Reflection and Refraction— Plane Waves and Plane Surfaces

CHAPTER 41

41–1 Reflection and Refraction

In Fig. 41–1*a* a light beam falling on a water surface is both reflected from the surface and bent (that is, *refracted*) as it enters the water. The incident beam is represented in Fig. 41–1*b* by a single line, the *incident ray*, parallel to the direction of propagation. The incident beam is assumed in Fig. 41–1*b* to be a *plane wave*, the wavefronts being normal to the incident ray. The reflected and refracted beams are also represented by rays. The angles of *incidence* (θ_1), of *reflection* (θ_1'), and of *refraction* (θ_2) are measured between the normal to the surface and the appropriate ray, as shown in the figure.

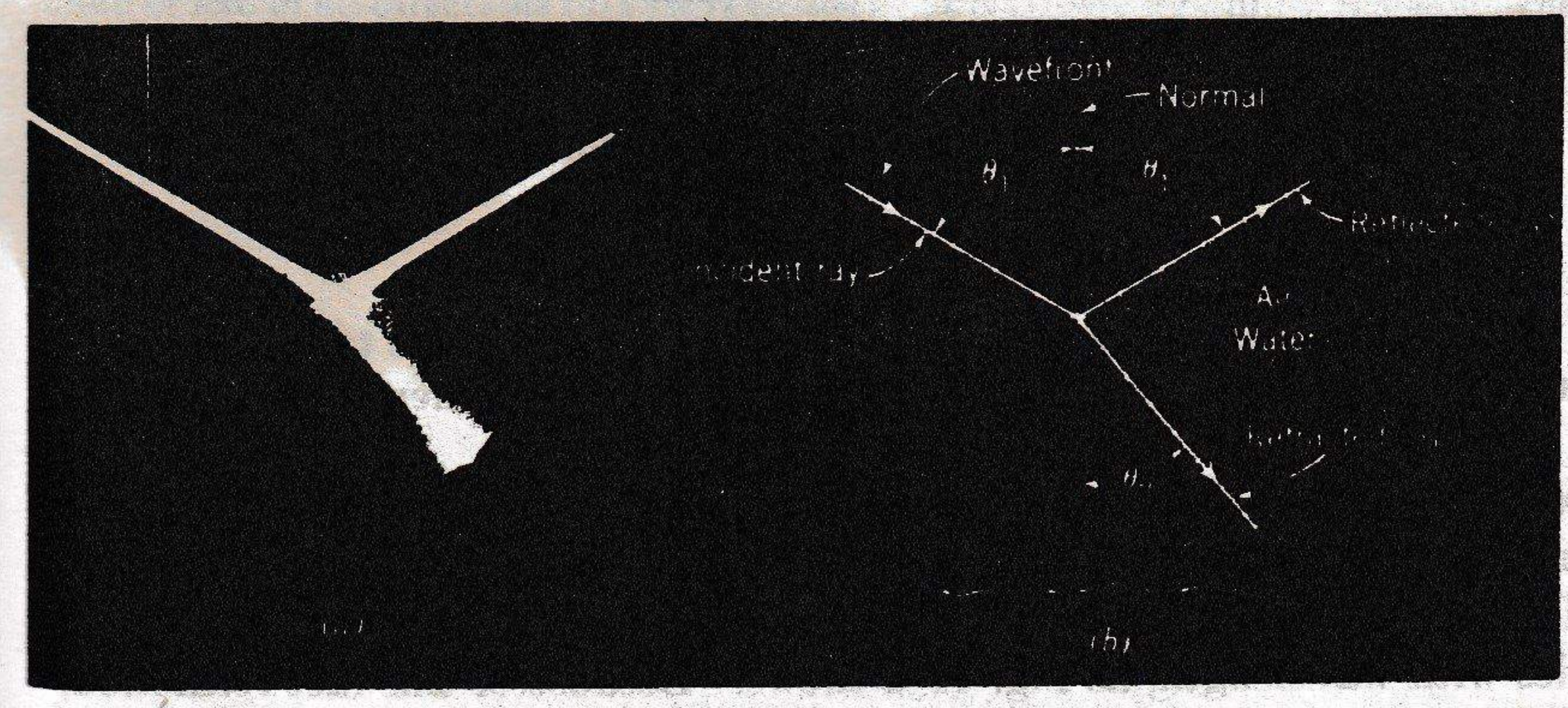


Fig. 41–1 (*a*) A photograph showing reflection and refraction at an air-water interface. (*b*) A representation using rays.

