUTION**

*Don't confuse Equations for interference with Eq. 1 for diffraction: note the differences*

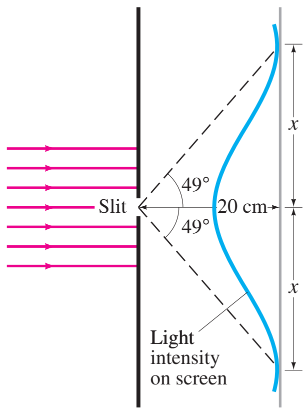



**FIGURE 5** Example 1.

**EXAMPLE 1** **Single-slit diffraction maximum.** Light of wavelength 750 nm passes through a slit $1.0 \times 10^{-3}$ mm wide. How wide is the central maximum (*a*) in degrees, and (*b*) in centimeters, on a screen 20 cm away?

**APPROACH** The width of the central maximum goes from the first minimum on one side to the first minimum on the other side. We use Eq. 1 to find the angular position of the first single-slit diffraction minimum.

**SOLUTION** (*a*) The first minimum occurs at

$$\sin\theta = \frac{\lambda}{D} = \frac{7.5 \times 10^{-7}\,\mathrm{m}}{1.0 \times 10^{-6}\,\mathrm{m}} = 0.75.$$

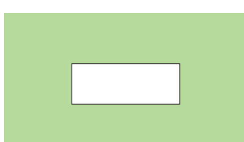
So $\theta = 49°$. This is the angle between the center and the first minimum, Fig. 5. The angle subtended by the whole central maximum, between the minima above and below the center, is twice this, or 98°.

(*b*) The width of the central maximum is $2x$, where $\tan\theta = x/20\,\mathrm{cm}$. So $2x = 2(20\,\mathrm{cm})(\tan 49°) = 46\,\mathrm{cm}$.

**NOTE** A large width of the screen will be illuminated, but it will not normally be very bright since the amount of light that passes through such a small slit will be small and it is spread over a large area. Note also that we *cannot* use the small-angle approximation here ($\theta \approx \sin\theta \approx \tan\theta$) because $\theta$ is large.

**EXERCISE A** In Example 1, red light ($\lambda = 750\,\mathrm{nm}$) was used. If instead yellow light at 575 nm had been used, would the central maximum be wider or narrower?

**FIGURE 6** Example 2.



**CONCEPTUAL EXAMPLE 2** **Diffraction spreads.** Light shines through a rectangular hole that is narrower in the vertical direction than the horizontal, Fig. 6. (*a*) Would you expect the diffraction pattern to be more spread out in the vertical direction or in the horizontal direction? (*b*) Should a rectangular loudspeaker horn at a stadium be high and narrow, or wide and flat?

**RESPONSE** (*a*) From Eq. 1 we can see that if we make the slit (width *D*) narrower, the pattern spreads out more. This is consistent with the study of waves. The diffraction through the rectangular hole will be wider vertically, since the opening is smaller in that direction. (*b*) For a loudspeaker, the sound pattern desired is one spread out horizontally, so the horn should be tall and narrow (rotate Fig. 6 by 90°).

## *2 Intensity in Single-Slit Diffraction Pattern

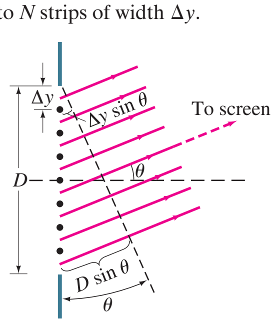
We have determined the positions of the minima in the diffraction pattern produced by light passing through a single slit, Eq. 2. We now discuss a method for predicting the amplitude and intensity at any point in the pattern using the phasor technique.

Let us consider the slit divided into $N$ very thin strips of width $\Delta y$ as indicated in Fig. 7. Each strip sends light in all directions toward a screen on the right. Again we take the rays heading for any particular point on the distant screen to be parallel, all making an angle $\theta$ with the horizontal as shown. We choose the strip width $\Delta y$ to be much smaller than the wavelength $\lambda$ of the monochromatic light falling on the slit, so all the light from a given strip is in phase. The strips are of equal size, and if the whole slit is uniformly illuminated, we can take the electric field wave amplitudes $\Delta E_0$ from each thin strip to be equal as long as $\theta$ is not too large. However, the separate amplitudes from the different strips will differ in phase. The phase difference in the light coming from adjacent strips will be

$$\Delta\beta = \frac{2\pi}{\lambda}\Delta y \sin\theta \qquad \textbf{(3)}$$

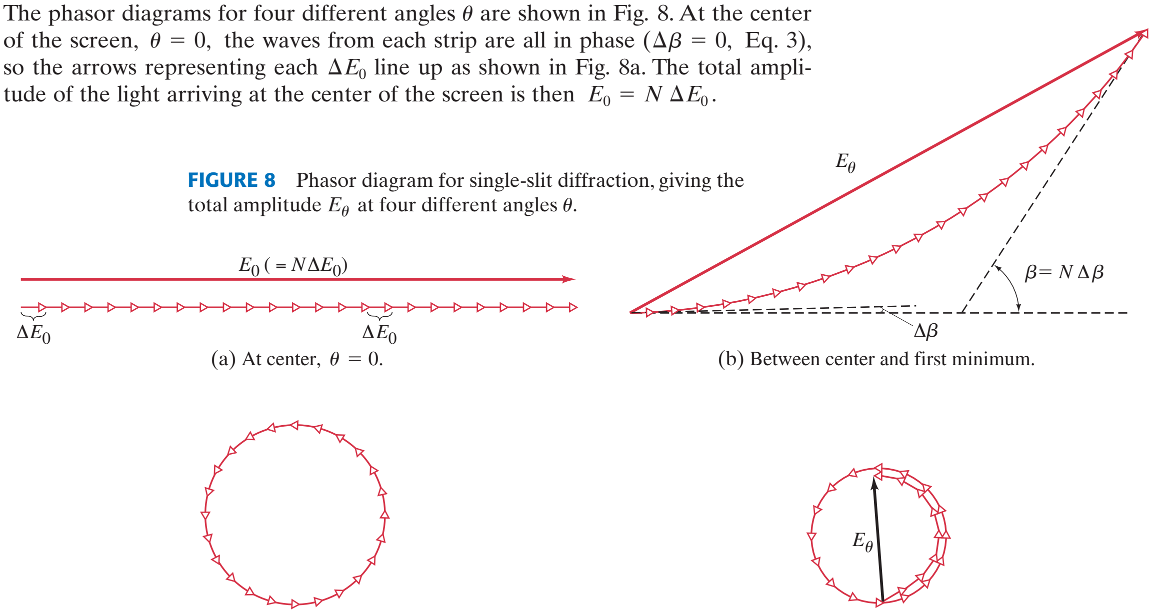
since the difference in path length is $\Delta y \sin\theta$.

**FIGURE 7** Slit of width *D* divided into *N* strips of width $\Delta y$.



The total amplitude on the screen at any angle $\theta$ will be the sum of the separate wave amplitudes due to each strip. These wavelets have the same amplitude $\Delta E_0$ but differ in phase. To obtain the total amplitude, we can use a phasor diagram. The phasor diagrams for four different angles $\theta$ are shown in Fig. 8. At the center of the screen, $\theta = 0$, the waves from each strip are all in phase ($\Delta\beta = 0$, Eq. 3), so the arrows representing each $\Delta E_0$ line up as shown in Fig. 8a. The total amplitude of the light arriving at the center of the screen is then $E_0 = N\,\Delta E_0$.

**FIGURE 8** Phasor diagram for single-slit diffraction, giving the total amplitude $E_\theta$ at four different angles $\theta$.




(a) At center, $\theta = 0$.

(b) Between center and first minimum.

(c) First minimum, $E_\theta = 0$ ($\beta = 2\pi = 360°$).

(d) Near secondary maximum.

At a small angle $\theta$, for a point on the distant screen not far from the center, Fig. 8b shows how the wavelets of amplitude $\Delta E_0$ add up to give $E_\theta$, the total amplitude on the screen at this angle $\theta$. Note that each wavelet differs in phase from the adjacent one by $\Delta\beta$. The phase difference between the wavelets from the top and bottom edges of the slit is

$$\beta = N\,\Delta\beta = \frac{2\pi}{\lambda} N\,\Delta y \sin\theta = \frac{2\pi}{\lambda} D \sin\theta \tag{4}$$

where $D = N\,\Delta y$ is the total width of the slit. Although the "arc" in Fig. 8b has length $N\,\Delta E_0$, and so would equal $E_0$ (total amplitude at $\theta = 0$), the amplitude of the total wave $E_\theta$ at angle $\theta$ is the *vector* sum of each wavelet amplitude and so is equal to the length of the chord as shown. The chord is shorter than the arc, so $E_\theta < E_0$.

For greater $\theta$, we eventually come to the case, illustrated in Fig. 8c, where the chain of arrows closes on itself. In this case the vector sum is zero, so $E_\theta = 0$ for this angle $\theta$. This corresponds to the first minimum. Since $\beta = N\,\Delta\beta$ is 360° or $2\pi$ in this case, we have from Eq. 3,

$$2\pi = N\,\Delta\beta = N\left(\frac{2\pi}{\lambda}\Delta y \sin\theta\right)$$

or, since the slit width $D = N\,\Delta y$,

$$\sin\theta = \frac{\lambda}{D}.$$

Thus the first minimum $(E_\theta = 0)$ occurs where $\sin\theta = \lambda/D$, which is the same result we obtained in the previous Section, Eq. 1.

For even greater values of $\theta$, the chain of arrows spirals beyond 360°. Figure 8d shows the case near the secondary maximum next to the first minimum. Here $\beta = N\,\Delta\beta \approx 360° + 180° = 540°$ or $3\pi$. When greater angles $\theta$ are considered, new maxima and minima occur. But since the total length of the coil remains constant, equal to $N\,\Delta E_0\,(= E_0)$, each succeeding maximum is smaller and smaller as the coil winds in on itself.

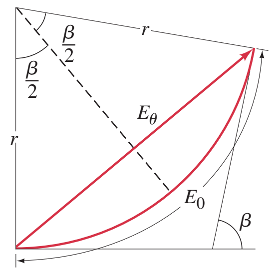



**FIGURE 9** Determining amplitude $E_\theta$ as a function of $\theta$ for single-slit diffraction.

To obtain a quantitative expression for the amplitude (and intensity) for any point on the screen (that is, for any angle $\theta$), we now consider the limit $N \rightarrow \infty$ so $\Delta y$ becomes the infinitesimal width $dy$. In this case, the diagrams of Fig. 8 become smooth curves, one of which is shown in Fig. 9. For any angle $\theta$, the wave amplitude on the screen is $E_\theta$, equal to the chord in Fig. 9. The length of the arc is $E_0$, as before. If $r$ is the radius of curvature of the arc, then

$$\frac{E_\theta}{2} = r \sin \frac{\beta}{2}.$$

Using radian measure for $\beta/2$, we also have

$$\frac{E_0}{2} = r \frac{\beta}{2}.$$

We combine these to obtain

$$E_\theta = E_0 \frac{\sin \beta/2}{\beta/2}. \tag{5}$$

The angle $\beta$ is the phase difference between the waves from the top and bottom edges of the slit. The path difference for these two rays is $D \sin \theta$ (see Fig. 7 as well as Eq. 4), so

$$\beta = \frac{2\pi}{\lambda} D \sin \theta. \tag{6}$$

Intensity is proportional to the square of the wave amplitude, so the intensity $I_\theta$ at any angle $\theta$ is, from Eq. 5,

$$I_\theta = I_0 \left( \frac{\sin \beta/2}{\beta/2} \right)^2 \tag{7}$$

where $I_0 \left( \propto E_0^2 \right)$ is the intensity at $\theta = 0$ (the central maximum). We can combine Eqs. 7 and 6 (although it is often simpler to leave them as separate equations) to obtain

$$I_\theta = I_0 \left[ \frac{\sin\left( \dfrac{\pi D \sin \theta}{\lambda} \right)}{\left( \dfrac{\pi D \sin \theta}{\lambda} \right)} \right]^2. \tag{8}$$

According to Eq. 8, minima $\left( I_\theta = 0 \right)$ occur where $\sin(\pi D \sin \theta/\lambda) = 0$, which means $\pi D \sin \theta/\lambda$ must be $\pi, 2\pi, 3\pi$, and so on, or

$$D \sin \theta = m\lambda, \qquad m = 1, 2, 3, \cdots \qquad \text{[minima]}$$

which is what we have obtained previously, Eq. 2. Notice that $m$ cannot be zero: when $\beta/2 = \pi D \sin \theta/\lambda = 0$, the denominator as well as the numerator in Eqs. 7 or 8 vanishes. We can evaluate the intensity in this case by taking the limit as $\theta \rightarrow 0$ (or $\beta \rightarrow 0$); for very small angles, $\sin \beta/2 \approx \beta/2$, so $(\sin \beta/2)/(\beta/2) \rightarrow 1$ and $I_\theta = I_0$, the *maximum* at the center of the pattern.

