CHAPTER

36 DIFFRACTION

36-1 WHAT IS PHYSICS?

One focus of physics in the study of light is to understand and put to use the diffraction of light as it passes through a narrow slit or (as we shall discuss) past either a narrow obstacle or an edge. We touched on this phenomenon in Chapter 35 when we looked at how light flared—diffracted—through the slits in Young's experiment. Diffraction through a given slit is more complicated than simple flaring, however, because the light also interferes with itself and produces an interference pattern. It is because of such complications that light is rich with application opportunities. Even though the diffraction of light as it passes through a slit or past an obstacle seems awfully academic, countless engineers and scientists make their living using this physics, and the total worth of diffraction applications worldwide is probably incalculable.

Before we can discuss some of these applications, we first must discuss why diffraction is due to the wave nature of light.

36-2 Diffraction and the Wave Theory of Light

In Chapter 35 we defined diffraction rather loosely as the flaring of light as it emerges from a narrow slit. More than just flaring occurs, however, because the light produces an interference pattern called a **diffraction pattern.** For example, when monochromatic light from a distant source (or a laser) passes through a narrow slit and is then intercepted by a viewing screen, the light produces on the screen a diffraction pattern like that in Fig. 36-1. This pattern consists of a broad and intense (very bright) central maximum plus a number of narrower and less intense maxima (called **secondary** or **side** maxima) to both sides. In between the maxima are minima. Light flares into those dark regions, but the light waves cancel out one another.

Such a pattern would be totally unexpected in geometrical optics: If light traveled in straight lines as rays, then the slit would allow some of those rays through to form a sharp rendition of the slit on the viewing screen instead of a pattern of bright and dark bands as we see in Fig. 36-1. As in Chapter 35, we must conclude that geometrical optics is only an approximation.

Diffraction is not limited to situations when light passes through a narrow opening (such as a slit or pinhole). It also occurs when light passes an edge, such as the edges of the razor blade whose diffraction pattern is shown in Fig. 36-2. Note the lines of maxima and minima that run approximately parallel to the edges, at both the inside edges of the blade and the outside edges. As the light passes, say, the vertical edge at the left, it flares left and right and undergoes interference, producing the pattern along the left edge. The rightmost portion of that pattern actually lies behind the blade, within what would be the blade's shadow if geometrical optics prevailed.

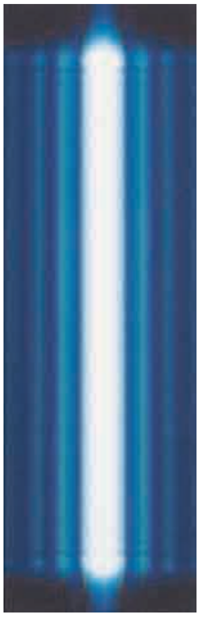


Fig. 36-1 This diffraction pattern appeared on a viewing screen when light that had passed through a narrow vertical slit reached the screen. Diffraction caused the light to flare out perpendicular to the long sides of the slit. That flaring produced an interference pattern consisting of a broad central maximum plus less intense and narrower secondary (or side) maxima, with minima between them. *(Ken Kay/Fundamental Photographs)*

You encounter a common example of diffraction when you look at a clear blue sky and see tiny specks and hairlike structures floating in your view. These *floaters,* as they are called, are produced when light passes the edges of tiny deposits in the vitreous humor, the transparent material filling most of the eyeball. What you are seeing when a floater is in your field of vision is the diffraction pattern produced on the retina by one of these deposits. If you sight through a pinhole in a piece of cardboard so as to make the light entering your eye approximately a plane wave, you can distinguish individual maxima and minima in the patterns.

Diffraction is a wave effect. That is, it occurs because light is a wave and it occurs with other types of waves as well. For example, you have probably seen diffraction in action at football games. When a cheerleader near the playing field yells up at several thousand noisy fans, the yell can hardly be heard because the sound waves diffract when they pass through the narrow opening of the cheerleader's mouth. This flaring leaves little of the waves traveling toward the fans in front of the cheerleader. To offset the diffraction, the cheerleader can yell through a megaphone. The sound waves then emerge from the much wider opening at the end of the megaphone. The flaring is thus reduced, and much more of the sound reaches the fans in front of the cheerleader.

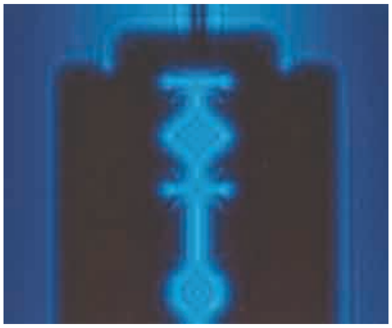


Fig. 36-2 The diffraction pattern produced by a razor blade in monochromatic light. Note the lines of alternating maximum and minimum intensity. *(Ken Kay/Fundamental Photographs)*

The Fresnel Bright Spot

Diffraction finds a ready explanation in the wave theory of light. However, this theory, originally advanced in the late 1600s by Huygens and used 123 years later by Young to explain double-slit interference, was very slow in being adopted, largely because it ran counter to Newton's theory that light was a stream of particles.

Newton's view was the prevailing view in French scientific circles of the early 19th century, when Augustin Fresnel was a young military engineer. Fresnel, who believed in the wave theory of light, submitted a paper to the French Academy of Sciences describing his experiments with light and his wave-theory explanations of them.

In 1819, the Academy, dominated by supporters of Newton and thinking to challenge the wave point of view, organized a prize competition for an essay on the subject of diffraction. Fresnel won. The Newtonians, however, were not swayed. One of them, S. D. Poisson, pointed out the "strange result" that if Fresnel's theories were correct, then light waves should flare into the shadow region of a sphere as they pass the edge of the sphere, producing a bright spot at the center of the shadow. The prize committee arranged a test of Poisson's prediction and discovered that the predicted *Fresnel bright spot,* as we call it today, was indeed there (Fig. 36-3). Nothing builds confidence in a theory so much as having one of its unexpected and counterintuitive predictions verified by experiment.

Fig. 36-3 A photograph of the diffraction pattern of a disk. Note the concentric diffraction rings and the Fresnel bright spot at the center of the pattern. This experiment is essentially identical to that arranged by the committee testing Fresnel's theories, because both the sphere they used and the disk used here have a cross section with a circular edge. *(Jearl Walker)*

36-3 Diffraction by a Single Slit: Locating the Minima

Let us now examine the diffraction pattern of plane waves of light of wavelength λ that are diffracted by a single long, narrow slit of width a in an otherwise opaque screen B, as shown in cross section in Fig. 36-4. (In that figure, the slit's length extends into and out of the page, and the incoming wavefronts are parallel to screen B.) When the diffracted light reaches viewing screen C, waves from different points within the slit undergo interference and produce a diffraction pattern of bright and dark fringes (interference maxima and minima) on the screen. To locate the fringes, we shall use a procedure somewhat similar to the one we used to locate the fringes in a two-slit interference pattern. However, diffraction is more mathematically challenging, and here we shall be able to find equations for only the dark fringes.

Before we do that, however, we can justify the central bright fringe seen in Fig. 36-1 by noting that the Huygens wavelets from all points in the slit travel about the same distance to reach the center of the pattern and thus are in phase there. As for the other bright fringes, we can say only that they are approximately halfway between adjacent dark fringes.

To find the dark fringes, we shall use a clever (and simplifying) strategy that involves pairing up all the rays coming through the slit and then finding what conditions cause the wavelets of the rays in each pair to cancel each other. We apply this strategy in Fig. 36-4 to locate the first dark fringe, at point P_1. First, we mentally divide the slit into two *zones* of equal widths $a/2$. Then we extend to P_1 a light ray r_1 from the top point of the top zone and a light ray r_2 from the top point of the bottom zone. We want the wavelets along these two rays to cancel each other when they arrive at P_1. Then any similar pairing of rays from the two zones will give cancellation. A central axis is drawn from the center of the slit to screen C, and P_1 is located at an angle θ to that axis.

The wavelets of the pair of rays r_1 and r_2 are in phase within the slit because they originate from the same wavefront passing through the slit, along the width of the slit. However, to produce the first dark fringe they must be out of phase by $\lambda/2$ when they reach P_1; this phase difference is due to their path length difference, with the path traveled by the wavelet of r_2 to reach P_1 being longer than the path traveled by the wavelet of r_1. To display this path length difference, we find a point b on ray r_2 such that the path length from b to P_1 matches the path length of ray r_1. Then the path length difference between the two rays is the distance from the center of the slit to b.

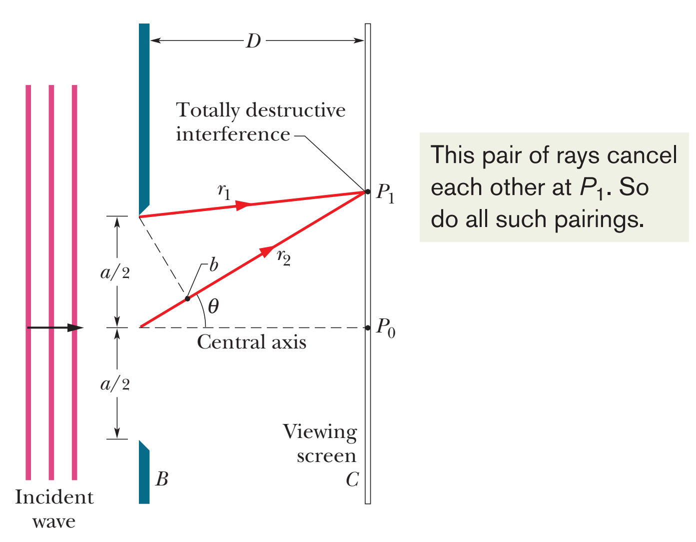


Fig. 36-4 Waves from the top points of two zones of width $a/2$ undergo fully destructive interference at point P_1 on viewing screen C.

When viewing screen C is near screen B, as in Fig. 36-4, the diffraction pattern on C is difficult to describe mathematically. However, we can simplify the mathematics considerably if we arrange for the screen separation D to be much larger than the slit width a. Then we can approximate rays r_1 and r_2 as being parallel, at angle θ to the central axis (Fig. 36-5). We can also approximate the triangle formed by point b, the top point of the slit, and the center point of the slit as being a right triangle, and one of the angles inside that triangle as being θ. The path length difference between rays r_1 and r_2 (which is still the distance from the center of the slit to point b) is then equal to $(a/2)\sin\theta$.

We can repeat this analysis for any other pair of rays originating at corresponding points in the two zones (say, at the midpoints of the zones) and extending to point P_1. Each such pair of rays has the same path length difference $(a/2)\sin\theta$. Setting this common path length difference equal to $\lambda/2$ (our condition for the first dark fringe), we have

$$\frac{a}{2}\sin\theta = \frac{\lambda}{2},$$

which gives us

$$a\sin\theta = \lambda \quad \text{(first minimum)}. \tag{36-1}$$

Given slit width a and wavelength λ, Eq. 36-1 tells us the angle θ of the first dark fringe above and (by symmetry) below the central axis.

Note that if we begin with $a > \lambda$ and then narrow the slit while holding the wavelength constant, we increase the angle at which the first dark fringes appear; that is, the extent of the diffraction (the extent of the flaring and the width of the pattern) is *greater* for a *narrower* slit. When we have reduced the slit width to the wavelength (that is, $a = \lambda$), the angle of the first dark fringes is 90°. Since the first dark fringes mark the two edges of the central bright fringe, that bright fringe must then cover the entire viewing screen.

We find the second dark fringes above and below the central axis as we found the first dark fringes, except that we now divide the slit into *four* zones of equal widths $a/4$, as shown in Fig. 36-6a. We then extend rays r_1, r_2, r_3, and r_4 from the top points of the zones to point P_2, the location of the second dark fringe above the central axis. To produce that fringe, the path length difference

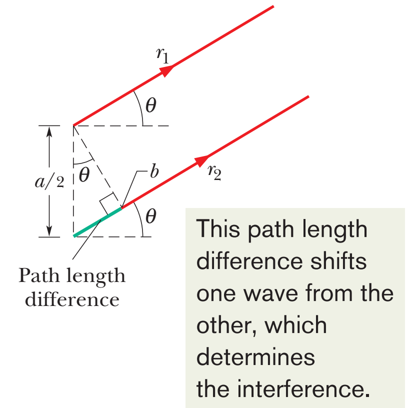


Fig. 36-5 For $D \gg a$, we can approximate rays r_1 and r_2 as being parallel, at angle θ to the central axis.