The laws governing reflection and refraction can easily be found from experiment:

1. The reflected and the refracted rays lie in the plane formed by the incident ray and the normal to the surface at the point of incidence, that is, the plane of Fig. 41–1*b*.

2. For reflection:

$$\theta_1' = \theta_1. \tag{41-1}$$

3. For refraction:

$$\frac{\sin \theta_1}{\sin \theta_2} = n_{21}, \tag{41-2}$$

where n_{21} is a constant called the *index of refraction* of medium 2 with respect to medium 1. Table 41–1 shows the indices of refraction for some common

Table 41–1

SOME INDICES OF REFRACTION *
(For λ = 5890 A)

Medium	Index of Refraction
Water	1.33
Ethyl alcohol	1.36
Carbon bisulfide	1.63
Air (1 atm and 20°C)	1.0003
Methylene iodide	1.74
Fused quartz	1.46
Glass, crown	1.52
Glass, dense flint	1.66
Sodium chloride	1.53
Polyethylene	1.50–1.54
Fluorite	1.43

* Measured with respect to a vacuum. The index with respect to air (except, of course, the index of air itself—see item 4) will be negligibly different in most cases.

substances with respect to a vacuum for a wavelength of 5890 A (sodium light).

The index of refraction of one medium with respect to another generally varies with wavelength, as Fig. 41–2 shows. Because of this fact refraction, unlike reflection, can be used to analyze a beam of light into its component wavelengths. Figure 41–3, taken from Newton's *Opticks*, shows how Newton, using glass prism ABC, formed a spectrum of sunlight entering his window through a small hole at F.