The intensity $I_\theta$ as a function of $\theta$, as given by Eq. 8, corresponds to the diagram of Fig. 4.

**EXAMPLE 3 ESTIMATE** **Intensity at secondary maxima.** Estimate the intensities of the first two secondary maxima to either side of the central maximum.

**APPROACH** The secondary maxima occur close to halfway between the minima, at about

$$\frac{\beta}{2} = \frac{\pi D \sin \theta}{\lambda} \approx \left(m + \tfrac{1}{2}\right)\pi. \qquad m = 1, 2, 3, \cdots$$

The actual maxima are not quite at these points—their positions can be determined by differentiating Eq. 7—but we are only seeking an estimate.

**SOLUTION** Using these values for $\beta$ in Eq. 7 or 8, with $\sin\left(m + \tfrac{1}{2}\right)\pi = 1$, gives

$$I_\theta = \frac{I_0}{\left(m + \tfrac{1}{2}\right)^2 \pi^2}. \qquad m = 1, 2, 3, \cdots$$

For $m = 1$ and 2, we get

$$I_\theta = \frac{I_0}{22.2} = 0.045 I_0 \qquad [m = 1]$$

$$I_\theta = \frac{I_0}{61.7} = 0.016 I_0. \qquad [m = 2]$$

The first maximum to the side of the central peak has only 1/22, or 4.5%, the intensity of the central peak, and succeeding ones are smaller still, just as we can see in Fig. 4 and the photo of Fig. 2c.

Diffraction by a circular opening produces a similar pattern (though circular rather than rectangular) and is of great practical importance, since lenses are essentially circular apertures through which light passes. We will discuss this in Section 4 and see how diffraction limits the resolution (or sharpness) of images.

## *3 Diffraction in the Double-Slit Experiment

When introducing Young's double-slit experiment, it is assumed that the central portion of the screen was uniformly illuminated. This is equivalent to assuming the slits are infinitesimally narrow, so that the central diffraction peak is spread out over the whole screen. This can never be the case for real slits; diffraction reduces the intensity of the bright interference fringes to the side of center so they are not all of the same height.

To calculate the intensity in a double-slit interference pattern, including diffraction, let us assume the slits have equal widths $D$ and their centers are separated by a distance $d$. Since the distance to the screen is large compared to the slit separation $d$, the wave amplitude due to each slit is essentially the same at each point on the screen. Then the total wave amplitude at any angle $\theta$ would not be

$$E_{\theta 0} = 2E_0 \cos\frac{\delta}{2}.$$

Rather, it must be modified, because of diffraction, by Eq. 5, so that

$$E_{\theta 0} = 2E_0\left(\frac{\sin \beta/2}{\beta/2}\right)\cos\frac{\delta}{2}.$$

Thus the intensity will be given by

$$I_\theta = I_0\left(\frac{\sin \beta/2}{\beta/2}\right)^2\left(\cos\frac{\delta}{2}\right)^2 \tag{9}$$

where $I_0 = 4E_0^2$, we have

$$\frac{\beta}{2} = \frac{\pi}{\lambda} D \sin\theta \qquad \text{and} \qquad \frac{\delta}{2} = \frac{\pi}{\lambda} d \sin\theta.$$

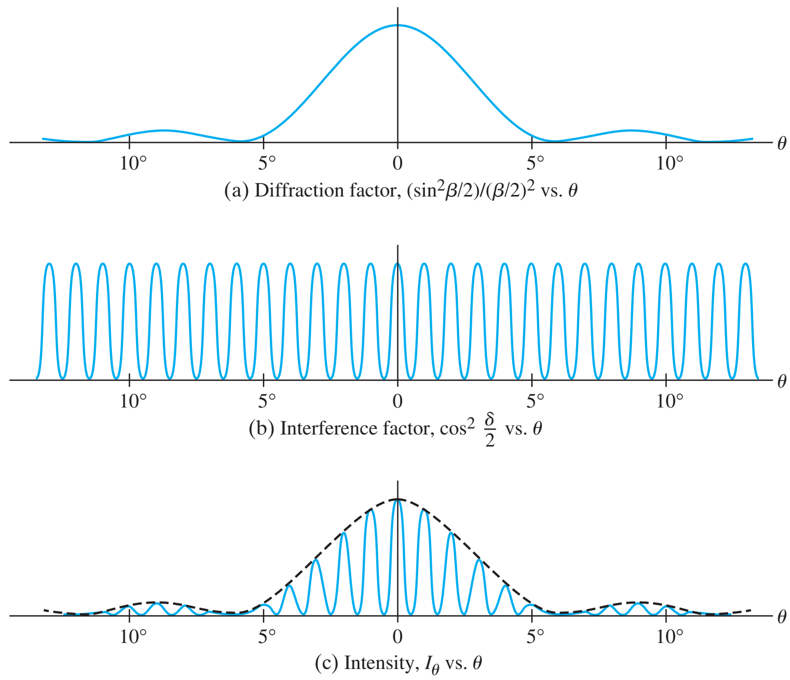


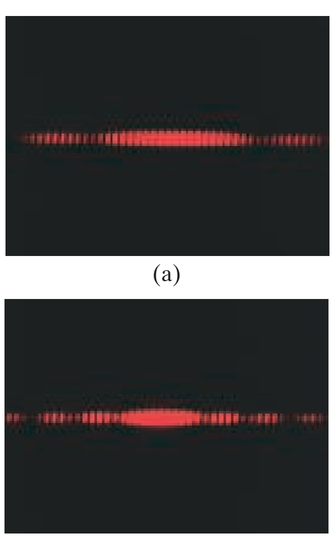

**FIGURE 10** (a) Diffraction factor, (b) interference factor, and (c) the resultant intensity $I_\theta$, plotted as a function of $\theta$ for $d = 6D = 60\lambda$.

Equation 9 for the intensity in a double-slit pattern, as we just saw, is

$$I_\theta = I_0\left(\frac{\sin\beta/2}{\beta/2}\right)^2\left(\cos\frac{\delta}{2}\right)^2. \qquad \textbf{(9)}$$

The first term in parentheses is sometimes called the "diffraction factor" and the second one the "interference factor." These two factors are plotted in Fig. 10a and b for the case when $d = 6D$, and $D = 10\lambda$. Figure 10c shows the product of these two curves (times $I_0$) which is the actual intensity as a function of $\theta$ (or as a function of position on the screen for $\theta$ not too large) as given by Eq. 9. As indicated by the dashed lines in Fig. 10c, the diffraction factor acts as a sort of envelope that limits the interference peaks.

**FIGURE 11** Photographs of a double-slit interference pattern using a laser beam showing effects of diffraction. In both cases $d = 0.50$ mm, whereas $D = 0.040$ mm in (a) and 0.080 mm in (b).



(a)

(b)

Richard Megna/Fundamental Photographs, NYC

**EXAMPLE 4** **Diffraction plus interference.** Show why the central diffraction peak in Fig. 10c contains 11 interference fringes.

**APPROACH** The first minimum in the diffraction pattern occurs where

$$\sin\theta = \frac{\lambda}{D}.$$

Since $d = 6D$,

$$d\sin\theta = 6D\left(\frac{\lambda}{D}\right) = 6\lambda.$$

**SOLUTION** Interference peaks (maxima) occur for $d\sin\theta = m\lambda$ where $m$ can be $0, 1, \cdots$ or any integer. Thus the diffraction minimum $(d\sin\theta = 6\lambda)$ coincides with $m = 6$ in the interference pattern, so the $m = 6$ peak won't appear. Hence the central diffraction peak encloses the central interference peak $(m = 0)$ and five peaks ($m = 1$ to 5) on each side for a total of 11. Since the sixth order doesn't appear, it is said to be a "missing order."

Notice from Example 4 that the number of interference fringes in the central diffraction peak depends only on the ratio $d/D$. It does not depend on wavelength $\lambda$. The actual spacing (in angle, or in position on the screen) does depend on $\lambda$. For the case illustrated, $D = 10\lambda$, and so the first diffraction minimum occurs at $\sin\theta = \lambda/D = 0.10$ or about 6°.

The decrease in intensity of the interference fringes away from the center, as graphed in Fig. 10, is shown in Fig. 11.

### Interference vs. Diffraction

The patterns due to interference and diffraction arise from the same phenomenon—the superposition of coherent waves of different phase. The distinction between them is thus not so much physical as for convenience of description, as in this Section where we analyzed the two-slit pattern in terms of interference and diffraction separately. In general, we use the word "diffraction" when referring to an analysis by superposition of many infinitesimal and usually contiguous sources, such as when we subdivide a source into infinitesimal parts. We use the term "interference" when we superpose the wave from a finite (and usually small) number of coherent sources.

## 4 Limits of Resolution; Circular Apertures

The ability of a lens to produce distinct images of two point objects very close together is called the **resolution** of the lens. The closer the two images can be and still be seen as distinct (rather than overlapping blobs), the higher the resolution. The resolution of a camera lens, for example, is often specified as so many dots or lines per millimeter.

Two principal factors limit the resolution of a lens. The first is lens aberrations. Because of spherical and other aberrations, a point object is not a point on the image but a tiny blob. Careful design of compound lenses can reduce aberrations significantly, but they cannot be eliminated entirely. The second factor that limits resolution is *diffraction*, which cannot be corrected for because it is a natural result of the wave nature of light. We discuss it now.

In Section 1, we saw that because light travels as a wave, light from a point source passing through a slit is spread out into a diffraction pattern (Figs. 2 and 4). A lens, because it has edges, acts like a round slit. When a lens forms the image of a point object, the image is actually a tiny diffraction pattern. Thus *an image would be blurred even if aberrations were absent*.

In the analysis that follows, we assume that the lens is free of aberrations, so we can concentrate on diffraction effects and how much they limit the resolution of a lens. In Fig. 4 we saw that the diffraction pattern produced by light passing through a rectangular slit has a central maximum in which most of the light falls. This central peak falls to a minimum on either side of its center at an angle $\theta$ given by $\sin\theta = \lambda/D$ (this is Eq. 1), where $D$ is the slit width and $\lambda$ the wavelength of light used. $\theta$ is the angular half-width of the central maximum, and for small angles can be written

$$\theta \approx \sin\theta = \frac{\lambda}{D}.$$

There are also low-intensity fringes beyond.

For a lens, or any circular hole, the image of a point object will consist of a *circular* central peak (called the *diffraction spot* or *Airy disk*) surrounded by faint circular fringes, as shown in Fig. 12a. The central maximum has an angular half width given by

$$\theta = \frac{1.22\lambda}{D},$$

where $D$ is the diameter of the circular opening. This is a theoretical result for a perfect circle or lens. For real lenses or circles, the factor is on the order of 1 to 2. This formula differs from that for a slit (Eq. 1) by the factor 1.22. This factor appears because the width of a circular hole is not uniform (like a rectangular slit) but varies from its diameter $D$ to zero. A mathematical analysis shows that the "average" width is $D/1.22$. Hence we get the equation above rather than Eq. 1. The intensity of light in the diffraction pattern from a point source of light passing through a circular opening is shown in Fig. 13. The image for a non-point source is a superposition of such patterns. For most purposes we need consider only the central spot, since the concentric rings are so much dimmer.