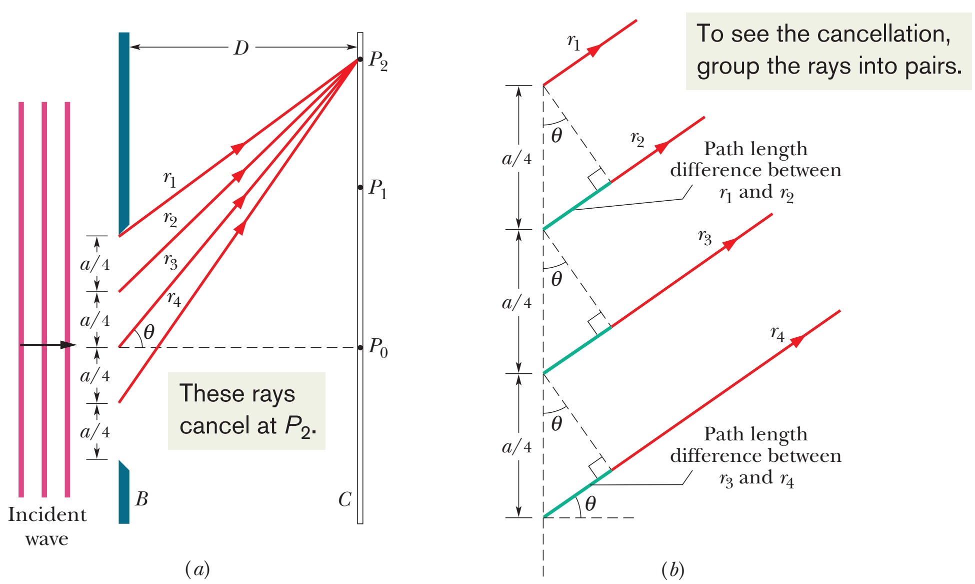


Fig. 36-6 (a) Waves from the top points of four zones of width $a/4$ undergo fully destructive interference at point P_2. (b) For $D \gg a$, we can approximate rays r_1, r_2, r_3, and r_4 as being parallel, at angle θ to the central axis.

between r_1 and r_2, that between r_2 and r_3, and that between r_3 and r_4 must all be equal to $\lambda/2$.

For $D \gg a$, we can approximate these four rays as being parallel, at angle θ to the central axis. To display their path length differences, we extend a perpendicular line through each adjacent pair of rays, as shown in Fig. 36-6*b*, to form a series of right triangles, each of which has a path length difference as one side. We see from the top triangle that the path length difference between r_1 and r_2 is $(a/4) \sin \theta$. Similarly, from the bottom triangle, the path length difference between r_3 and r_4 is also $(a/4) \sin \theta$. In fact, the path length difference for any two rays that originate at corresponding points in two adjacent zones is $(a/4) \sin \theta$. Since in each such case the path length difference is equal to $\lambda/2$, we have

$$\frac{a}{4} \sin \theta = \frac{\lambda}{2},$$

which gives us

$$a \sin \theta = 2\lambda \qquad \text{(second minimum)}. \tag{36-2}$$

We could now continue to locate dark fringes in the diffraction pattern by splitting up the slit into more zones of equal width. We would always choose an even number of zones so that the zones (and their waves) could be paired as we have been doing. We would find that the dark fringes above and below the central axis can be located with the general equation

$$a \sin \theta = m\lambda, \qquad \text{for } m = 1, 2, 3, \ldots \quad \text{(minima—dark fringes)}. \tag{36-3}$$

You can remember this result in the following way. Draw a triangle like the one in Fig. 36-5, but for the full slit width a, and note that the path length difference between the top and bottom rays equals $a \sin \theta$. Thus, Eq. 36-3 says:

In a single-slit diffraction experiment, dark fringes are produced where the path length differences ($a \sin \theta$) between the top and bottom rays are equal to $\lambda, 2\lambda, 3\lambda, \ldots$.

This may seem to be wrong because the waves of those two particular rays will be exactly in phase with each other when their path length difference is an integer number of wavelengths. However, they each will still be part of a pair of waves that are exactly out of phase with each other; thus, *each* wave will be canceled by some other wave, resulting in darkness. (Two light waves that are exactly out of phase will always cancel each other, giving a net wave of zero, even if they happen to be exactly in phase with other light waves.)

Equations 36-1, 36-2, and 36-3 are derived for the case of $D \gg a$. However, they also apply if we place a converging lens between the slit and the viewing screen and then move the screen in so that it coincides with the focal plane of the lens. The lens ensures that rays which now reach any point on the screen are *exactly* parallel (rather than approximately) back at the slit. They are like the initially parallel rays of Fig. 34-14*a* that are directed to the focal point by a converging lens.

CHECKPOINT 1

We produce a diffraction pattern on a viewing screen by means of a long narrow slit illuminated by blue light. Does the pattern expand away from the bright center (the maxima and minima shift away from the center) or contract toward it if we (a) switch to yellow light or (b) decrease the slit width?

Sample Problem

Single-slit diffraction pattern with white light

A slit of width a is illuminated by white light.

(a) For what value of a will the first minimum for red light of wavelength $\lambda = 650$ nm appear at $\theta = 15°$?

KEY IDEA

Diffraction occurs separately for each wavelength in the range of wavelengths passing through the slit, with the locations of the minima for each wavelength given by Eq. 36-3 ($a \sin \theta = m\lambda$).

Calculation: When we set $m = 1$ (for the first minimum) and substitute the given values of θ and λ, Eq. 36-3 yields

$$a = \frac{m\lambda}{\sin \theta} = \frac{(1)(650 \text{ nm})}{\sin 15°}$$
$$= 2511 \text{ nm} \approx 2.5\ \mu\text{m}. \qquad \text{(Answer)}$$

For the incident light to flare out that much ($\pm 15°$ to the first minima) the slit has to be very fine indeed—in this case, a mere four times the wavelength. For comparison, note that a fine human hair may be about 100 μm in diameter.

(b) What is the wavelength λ' of the light whose first side diffraction maximum is at 15°, thus coinciding with the first minimum for the red light?

KEY IDEA

The first side maximum for any wavelength is about halfway between the first and second minima for that wavelength.

Calculations: Those first and second minima can be located with Eq. 36-3 by setting $m = 1$ and $m = 2$, respectively. Thus, the first side maximum can be located *approximately* by setting $m = 1.5$. Then Eq. 36-3 becomes

$$a \sin \theta = 1.5\lambda'.$$

Solving for λ' and substituting known data yield

$$\lambda' = \frac{a \sin \theta}{1.5} = \frac{(2511 \text{ nm})(\sin 15°)}{1.5}$$
$$= 430 \text{ nm}. \qquad \text{(Answer)}$$

Light of this wavelength is violet (far blue, near the short-wavelength limit of the human range of visible light). From the two equations we used, can you see that the first side maximum for light of wavelength 430 nm will always coincide with the first minimum for light of wavelength 650 nm, no matter what the slit width is? However, the angle θ at which this overlap occurs does depend on slit width. If the slit is relatively narrow, the angle will be relatively large, and conversely.

36-4 Intensity in Single-Slit Diffraction, Qualitatively

In Section 36-3 we saw how to find the positions of the minima and the maxima in a single-slit diffraction pattern. Now we turn to a more general problem: find an expression for the intensity I of the pattern as a function of θ, the angular position of a point on a viewing screen.

To do this, we divide the slit of Fig. 36-4 into N zones of equal widths Δx small enough that we can assume each zone acts as a source of Huygens wavelets. We wish to superimpose the wavelets arriving at an arbitrary point P on the viewing screen, at angle θ to the central axis, so that we can determine the amplitude E_θ of the electric component of the resultant wave at P. The intensity of the light at P is then proportional to the square of that amplitude.

To find E_θ, we need the phase relationships among the arriving wavelets. The phase difference between wavelets from adjacent zones is given by

$$\left(\begin{array}{c}\text{phase}\\ \text{difference}\end{array}\right) = \left(\frac{2\pi}{\lambda}\right)\left(\begin{array}{c}\text{path length}\\ \text{difference}\end{array}\right).$$

For point P at angle θ, the path length difference between wavelets from adjacent zones is $\Delta x \sin \theta$; so the phase difference $\Delta\phi$ between wavelets from adjacent zones is

$$\Delta\phi = \left(\frac{2\pi}{\lambda}\right)(\Delta x \sin \theta). \tag{36-4}$$

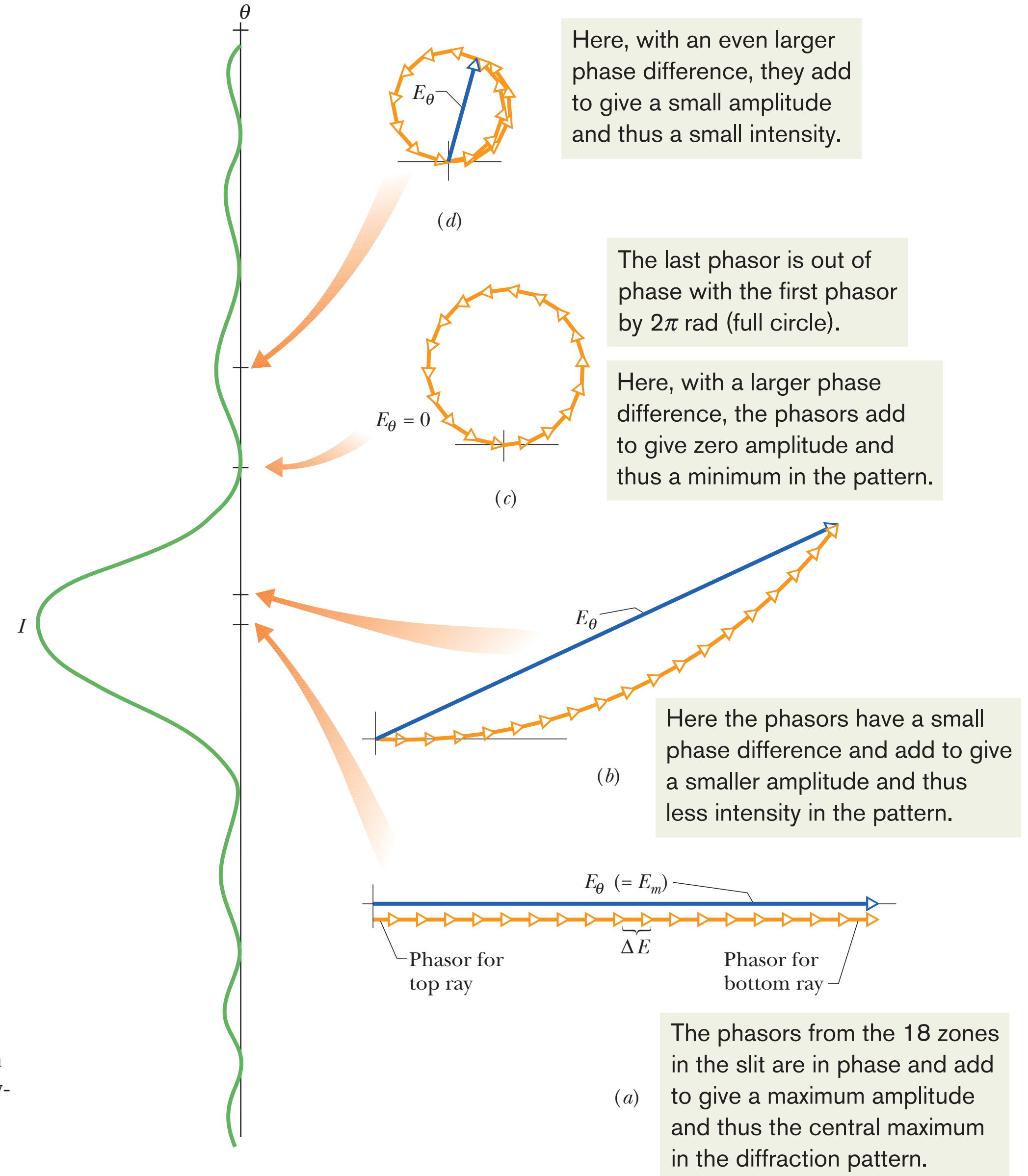


Fig. 36-7 Phasor diagrams for $N = 18$ phasors, corresponding to the division of a single slit into 18 zones. Resultant amplitudes E_θ are shown for (*a*) the central maximum at $\theta = 0$, (*b*) a point on the screen lying at a small angle θ to the central axis, (*c*) the first minimum, and (*d*) the first side maximum.