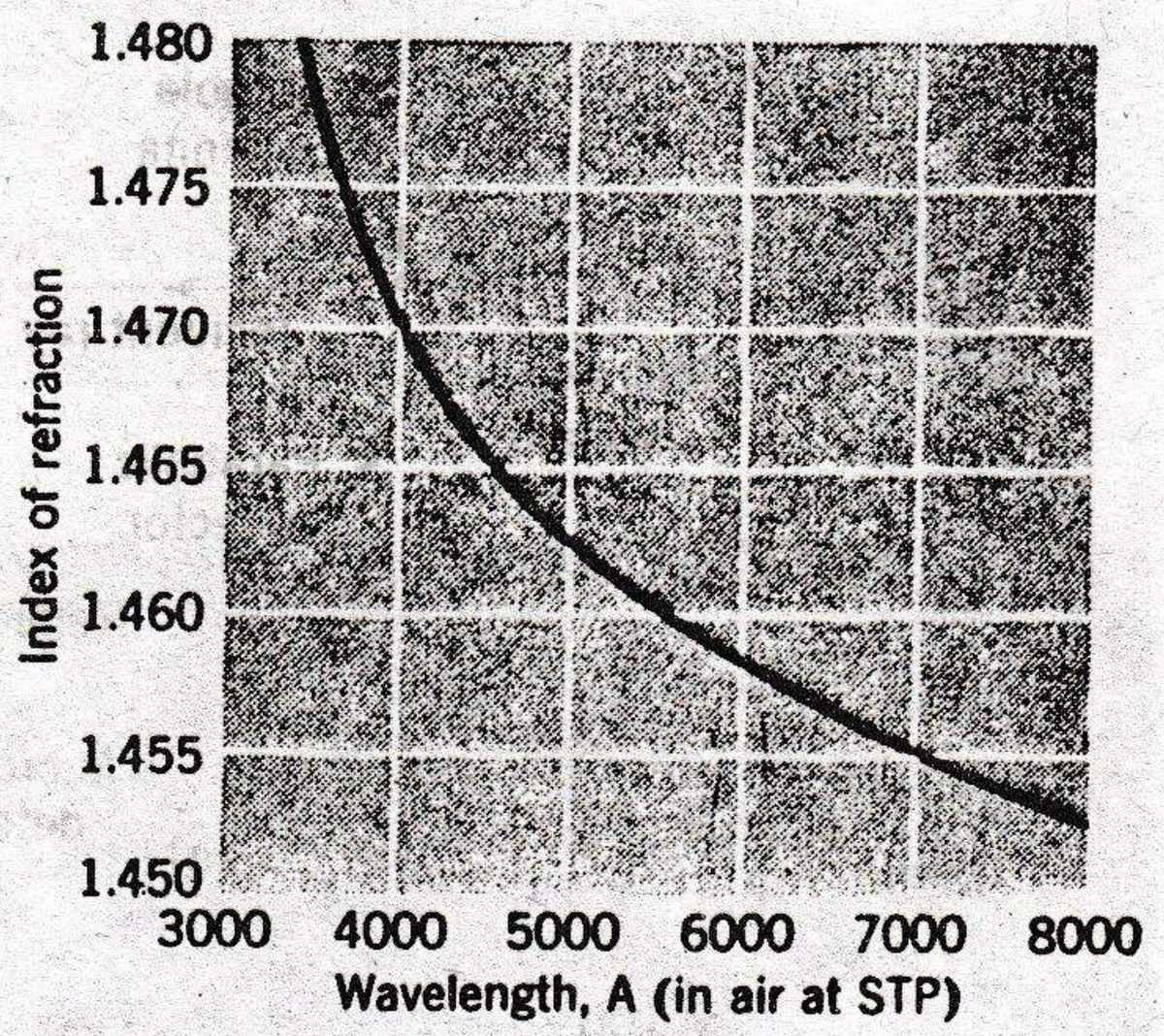


Fig. 41–2 The index of refraction of fused quartz at 18°C, with respect to a vacuum.

The law of reflection was known to Euclid. That of refraction was discovered experimentally by Willebrod Snell (1591–1626) and deduced from the early corpuscular theory of light by René Descartes (1596–1650). The law of refraction is known as Snell's law or (in France) as Descartes' law.

The laws of reflection and refraction can be derived from Maxwell's equations, which means that these laws should hold for all regions of the electromagnetic spectrum. Figure 41–4*a* shows an experimental setup for investigating the reflection of microwaves from a large metal sheet. Figure 41–4*b* shows the reading of the detector as a function of the angular position of the mirror. The existence of a reflected beam at the proper angle confirms the law of reflection for microwaves. There is ample experimental evidence that Eqs. 41–1 and 41–2 correctly describe the behavior of reflected and refracted beams in all parts of the electromagnetic spectrum.

It is common knowledge that a polished steel surface will form a well-defined reflected beam if an incident beam falls on it, but a sheet of paper will reflect light more or less in all directions (*diffuse reflection*). It is largely by diffuse reflection that we see nonluminous objects around us. The difference between diffuse and *specular* (that is, mirror-like) reflection is a matter of surface roughness; a reflected beam will