If two point objects are very close, the diffraction patterns of their images will overlap as shown in Fig. 12b. As the objects are moved closer, a separation is

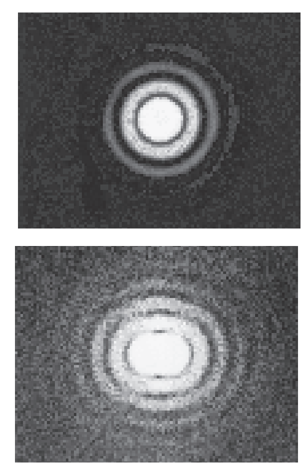
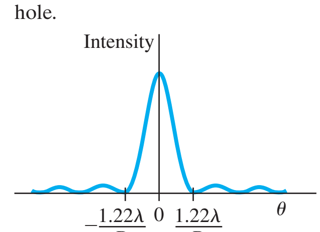




**FIGURE 12** Photographs of images (greatly magnified) formed by a lens, showing the diffraction pattern of an image for: (a) a single point object; (b) two point objects whose images are barely resolved.

**FIGURE 13** Intensity of light across the diffraction pattern of a circular hole.



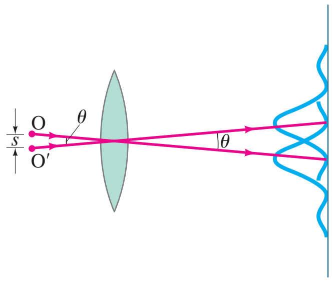



**FIGURE 14** The *Rayleigh criterion*. Two images are just resolvable when the center of the diffraction peak of one is directly over the first minimum in the diffraction pattern of the other. The two point objects O and O′ subtend an angle $\theta$ at the lens; only one ray (it passes through the center of the lens) is drawn for each object, to indicate the center of the diffraction pattern of its image.

reached where you can't tell if there are two overlapping images or a single image. The separation at which this happens may be judged differently by different observers. However, a generally accepted criterion is that proposed by Lord Rayleigh (1842–1919). This **Rayleigh criterion** states that *two images are just resolvable when the center of the diffraction disk of one image is directly over the first minimum in the diffraction pattern of the other*. This is shown in Fig. 14. Since the first minimum is at an angle $\theta = 1.22\lambda/D$ from the central maximum, Fig. 14 shows that two objects can be considered *just resolvable* if they are separated by at least an angle $\theta$ given by

$$\theta = \frac{1.22\lambda}{D}. \qquad [\theta \text{ in radians}] \quad \textbf{(10)}$$

In this equation, $D$ is the diameter of the lens, and applies also to a mirror diameter. This is the limit on resolution set by the wave nature of light due to diffraction. A smaller angle means better resolution: you can make out closer objects. We see from Eq. 10 that using a shorter wavelength $\lambda$ can reduce $\theta$ and thus increase resolution.

**EXERCISE B** Green light (550 nm) passes through a 25-mm-diameter camera lens. What is the angular half-width of the resulting diffraction pattern? (*a*) $2.7 \times 10^{-5}$ degrees, (*b*) $1.5 \times 10^{-3}$ degrees, (*c*) 3.2°, (*d*) 27°, (*e*) $1.5 \times 10^{3}$ degrees.

**FIGURE 15** Hubble Space Telescope, with Earth in the background. The flat orange panels are solar cells that collect energy from the Sun.

Space Telescope Science Institute

**EXAMPLE 5 Hubble Space Telescope.** The Hubble Space Telescope (HST) is a reflecting telescope that was placed in orbit above the Earth's atmosphere, so its resolution would not be limited by turbulence in the atmosphere (Fig. 15). Its objective diameter is 2.4 m. For visible light, say $\lambda = 550$ nm, estimate the improvement in resolution the Hubble offers over Earth-bound telescopes, which are limited in resolution by movement of the Earth's atmosphere to about half an arc second. (Each degree is divided into 60 minutes each containing 60 seconds, so $1° = 3600$ arc seconds.)

**APPROACH** Angular resolution for the Hubble is given (in radians) by Eq. 10. The resolution for Earth telescopes is given, and we first convert it to radians so we can compare.

**SOLUTION** Earth-bound telescopes are limited to an angular resolution of

$$\theta = \tfrac{1}{2}\left(\frac{1}{3600}\right)^{\circ}\left(\frac{2\pi \text{ rad}}{360°}\right) = 2.4 \times 10^{-6}\,\text{rad}.$$

The Hubble, on the other hand, is limited by diffraction (Eq. 10) which for $\lambda = 550$ nm is

$$\theta = \frac{1.22\lambda}{D} = \frac{1.22(550 \times 10^{-9}\,\text{m})}{2.4\,\text{m}} = 2.8 \times 10^{-7}\,\text{rad},$$

thus giving almost ten times better resolution $(2.4 \times 10^{-6}\,\text{rad}/2.8 \times 10^{-7}\,\text{rad} \approx 9\times)$.

**EXAMPLE 6 ESTIMATE Eye resolution.** You are in an airplane at an altitude of 10,000 m. If you look down at the ground, estimate the minimum separation $s$ between objects that you could distinguish. Could you count cars in a parking lot? Consider only diffraction, and assume your pupil is about 3.0 mm in diameter and $\lambda = 550$ nm.

**PHYSICS APPLIED**
*How well the eye can see*

**APPROACH** We use the Rayleigh criterion, Eq. 10, to estimate $\theta$. The separation $s$ of objects is $s = \ell\theta$, where $\ell = 10^4$ m and $\theta$ is in radians.

**SOLUTION** In Eq. 10, we set $D = 3.0$ mm for the opening of the eye:

$$s = \ell\theta = \ell\frac{1.22\lambda}{D} = \frac{(10^4\,\text{m})(1.22)(550 \times 10^{-9}\,\text{m})}{3.0 \times 10^{-3}\,\text{m}} = 2.2\,\text{m}.$$

Yes, you could just resolve a car (roughly 2 m wide by 3 or 4 m long) and count them.

# 5 Resolution of Telescopes and Microscopes; the $\lambda$ Limit

You might think that a microscope or telescope could be designed to produce any desired magnification, depending on the choice of focal lengths and quality of the lenses. But this is not possible, because of diffraction. An increase in magnification above a certain point merely results in magnification of the diffraction patterns. This can be highly misleading since we might think we are seeing details of an object when we are really seeing details of the diffraction pattern. To examine this problem, we apply the Rayleigh criterion: two objects (or two nearby points on one object) are just resolvable if they are separated by an angle $\theta$ (Fig. 14) given by Eq. 10:

$$\theta = \frac{1.22\lambda}{D}.$$

This formula is valid for either a microscope or a telescope, where $D$ is the diameter of the objective lens or mirror. For a telescope, the resolution is specified by stating $\theta$ as given by this equation.†

**FIGURE 16** The 300-meter radiotelescope in Arecibo, Puerto Rico, uses radio waves instead of visible light.

David Parker/Photo Researchers, Inc.

**EXAMPLE 7 Telescope resolution (radio wave vs. visible light).** What is the theoretical minimum angular separation of two stars that can just be resolved by (*a*) the 200-inch (mirror diameter) Hale telescope on Palomar Mountain in California; and (*b*) the Arecibo radiotelescope (Fig. 16), whose diameter is 300 m and whose radius of curvature is also 300 m. Assume $\lambda = 550\,\text{nm}$ for the visible-light telescope in part (*a*), and $\lambda = 4\,\text{cm}$ (the shortest wavelength at which the radiotelescope has been operated) in part (*b*).

**APPROACH** We apply the Rayleigh criterion (Eq. 10) for each telescope.

**SOLUTION** (*a*) Since $D = 200\,\text{in.} = 5.1\,\text{m}$, we have from Eq. 10 that

$$\theta = \frac{1.22\lambda}{D} = \frac{(1.22)(5.50 \times 10^{-7}\,\text{m})}{(5.1\,\text{m})} = 1.3 \times 10^{-7}\,\text{rad},$$

or $0.75 \times 10^{-5}$ deg. (Note that this is equivalent to resolving two points less than 1 cm apart from a distance of 100 km!)

(*b*) For radio waves with $\lambda = 0.04\,\text{m}$ emitted by stars, the resolution is

$$\theta = \frac{(1.22)(0.04\,\text{m})}{(300\,\text{m})} = 1.6 \times 10^{-4}\,\text{rad}.$$

The resolution is less because the wavelength is so much larger, but the larger objective collects more radiation and thus detects fainter objects.

**NOTE** In both cases, we determined the limit set by diffraction. The resolution for a visible-light Earth-bound telescope is not this good because of aberrations and, more importantly, turbulence in the atmosphere. In fact, large-diameter objectives are not justified by increased resolution, but by their greater light-gathering ability—they allow more light in, so fainter objects can be seen. Radiotelescopes are not hindered by atmospheric turbulence, and the resolution found in (*b*) is a good estimate.

†Earth-bound telescopes with large-diameter objectives are usually limited not by diffraction but by other effects such as turbulence in the atmosphere. The resolution of a high-quality microscope, on the other hand, normally *is* limited by diffraction; microscope objectives are complex compound lenses containing many elements of small diameter (since $f$ is small), thus reducing aberrations.

For a microscope, it is more convenient to specify the actual distance, $s$, between two points that are just barely resolvable: see Fig. 14. Since objects are normally placed near the focal point of the microscope objective, the angle subtended by two objects is $\theta = s/f$, so $s = f\theta$. If we combine this with Eq. 10, we obtain the **resolving power** (**RP**) of a microscope

$$\mathrm{RP} = s = f\theta = \frac{1.22\lambda f}{D}, \tag{11}$$

where $f$ is the objective lens' focal length (not frequency). This distance $s$ is called the resolving power of the lens because it is the minimum separation of two object points that can just be resolved—assuming the highest quality lens since this limit is imposed by the wave nature of light. A smaller RP means better resolution, better detail.

**EXERCISE C** What is the resolving power of a microscope with a 5-mm-diameter objective which has $f = 9\,\mathrm{mm}$? (*a*) 550 nm, (*b*) 750 nm, (*c*) 1200 nm, (*d*) 0.05 nm, (*e*) 0.005 nm.

Diffraction sets an ultimate limit on the detail that can be seen on any object. In Eq. 11 for resolving power of a microscope, the focal length of the lens cannot practically be made less than (approximately) the radius of the lens, and even that is very difficult. In this best case, Eq. 11 gives, with $f \approx D/2$.

$$\mathrm{RP} \approx \frac{\lambda}{2}. \tag{12}$$

Thus we can say, to within a factor of 2 or so, that

*$\lambda$ limits resolution*

**it is not possible to resolve detail of objects smaller than the wavelength of the radiation being used.**

This is an important and useful rule of thumb.

Compound lenses in microscopes are now designed so well that the actual limit on resolution is often set by diffraction—that is, by the wavelength of the light used. To obtain greater detail, one must use radiation of shorter wavelength. The use of UV radiation can increase the resolution by a factor of perhaps 2. Far more important, however, was the discovery in the early twentieth century that electrons have wave properties and that their wavelengths can be very small. The wave nature of electrons is utilized in the electron microscope, which can magnify 100 to 1000 times more than a visible-light microscope because of the much shorter wavelengths. X-rays, too, have very short wavelengths and are often used to study objects in great detail (Section 10).

## *6 Resolution of the Human Eye and Useful Magnification

The resolution of the human eye is limited by several factors, all of roughly the same order of magnitude. The resolution is best at the fovea, where the cone spacing is smallest, about $3\,\mu\mathrm{m}\ (= 3000\,\mathrm{nm})$. The diameter of the pupil varies from about 0.1 cm to about 0.8 cm. So for $\lambda = 550\,\mathrm{nm}$ (where the eye's sensitivity is greatest), the diffraction limit is about $\theta \approx 1.22\lambda/D \approx 8 \times 10^{-5}\,\mathrm{rad}$ to $6 \times 10^{-4}\,\mathrm{rad}$. The eye is about 2 cm long, giving a resolving power (Eq. 11) of $s \approx (2 \times 10^{-2}\,\mathrm{m})(8 \times 10^{-5}\,\mathrm{rad}) \approx 2\,\mu\mathrm{m}$ at best, to about $10\,\mu\mathrm{m}$ at worst (pupil small). Spherical and chromatic aberration also limit the resolution to about $10\,\mu\mathrm{m}$. The net result is that the eye can just resolve objects whose angular separation is around

$$5 \times 10^{-4}\,\mathrm{rad}. \qquad \begin{bmatrix}\text{best eye}\\ \text{resolution}\end{bmatrix}$$

This corresponds to objects separated by 1 cm at a distance of about 20 m.