We assume that the wavelets arriving at P all have the same amplitude ΔE. To find the amplitude E_θ of the resultant wave at P, we add the amplitude ΔE via phasors. To do this, we construct a diagram of N phasors, one corresponding to the wavelet from each zone in the slit.

For point P_0 at $\theta = 0$ on the central axis of Fig. 36-4, Eq. 36-4 tells us that the phase difference $\Delta\phi$ between the wavelets is zero; that is, the wavelets all arrive in phase. Figure 36-7*a* is the corresponding phasor diagram; adjacent phasors represent wavelets from adjacent zones and are arranged head to tail. Because there is zero phase difference between the wavelets, there is zero angle between each pair of adjacent phasors. The amplitude E_θ of the net wave at P_0 is the vector sum of these phasors. This arrangement of the phasors turns out to be the one that gives the greatest value for the amplitude E_θ. We call this value E_m; that is, E_m is the value of E_θ for $\theta = 0$.

We next consider a point P that is at a small angle θ to the central axis. Equation 36-4 now tells us that the phase difference $\Delta\phi$ between wavelets from adjacent zones is no longer zero. Figure 36-7*b* shows the corresponding phasor diagram; as before, the phasors are arranged head to tail, but now there is an angle $\Delta\phi$ between adjacent phasors. The amplitude E_θ at this new point is still the vector sum of the phasors, but it is smaller than that in Fig. 36-7*a*, which means that the intensity of the light is less at this new point P than at P_0.

If we continue to increase θ, the angle $\Delta\phi$ between adjacent phasors increases, and eventually the chain of phasors curls completely around so that the head of the last phasor just reaches the tail of the first phasor (Fig. 36-7*c*). The amplitude E_θ is now zero, which means that the intensity of the light is also zero. We have reached the first minimum, or dark fringe, in the diffraction pattern. The first and last phasors now have a phase difference of 2π rad, which means that the path length difference between the top and bottom rays through the slit equals one wavelength. Recall that this is the condition we determined for the first diffraction minimum.

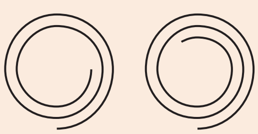
As we continue to increase θ, the angle $\Delta\phi$ between adjacent phasors continues to increase, the chain of phasors begins to wrap back on itself, and the resulting coil begins to shrink. Amplitude E_θ now increases until it reaches a maximum value in the arrangement shown in Fig. 36-7*d*. This arrangement corresponds to the first side maximum in the diffraction pattern.

If we increase θ a bit more, the resulting shrinkage of the coil decreases E_θ, which means that the intensity also decreases. When θ is increased enough, the head of the last phasor again meets the tail of the first phasor. We have then reached the second minimum.

We could continue this qualitative method of determining the maxima and minima of the diffraction pattern but, instead, we shall now turn to a quantitative method.

CHECKPOINT 2

The figures represent, in smoother form (with more phasors) than Fig. 36-7, the phasor diagrams for two points of a diffraction pattern that are on opposite sides of a certain diffraction maximum. (a) Which maximum is it? (b) What is the approximate value of m (in Eq. 36-3) that corresponds to this maximum?



(*a*) (*b*)

36-5 Intensity in Single-Slit Diffraction, Quantitatively

Equation 36-3 tells us how to locate the minima of the single-slit diffraction pattern on screen C of Fig. 36-4 as a function of the angle θ in that figure. Here we wish to derive an expression for the intensity $I(\theta)$ of the pattern as a function of θ. We state, and shall prove below, that the intensity is given by

$$I(\theta) = I_m\left(\frac{\sin\alpha}{\alpha}\right)^2, \tag{36-5}$$

where

$$\alpha = \tfrac{1}{2}\phi = \frac{\pi a}{\lambda}\sin\theta. \tag{36-6}$$

The symbol α is just a convenient connection between the angle θ that locates a point on the viewing screen and the light intensity $I(\theta)$ at that point. The intensity I_m is the greatest value of the intensities $I(\theta)$ in the pattern and occurs at the central maximum (where $\theta = 0$), and ϕ is the phase difference (in radians) between the top and bottom rays from the slit of width a.

Study of Eq. 36-5 shows that intensity minima will occur where

$$\alpha = m\pi, \quad \text{for } m = 1, 2, 3, \ldots. \tag{36-7}$$

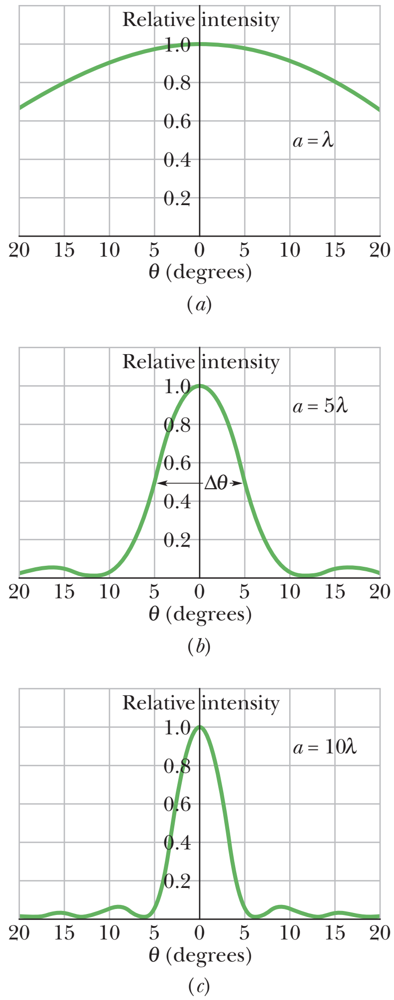


Fig. 36-8 The relative intensity in single-slit diffraction for three values of the ratio a/λ. The wider the slit is, the narrower is the central diffraction maximum.

If we put this result into Eq. 36-6, we find

$$m\pi = \frac{\pi a}{\lambda}\sin\theta, \qquad \text{for } m = 1, 2, 3, \ldots,$$

or

$$a\sin\theta = m\lambda, \qquad \text{for } m = 1, 2, 3, \ldots \quad \text{(minima—dark fringes)}, \tag{36-8}$$

which is exactly Eq. 36-3, the expression that we derived earlier for the location of the minima.

Figure 36-8 shows plots of the intensity of a single-slit diffraction pattern, calculated with Eqs. 36-5 and 36-6 for three slit widths: $a = \lambda$, $a = 5\lambda$, and $a = 10\lambda$. Note that as the slit width increases (relative to the wavelength), the width of the *central diffraction maximum* (the central hill-like region of the graphs) decreases; that is, the light undergoes less flaring by the slit. The secondary maxima also decrease in width (and become weaker). In the limit of slit width a being much greater than wavelength λ, the secondary maxima due to the slit disappear; we then no longer have single-slit diffraction (but we still have diffraction due to the edges of the wide slit, like that produced by the edges of the razor blade in Fig. 36-2).

Proof of Eqs. 36-5 and 36-6

To find an expression for the intensity at a point in the diffraction pattern, we need to divide the slit into many zones and then add the phasors corresponding to those zones, as we did in Fig. 36-7. The arc of phasors in Fig. 36-9 represents the wavelets that reach an arbitrary point P on the viewing screen of Fig. 36-4, corresponding to a particular small angle θ. The amplitude E_θ of the resultant wave at P is the vector sum of these phasors. If we divide the slit of Fig. 36-4 into infinitesimal zones of width Δx, the arc of phasors in Fig. 36-9 approaches the arc of a circle; we call its radius R as indicated in that figure. The length of the arc must be E_m, the amplitude at the center of the diffraction pattern, because if we straightened out the arc we would have the phasor arrangement of Fig. 36-7*a* (shown lightly in Fig. 36-9).

The angle ϕ in the lower part of Fig. 36-9 is the difference in phase between the infinitesimal vectors at the left and right ends of arc E_m. From the geometry, ϕ is also the angle between the two radii marked R in Fig. 36-9. The dashed line in that figure, which bisects ϕ, then forms two congruent right triangles. From either triangle we can write

$$\sin\tfrac{1}{2}\phi = \frac{E_\theta}{2R}. \tag{36-9}$$

In radian measure, ϕ is (with E_m considered to be a circular arc)

$$\phi = \frac{E_m}{R}.$$

Solving this equation for R and substituting in Eq. 36-9 lead to

$$E_\theta = \frac{E_m}{\frac{1}{2}\phi}\sin\tfrac{1}{2}\phi. \tag{36-10}$$

In Section 33-5 we saw that the intensity of an electromagnetic wave is proportional to the square of the amplitude of its electric field. Here, this means that the maximum intensity I_m (which occurs at the center of the diffraction pattern) is proportional to E_m^2 and the intensity $I(\theta)$ at angle θ is proportional to E_θ^2. Thus, we may write

$$\frac{I(\theta)}{I_m} = \frac{E_\theta^2}{E_m^2}. \tag{36-11}$$

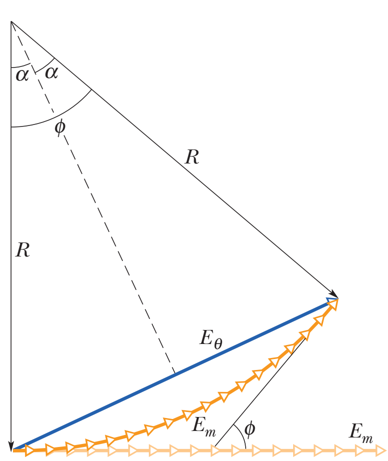


Fig. 36-9 A construction used to calculate the intensity in single-slit diffraction. The situation shown corresponds to that of Fig. 36-7*b*.

Substituting for E_θ with Eq. 36-10 and then substituting $\alpha = \frac{1}{2}\phi$, we are led to the following expression for the intensity as a function of θ:

$$I(\theta) = I_m\left(\frac{\sin\alpha}{\alpha}\right)^2.$$

This is exactly Eq. 36-5, one of the two equations we set out to prove.

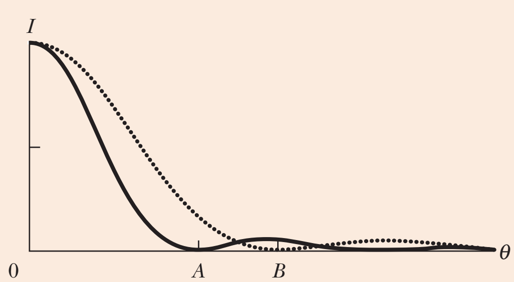
The second equation we wish to prove relates α to θ. The phase difference ϕ between the rays from the top and bottom of the entire slit may be related to a path length difference with Eq. 36-4; it tells us that

$$\phi = \left(\frac{2\pi}{\lambda}\right)(a\sin\theta),$$

where a is the sum of the widths Δx of the infinitesimal zones. However, $\phi = 2\alpha$, so this equation reduces to Eq. 36-6.

CHECKPOINT 3

Two wavelengths, 650 and 430 nm, are used separately in a single-slit diffraction experiment. The figure shows the results as graphs of intensity I versus angle θ for the two diffraction patterns. If both wavelengths are then used simultaneously, what color will be seen in the combined diffraction pattern at (a) angle A and (b) angle B?



Sample Problem

Intensities of the maxima in a single-slit interference pattern

Find the intensities of the first three secondary maxima (side maxima) in the single-slit diffraction pattern of Fig. 36-1, measured as a percentage of the intensity of the central maximum.

KEY IDEAS

The secondary maxima lie approximately halfway between the minima, whose angular locations are given by Eq. 36-7 ($\alpha = m\pi$). The locations of the secondary maxima are then given (approximately) by

$$a = (m + \tfrac{1}{2})\pi, \qquad \text{for } m = 1, 2, 3, \ldots,$$

with α in radian measure. We can relate the intensity I at any point in the diffraction pattern to the intensity I_m of the central maximum via Eq. 36-5.