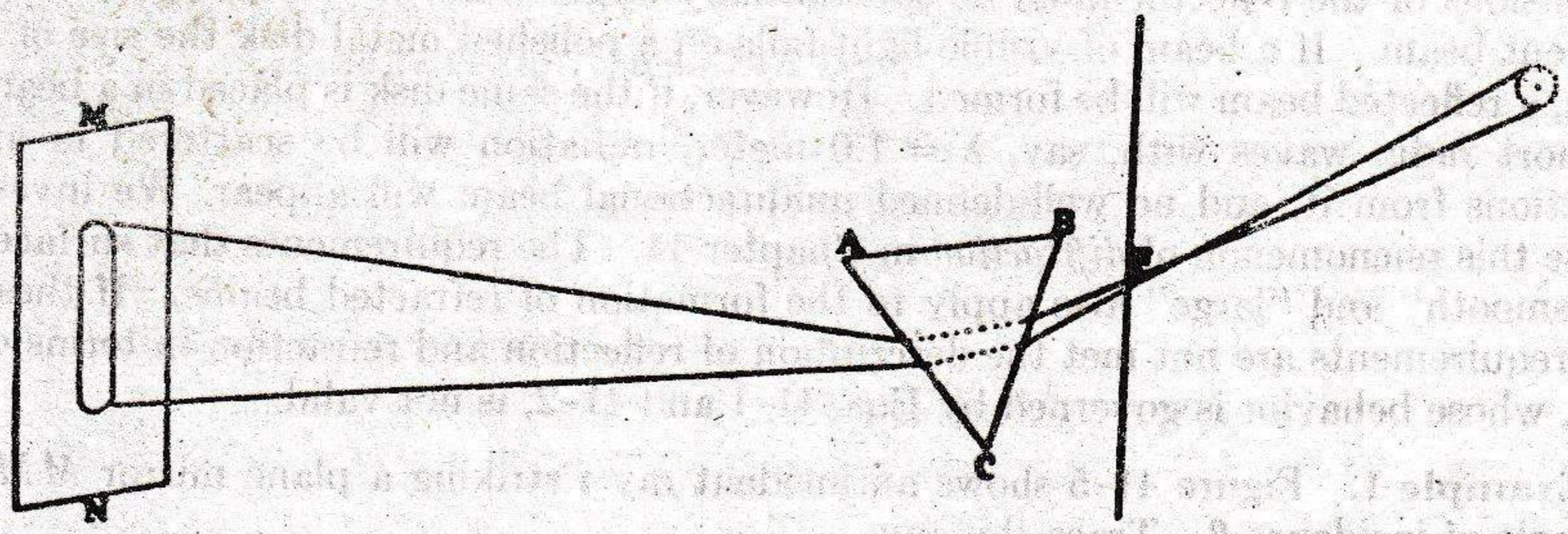


Fig. 41–3 Sunlight from a hole F in a screen is refracted by prism ABC, forming a spectrum on screen MN; from Newton's *Opticks* (1704).