The typical near point of a human eye is about 25 cm. At this distance, the eye can just resolve objects that are $(25\,\mathrm{cm})(5 \times 10^{-4}\,\mathrm{rad}) \approx 10^{-4}\,\mathrm{m} = \frac{1}{10}\,\mathrm{mm}$ apart. Since the best light microscopes can resolve objects no smaller than about 200 nm at best (Eq. 12 for violet light, $\lambda = 400\,\mathrm{nm}$), the useful magnification

[= (resolution by naked eye)/(resolution by microscope)] is limited to about

$$\frac{10^{-4}\,\mathrm{m}}{200 \times 10^{-9}\,\mathrm{m}} \approx 500\times. \qquad \begin{bmatrix}\text{maximum useful} \\ \text{microscope magnification}\end{bmatrix}$$

In practice, magnifications of about 1000× are often used to minimize eyestrain. Any greater magnification would simply make visible the diffraction pattern produced by the microscope objective lens.

Now you have the answers to the Chapter-Opening Question: (*c*), by the equation above, and (*c*) by the $\lambda$ rule.

# 7 Diffraction Grating

A large number of equally spaced parallel slits is called a **diffraction grating**, although the term "interference grating" might be as appropriate. Gratings can be made by precision machining of very fine parallel lines on a glass plate. The untouched spaces between the lines serve as the slits. Photographic transparencies of an original grating serve as inexpensive gratings. Gratings containing 10,000 lines per centimeter are common, and are very useful for precise measurements of wavelengths. A diffraction grating containing slits is called a **transmission grating**. Another type of diffraction grating is the **reflection grating**, made by ruling fine lines on a metallic or glass surface from which light is reflected and analyzed. The analysis is basically the same as for a transmission grating, which we now discuss.

The analysis of a diffraction grating is much like that of Young's double-slit experiment. We assume parallel rays of light are incident on the grating as shown in Fig. 17. We also assume that the slits are narrow enough so that diffraction by each of them spreads light over a very wide angle on a distant screen beyond the grating, and interference can occur with light from all the other slits. Light rays that pass through each slit without deviation $(\theta = 0°)$ interfere constructively to produce a bright line at the center of the screen. Constructive interference also occurs at an angle $\theta$ such that rays from adjacent slits travel an extra distance of $\Delta\ell = m\lambda,$ where $m$ is an integer. If $d$ is the distance between slits, then we see from Fig. 17 that $\Delta\ell = d\sin\theta,$ and

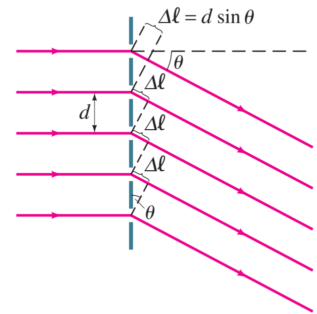



**FIGURE 17** Diffraction grating.

$$\sin\theta = \frac{m\lambda}{d}, \qquad m = 0, 1, 2, \cdots \qquad \begin{bmatrix}\text{diffraction grating,} \\ \text{principal maxima}\end{bmatrix} \qquad \textbf{(13)}$$

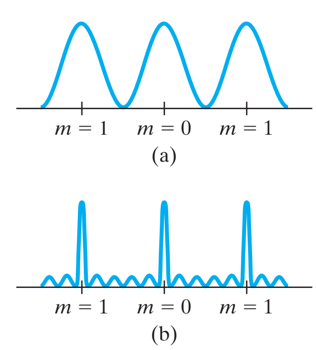
is the criterion to have a brightness maximum. This is the same equation as for the double-slit situation, and again $m$ is called the order of the pattern.

**CAUTION**

*Diffraction grating is analyzed using interference formulas, not diffraction formulas*

There is an important difference between a double-slit and a multiple-slit pattern. The bright maxima are much *sharper* and *narrower* for a grating. Why? Suppose that the angle $\theta$ is increased just slightly beyond that required for a maximum. In the case of only two slits, the two waves will be only slightly out of phase, so nearly full constructive interference occurs. This means the maxima are wide. For a grating, the waves from two adjacent slits will also not be significantly out of phase. But waves from one slit and those from a second one a few hundred slits away may be exactly out of phase; all or nearly all the light can cancel in pairs in this way. For example, suppose the angle $\theta$ is very slightly different from its first-order maximum, so that the extra path length for a pair of adjacent slits is not exactly $\lambda$ but rather $1.0010\lambda$. The wave through one slit and another one 500 slits below will have a path difference of $1\lambda + (500)(0.0010\lambda) = 1.5000\lambda,$ or $1\frac{1}{2}$ wavelengths, so the two will cancel. A pair of slits, one below each of these, will also cancel. That is, the light from slit 1 cancels with that from slit 501; light from slit 2 cancels with that from slit 502, and so on. Thus even for a tiny angle† corresponding to an extra path length of $\frac{1}{1000}\lambda$, there is much destructive interference, and so the maxima are very narrow. The more lines there are in a grating, the sharper will be the peaks (see Fig. 18). Because a grating produces much sharper lines than two slits alone can (and much brighter lines because there are many more slits), a grating is a far more precise device for measuring wavelengths.

**FIGURE 18** Intensity as a function of viewing angle $\theta$ (or position on the screen) for (a) two slits, (b) six slits. For a diffraction grating, the number of slits is very large $(\approx 10^4)$ and the peaks are narrower still.




†Depending on the total number of slits, there may or may not be complete cancellation for such an angle, so there will be very tiny peaks between the main maxima (see Fig. 18b), but they are usually much too small to be seen.

Suppose the light striking a diffraction grating is not monochromatic, but consists of two or more distinct wavelengths. Then for all orders other than $m = 0$, each wavelength will produce a maximum at a different angle (Fig. 19a), just as for a double slit. If white light strikes a grating, the central $(m = 0)$ maximum will be a sharp white peak. But for all other orders, there will be a distinct spectrum of colors spread out over a certain angular width, Fig. 19b. Because a diffraction grating spreads out light into its component wavelengths, the resulting pattern is called a **spectrum**.

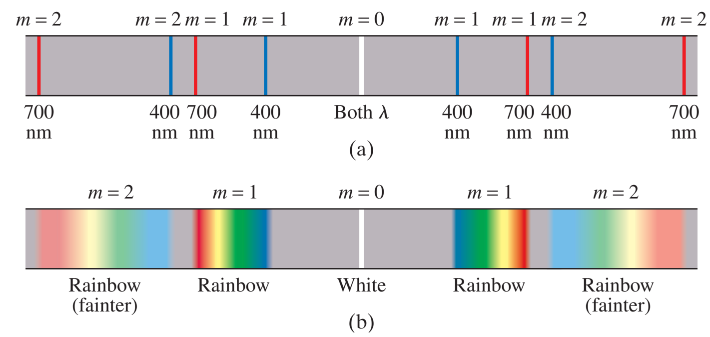



**FIGURE 19** Spectra produced by a grating: (a) two wavelengths, 400 nm and 700 nm; (b) white light. The second order will normally be dimmer than the first order. (Higher orders are not shown.) If grating spacing is small enough, the second and higher orders will be missing.

**EXAMPLE 8** **Diffraction grating: lines.** Determine the angular positions of the first- and second-order maxima for light of wavelength 400 nm and 700 nm incident on a grating containing 10,000 lines/cm.

**APPROACH** First we find the distance $d$ between grating lines: if the grating has $N$ lines in 1 m, then the distance between lines must be $d = 1/N$ meters. Then we use Eq. 13, $\sin\theta = m\lambda/d$, to find the angles for the two wavelengths for $m = 1$ and 2.

**SOLUTION** The grating contains $1.00 \times 10^4$ lines/cm $= 1.00 \times 10^6$ lines/m, which means the distance between lines is $d = (1/1.00 \times 10^6)\,\text{m} = 1.00 \times 10^{-6}\,\text{m} = 1.00\,\mu\text{m}$. In first order $(m = 1)$, the angles are

$$\sin\theta_{400} = \frac{m\lambda}{d} = \frac{(1)(4.00 \times 10^{-7}\,\text{m})}{1.00 \times 10^{-6}\,\text{m}} = 0.400$$

$$\sin\theta_{700} = \frac{(1)(7.00 \times 10^{-7}\,\text{m})}{1.00 \times 10^{-6}\,\text{m}} = 0.700$$

so $\theta_{400} = 23.6°$ and $\theta_{700} = 44.4°$. In second order,

$$\sin\theta_{400} = \frac{2\lambda}{d} = \frac{(2)(4.00 \times 10^{-7}\,\text{m})}{1.00 \times 10^{-6}\,\text{m}} = 0.800$$

$$\sin\theta_{700} = \frac{(2)(7.00 \times 10^{-7}\,\text{m})}{1.00 \times 10^{-6}\,\text{m}} = 1.40$$

so $\theta_{400} = 53.1°$. But the second order does not exist for $\lambda = 700\,\text{nm}$ because $\sin\theta$ cannot exceed 1. No higher orders will appear.

**EXAMPLE 9** **Spectra overlap.** White light containing wavelengths from 400 nm to 750 nm strikes a grating containing 4000 lines/cm. Show that the blue at $\lambda = 450\,\text{nm}$ of the third-order spectrum overlaps the red at 700 nm of the second order.

**APPROACH** We use $\sin\theta = m\lambda/d$ to calculate the angular positions of the $m = 3$ blue maximum and the $m = 2$ red one.

**SOLUTION** The grating spacing is $d = (1/4000)\text{ cm} = 2.50 \times 10^{-6}\text{ m}$. The blue of the third order occurs at an angle $\theta$ given by

$$\sin\theta = \frac{m\lambda}{d} = \frac{(3)(4.50 \times 10^{-7}\text{ m})}{(2.50 \times 10^{-6}\text{ m})} = 0.540.$$

Red in second order occurs at

$$\sin\theta = \frac{(2)(7.00 \times 10^{-7}\text{ m})}{(2.50 \times 10^{-6}\text{ m})} = 0.560,$$

which is a greater angle; so the second order overlaps into the beginning of the third-order spectrum.

**CONCEPTUAL EXAMPLE 10** **Compact disk.** When you look at the surface of a music CD (Fig. 20), you see the colors of a rainbow. (*a*) Estimate the distance between the curved lines (they are read by a laser). (*b*) Estimate the distance between lines, noting that a CD contains at most 80 min of music, that it rotates at speeds from 200 to 500 rev/min, and that $\frac{2}{3}$ of its 6 cm radius contains the lines.

**RESPONSE** (*a*) The CD acts like a reflection diffraction grating. To satisfy Eq. 13, we might estimate the line spacing as one or a few (2 or 3) wavelengths ($\lambda \approx 550\text{ nm}$) or 0.5 to 1.5 $\mu$m. (*b*) Average rotation speed of 350 rev/min times 80 min gives 28,000 total rotations or 28,000 lines, which are spread over $\left(\frac{2}{3}\right)(6\text{ cm}) = 4\text{ cm}$. So we have a sort of reflection diffraction grating with about $(28{,}000\text{ lines})/(4\text{ cm}) = 7000\text{ lines/cm}$. The distance $d$ between lines is roughly $1\text{ cm}/7000\text{ lines} \approx 1.4 \times 10^{-6}\text{ m} = 1.4\,\mu\text{m}$. Our results in (*a*) and (*b*) agree.

**FIGURE 20** A compact disk, Example 10.

Spike Mafford/Photodisc/Getty Images

## 8 The Spectrometer and Spectroscopy

A **spectrometer** or **spectroscope**, Fig. 21, is a device to measure wavelengths accurately using a diffraction grating (or a prism) to separate different wavelengths of light. Light from a source passes through a narrow slit S in the "collimator." The slit is at the focal point of the lens L, so parallel light falls on the grating. The movable telescope can bring the rays to a focus. Nothing will be seen in the viewing telescope unless it is positioned at an angle $\theta$ that corresponds to a diffraction peak (first order is usually used) of a wavelength emitted by the source. The angle $\theta$ can be measured to very high accuracy, so the wavelength of a line can be determined to high accuracy using Eq. 13:

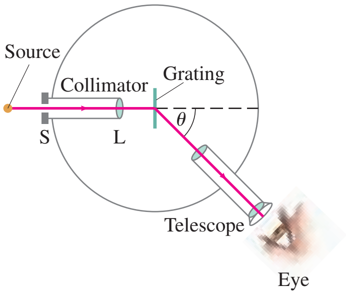
$$\lambda = \frac{d}{m}\sin\theta,$$

where $m$ is an integer representing the order, and $d$ is the distance between grating lines. The line you see in a spectrometer corresponding to each wavelength is actually an image of the slit S. A narrower slit results in dimmer light but we can measure the angular positions more precisely. If the light contains a continuous range of wavelengths, then a continuous spectrum is seen in the spectroscope.