Calculations: Substituting the approximate values of α for the secondary maxima into Eq. 36-5 to obtain the relative intensities at those maxima, we get

$$\frac{I}{I_m} = \left(\frac{\sin\alpha}{\alpha}\right)^2 = \left(\frac{\sin(m+\frac{1}{2})\pi}{(m+\frac{1}{2})\pi}\right)^2, \quad \text{for } m = 1, 2, 3, \ldots.$$

The first of the secondary maxima occurs for $m = 1$, and its relative intensity is

$$\frac{I_1}{I_m} = \left(\frac{\sin(1+\frac{1}{2})\pi}{(1+\frac{1}{2})\pi}\right)^2 = \left(\frac{\sin 1.5\pi}{1.5\pi}\right)^2$$

$$= 4.50 \times 10^{-2} \approx 4.5\%. \qquad \text{(Answer)}$$

For $m = 2$ and $m = 3$ we find that

$$\frac{I_2}{I_m} = 1.6\% \quad \text{and} \quad \frac{I_3}{I_m} = 0.83\%. \qquad \text{(Answer)}$$

As you can see from these results, successive secondary maxima decrease rapidly in intensity. Figure 36-1 was deliberately overexposed to reveal them.

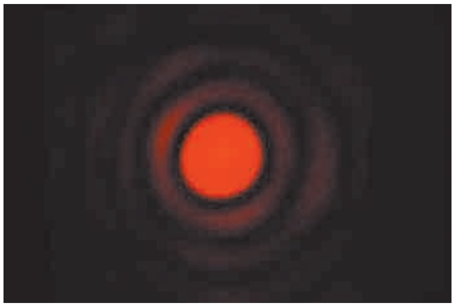


Fig. 36-10 The diffraction pattern of a circular aperture. Note the central maximum and the circular secondary maxima. The figure has been overexposed to bring out these secondary maxima, which are much less intense than the central maximum. *(Jearl Walker)*

36-6 Diffraction by a Circular Aperture

Here we consider diffraction by a circular aperture—that is, a circular opening, such as a circular lens, through which light can pass. Figure 36-10 shows the image formed by light from a laser that was directed onto a circular aperture with a very small diameter. This image is not a point, as geometrical optics would suggest, but a circular disk surrounded by several progressively fainter secondary rings. Comparison with Fig. 36-1 leaves little doubt that we are dealing with a diffraction phenomenon. Here, however, the aperture is a circle of diameter d rather than a rectangular slit.

The (complex) analysis of such patterns shows that the first minimum for the diffraction pattern of a circular aperture of diameter d is located by

$$\sin\theta = 1.22\frac{\lambda}{d} \quad \text{(first minimum—circular aperture).} \tag{36-12}$$

The angle θ here is the angle from the central axis to any point on that (circular) minimum. Compare this with Eq. 36-1,

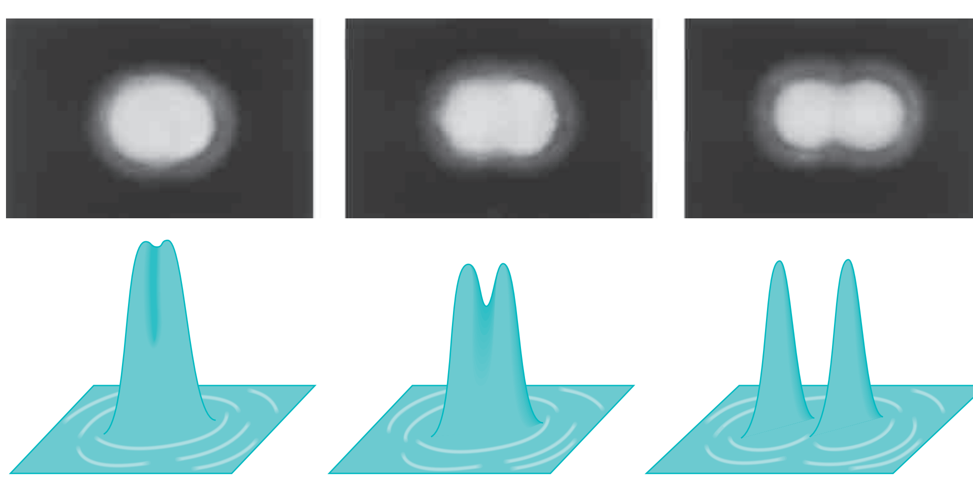
$$\sin\theta = \frac{\lambda}{a} \quad \text{(first minimum—single slit),} \tag{36-13}$$

which locates the first minimum for a long narrow slit of width a. The main difference is the factor 1.22, which enters because of the circular shape of the aperture.

Resolvability

The fact that lens images are diffraction patterns is important when we wish to *resolve* (distinguish) two distant point objects whose angular separation is small. Figure 36-11 shows, in three different cases, the visual appearance and corresponding intensity pattern for two distant point objects (stars, say) with small angular separation. In Figure 36-11*a*, the objects are not resolved because of diffraction; that is, their diffraction patterns (mainly their central maxima) overlap so much that the two objects cannot be distinguished from a single point object. In Fig. 36-11*b* the objects are barely resolved, and in Fig. 36-11*c* they are fully resolved.

In Fig. 36-11*b* the angular separation of the two point sources is such that the central maximum of the diffraction pattern of one source is centered on the first minimum of the diffraction pattern of the other, a condition called **Rayleigh's criterion** for resolvability. From Eq. 36-12, two objects that are barely resolvable



(*a*) (*b*) (*c*)

Fig. 36-11 At the top, the images of two point sources (stars) formed by a converging lens. At the bottom, representations of the image intensities. In (*a*) the angular separation of the sources is too small for them to be distinguished, in (*b*) they can be marginally distinguished, and in (*c*) they are clearly distinguished. Rayleigh's criterion is satisfied in (*b*), with the central maximum of one diffraction pattern coinciding with the first minimum of the other.

by this criterion must have an angular separation θ_R of

$$\theta_R = \sin^{-1}\frac{1.22\lambda}{d}.$$

Since the angles are small, we can replace $\sin\theta_R$ with θ_R expressed in radians:

$$\theta_R = 1.22\frac{\lambda}{d} \quad \text{(Rayleigh's criterion).} \tag{36-14}$$

Applying Rayleigh's criterion for resolvability to human vision is only an approximation because visual resolvability depends on many factors, such as the relative brightness of the sources and their surroundings, turbulence in the air between the sources and the observer, and the functioning of the observer's visual system. Experimental results show that the least angular separation that can actually be resolved by a person is generally somewhat greater than the value given by Eq. 36-14. However, for calculations here, we shall take Eq. 36-14 as being a precise criterion: If the angular separation θ between the sources is greater than θ_R, we can visually resolve the sources; if it is less, we cannot.

Fig. 36-12 The pointillistic painting *The Seine at Herblay* by Maximilien Luce consists of thousands of colored dots. With the viewer very close to the canvas, the dots and their true colors are visible. At normal viewing distances, the dots are irresolvable and thus blend. *(Maximilien Luce, The Seine at Herblay, 1890. Musee d'Orsay, Paris, France. Photo by Erich Lessing/Art Resource)*

Rayleigh's criterion can explain the arresting illusions of color in the style of painting known as pointillism (Fig. 36-12). In this style, a painting is made not with brush strokes in the usual sense but rather with a myriad of small colored dots. One fascinating aspect of a pointillistic painting is that when you change your distance from it, the colors shift in subtle, almost subconscious ways. This color shifting has to do with whether you can resolve the colored dots. When you stand close enough to the painting, the angular separations θ of adjacent dots are greater than θ_R and thus the dots can be seen individually. Their colors are the true colors of the paints used. However, when you stand far enough from the painting, the angular separations θ are less than θ_R and the dots cannot be seen individually. The resulting blend of colors coming into your eye from any group of dots can then cause your brain to "make up" a color for that group—a color that may not actually exist in the group. In this way, a pointillistic painter uses your visual system to create the colors of the art.

When we wish to use a lens instead of our visual system to resolve objects of small angular separation, it is desirable to make the diffraction pattern as small as possible. According to Eq. 36-14, this can be done either by increasing the lens diameter or by using light of a shorter wavelength. For this reason ultraviolet light is often used with microscopes because its wavelength is shorter than a visible light wavelength.

CHECKPOINT 4

Suppose that you can barely resolve two red dots because of diffraction by the pupil of your eye. If we increase the general illumination around you so that the pupil decreases in diameter, does the resolvability of the dots improve or diminish? Consider only diffraction. (You might experiment to check your answer.)

Sample Problem

Pointillistic paintings use the diffraction of your eye

Figure 36-13*a* is a representation of the colored dots on a pointillistic painting. Assume that the average center-to-center separation of the dots is $D = 2.0$ mm. Also assume that the diameter of the pupil of your eye is $d = 1.5$ mm and that the least angular separation between dots you can resolve is set only by Rayleigh's criterion. What is the least viewing distance from which you cannot distinguish any dots on the painting?

KEY IDEA

Consider any two adjacent dots that you can distinguish when you are close to the painting. As you move away, you continue to distinguish the dots until their angular separation θ (in your view) has decreased to the angle given by

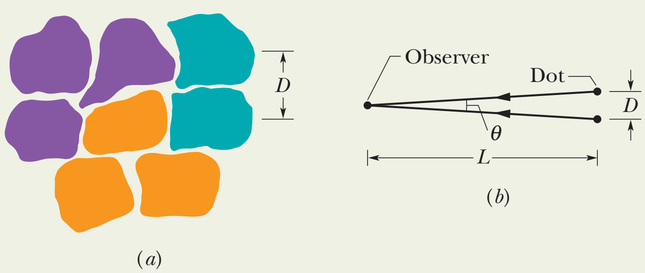


Fig. 36-13 (*a*) Representation of some dots on a pointillistic painting, showing an average center-to-center separation D. (*b*) The arrangement of separation D between two dots, their angular separation θ, and the viewing distance L.

Rayleigh's criterion:

$$\theta_R = 1.22\frac{\lambda}{d}. \tag{36-15}$$

Calculations: Figure 36-13*b* shows, from the side, the angular separation θ of the dots, their center-to-center separation D, and your distance L from them. Because D/L is small, angle θ is also small and we can make the approximation

$$\theta = \frac{D}{L}. \tag{36-16}$$

Setting θ of Eq. 36-16 equal to θ_R of Eq. 36-15 and solving for L, we then have

$$L = \frac{Dd}{1.22\lambda}. \tag{36-17}$$

Equation 36-17 tells us that L is larger for smaller λ. Thus, as you move away from the painting, adjacent red dots (long wavelengths) become indistinguishable before adjacent blue dots do. To find the least distance L at which *no* colored dots are distinguishable, we substitute $\lambda = 400$ nm (blue or violet light) into Eq. 36-17:

$$L = \frac{(2.0\times10^{-3}\text{ m})(1.5\times10^{-3}\text{ m})}{(1.22)(400\times10^{-9}\text{ m})} = 6.1\text{ m}. \quad \text{(Answer)}$$

At this or a greater distance, the color you perceive at any given spot on the painting is a blended color that may not actually exist there.

Sample Problem

Rayleigh's criterion for resolving two distant objects

A circular converging lens, with diameter $d = 32$ mm and focal length $f = 24$ cm, forms images of distant point objects in the focal plane of the lens. The wavelength is $\lambda = 550$ nm.

(a) Considering diffraction by the lens, what angular separation must two distant point objects have to satisfy Rayleigh's criterion?

KEY IDEA

Figure 36-14 shows two distant point objects P_1 and P_2, the lens, and a viewing screen in the focal plane of the lens. It also shows, on the right, plots of light intensity I versus position on the screen for the central maxima of the images formed by the lens. Note that the angular separation θ_o of the objects equals the angular separation θ_i of the images. Thus, if the images are to satisfy Rayleigh's criterion

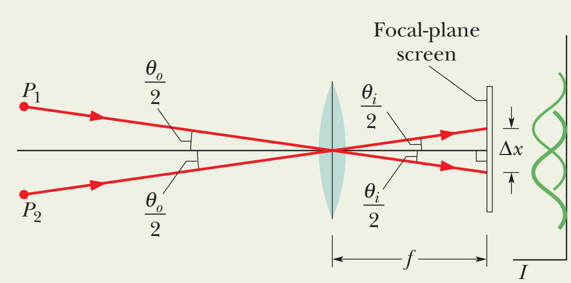


Fig. 36-14 Light from two distant point objects P_1 and P_2 passes through a converging lens and forms images on a viewing screen in the focal plane of the lens. Only one representative ray from each object is shown. The images are not points but diffraction patterns, with intensities approximately as plotted at the right. The angular separation of the objects is θ_o and that of the images is θ_i; the central maxima of the images have a separation Δx.

for resolvability, the angular separations on both sides of the lens must be given by Eq. 36-14 (assuming small angles).