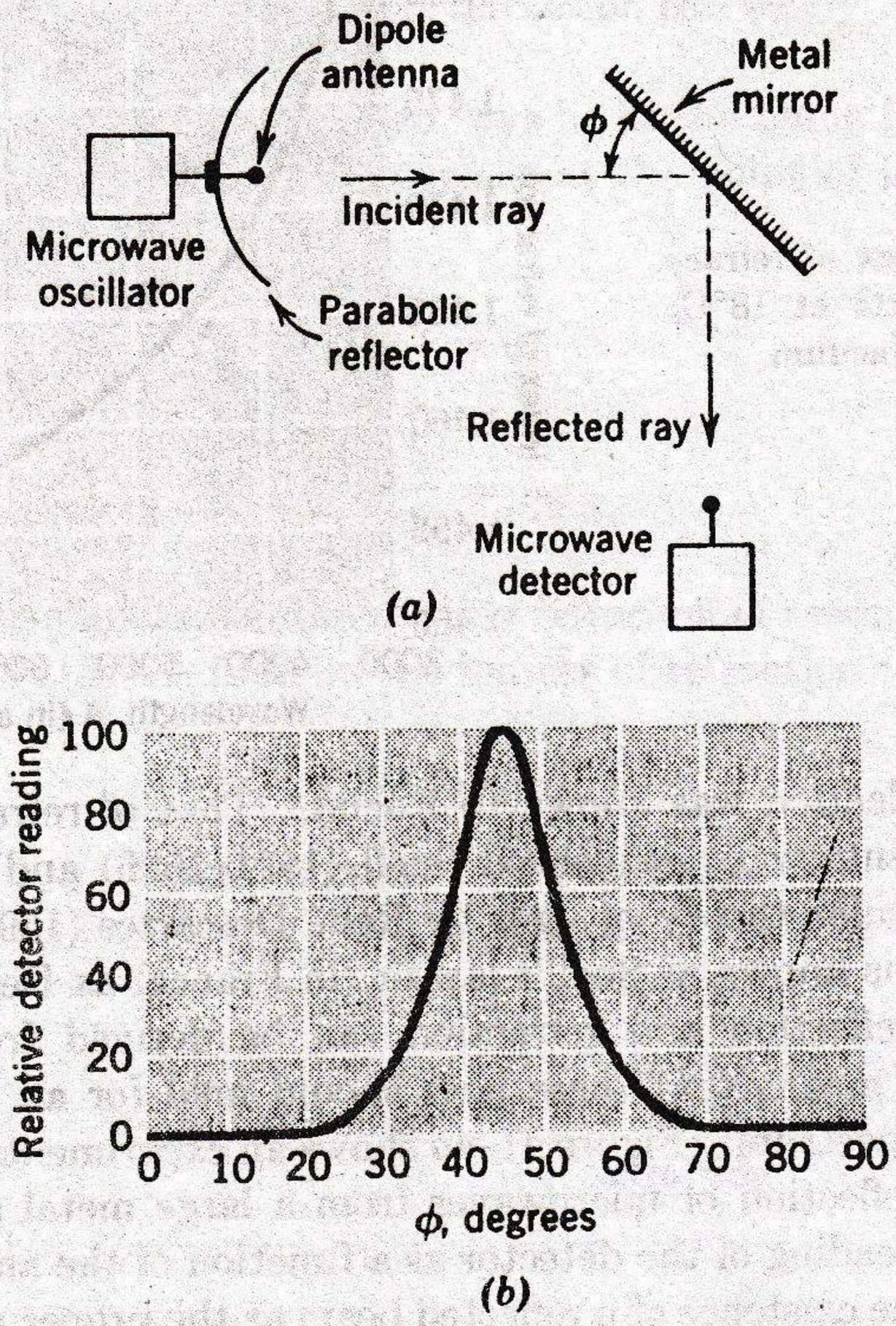


Fig. 41–4 (*a*) An apparatus to test the law of reflection for microwaves. (*b*) A reflected beam (for $\lambda \cong 10$ cm) appears for the expected orientation of the mirror.

be formed only if the average depth of the surface irregularities of the reflector is substantially less than the wavelength of the incident light. This criterion of surface roughness has different implications in different regions of the electromagnetic spectrum. The bottom of a cast-iron skillet for example is a good reflector for microwaves of wavelength 0.5 cm but it is not a good reflector for visible light (that is, one cannot shave by it).

A second requirement for the existence of a reflected beam is that the transverse dimensions of the reflector must be substantially larger than the wavelength of the incident beam. If a beam of visible light falls on a polished metal disk the size of a dime, a reflected beam will be formed. However, if the same disk is placed in a beam of short radio waves with, say, $\lambda = 1.0$ meter, radiation will be scattered in all directions from it, and no well-defined unidirectional beam will appear. We investigate this phenomenon of *diffraction* in Chapter 44. The requirements that surfaces be "smooth" and "large" also apply to the formation of refracted beams. If these two requirements are not met the description of reflection and refraction in terms of rays, whose behavior is governed by Eqs. 41–1 and 41–2, is not valid.

▶ **Example 1.** Figure 41–5 shows an incident ray i striking a plane mirror MM' at angle of incidence θ. Trace this ray.

The reflected ray makes an angle θ with the normal at b and falls as an incident ray on mirror $M'M''$. Its angle of incidence θ' on this mirror is $\pi/2 - \theta$. A second

reflected ray r' makes an angle θ' with the normal erected at b'. Rays i and r' are antiparallel for any value of θ. To see this, note that

$$\phi = \pi - 2\theta' = \pi - 2\left(\frac{\pi}{2} - \theta\right) = 2\theta.$$

Two lines are parallel if their opposite interior angles for an intersecting line (ϕ and 2θ) are equal.

Repeat the problem if the angle between the mirrors is 120° rather than 90°.