The spectrometer in Fig. 21 uses a transmission grating. Others may use a reflection grating, or sometimes a prism. A prism works because of dispersion, bending light of different wavelengths into different angles. (A prism is not a linear device and must be calibrated.)

An important use of a spectrometer is for the identification of atoms or molecules. When a gas is heated or an electric current is passed through it, the gas emits a characteristic **line spectrum**. That is, only certain discrete wavelengths of light are emitted, and these are different for different elements and compounds.† Figure 22 shows the line spectra for a number of elements in the gas state. Line spectra occur only for gases at high temperatures and low pressure and density. The light from heated solids, such as a lightbulb filament, and even from a dense gaseous object such as the Sun, produces a **continuous spectrum** including a wide range of wavelengths.

**FIGURE 21** Spectrometer or spectroscope.




†Why atoms and molecules emit line spectra was a great mystery for many years and played a central role in the development of modern quantum theory.

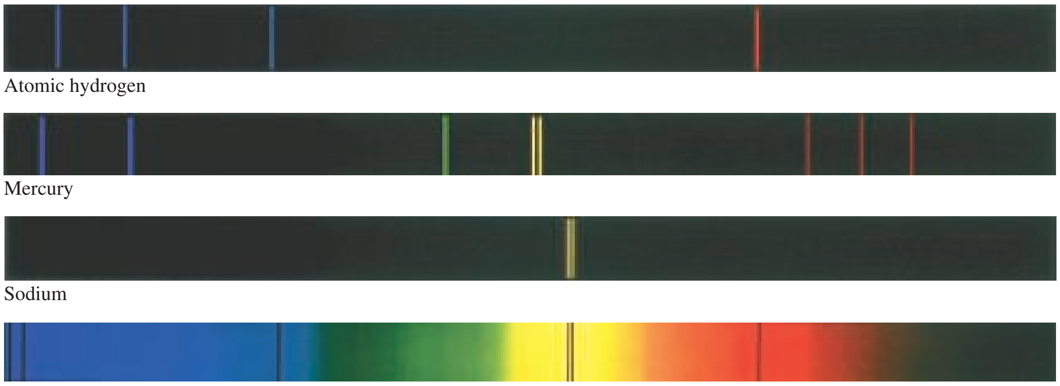



**FIGURE 22** Line spectra for the gases indicated, and the spectrum from the Sun showing absorption lines.

Figure 22 also shows the Sun's "continuous spectrum," which contains a number of *dark* lines (only the most prominent are shown), called **absorption lines**. Atoms and molecules can absorb light at the same wavelengths at which they emit light. The Sun's absorption lines are due to absorption by atoms and molecules in the cooler outer atmosphere of the Sun, as well as by atoms and molecules in the Earth's atmosphere. A careful analysis of all these thousands of lines reveals that at least two-thirds of all elements are present in the Sun's atmosphere. The presence of elements in the atmosphere of other planets, in interstellar space, and in stars, is also determined by spectroscopy.

**PHYSICS APPLIED**
*Chemical and biochemical analysis by spectroscopy*

Spectroscopy is useful for determining the presence of certain types of molecules in laboratory specimens where chemical analysis would be difficult. For example, biological DNA and different types of protein absorb light in particular regions of the spectrum (such as in the UV). The material to be examined, which is often in solution, is placed in a monochromatic light beam whose wavelength is selected by the placement angle of a diffraction grating or prism. The amount of absorption, as compared to a standard solution without the specimen, can reveal not only the presence of a particular type of molecule, but also its concentration.

Light emission and absorption also occur outside the visible part of the spectrum, such as in the UV and IR regions. Glass absorbs light in these regions, so reflection gratings and mirrors (in place of lenses) are used. Special types of film or sensors are used for detection.

**EXAMPLE 11** **Hydrogen spectrum.** Light emitted by hot hydrogen gas is observed with a spectroscope using a diffraction grating having $1.00 \times 10^4$ lines/cm. The spectral lines nearest to the center $(0°)$ are a violet line at 24.2°, a blue line at 25.7°, a blue-green line at 29.1°, and a red line at 41.0° from the center. What are the wavelengths of these spectral lines of hydrogen?

**APPROACH** The wavelengths can be determined from the angles by using $\lambda = (d/m)\sin\theta$ where $d$ is the spacing between slits, and $m$ is the order of the spectrum (Eq. 13).

**SOLUTION** Since these are the closest lines to $\theta = 0°$, this is the first-order spectrum $(m = 1)$. The slit spacing is $d = 1/(1.00 \times 10^4\,\text{cm}^{-1}) = 1.00 \times 10^{-6}\,\text{m}$. The wavelength of the violet line is

$$\lambda = \left(\frac{d}{m}\right)\sin\theta = \left(\frac{1.00 \times 10^{-6}\,\text{m}}{1}\right)\sin 24.2° = 4.10 \times 10^{-7}\,\text{m} = 410\,\text{nm}.$$

The other wavelengths are:

blue: $\lambda = (1.00 \times 10^{-6}\,\text{m})\sin 25.7° = 434\,\text{nm},$

blue-green: $\lambda = (1.00 \times 10^{-6}\,\text{m})\sin 29.1° = 486\,\text{nm},$

red: $\lambda = (1.00 \times 10^{-6}\,\text{m})\sin 41.0° = 656\,\text{nm}.$

**NOTE** In an unknown mixture of gases, these four spectral lines need to be seen to identify that the mixture contains hydrogen.

## *9 Peak Widths and Resolving Power for a Diffraction Grating

We now look at the pattern of maxima produced by a multiple-slit grating using phasor diagrams. We can determine a formula for the width of each peak, and we will see why there are tiny maxima between the principal maxima, as indicated in Fig. 18b. First of all, it should be noted that the two-slit and six-slit patterns shown in Fig. 18 were drawn assuming very narrow slits so that diffraction does not limit the height of the peaks. For real diffraction gratings, this is not normally the case: the slit width $D$ is often not much smaller than the slit separation $d$, and diffraction thus limits the intensity of the peaks so the central peak $(m = 0)$ is brighter than the side peaks. We won't worry about this effect on intensity except to note that if a diffraction minimum coincides with a particular order of the interference pattern, that order will not appear. (For example, if $d = 2D$, all the even orders, $m = 2, 4, \cdots,$ will be missing. Can you see why? Hint: See Example 4.)

Figures 23 and 24 show phasor diagrams for a two-slit and a six-slit grating, respectively. Each short arrow represents the amplitude of a wave from a single slit, and their vector sum (as phasors) represents the total amplitude for a given viewing angle $\theta$. Part (a) of each Figure shows the phasor diagram at $\theta = 0°$, at the center of the pattern, which is the central maximum $(m = 0)$. Part (b) of each Figure shows the condition for the adjacent minimum: where the arrows first close on themselves (add to zero) so the amplitude $E_\theta$ is zero. For two slits, this occurs when the two separate amplitudes are 180° out of phase. For six slits, it occurs when each amplitude makes a 60° angle with its neighbor. For two slits, the minimum occurs when the phase between slits is $\delta = 2\pi/2$ (in radians); for six slits it occurs when the phase $\delta$ is $2\pi/6$; and in the general case of $N$ slits, the minimum occurs for a phase difference between adjacent slits of

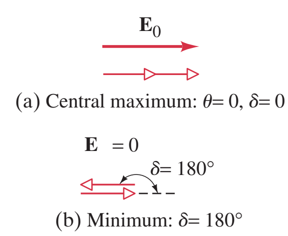


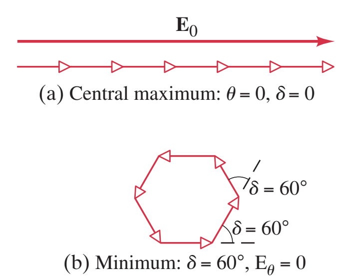

**FIGURE 23** Phasor diagram for two slits (a) at the central maximum, (b) at the nearest minimum.

$$\delta = \frac{2\pi}{N}. \tag{14}$$

What does this correspond to in $\theta$? First note that $\delta$ is related to $\theta$ by

$$\frac{\delta}{2\pi} = \frac{d \sin\theta}{\lambda} \quad \text{or} \quad \delta = \frac{2\pi}{\lambda} d \sin\theta. \tag{15}$$

**FIGURE 24** Phasor diagram for six slits (a) at the central maximum, (b) at the nearest minimum.




Let us call $\Delta\theta_0$ the angular position of the minimum next to the peak at $\theta = 0$. Then

$$\frac{\delta}{2\pi} = \frac{d \sin\Delta\theta_0}{\lambda}.$$

We insert Eq. 14 for $\delta$ and find

$$\sin\Delta\theta_0 = \frac{\lambda}{Nd}. \tag{16a}$$

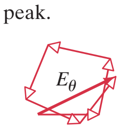
Since $\Delta\theta_0$ is usually small ($N$ is usually very large for a grating), $\sin\Delta\theta_0 \approx \Delta\theta_0$, so in the small angle limit we can write

$$\Delta\theta_0 = \frac{\lambda}{Nd}. \tag{16b}$$

It is clear from either of the last two relations that the larger $N$ is, the narrower will be the central peak. (For $N = 2$, $\sin\Delta\theta_0 = \lambda/2d$, which is what we obtain for the double slit, with $m = 0$.)

Either of Eqs. 16 shows why the peaks become narrower for larger $N$. The origin of the small secondary maxima between the principal peaks (see Fig. 18b) can be deduced from the diagram of Fig. 25. This is just a continuation of Fig. 24b (where $\delta = 60°$); but now the phase $\delta$ has been increased to almost 90°, where $E_\theta$ is a relative maximum. Note that $E_\theta$ is much less than $E_0$ (Fig. 24a), so the intensity in this secondary maximum is much smaller than in a principal peak. As $\delta$ (and $\theta$) is increased further, $E_\theta$ again decreases to zero (a "double circle"), then reaches another tiny maximum, and so on. Eventually the diagram unfolds again and when $\delta = 360°$, all the amplitudes again lie in a straight line (as in Fig. 24a) corresponding to the next principal maximum ($m = 1$ in Eq. 13).

**FIGURE 25** Phasor diagram for the secondary peak.



Equation 16 b gives the half-width of the central $(m = 0)$ peak. To determine the half-width of higher order peaks, $\Delta\theta_m$ for order $m$, we differentiate Eq. 15 so as to relate the change $\Delta\delta$ in $\delta$, to the change $\Delta\theta$ in the angle $\theta$:

$$\Delta\delta \approx \frac{d\delta}{d\theta}\Delta\theta = \frac{2\pi d}{\lambda}\cos\theta\,\Delta\theta.$$

If $\Delta\theta_m$ represents the half-width of a peak of order $m$ $(m = 1, 2, \cdots)$—that is, the angle between the peak maximum and the minimum to either side—then $\Delta\delta = 2\pi/N$ as given by Eq. 14. We insert this into the above relation and find

$$\Delta\theta_m = \frac{\lambda}{Nd\cos\theta_m}, \tag{17}$$

where $\theta_m$ is the angular position of the $m^{\text{th}}$ peak as given by Eq. 13. This derivation is valid, of course, only for small $\Delta\delta\ (= 2\pi/N)$ which is indeed the case for real gratings since $N$ is on the order of $10^4$ or more.

An important property of any diffraction grating used in a spectrometer is its ability to resolve two very closely spaced wavelengths (wavelength difference = $\Delta\lambda$). The **resolving power** R of a grating is defined as

$$\mathrm{R} = \frac{\lambda}{\Delta\lambda}. \tag{18}$$

With a little work, using Eq. 17, we can show that $\Delta\lambda = \lambda/Nm$ where $N$ is the total number of grating lines and $m$ is the order. Then we have

$$\mathrm{R} = \frac{\lambda}{\Delta\lambda} = Nm. \tag{19}$$

The larger the value of R, the closer two wavelengths can be resolvable. If R is given, the minimum separation $\Delta\lambda$ between two wavelengths near $\lambda$, is (by Eq. 18)

$$\Delta\lambda = \frac{\lambda}{\mathrm{R}}.$$

**EXAMPLE 12** **Resolving two close lines.** Yellow sodium light, which consists of two wavelengths, $\lambda_1 = 589.00\ \mathrm{nm}$ and $\lambda_2 = 589.59\ \mathrm{nm}$, falls on a 7500-line/cm diffraction grating. Determine ($a$) the maximum order $m$ that will be present for sodium light, ($b$) the width of grating necessary to resolve the two sodium lines.