Calculations: From Eq. 36-14, we obtain

$$\theta_o = \theta_i = \theta_R = 1.22\frac{\lambda}{d}$$

$$= \frac{(1.22)(550 \times 10^{-9}\text{ m})}{32 \times 10^{-3}\text{ m}} = 2.1 \times 10^{-5}\text{ rad. (Answer)}$$

At this angular separation, each central maximum in the two intensity curves of Fig. 36-14 is centered on the first minimum of the other curve.

(b) What is the separation Δx of the centers of the *images* in the focal plane? (That is, what is the separation of the *central* peaks in the two intensity-versus-position curves?)

Calculations: From either triangle between the lens and the screen in Fig. 36-14, we see that tan $\theta_i/2 = \Delta x/2f$. Rearranging this equation and making the approximation tan $\theta \approx \theta$, we find

$$\Delta x = f\theta_i, \tag{36-18}$$

where θ_i is in radian measure. Substituting known data then yields

$$\Delta x = (0.24\text{ m})(2.1 \times 10^{-5}\text{ rad}) = 5.0\ \mu\text{m}. \qquad \text{(Answer)}$$

WILEY PLUS Additional examples, video, and practice available at *WileyPLUS*

36-7 Diffraction by a Double Slit

In the double-slit experiments of Chapter 35, we implicitly assumed that the slits were much narrower than the wavelength of the light illuminating them; that is, $a \ll \lambda$. For such narrow slits, the central maximum of the diffraction pattern of either slit covers the entire viewing screen. Moreover, the interference of light from the two slits produces bright fringes with approximately the same intensity (Fig. 35-12).

In practice with visible light, however, the condition $a \ll \lambda$ is often not met. For relatively wide slits, the interference of light from two slits produces bright fringes that do not all have the same intensity. That is, the intensities of the fringes produced by double-slit interference (as discussed in Chapter 35) are modified by diffraction of the light passing through each slit (as discussed in this chapter).

As an example, the intensity plot of Fig. 36-15*a* suggests the double-slit interference pattern that would occur if the slits were infinitely narrow (and thus $a \ll \lambda$); all the bright interference fringes would have the same intensity. The intensity plot of Fig. 36-15*b* is that for diffraction by a single actual slit; the diffraction pattern has a broad central maximum and weaker secondary maxima at $\pm 17°$. The plot of Fig. 36-15*c* suggests the interference pattern for two actual slits. That plot was constructed by using the curve of Fig. 36-15*b* as an *envelope* on the intensity plot in Fig. 36-15*a*. The positions of the fringes are not changed; only the intensities are affected.

Figure 36-16*a* shows an actual pattern in which both double-slit interference and diffraction are evident. If one slit is covered, the single-slit diffraction pattern of Fig. 36-16*b* results. Note the correspondence between Figs. 36-16*a* and 36-15*c*, and between Figs. 36-16*b* and 36-15*b*. In comparing these figures, bear in mind that Fig. 36-16 has been deliberately overexposed to bring

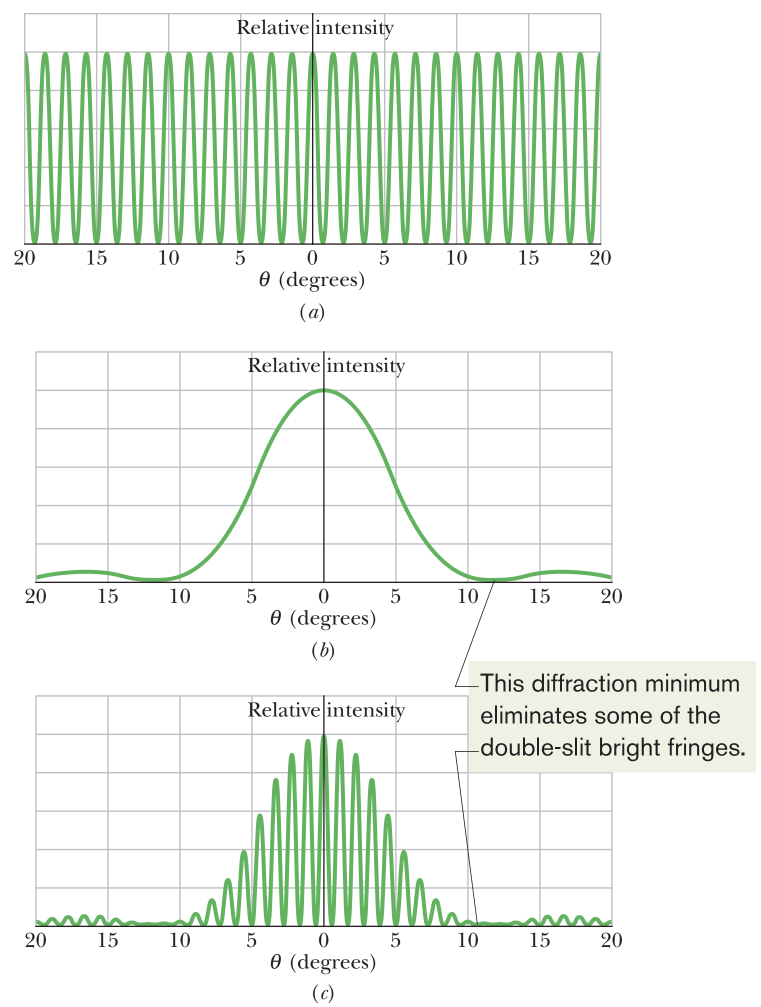


Fig. 36-15 (*a*) The intensity plot to be expected in a double-slit interference experiment with vanishingly narrow slits. (*b*) The intensity plot for diffraction by a typical slit of width a (not vanishingly narrow). (*c*) The intensity plot to be expected for two slits of width a. The curve of (*b*) acts as an envelope, limiting the intensity of the double-slit fringes in (*a*). Note that the first minima of the diffraction pattern of (*b*) eliminate the double-slit fringes that would occur near 12° in (*c*).

(*a*)

(*b*)

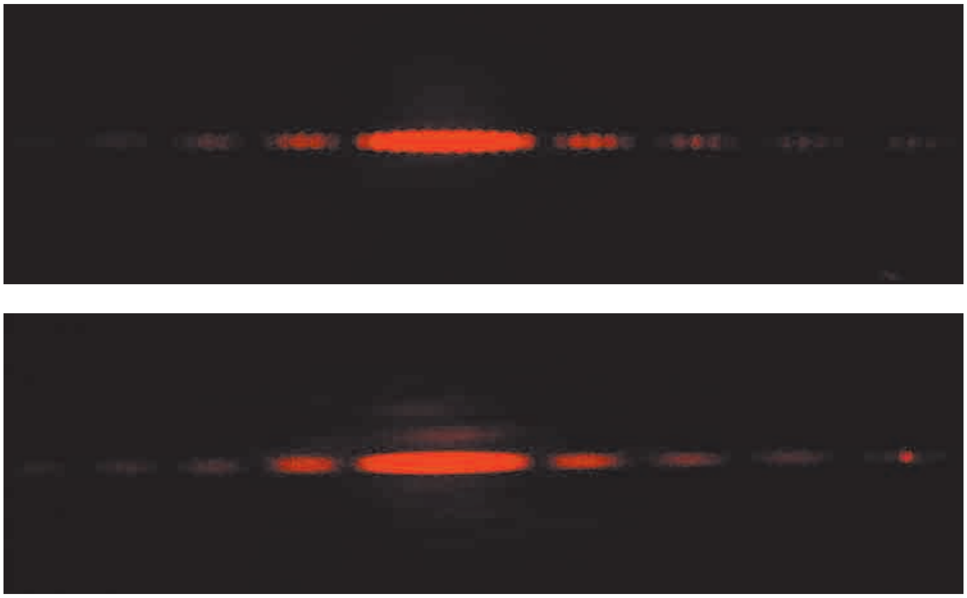


Fig. 36-16 (*a*) Interference fringes for an actual double-slit system; compare with Fig. 36-15*c*. (*b*) The diffraction pattern of a single slit; compare with Fig. 36-15*b*. *(Jearl Walker)*

out the faint secondary maxima and that several secondary maxima (rather than one) are shown.

With diffraction effects taken into account, the intensity of a double-slit interference pattern is given by

$$I(\theta) = I_m(\cos^2 \beta)\left(\frac{\sin \alpha}{\alpha}\right)^2 \qquad \text{(double slit)}, \qquad (36\text{-}19)$$

in which

$$\beta = \frac{\pi d}{\lambda} \sin \theta \qquad (36\text{-}20)$$

and

$$\alpha = \frac{\pi a}{\lambda} \sin \theta. \qquad (36\text{-}21)$$

Here d is the distance between the centers of the slits and a is the slit width. Note carefully that the right side of Eq. 36-19 is the product of I_m and two factors. (1) The *interference factor* $\cos^2 \beta$ is due to the interference between two slits with slit separation d (as given by Eqs. 35-22 and 35-23). (2) The *diffraction factor* $[(\sin \alpha)/\alpha]^2$ is due to diffraction by a single slit of width a (as given by Eqs. 36-5 and 36-6).

Let us check these factors. If we let $a \to 0$ in Eq. 36-21, for example, then $\alpha \to 0$ and $(\sin \alpha)/\alpha \to 1$. Equation 36-19 then reduces, as it must, to an equation describing the interference pattern for a pair of vanishingly narrow slits with slit separation d. Similarly, putting $d = 0$ in Eq. 36-20 is equivalent physically to causing the two slits to merge into a single slit of width a. Then Eq. 36-20 yields $\beta = 0$ and $\cos^2 \beta = 1$. In this case Eq. 36-19 reduces, as it must, to an equation describing the diffraction pattern for a single slit of width a.

The double-slit pattern described by Eq. 36-19 and displayed in Fig. 36-16*a* combines interference and diffraction in an intimate way. Both are superposition effects, in that they result from the combining of waves with different phases at a given point. If the combining waves originate from a small number of elementary coherent sources—as in a double-slit experiment with $a \ll \lambda$—we call the process *interference.* If the combining waves originate in a single wavefront—as in a single-slit experiment—we call the process *diffraction.* This distinction between interference and diffraction (which is somewhat arbitrary and not always adhered to) is a convenient one, but we should not forget that both are superposition effects and usually both are present simultaneously (as in Fig. 36-16*a*).

Sample Problem

Double-slit experiment with diffraction of each slit included

In a double-slit experiment, the wavelength λ of the light source is 405 nm, the slit separation d is 19.44 μm, and the slit width a is 4.050 μm. Consider the interference of the light from the two slits and also the diffraction of the light through each slit.

(a) How many bright interference fringes are within the central peak of the diffraction envelope?

KEY IDEAS

We first analyze the two basic mechanisms responsible for the optical pattern produced in the experiment:

1. *Single-slit diffraction:* The limits of the central peak are the first minima in the diffraction pattern due to either slit individually. (See Fig. 36-15.) The angular locations of those minima are given by Eq. 36-3 ($a \sin \theta = m\lambda$). Here let us rewrite this equation as $a \sin \theta = m_1\lambda$, with the subscript 1 referring to the one-slit diffraction. For the first minima in the diffraction pattern, we substitute $m_1 = 1$, obtaining
$$a \sin \theta = \lambda. \qquad (36\text{-}22)$$
2. *Double-slit interference:* The angular locations of the bright fringes of the double-slit interference pattern are given by Eq. 35-14, which we can write as
$$d \sin \theta = m_2\lambda, \qquad \text{for } m_2 = 0, 1, 2, \ldots . \qquad (36\text{-}23)$$
Here the subscript 2 refers to the double-slit interference.

Calculations: We can locate the first diffraction minimum within the double-slit fringe pattern by dividing Eq. 36-23 by Eq. 36-22 and solving for m_2. By doing so and then substituting the given data, we obtain

$$m_2 = \frac{d}{a} = \frac{19.44\ \mu\text{m}}{4.050\ \mu\text{m}} = 4.8.$$

This tells us that the bright interference fringe for $m_2 = 4$ fits into the central peak of the one-slit diffraction pattern, but the fringe for $m_2 = 5$ does not fit. Within the central diffraction peak we have the central bright fringe ($m_2 = 0$), and four bright fringes (up to $m_2 = 4$) on each side of it. Thus, a total of nine bright fringes of the double-slit interference pattern are within the central peak of the diffraction envelope.