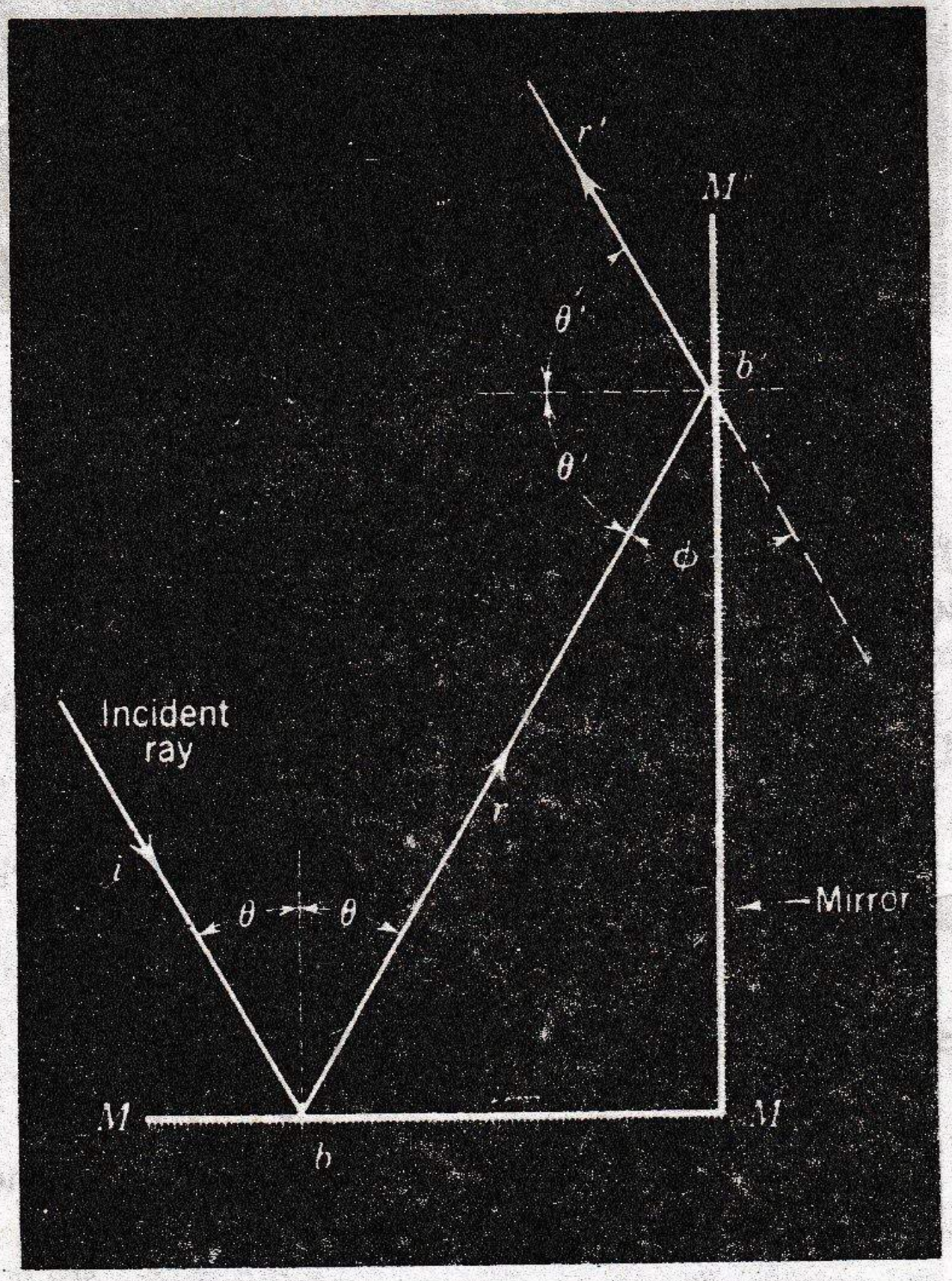


Fig. 41–5 Example 1.

Example 2. An incident ray in air falls on the plane surface of a block of quartz and makes an angle of 30° with the normal. This beam contains two wavelengths, 4000 and 5000 A. The indices of refraction for quartz with respect to air (n_{qa}) at these wavelengths are 1.4702 and 1.4624, respectively. What is the angle between the two refracted beams?

From Eq. 41-2 we have, for the 4000-A beam,

$$\sin\theta_1 = n_{qa}\sin\theta_2,$$

or
$$\sin 30° = (1.4702)\sin\theta_2,$$

which leads to
$$\theta_2 = 19.88°.$$

For the 5000-A beam we have

$$\sin 30° = (1.4624)\sin\theta_2',$$

or
$$\theta_2' = 19.99°.$$

The angle $\Delta\theta$ between the beams is 0.11°, the shorter wavelength component being bent through the larger angle, that is, having the smaller angle of refraction.