**APPROACH** We first find $d = 1\ \mathrm{cm}/7500 = 1.33 \times 10^{-6}\ \mathrm{m}$, and then use Eq. 13 to find $m$. For ($b$) we use Eqs. 18 and 19.

**SOLUTION** ($a$) The maximum value of $m$ at $\lambda = 589\ \mathrm{nm}$, using Eq. 13 with $\sin\theta \leq 1$, is

$$m = \frac{d}{\lambda}\sin\theta \leq \frac{d}{\lambda} = \frac{1.33 \times 10^{-6}\ \mathrm{m}}{5.89 \times 10^{-7}\ \mathrm{m}} = 2.26,$$

so $m = 2$ is the maximum order present.

($b$) The resolving power needed is

$$\mathrm{R} = \frac{\lambda}{\Delta\lambda} = \frac{589\ \mathrm{nm}}{0.59\ \mathrm{nm}} = 1000.$$

From Eq. 19, the total number $N$ of lines needed for the $m = 2$ order is $N = \mathrm{R}/m = 1000/2 = 500$, so the grating need only be $500/7500\ \mathrm{cm}^{-1} = 0.0667\ \mathrm{cm}$ wide. A typical grating is a few centimeters wide, and so will easily resolve the two lines.

## 10 X-Rays and X-Ray Diffraction

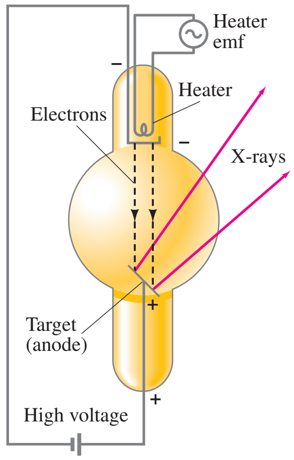
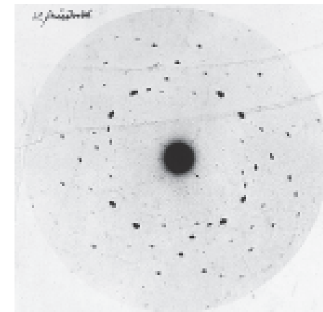



**FIGURE 26** X-ray tube. Electrons emitted by a heated filament in a vacuum tube are accelerated by a high voltage. When they strike the surface of the anode, the "target," X-rays are emitted.

In 1895, W. C. Roentgen (1845–1923) discovered that when electrons were accelerated by a high voltage in a vacuum tube and allowed to strike a glass or metal surface inside the tube, fluorescent minerals some distance away would glow, and photographic film would become exposed. Roentgen attributed these effects to a new type of radiation (different from cathode rays). They were given the name **X-rays** after the algebraic symbol $x$, meaning an unknown quantity. He soon found that X-rays penetrated through some materials better than through others, and within a few weeks he presented the first X-ray photograph (of his wife's hand). The production of X-rays today is usually done in a tube (Fig. 26) similar to Roentgen's, using voltages of typically 30 kV to 150 kV.

Investigations into the nature of X-rays indicated they were not charged particles (such as electrons) since they could not be deflected by electric or magnetic fields. It was suggested that they might be a form of invisible light. However, they showed no diffraction or interference effects using ordinary gratings. Indeed, if their wavelengths were much smaller than the typical grating spacing of $10^{-6}\,\mathrm{m}\ (= 10^3\,\mathrm{nm})$, no effects would be expected. Around 1912, Max von Laue (1879–1960) suggested that if the atoms in a crystal were arranged in a regular array (see Fig. 17–2a), such a crystal might serve as a diffraction grating for very short wavelengths on the order of the spacing between atoms, estimated to be about $10^{-10}\,\mathrm{m}\ (= 10^{-1}\,\mathrm{nm})$. Experiments soon showed that X-rays scattered from a crystal did indeed show the peaks and valleys of a diffraction pattern (Fig. 27). Thus it was shown, in a single blow, that X-rays have a wave nature and that atoms are arranged in a regular way in crystals. Today, X-rays are recognized as electromagnetic radiation with wavelengths in the range of about $10^{-2}$ nm to 10 nm, the range readily produced in an X-ray tube.

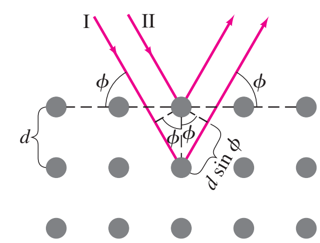


Burndy Library

**FIGURE 27** This X-ray diffraction pattern is one of the first observed by Max von Laue in 1912 when he aimed a beam of X-rays at a zinc sulfide crystal. The diffraction pattern was detected directly on a photographic plate.

We saw in Section 5 that light of shorter wavelength provides greater resolution when we are examining an object microscopically. Since X-rays have much shorter wavelengths than visible light, they should in principle offer much greater resolution. However, there seems to be no effective material to use as lenses for the very short wavelengths of X-rays. Instead, the clever but complicated technique of **X-ray diffraction** (or **crystallography**) has proved very effective for examining the microscopic world of atoms and molecules. In a simple crystal such as NaCl, the atoms are arranged in an orderly cubical fashion, Fig. 28, with atoms spaced a distance $d$ apart. Suppose that a beam of X-rays is incident on the crystal at an angle $\phi$ to the surface, and that the two rays shown are reflected from two subsequent planes of atoms as shown. The two rays will constructively interfere if the extra distance ray I travels is a whole number of wavelengths farther than the distance ray II travels. This extra distance is $2d\sin\phi$. Therefore, constructive interference will occur when

**FIGURE 28** X-ray diffraction by a crystal.




$$m\lambda = 2d\sin\phi, \qquad m = 1, 2, 3, \cdots, \tag{20}$$

where $m$ can be any integer. (Notice that $\phi$ is *not* the angle with respect to the normal to the surface.) This is called the **Bragg equation** after W. L. Bragg (1890–1971), who derived it and who, together with his father W. H. Bragg (1862–1942), developed the theory and technique of X-ray diffraction by crystals in 1912–1913. If the X-ray wavelength is known and the angle $\phi$ is measured, the distance $d$ between atoms can be obtained. This is the basis for X-ray crystallography.

**EXERCISE D** When X-rays of wavelength $0.10 \times 10^{-9}$ m are scattered from a sodium chloride crystal, a second-order diffraction peak is observed at 21°. What is the spacing between the planes of atoms for this scattering?

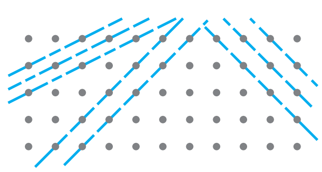
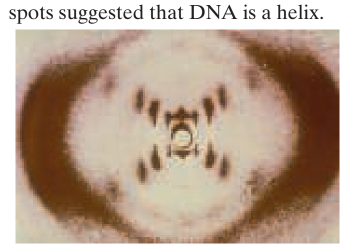


**FIGURE 29** X-rays can be diffracted from many possible planes within a crystal.

Actual X-ray diffraction patterns are quite complicated. First of all, a crystal is a three-dimensional object, and X-rays can be diffracted from different planes at different angles within the crystal, as shown in Fig. 29. Although the analysis is complex, a great deal can be learned about any substance that can be put in crystalline form.

X-ray diffraction has also been very useful in determining the structure of biologically important molecules, such as the double helix structure of DNA, worked out by James Watson and Francis Crick in 1953 (see Fig. 30). Around 1960, the first detailed structure of a protein molecule, myoglobin, was elucidated with the aid of X-ray diffraction. Soon the structure of an important constituent of blood, hemoglobin, was worked out, and since then the structures of a great many molecules have been determined with the help of X-rays.

**FIGURE 30** X-ray diffraction photo of DNA molecules taken by Rosalind Franklin in the early 1950s. The cross of spots suggested that DNA is a helix.



Rosalind Franklin/Photo Researchers, Inc.

# 11 Polarization

An important and useful property of light is that it can be *polarized*. To see what this means, let us examine waves traveling on a rope. A rope can be set into oscillation in a vertical plane as in Fig. 31a, or in a horizontal plane as in Fig. 31b. In either case, the wave is said to be **linearly polarized** or **plane-polarized**—that is, the oscillations are in a plane.

If we now place an obstacle containing a vertical slit in the path of the wave, Fig. 32, a vertically polarized wave passes through the vertical slit, but a horizontally polarized wave will not. If a horizontal slit were used, the vertically polarized wave would be stopped. If both types of slit were used, both types of wave would be stopped by one slit or the other. Note that polarization can exist *only* for *transverse waves*, and not for longitudinal waves such as sound. The latter oscillate only along the direction of motion, and neither orientation of slit would stop them.

Light is not necessarily polarized. It can also be **unpolarized**, which means that the source has oscillations in many planes at once, as shown in Fig. 33. An ordinary incandescent lightbulb emits unpolarized light, as does the Sun.

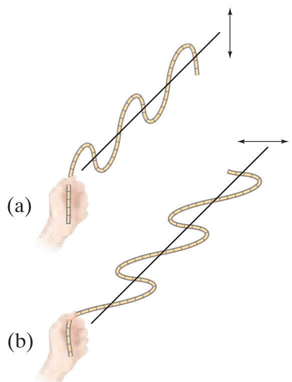



**FIGURE 31** Transverse waves on a rope polarized (a) in a vertical plane and (b) in a horizontal plane.

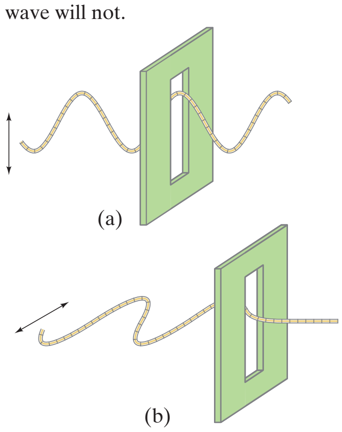
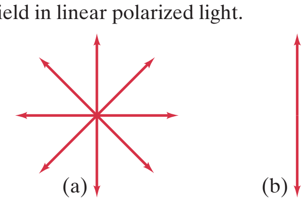



**FIGURE 32** (a) Vertically polarized wave passes through a vertical slit, but (b) a horizontally polarized wave will not.

**FIGURE 33** (below) (a) Oscillation of the electric field vectors in unpolarized light. The light is traveling into or out of the page. (b) Electric field in linear polarized light.




## Polaroids (Polarization by Absorption)

Plane-polarized light can be obtained from unpolarized light using certain crystals such as tourmaline. Or, more commonly, we use a **Polaroid sheet**. (Polaroid materials were invented in 1929 by Edwin Land.) A Polaroid sheet consists of long complex molecules arranged parallel to one another. Such a Polaroid acts like a series of parallel slits to allow one orientation of polarization to pass through nearly undiminished. This direction is called the *transmission axis* of the Polaroid. Polarization perpendicular to this direction is absorbed almost completely by the Polaroid.

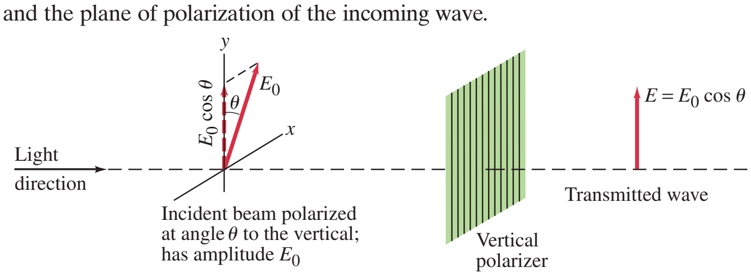
Absorption by a Polaroid can be explained at the molecular level. An electric field $\vec{\mathbf{E}}$ that oscillates parallel to the long molecules can set electrons into motion along the molecules, thus doing work on them and transferring energy. Hence, if $\vec{\mathbf{E}}$ is parallel to the molecules, it gets absorbed. An electric field $\vec{\mathbf{E}}$ perpendicular to the long molecules does not have this possibility of doing work and transferring its energy, and so passes through freely. When we speak of the *transmission axis* of a Polaroid, we mean the direction for which $\vec{\mathbf{E}}$ is passed, so a Polaroid axis is *perpendicular* to the long molecules. If we want to think of there being slits between the parallel molecules in the sense of Fig. 32, then Fig. 32 would apply for the $\vec{\mathbf{B}}$ field in the EM wave, not the $\vec{\mathbf{E}}$ field.