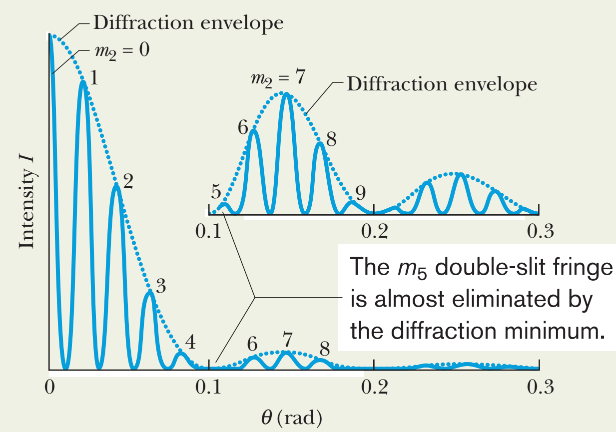


Fig. 36-17 One side of the intensity plot for a two-slit interference experiment. The inset shows (vertically expanded) the plot within the first and second side peaks of the diffraction envelope.

The bright fringes to one side of the central bright fringe are shown in Fig. 36-17.

(b) How many bright fringes are within either of the first side peaks of the diffraction envelope?

KEY IDEA

The outer limits of the first side diffraction peaks are the second diffraction minima, each of which is at the angle θ given by $a \sin \theta = m_1\lambda$ with $m_1 = 2$:

$$a \sin \theta = 2\lambda. \qquad (36\text{-}24)$$

Calculation: Dividing Eq. 36-23 by Eq. 36-24, we find

$$m_2 = \frac{2d}{a} = \frac{(2)(19.44\ \mu\text{m})}{4.050\ \mu\text{m}} = 9.6.$$

This tells us that the second diffraction minimum occurs just before the bright interference fringe for $m_2 = 10$ in Eq. 36-23. Within either first side diffraction peak we have the fringes from $m_2 = 5$ to $m_2 = 9$, for a total of five bright fringes of the double-slit interference pattern (shown in the inset of Fig. 36-17). However, if the $m_2 = 5$ bright fringe, which is almost eliminated by the first diffraction minimum, is considered too dim to count, then only four bright fringes are in the first side diffraction peak.

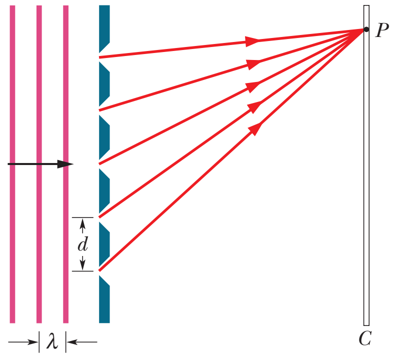


Fig. 36-18 An idealized diffraction grating, consisting of only five rulings, that produces an interference pattern on a distant viewing screen *C*.

36-8 Diffraction Gratings

One of the most useful tools in the study of light and of objects that emit and absorb light is the **diffraction grating.** This device is somewhat like the double-slit arrangement of Fig. 35-10 but has a much greater number N of slits, often called *rulings,* perhaps as many as several thousand per millimeter. An idealized grating consisting of only five slits is represented in Fig. 36-18. When monochromatic light is sent through the slits, it forms narrow interference fringes that can be analyzed to determine the wavelength of the light. (Diffraction gratings can also be opaque surfaces with narrow parallel grooves arranged like the slits in Fig. 36-18. Light then scatters back from the grooves to form interference fringes rather than being transmitted through open slits.)

With monochromatic light incident on a diffraction grating, if we gradually increase the number of slits from two to a large number N, the intensity plot changes from the typical double-slit plot of Fig. 36-15*c* to a much more complicated one and then eventually to a simple graph like that shown in Fig. 36-19*a*. The pattern you would see on a viewing screen using monochromatic red light from, say, a helium–neon laser is shown in Fig. 36-19*b*. The maxima are now very narrow (and so are called *lines*); they are separated by relatively wide dark regions.

We use a familiar procedure to find the locations of the bright lines on the viewing screen. We first assume that the screen is far enough from the grating so that the rays reaching a particular point P on the screen are approximately parallel when they leave the grating (Fig. 36-20). Then we apply to each pair of adjacent rulings the same reasoning we used for double-slit interference. The separation d between rulings is called the *grating spacing.* (If N rulings occupy a total width w, then $d = w/N$.) The path length difference between adjacent rays is again $d \sin\theta$ (Fig. 36-20), where θ is the angle from the central axis of the grating (and of the diffraction pattern) to point P. A line will be located at P if the path length difference between adjacent rays is an integer number of wavelengths—that is, if

$$d \sin\theta = m\lambda, \quad \text{for } m = 0, 1, 2, \ldots \quad \text{(maxima—lines)}, \tag{36-25}$$

where λ is the wavelength of the light. Each integer m represents a different line; hence these integers can be used to label the lines, as in Fig. 36-19. The integers are then called the *order numbers,* and the lines are called the zeroth-order line (the central line, with $m = 0$), the first-order line ($m = 1$), the second-order line ($m = 2$), and so on.

If we rewrite Eq. 36-25 as $\theta = \sin^{-1}(m\lambda/d)$, we see that, for a given diffraction grating, the angle from the central axis to any line (say, the third-order line) depends on the wavelength of the light being used. Thus, when light of an unknown wavelength is sent through a diffraction grating, measurements of the angles to the higher-order lines can be used in Eq. 36-25 to determine the wavelength. Even light of several unknown wavelengths can be distinguished and identified in this way. We cannot do that with the double-slit arrangement of Section 35-4, even though the same equation and wavelength dependence apply there. In double-slit interference, the bright fringes due to different wavelengths overlap too much to be distinguished.

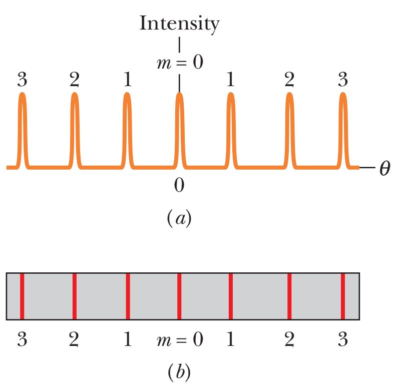


Fig. 36-19 (*a*) The intensity plot produced by a diffraction grating with a great many rulings consists of narrow peaks, here labeled with their order numbers m. (*b*) The corresponding bright fringes seen on the screen are called lines and are here also labeled with order numbers m.

Width of the Lines

A grating's ability to resolve (separate) lines of different wavelengths depends on the width of the lines. We shall here derive an expression for the *half-width* of the central line (the line for which $m = 0$) and then state an expression for the half-widths of the higher-order lines. We define the **half-width** of the central line as being the angle $\Delta\theta_{hw}$ from the center of the line at $\theta = 0$ outward to where the line effectively ends and darkness effectively begins with the first minimum

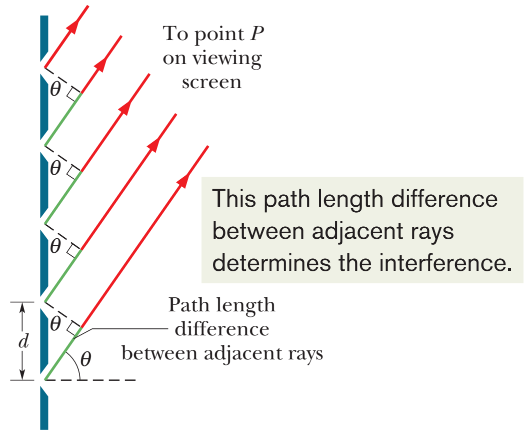


Fig. 36-20 The rays from the rulings in a diffraction grating to a distant point P are approximately parallel. The path length difference between each two adjacent rays is $d \sin \theta$, where θ is measured as shown. (The rulings extend into and out of the page.)

(Fig. 36-21). At such a minimum, the N rays from the N slits of the grating cancel one another. (The actual width of the central line is, of course, $2(\Delta\theta_{hw})$, but line widths are usually compared via half-widths.)

In Section 36-3 we were also concerned with the cancellation of a great many rays, there due to diffraction through a single slit. We obtained Eq. 36-3, which, because of the similarity of the two situations, we can use to find the first minimum here. It tells us that the first minimum occurs where the path length difference between the top and bottom rays equals λ. For single-slit diffraction, this difference is $a \sin \theta$. For a grating of N rulings, each separated from the next by distance d, the distance between the top and bottom rulings is Nd (Fig. 36-22), and so the path length difference between the top and bottom rays here is $Nd \sin \Delta\theta_{hw}$. Thus, the first minimum occurs where

$$Nd \sin \Delta\theta_{hw} = \lambda. \tag{36-26}$$

Because $\Delta\theta_{hw}$ is small, $\sin \Delta\theta_{hw} = \Delta\theta_{hw}$ (in radian measure). Substituting this in Eq. 36-26 gives the half-width of the central line as

$$\Delta\theta_{hw} = \frac{\lambda}{Nd} \quad \text{(half-width of central line).} \tag{36-27}$$

We state without proof that the half-width of any other line depends on its location relative to the central axis and is

$$\Delta\theta_{hw} = \frac{\lambda}{Nd \cos \theta} \quad \text{(half-width of line at } \theta\text{).} \tag{36-28}$$

Note that for light of a given wavelength λ and a given ruling separation d, the widths of the lines decrease with an increase in the number N of rulings. Thus, of two diffraction gratings, the grating with the larger value of N is better able to distinguish between wavelengths because its diffraction lines are narrower and so produce less overlap.

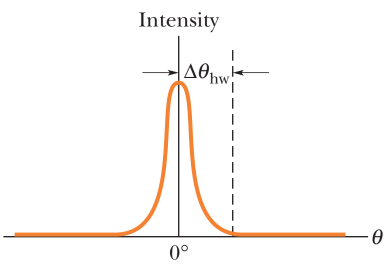


Fig. 36-21 The half-width $\Delta\theta_{hw}$ of the central line is measured from the center of that line to the adjacent minimum on a plot of I versus θ like Fig. 36-19*a*.

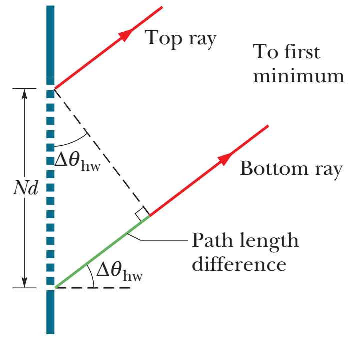


Fig. 36-22 The top and bottom rulings of a diffraction grating of N rulings are separated by Nd. The top and bottom rays passing through these rulings have a path length difference of $Nd \sin \Delta\theta_{hw}$, where $\Delta\theta_{hw}$ is the angle to the first minimum. (The angle is here greatly exaggerated for clarity.)

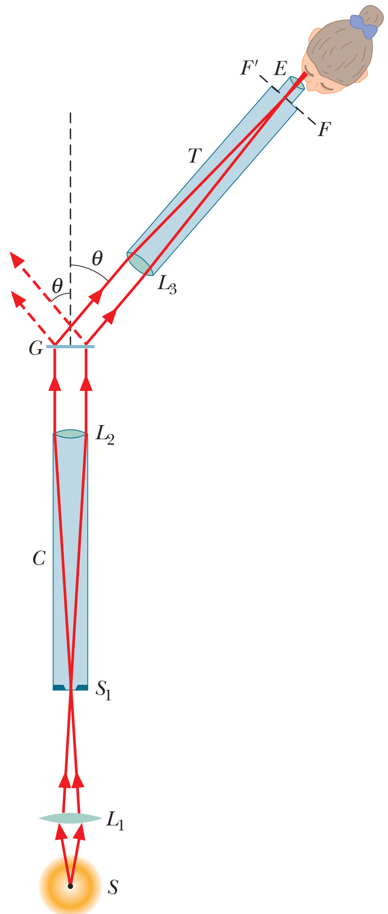


Fig. 36-23 A simple type of grating spectroscope used to analyze the wavelengths of light emitted by source S.

Grating Spectroscope

Diffraction gratings are widely used to determine the wavelengths that are emitted by sources of light ranging from lamps to stars. Figure 36-23 shows a simple *grating spectroscope* in which a grating is used for this purpose. Light from source S is focused by lens L_1 on a vertical slit S_1 placed in the focal plane of lens L_2. The light emerging from tube C (called a *collimator*) is a plane wave and is incident perpendicularly on grating G, where it is diffracted into a diffraction pattern, with the $m = 0$ order diffracted at angle $\theta = 0$ along the central axis of the grating.

We can view the diffraction pattern that would appear on a viewing screen at any angle θ simply by orienting telescope T in Fig. 36-23 to that angle. Lens L_3 of the telescope then focuses the light diffracted at angle θ (and at slightly smaller and larger angles) onto a focal plane FF' within the telescope. When we look through eyepiece E, we see a magnified view of this focused image.

By changing the angle θ of the telescope, we can examine the entire diffraction pattern. For any order number other than $m = 0$, the original light is spread out according to wavelength (or color) so that we can determine, with Eq. 36-25, just what wavelengths are being emitted by the source. If the source emits discrete wavelengths, what we see as we rotate the telescope horizontally through the angles corresponding to an order m is a vertical line of color for each wavelength, with the shorter-wavelength line at a smaller angle θ than the longer-wavelength line.

For example, the light emitted by a hydrogen lamp, which contains hydrogen gas, has four discrete wavelengths in the visible range. If our eyes intercept this light directly, it appears to be white. If, instead, we view it through a grating spectroscope, we can distinguish, in several orders, the lines of the four colors corresponding to these visible wavelengths. (Such lines are called *emission lines.*) Four orders are represented in Fig. 36-24. In the central order ($m = 0$), the lines corresponding to all four wavelengths are superimposed, giving a single white line at $\theta = 0$. The colors are separated in the higher orders.

The third order is not shown in Fig. 36-24 for the sake of clarity; it actually overlaps the second and fourth orders. The fourth-order red line is missing because it is not formed by the grating used here. That is, when we attempt to solve Eq. 36-25 for the angle θ for the red wavelength when $m = 4$, we find that $\sin\theta$ is greater than unity, which is not possible. The fourth order is then said to be *incomplete* for this grating; it might not be incomplete for a grating with greater spacing d, which will spread the lines less than in Fig. 36-24. Figure 36-25 is a photograph of the visible emission lines produced by cadmium.

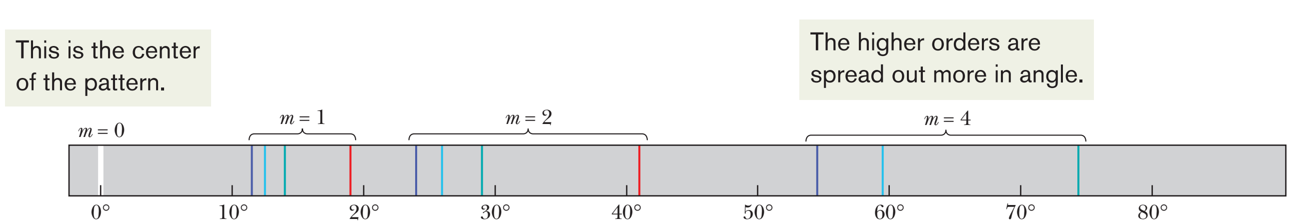
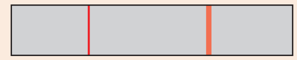


Fig. 36-24 The zeroth, first, second, and fourth orders of the visible emission lines from hydrogen. Note that the lines are farther apart at greater angles. (They are also dimmer and wider, although that is not shown here.)

Fig. 36-25 The visible emission lines of cadmium, as seen through a grating spectroscope. *(Department of Physics, Imperial College/Science Photo Library/Photo Researchers)*

CHECKPOINT 5

The figure shows lines of different orders produced by a diffraction grating in monochromatic red light. (a) Is the center of the pattern to the left or right? (b) In monochromatic green light, are the half-widths of the lines produced in the same orders greater than, less than, or the same as the half-widths of the lines shown?



36-9 Gratings: Dispersion and Resolving Power

Dispersion

To be useful in distinguishing wavelengths that are close to each other (as in a grating spectroscope), a grating must spread apart the diffraction lines associated with the various wavelengths. This spreading, called **dispersion,** is defined as

$$D = \frac{\Delta\theta}{\Delta\lambda} \qquad \text{(dispersion defined).} \qquad (36\text{-}29)$$

Here $\Delta\theta$ is the angular separation of two lines whose wavelengths differ by $\Delta\lambda$. The greater D is, the greater is the distance between two emission lines whose wavelengths differ by $\Delta\lambda$. We show below that the dispersion of a grating at angle θ is given by

$$D = \frac{m}{d\cos\theta} \qquad \text{(dispersion of a grating).} \qquad (36\text{-}30)$$

Thus, to achieve higher dispersion we must use a grating of smaller grating spacing d and work in a higher-order m. Note that the dispersion does not depend on the number of rulings N in the grating. The SI unit for D is the degree per meter or the radian per meter.

The fine rulings, each 0.5 μm wide, on a compact disc function as a diffraction grating. When a small source of white light illuminates a disc, the diffracted light forms colored "lanes" that are the composite of the diffraction patterns from the rulings. *(Kristen Brochmann/Fundamental Photographs)*

Resolving Power

To *resolve* lines whose wavelengths are close together (that is, to make the lines distinguishable), the line should also be as narrow as possible. Expressed otherwise, the grating should have a high **resolving power** R, defined as

$$R = \frac{\lambda_{avg}}{\Delta\lambda} \qquad \text{(resolving power defined).} \qquad (36\text{-}31)$$

Here λ_{avg} is the mean wavelength of two emission lines that can barely be recognized as separate, and $\Delta\lambda$ is the wavelength difference between them. The greater R is, the closer two emission lines can be and still be resolved. We shall show below that the resolving power of a grating is given by the simple expression

$$R = Nm \qquad \text{(resolving power of a grating).} \qquad (36\text{-}32)$$

To achieve high resolving power, we must use many rulings (large N).

Proof of Eq. 36-30

Let us start with Eq. 36-25, the expression for the locations of the lines in the diffraction pattern of a grating:

$$d\sin\theta = m\lambda.$$

Let us regard θ and λ as variables and take differentials of this equation. We find

$$d(\cos\theta)\, d\theta = m\, d\lambda.$$

For small enough angles, we can write these differentials as small differences, obtaining

$$d(\cos\theta)\,\Delta\theta = m\,\Delta\lambda \tag{36-33}$$

or

$$\frac{\Delta\theta}{\Delta\lambda} = \frac{m}{d\cos\theta}.$$

The ratio on the left is simply D (see Eq. 36-29), and so we have indeed derived Eq. 36-30.

Proof of Eq. 36-32

We start with Eq. 36-33, which was derived from Eq. 36-25, the expression for the locations of the lines in the diffraction pattern formed by a grating. Here $\Delta\lambda$ is the small wavelength difference between two waves that are diffracted by the grating, and $\Delta\theta$ is the angular separation between them in the diffraction pattern. If $\Delta\theta$ is to be the smallest angle that will permit the two lines to be resolved, it must (by Rayleigh's criterion) be equal to the half-width of each line, which is given by Eq. 36-28:

$$\Delta\theta_{\text{hw}} = \frac{\lambda}{Nd\cos\theta}.$$

If we substitute $\Delta\theta_{\text{hw}}$ as given here for $\Delta\theta$ in Eq. 36-33, we find that

$$\frac{\lambda}{N} = m\,\Delta\lambda,$$

from which it readily follows that

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

This is Eq. 36-32, which we set out to derive.

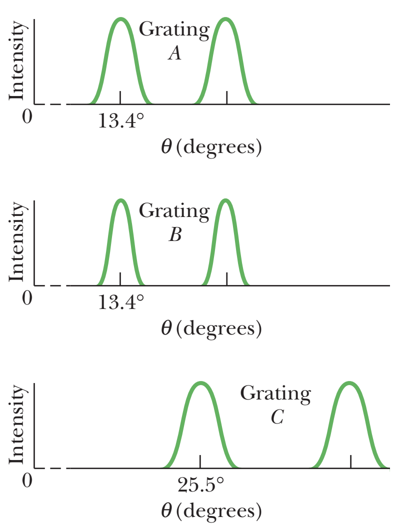


Fig. 36-26 The intensity patterns for light of two wavelengths sent through the gratings of Table 36-1. Grating B has the highest resolving power, and grating C the highest dispersion.

Dispersion and Resolving Power Compared

The resolving power of a grating must not be confused with its dispersion. Table 36-1 shows the characteristics of three gratings, all illuminated with light of wavelength $\lambda = 589$ nm, whose diffracted light is viewed in the first order ($m = 1$ in Eq. 36-25). You should verify that the values of D and R as given in the table can be calculated with Eqs. 36-30 and 36-32, respectively. (In the calculations for D, you will need to convert radians per meter to degrees per micrometer.)

For the conditions noted in Table 36-1, gratings A and B have the same *dispersion* D and A and C have the same *resolving power* R.

Figure 36-26 shows the intensity patterns (also called *line shapes*) that would be produced by these gratings for two lines of wavelengths λ_1 and λ_2, in the vicinity of $\lambda = 589$ nm. Grating B, with the higher resolving power, produces narrower lines and thus is capable of distinguishing lines that are much closer together in wavelength than those in the figure. Grating C, with the higher dispersion, produces the greater angular separation between the lines.

Table 36-1

Three Gratings[a]

Grating	N	d (nm)	θ	D (°/μm)	R
A	10 000	2540	13.4°	23.2	10 000
B	20 000	2540	13.4°	23.2	20 000
C	10 000	1360	25.5°	46.3	10 000

[a]Data are for $\lambda = 589$ nm and $m = 1$.

Sample Problem

Dispersion and resolving power of a diffraction grating

A diffraction grating has 1.26×10^4 rulings uniformly spaced over width $w = 25.4$ mm. It is illuminated at normal incidence by yellow light from a sodium vapor lamp. This light contains two closely spaced emission lines (known as the sodium doublet) of wavelengths 589.00 nm and 589.59 nm.

(a) At what angle does the first-order maximum occur (on either side of the center of the diffraction pattern) for the wavelength of 589.00 nm?

KEY IDEA

The maxima produced by the diffraction grating can be determined with Eq. 36-25 ($d \sin \theta = m\lambda$).

Calculations: The grating spacing d is

$$d = \frac{w}{N} = \frac{25.4 \times 10^{-3}\ \text{m}}{1.26 \times 10^4} = 2.016 \times 10^{-6}\ \text{m} = 2016\ \text{nm}.$$

The first-order maximum corresponds to $m = 1$. Substituting these values for d and m into Eq. 36-25 leads to

$$\theta = \sin^{-1} \frac{m\lambda}{d} = \sin^{-1} \frac{(1)(589.00\ \text{nm})}{2016\ \text{nm}} = 16.99^\circ \approx 17.0^\circ. \qquad \text{(Answer)}$$

(b) Using the dispersion of the grating, calculate the angular separation between the two lines in the first order.

KEY IDEAS

(1) The angular separation $\Delta\theta$ between the two lines in the first order depends on their wavelength difference $\Delta\lambda$ and the dispersion D of the grating, according to Eq. 36-29 ($D = \Delta\theta/\Delta\lambda$). (2) The dispersion D depends on the angle θ at which it is to be evaluated.

Calculations: We can assume that, in the first order, the two sodium lines occur close enough to each other for us to evaluate D at the angle $\theta = 16.99^\circ$ we found in part (a) for one of those lines. Then Eq. 36-30 gives the dispersion as

$$D = \frac{m}{d \cos \theta} = \frac{1}{(2016\ \text{nm})(\cos 16.99^\circ)} = 5.187 \times 10^{-4}\ \text{rad/nm}.$$

From Eq. 36-29 and with $\Delta\lambda$ in nanometers, we then have

$$\Delta\theta = D\,\Delta\lambda = (5.187 \times 10^{-4}\ \text{rad/nm})(589.59 - 589.00) = 3.06 \times 10^{-4}\ \text{rad} = 0.0175^\circ. \qquad \text{(Answer)}$$

You can show that this result depends on the grating spacing d but not on the number of rulings there are in the grating.

(c) What is the least number of rulings a grating can have and still be able to resolve the sodium doublet in the first order?

KEY IDEAS

(1) The resolving power of a grating in any order m is physically set by the number of rulings N in the grating according to Eq. 36-32 ($R = Nm$). (2) The smallest wavelength difference $\Delta\lambda$ that can be resolved depends on the average wavelength involved and on the resolving power R of the grating, according to Eq. 36-31 ($R = \lambda_{avg}/\Delta\lambda$).