If a beam of plane-polarized light strikes a Polaroid whose transmission axis is at an angle $\theta$ to the incident polarization direction, the beam will emerge plane-polarized parallel to the Polaroid transmission axis, and the amplitude of $E$ will be reduced to $E\cos\theta$, Fig. 34. Thus, a Polaroid passes only that component of polarization (the electric field vector, $\vec{\mathbf{E}}$) that is parallel to its transmission axis. Because the intensity of a light beam is proportional to the square of the amplitude, we see that the intensity of a plane-polarized beam transmitted by a polarizer is

$$I = I_0 \cos^2\theta, \qquad \left[\begin{array}{c}\text{intensity of plane polarized}\\ \text{wave reduced by polarizer}\end{array}\right] \qquad \textbf{(21)}$$

where $I_0$ is the incoming intensity and $\theta$ is the angle between the polarizer transmission axis and the plane of polarization of the incoming wave.

**FIGURE 34** (right) Vertical Polaroid transmits only the vertical component of a wave (electric field) incident upon it.



A Polaroid can be used as a **polarizer** to produce plane-polarized light from unpolarized light, since only the component of light parallel to the axis is transmitted. A Polaroid can also be used as an **analyzer** to determine (1) if light is polarized and (2) the plane of polarization. A Polaroid acting as an analyzer will pass the same amount of light independent of the orientation of its axis if the light is unpolarized; try rotating one lens of a pair of Polaroid sunglasses while looking through it at a lightbulb. If the light is polarized, however, when you rotate the Polaroid the transmitted light will be a maximum when the plane of polarization is parallel to the Polaroid's axis, and a minimum when perpendicular to it. If you do this while looking at the sky, preferably at right angles to the Sun's direction, you will see that skylight is polarized. (Direct sunlight is unpolarized, but don't look directly at the Sun, even through a polarizer, for damage to the eye may occur.) If the light transmitted by an analyzer Polaroid falls to zero at one orientation, then the light is 100% plane-polarized. If it merely reaches a minimum, the light is *partially polarized*.

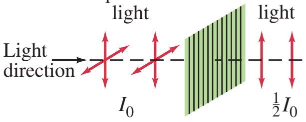



**FIGURE 35** Unpolarized light has equal intensity vertical and horizontal components. After passing through a polarizer, one of these components is eliminated. The intensity of the light is reduced to half.

Unpolarized light consists of light with random directions of polarization. Each of these polarization directions can be resolved into components along two mutually perpendicular directions. On average, an unpolarized beam can be thought of as two plane-polarized beams of equal magnitude perpendicular to one another. When unpolarized light passes through a polarizer, one component is eliminated. So the intensity of the light passing through is reduced by half since half the light is eliminated: $I = \frac{1}{2}I_0$ (Fig. 35).

When two Polaroids are *crossed*—that is, their polarizing axes are perpendicular to one another—unpolarized light can be entirely stopped. As shown in Fig. 36, unpolarized light is made plane-polarized by the first Polaroid (the polarizer).

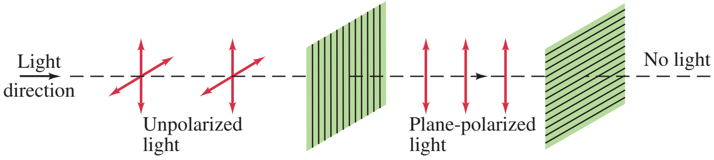



**FIGURE 36** Crossed Polaroids completely eliminate light.

The second Polaroid, the analyzer, then eliminates this component since its transmission axis is perpendicular to the first. You can try this with Polaroid sunglasses (Fig. 37). Note that Polaroid sunglasses eliminate 50% of unpolarized light because of their polarizing property; they absorb even more because they are colored.

**EXAMPLE 13** **Two Polaroids at 60°.** Unpolarized light passes through two Polaroids; the axis of one is vertical and that of the other is at 60° to the vertical. Describe the orientation and intensity of the transmitted light.

**APPROACH** Half of the unpolarized light is absorbed by the first Polaroid, and the remaining light emerges plane polarized. When that light passes through the second Polaroid, the intensity is further reduced according to Eq. 21, and the plane of polarization is then along the axis of the second Polaroid.

**SOLUTION** The first Polaroid eliminates half the light, so the intensity is reduced by half: $I_1 = \frac{1}{2}I_0$. The light reaching the second polarizer is vertically polarized and so is reduced in intensity (Eq. 21) to

$$I_2 = I_1(\cos 60°)^2 = \tfrac{1}{4}I_1.$$

Thus, $I_2 = \frac{1}{8}I_0$. The transmitted light has an intensity one-eighth that of the original and is plane-polarized at a 60° angle to the vertical.

**FIGURE 37** Crossed Polaroids. When the two polarized sunglass lenses overlap, with axes perpendicular, almost no light passes through.

Diane Schiumo/Fundamental Photographs, NYC

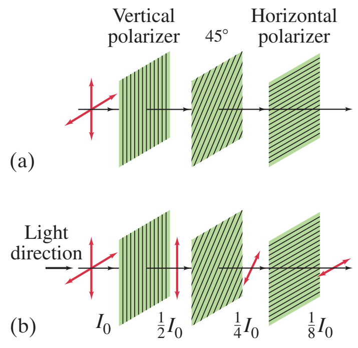



**FIGURE 38** Example 14.

**CONCEPTUAL EXAMPLE 14** **Three Polaroids.** We saw in Fig. 36 that when unpolarized light falls on two crossed Polaroids (axes at 90°), no light passes through. What happens if a third Polaroid, with axis at 45° to each of the other two, is placed between them (Fig. 38a)?

**RESPONSE** We start just as in Example 13 and recall again that light emerging from each Polaroid is polarized parallel to that Polaroid's axis. Thus the angle in Eq. 21 is that between the transmission axes of each pair of Polaroids taken in turn. The first Polaroid changes the unpolarized light to plane-polarized and reduces the intensity from $I_0$ to $I_1 = \frac{1}{2}I_0$. The second polarizer further reduces the intensity by $(\cos 45°)^2$, Eq. 21:

$$I_2 = I_1(\cos 45°)^2 = \tfrac{1}{2}I_1 = \tfrac{1}{4}I_0.$$

The light leaving the second polarizer is plane polarized at 45° (Fig. 38b) relative to the third polarizer, so the third one reduces the intensity to

$$I_3 = I_2(\cos 45°)^2 = \tfrac{1}{2}I_2,$$

or $I_3 = \frac{1}{8}I_0$. Thus $\frac{1}{8}$ of the original intensity gets transmitted.

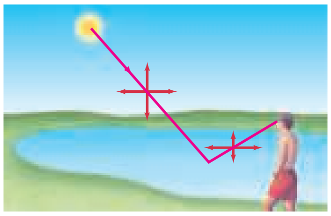
**NOTE** If we don't insert the 45° Polaroid, zero intensity results (Fig. 36).

**EXERCISE E** How much light would pass through if the 45° polarizer in Example 14 was placed not between the other two polarizers but (*a*) before the vertical (first) polarizer, or (*b*) after the horizontal polarizer?

## Polarization by Reflection

Another means of producing polarized light from unpolarized light is by reflection. When light strikes a nonmetallic surface at any angle other than perpendicular, the reflected beam is polarized preferentially in the plane parallel to the surface, Fig. 39. In other words, the component with polarization in the plane perpendicular to the surface is preferentially transmitted or absorbed. You can check this by rotating Polaroid sunglasses while looking through them at a flat surface of a lake or road. Since most outdoor surfaces are horizontal, Polaroid sunglasses are made with their axes vertical to eliminate the more strongly reflected horizontal component, and thus reduce glare. People who go fishing wear Polaroids to eliminate reflected glare from the surface of a lake or stream and thus see beneath the water more clearly (Fig. 40).

**FIGURE 39** Light reflected from a nonmetallic surface, such as the smooth surface of water in a lake, is partially polarized parallel to the surface.



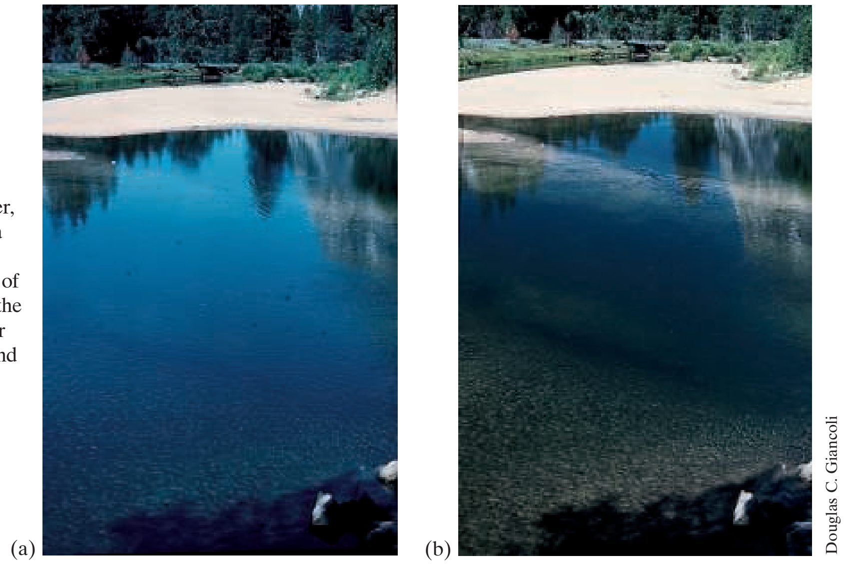



**FIGURE 40** Photographs of a river, (a) allowing all light into the camera lens, and (b) using a polarizer. The polarizer is adjusted to absorb most of the (polarized) light reflected from the water's surface, allowing the dimmer light from the bottom of the river, and any fish lying there, to be seen more readily.

The amount of polarization in the reflected beam depends on the angle, varying from no polarization at normal incidence to 100% polarization at an angle known as the **polarizing angle** $\theta_p$.[†] This angle is related to the index of refraction of the two materials on either side of the boundary by the equation

$$\tan\theta_p = \frac{n_2}{n_1}, \tag{22a}$$

where $n_1$ is the index of refraction of the material in which the beam is traveling, and $n_2$ is that of the medium beyond the reflecting boundary. If the beam is traveling in air, $n_1 = 1$, and Eq. 22a becomes

$$\tan\theta_p = n. \tag{22b}$$

The polarizing angle $\theta_p$ is also called **Brewster's angle**, and Eqs. 22 *Brewster's law*, after the Scottish physicist David Brewster (1781–1868), who worked it out experimentally in 1812. Equations 22 can be derived from the electromagnetic wave theory of light. It is interesting that at Brewster's angle, the reflected ray and the transmitted (refracted) ray make a 90° angle to each other; that is, $\theta_p + \theta_r = 90°$, where $\theta_r$ is the refraction angle (Fig. 41). This can be seen by substituting Eq. 22a, $n_2 = n_1\tan\theta_p = n_1\sin\theta_p/\cos\theta_p$, into Snell's law, $n_1\sin\theta_p = n_2\sin\theta_r$, which gives $\cos\theta_p = \sin\theta_r$ which can only hold if $\theta_p = 90° - \theta_r$.

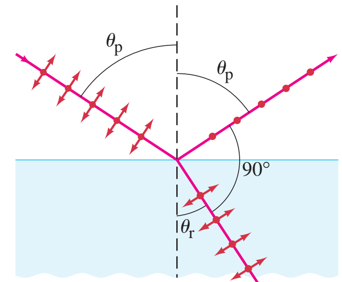



**FIGURE 41** At $\theta_p$ the reflected light is plane-polarized parallel to the surface, and $\theta_p + \theta_r = 90°$, where $\theta_r$ is the refraction angle. (The large dots represent vibrations perpendicular to the page.)

**EXAMPLE 15 Polarizing angle.** (*a*) At what incident angle is sunlight reflected from a lake plane-polarized? (*b*) What is the refraction angle?

**APPROACH** The polarizing angle at the surface is Brewster's angle, Eq. 22b. We find the angle of refraction from Snell's law.