Calculation: For the sodium doublet to be barely resolved, $\Delta\lambda$ must be their wavelength separation of 0.59 nm, and λ_{avg} must be their average wavelength of 589.30 nm. Thus, we find that the smallest number of rulings for a grating to resolve the sodium doublet is

$$N = \frac{R}{m} = \frac{\lambda_{avg}}{m\,\Delta\lambda} = \frac{589.30\ \text{nm}}{(1)(0.59\ \text{nm})} = 999\ \text{rulings}. \qquad \text{(Answer)}$$

WILEY PLUS Additional examples, video, and practice available at *WileyPLUS*

36-10 X-Ray Diffraction

X rays are electromagnetic radiation whose wavelengths are of the order of 1 Å ($= 10^{-10}$ m). Compare this with a wavelength of 550 nm ($= 5.5 \times 10^{-7}$ m) at the center of the visible spectrum. Figure 36-27 shows that x rays are produced when electrons escaping from a heated filament F are accelerated by a potential difference V and strike a metal target T.

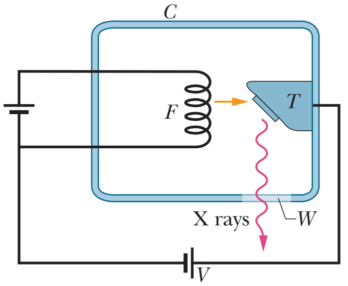


Fig. 36-27 X rays are generated when electrons leaving heated filament F are accelerated through a potential difference V and strike a metal target T. The "window" W in the evacuated chamber C is transparent to x rays.

A standard optical diffraction grating cannot be used to discriminate between different wavelengths in the x-ray wavelength range. For $\lambda = 1$ Å ($= 0.1$ nm) and $d = 3000$ nm, for example, Eq. 36-25 shows that the first-order maximum occurs at

$$\theta = \sin^{-1}\frac{m\lambda}{d} = \sin^{-1}\frac{(1)(0.1\text{ nm})}{3000\text{ nm}} = 0.0019°.$$

This is too close to the central maximum to be practical. A grating with $d \approx \lambda$ is desirable, but, because x-ray wavelengths are about equal to atomic diameters, such gratings cannot be constructed mechanically.

In 1912, it occurred to German physicist Max von Laue that a crystalline solid, which consists of a regular array of atoms, might form a natural three-dimensional "diffraction grating" for x rays. The idea is that, in a crystal such as sodium chloride (NaCl), a basic unit of atoms (called the *unit cell*) repeats itself throughout the array. Figure 36-28*a* represents a section through a crystal of NaCl and identifies this basic unit. The unit cell is a cube measuring a_0 on each side.

When an x-ray beam enters a crystal such as NaCl, x rays are *scattered*—that is, redirected—in all directions by the crystal structure. In some directions the scattered waves undergo destructive interference, resulting in intensity minima; in other directions the interference is constructive, resulting in intensity maxima. This process of scattering and interference is a form of diffraction.

Although the process of diffraction of x rays by a crystal is complicated, the maxima turn out to be in directions *as if* the x rays were reflected by a family of parallel *reflecting planes* (or *crystal planes*) that extend through the atoms within the crystal and that contain regular arrays of the atoms. (The x rays are not actually reflected; we use these fictional planes only to simplify the analysis of the actual diffraction process.)

Figure 36-28*b* shows three reflecting planes (part of a family containing many parallel planes) with *interplanar spacing d,* from which the incident rays shown are said to reflect. Rays 1, 2, and 3 reflect from the first, second, and third planes, respectively. At each reflection the angle of incidence and the angle of reflection are represented with θ. Contrary to the custom in optics, these angles are defined relative to the *surface* of the reflecting plane rather than a normal to that surface. For the situation of Fig. 36-28*b*, the interplanar spacing happens to be equal to the unit cell dimension a_0.

Figure 36-28*c* shows an edge-on view of reflection from an adjacent pair of planes. The waves of rays 1 and 2 arrive at the crystal in phase. After they are reflected, they must again be in phase because the reflections and the reflecting planes have been defined solely to explain the intensity maxima in the diffraction of x rays by a crystal. Unlike light rays, the x rays do not refract upon entering the crystal; moreover, we do not define an index of refraction for this situation. Thus, the relative phase between the waves of rays 1 and 2 as they leave the crystal is set solely by their path length difference. For these rays to be in phase, the path length difference must be equal to an integer multiple of the wavelength λ of the x rays.

By drawing the dashed perpendiculars in Fig. 36-28*c*, we find that the path length difference is $2d \sin\theta$. In fact, this is true for any pair of adjacent planes in the family of planes represented in Fig. 36-28*b*. Thus, we have, as the criterion for intensity maxima for x-ray diffraction,

$$2d\sin\theta = m\lambda, \qquad \text{for } m = 1, 2, 3, \ldots \quad \text{(Bragg's law)}, \qquad (36\text{-}34)$$

where m is the order number of an intensity maximum. Equation 36-34 is called **Bragg's law** after British physicist W. L. Bragg, who first derived it. (He and his father shared the 1915 Nobel Prize in physics for their use of x rays to study the structures of crystals.) The angle of incidence and reflection in Eq. 36-34 is called a *Bragg angle.*

Regardless of the angle at which x rays enter a crystal, there is always a family of planes from which they can be said to reflect so that we can apply Bragg's law. In Fig. 36-28*d*, notice that the crystal structure has the same orientation as it does in Fig. 36-28*a*, but the angle at which the beam enters the structure differs

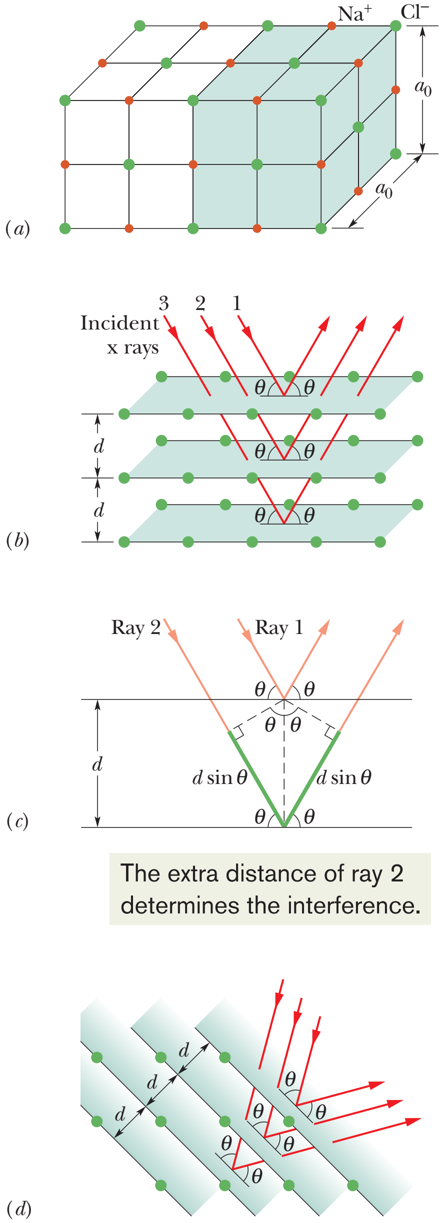


Fig. 36-28 (*a*) The cubic structure of NaCl, showing the sodium and chlorine ions and a unit cell (shaded). (*b*) Incident x rays undergo diffraction by the structure of (*a*). The x rays are diffracted as if they were reflected by a family of parallel planes, with the angle of reflection equal to the angle of incidence, both angles measured relative to the planes (not relative to a normal as in optics). (*c*) The path length difference between waves effectively reflected by two adjacent planes is $2d \sin\theta$. (*d*) A different orientation of the incident x rays relative to the structure. A different family of parallel planes now effectively reflects the x rays.

from that shown in Fig. 36-28*b*. This new angle requires a new family of reflecting planes, with a different interplanar spacing d and different Bragg angle θ, in order to explain the x-ray diffraction via Bragg's law.

Figure 36-29 shows how the interplanar spacing d can be related to the unit cell dimension a_0. For the particular family of planes shown there, the Pythagorean theorem gives

$$5d = \sqrt{\tfrac{5}{4}a_0^2},$$

or

$$d = \frac{a_0}{\sqrt{20}} = 0.2236a_0. \qquad (36\text{-}35)$$

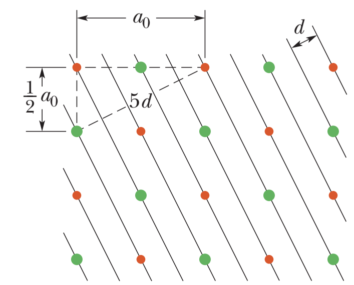


Fig. 36-29 A family of planes through the structure of Fig. 36-28*a*, and a way to relate the edge length a_0 of a unit cell to the interplanar spacing d.

Figure 36-29 suggests how the dimensions of the unit cell can be found once the interplanar spacing has been measured by means of x-ray diffraction.

X-ray diffraction is a powerful tool for studying both x-ray spectra and the arrangement of atoms in crystals. To study spectra, a particular set of crystal planes, having a known spacing d, is chosen. These planes effectively reflect different wavelengths at different angles. A detector that can discriminate one angle from another can then be used to determine the wavelength of radiation reaching it. The crystal itself can be studied with a monochromatic x-ray beam, to determine not only the spacing of various crystal planes but also the structure of the unit cell.

REVIEW & SUMMARY

Diffraction When waves encounter an edge, an obstacle, or an aperture the size of which is comparable to the wavelength of the waves, those waves spread out as they travel and, as a result, undergo interference. This is called **diffraction.**

Single-Slit Diffraction Waves passing through a long narrow slit of width a produce, on a viewing screen, a **single-slit diffraction pattern** that includes a central maximum and other maxima, separated by minima located at angles θ to the central axis that satisfy

$$a \sin\theta = m\lambda, \quad \text{for } m = 1, 2, 3, \ldots \quad \text{(minima)}. \qquad (36\text{-}3)$$

The intensity of the diffraction pattern at any given angle θ is

$$I(\theta) = I_m\left(\frac{\sin\alpha}{\alpha}\right)^2, \quad \text{where } \alpha = \frac{\pi a}{\lambda}\sin\theta \qquad (36\text{-}5, 36\text{-}6)$$

and I_m is the intensity at the center of the pattern.

Circular-Aperture Diffraction Diffraction by a circular aperture or a lens with diameter d produces a central maximum and concentric maxima and minima, with the first minimum at an angle θ given by

$$\sin\theta = 1.22\frac{\lambda}{d} \quad \text{(first minimum—circular aperture)}. \qquad (36\text{-}12)$$

Rayleigh's Criterion *Rayleigh's criterion* suggests that two objects are on the verge of resolvability if the central diffraction maximum of one is at the first minimum of the other. Their angular separation must then be at least

$$\theta_R = 1.22\frac{\lambda}{d} \quad \text{(Rayleigh's criterion)}, \qquad (36\text{-}14)$$

in which d is the diameter of the aperture through which the light passes.

Double-Slit Diffraction Waves passing through two slits, each of width a, whose centers are a distance d apart, display diffraction patterns whose intensity I at angle θ is

$$I(\theta) = I_m(\cos^2\beta)\left(\frac{\sin\alpha}{\alpha}\right)^2 \quad \text{(double slit)}, \qquad (36\text{-}19)$$

with $\beta = (\pi d/\lambda)\sin\theta$ and α as for single-slit diffraction.

Diffraction Gratings A *diffraction grating* is a series of "slits" used to separate an incident wave into its component wavelengths by separating and displaying their diffraction maxima. Diffraction by N (multiple) slits results in maxima (lines) at angles θ such that

$$d\sin\theta = m\lambda, \quad \text{for } m = 0, 1, 2, \ldots \quad \text{(maxima)}, \qquad (36\text{-}25)$$

with the **half-widths** of the lines given by

$$\Delta\theta_{hw} = \frac{\lambda}{Nd\cos\theta} \quad \text{(half-widths)}. \qquad (36\text{-}28)$$

The dispersion D and resolving power R are given by

$$D = \frac{\Delta\theta}{\Delta\lambda} = \frac{m}{d\cos\theta} \qquad (36\text{-}29, 36\text{-}30)$$

and

$$R = \frac{\lambda_{avg}}{\Delta\lambda} = Nm. \qquad (36\text{-}31, 36\text{-}32)$$

X-Ray Diffraction The regular array of atoms in a crystal is a three-dimensional diffraction grating for short-wavelength waves such as x rays. For analysis purposes, the atoms can be visualized as being arranged in planes with characteristic interplanar spacing d. Diffraction maxima (due to constructive interference) occur if the incident direction of the wave, measured from the surfaces of these planes, and the wavelength λ of the radiation satisfy **Bragg's law:**

$$2d\sin\theta = m\lambda, \quad \text{for } m = 1, 2, 3, \ldots \quad \text{(Bragg's law)}. \qquad (36\text{-}34)$$