**SOLUTION** (*a*) We use Eq. 22b with $n = 1.33$, so $\tan\theta_p = 1.33$ giving $\theta_p = 53.1°$.
(*b*) From Snell's law, $\sin\theta_r = \sin\theta_p/n = \sin 53.1°/1.33 = 0.601$ giving $\theta_r = 36.9°$.

**NOTE** $\theta_p + \theta_r = 53.1° + 36.9° = 90.0°$, as expected.

## *12 Liquid Crystal Displays (LCD)

A wonderful use of polarization is in a **liquid crystal display** (LCD). LCDs are used as the display in hand-held calculators, digital wrist watches, cell phones, and in beautiful color flat-panel computer and television screens.

A liquid crystal display is made up of many tiny rectangles called **pixels**, or "picture elements." The picture you see depends on which pixels are dark or light and of what color, as suggested in Fig. 42 for a simple black and white picture.

Liquid crystals are organic materials that at room temperature exist in a phase that is neither fully solid nor fully liquid. They are sort of gooey, and their molecules display a randomness of position characteristic of liquids. They also show some of the orderliness of a solid crystal, but only in one dimension. The liquid crystals we find useful are made up of relatively rigid rod-like molecules that interact weakly with each other and tend to align parallel to each other, as shown in Fig. 43.

**FIGURE 42** Example of an image made up of many small squares or *pixels* (picture elements). This one has rather poor resolution.

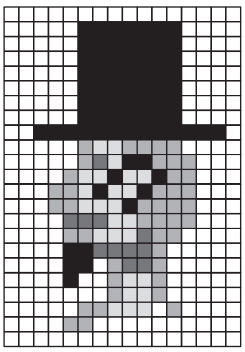


**FIGURE 43** Liquid crystal molecules tend to align in one dimension (parallel to each other) but have random positions (left-right, up-down).

[†]Only a fraction of the incident light is reflected at the surface of a transparent medium. Although this reflected light is 100% polarized (if $\theta = \theta_p$), the remainder of the light, which is transmitted into the new medium, is only partially polarized.

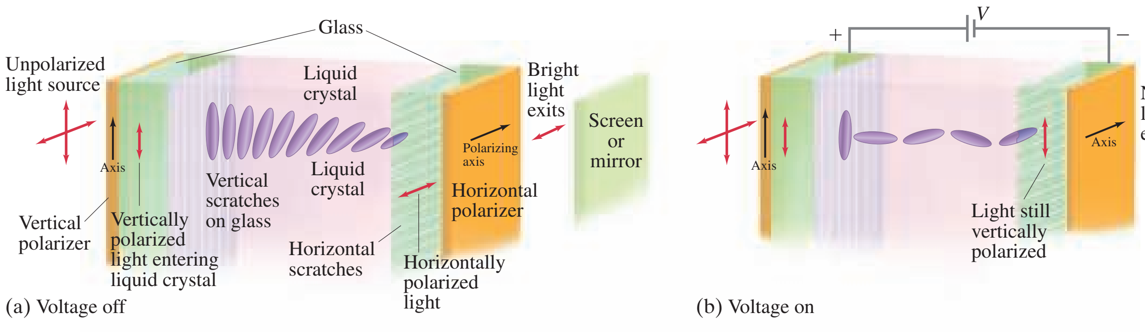



**FIGURE 44** (a) "Twisted" form of liquid crystal. Light polarization plane is rotated 90°, and so is transmitted by the horizontal polarizer. Only one line of molecules is shown. (b) Molecules disoriented by electric field. Plane of polarization is not changed, so light does not pass through the horizontal polarizer. (The transparent electrodes are not shown.)

In a simple LCD, each pixel (picture element) contains a liquid crystal sandwiched between two glass plates whose inner surfaces have been brushed to form nanometer-wide parallel scratches. The rod-like liquid crystal molecules in contact with the scratches tend to line up along the scratches. The two plates typically have their scratches at 90° to each other, and the weak forces between the rod-like molecules tend to keep them nearly aligned with their nearest neighbors, resulting in the twisted pattern shown in Fig. 44a.

The outer surfaces of the glass plates each have a thin film polarizer, they too oriented at 90° to each other. Unpolarized light incident from the left becomes plane-polarized and the liquid crystal molecules keep this polarization aligned with their rod-like shape. That is, the plane of polarization of the light rotates with the molecules as the light passes through the liquid crystal. The light emerges with its plane of polarization rotated by 90°, and passes through the second polarizer readily (Fig. 44a). A tiny LCD pixel in this situation will appear bright.

Now suppose a voltage is applied to transparent electrodes on each glass plate of the pixel. The rod-like molecules are polar (or can acquire an internal separation of charge due to the applied electric field). The applied voltage tends to align the molecules and they no longer follow the twisted pattern shown in Fig. 44a, with the end molecules always lying in a plane parallel to the glass plates. Instead the applied electric field tends to align the molecules flat, left to right (perpendicular to the glass plates), and they don't affect the light polarization significantly. The entering plane-polarized light no longer has its plane of polarization rotated as it passes through, and no light can exit through the second (horizontal) polarizer (Fig. 44b). With the voltage on, the pixel appears dark.†

The simple display screens of watches and calculators use ambient light as the source (you can't see the display in the dark), and a mirror behind the LCD to reflect the light back. There are only a few pixels, corresponding to the elongated segments needed to form the numbers from 0 to 9 (and letters in some displays), as seen in Fig. 45. Any pixels to which a voltage is applied appear dark and form part of a number. With no voltage, pixels pass light through the polarizers to the mirror and back out, which forms a bright background to the dark numbers on the display.

**FIGURE 45** Calculator LCD display. The black segments or pixels have a voltage applied to them. Note that the 8 uses all seven segments (pixels), whereas other numbers use fewer.


Texas Instruments Inc.

Color television and computer LCDs are more sophisticated. A color pixel consists of three cells, or subpixels, each covered with a red, green, or blue filter. Varying brightnesses of these three primary colors can yield almost any natural color. A good-quality screen consists of a million or more pixels. Behind this array of pixels is a light source, often thin fluorescent tubes the diameter of a straw. The light passes through the pixels, or not, depending on the voltage applied to each subpixel, as in Fig. 44a and b.

†In some displays, the polarizers are parallel to each other (the scratches remain at 90° to maintain the twist). Then voltage off results in black (no light), and voltage on results in bright light.

# *13 Scattering of Light by the Atmosphere

Sunsets are red, the sky is blue, and skylight is polarized (at least partially). These phenomena can be explained on the basis of the *scattering* of light by the molecules of the atmosphere. In Fig. 46 we see unpolarized light from the Sun impinging on a molecule of the Earth's atmosphere. The electric field of the EM wave sets the electric charges within the molecule into oscillation, and the molecule absorbs some of the incident radiation. But the molecule quickly reemits this light since the charges are oscillating. Oscillating electric charges produce EM waves. The intensity is strongest along the direction perpendicular to the oscillation, and drops to zero along the line of oscillation. In Fig. 46 the motion of the charges is resolved into two components. An observer at right angles to the direction of the sunlight, as shown, will see plane-polarized light because no light is emitted along the line of the other component of the oscillation. (When viewing along the line of an oscillation, you don't see that oscillation, and hence see no waves made by it.) At other viewing angles, both components will be present; one will be stronger, however, so the light appears partially polarized. Thus, the process of scattering explains the polarization of skylight.

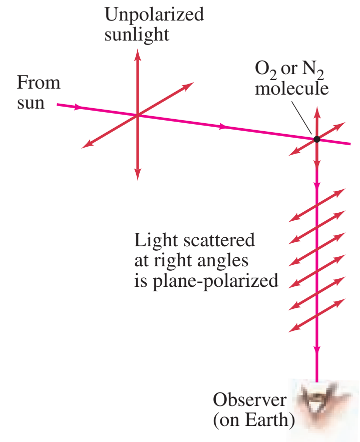



**FIGURE 46** Unpolarized sunlight scattered by molecules of the air. An observer at right angles sees plane-polarized light, since the component of oscillation along the line of sight emits no light along that line.

Scattering of light by the Earth's atmosphere depends on wavelength $\lambda$. For particles much smaller than the wavelength of light (such as molecules of air), the particles will be less of an obstruction to long wavelengths than to short ones. The scattering decreases, in fact, as $1/\lambda^4$. Blue and violet light are thus scattered much more than red and orange, which is why the sky looks blue. At sunset, the Sun's rays pass through a maximum length of atmosphere. Much of the blue has been taken out by scattering. The light that reaches the surface of the Earth, and reflects off clouds and haze, is thus lacking in blue. That is why sunsets appear reddish.

**PHYSICS APPLIED**
*Why the sky is blue*
*Why sunsets are red*

The dependence of scattering on $1/\lambda^4$ is valid only if the scattering objects are much smaller than the wavelength of the light. This is valid for oxygen and nitrogen molecules whose diameters are about 0.2 nm. Clouds, however, contain water droplets or crystals that are much larger than $\lambda$. They scatter all frequencies of light nearly uniformly. Hence clouds appear white (or gray, if shadowed).

**PHYSICS APPLIED**
*Why clouds are white*

## Summary

**Diffraction** refers to the fact that light, like other waves, bends around objects it passes, and spreads out after passing through narrow slits. This bending gives rise to a **diffraction pattern** due to interference between rays of light that travel different distances.

Light passing through a very narrow slit of width $D$ (on the order of the wavelength $\lambda$) will produce a pattern with a bright central maximum of half-width $\theta$ given by

$$\sin\theta = \frac{\lambda}{D}, \tag{1}$$

flanked by fainter lines to either side.

The minima in the diffraction pattern occur at

$$D\sin\theta = m\lambda \tag{2}$$

where $m = 1, 2, 3, \cdots$, but not $m = 0$ (for which the pattern has its strongest maximum).

The **intensity** at any point in the single-slit diffraction pattern can be calculated using **phasor** diagrams. The same technique can be used to determine the intensity of the pattern produced by two slits.

The pattern for two-slit interference can be described as a series of maxima due to interference of light from the two slits, modified by an "envelope" due to diffraction at each slit.

The wave nature of light limits the sharpness or **resolution** of images. Because of diffraction, it is not possible to *discern details smaller than the wavelength* of the radiation being used. The useful magnification of a light microscope is limited by diffraction to about 500×.

A **diffraction grating** consists of many parallel slits or lines, each separated from its neighbors by a distance $d$. The peaks of constructive interference occur at angles $\theta$ given by

$$\sin\theta = \frac{m\lambda}{d}, \tag{13}$$

where $m = 0, 1, 2, \cdots$. The peaks of constructive interference are much brighter and sharper for a diffraction grating than for a simple two-slit apparatus. [*Peak width is inversely proportional to the total number of lines in the grating.]

[*A diffraction grating (or a prism) is used in a **spectrometer** to separate different colors or to observe **line spectra**. For a given order $m$, $\theta$ depends on $\lambda$. Precise determination of wavelength can be done with a spectrometer by careful measurement of $\theta$.]

*__X-rays__ are a form of electromagnetic radiation of very short wavelength. They are produced when high-speed electrons, accelerated by high voltage in an evacuated tube, strike a glass or metal target.

In **unpolarized light**, the electric field vectors oscillate in all transverse directions. If the electric vector oscillates only in one plane, the light is said to be **plane-polarized**. Light can also be partially polarized.

When an unpolarized light beam passes through a **Polaroid** sheet, the emerging beam is plane-polarized. When a light beam is polarized and passes through a Polaroid, the intensity varies as the Polaroid is rotated. Thus a Polaroid can act as a **polarizer** or as an **analyzer**.

The intensity $I$ of a plane-polarized light beam incident on a Polaroid is reduced by the factor

$$I = I_0 \cos^2 \theta \tag{21}$$

where $\theta$ is the angle between the axis of the Polaroid and the initial plane of polarization.

Light can also be partially or fully **polarized by reflection**. If light traveling in air is reflected from a medium of index of refraction $n$, the reflected beam will be *completely* plane-polarized if the incident angle $\theta_p$ is given by

$$\tan \theta_p = n. \tag{22b}$$

The fact that light can be polarized shows that it must be a transverse wave.

## Answers to Exercises

**A:** Narrower.

**B:** (*b*).

**C:** (c).

**D:** 0.28 nm.

**E:** Zero for both (*a*) and (*b*), because the two successive polarizers at 90° cancel all light. The 45° Polaroid must be inserted *between* the other two if any transmission is to occur.