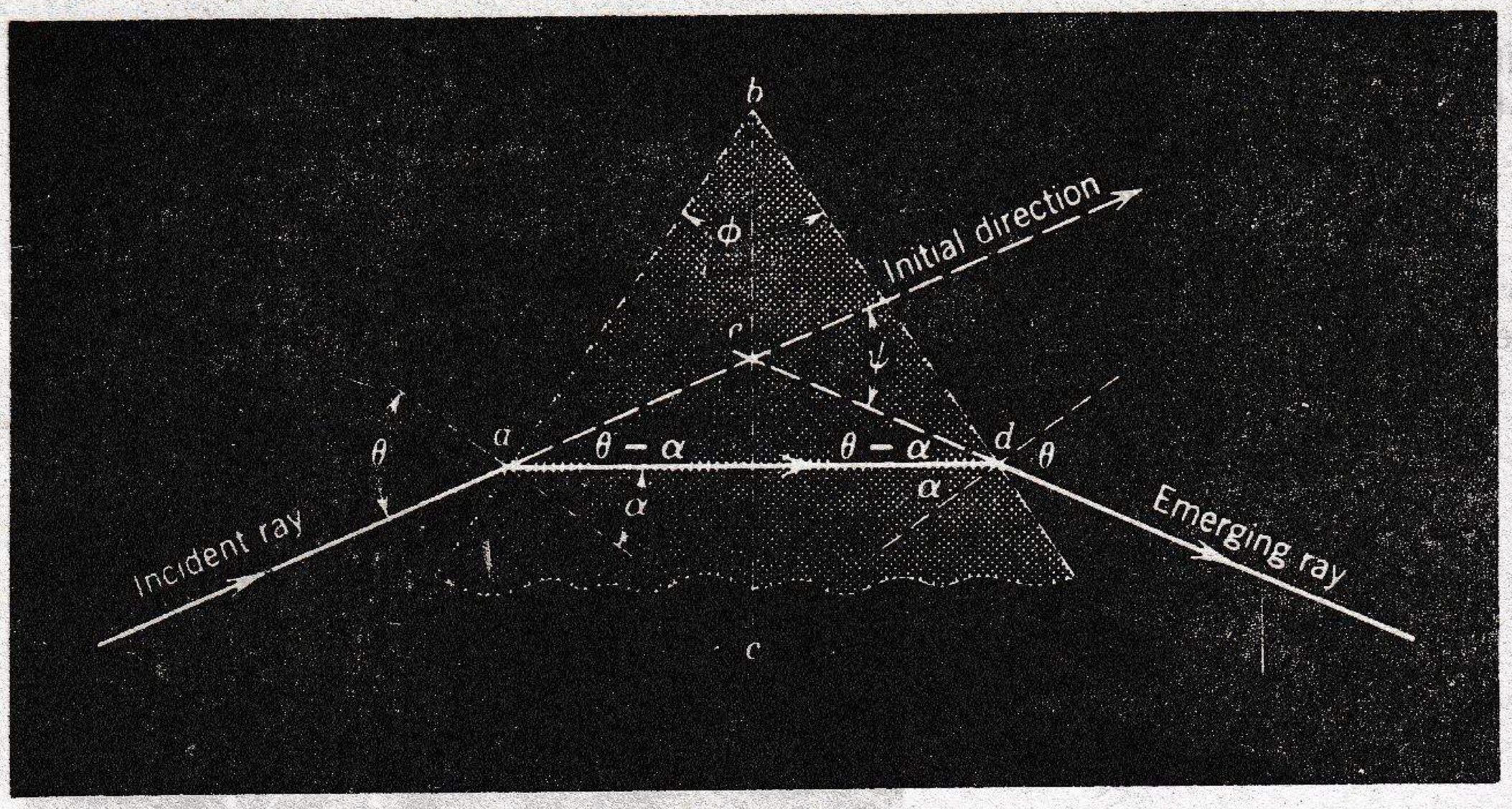


Fig. 41–6 Example 3.

Example 3. An incident ray falls on one face of a glass prism in air as in Fig. 41–6. The angle θ is so chosen that the emerging ray also makes an angle θ with the normal to the other face. Derive an expression for the index of refraction of the prism material with respect to air.

Note that $\angle\, abc = \alpha$, the two angles having their sides mutually perpendicular. Therefore

$$\alpha = \tfrac{1}{2}\phi. \tag{41–3}$$

The *deviation angle* ψ is the sum of the two opposite interior angles in triangle *aed*, or

$$\psi = 2(\theta - \alpha).$$

Substituting $\frac{1}{2}\phi$ for α and solving for θ yields

$$\theta = \tfrac{1}{2}(\psi + \phi). \tag{41–4}$$

At point a, θ is the angle of incidence and α the angle of refraction. The law of refraction (see Eq. 41–2) is

$$\sin\theta = n_{ga}\sin\alpha,$$

in which n_{ga} is the index of refraction of the glass with respect to air.

From Eqs. 41–3 and 41–4 this yields

$$\sin\frac{\psi+\phi}{2} = n_{ga}\sin\frac{\phi}{2}$$

or

$$n_{ga} = \frac{\sin\frac{1}{2}(\psi+\phi)}{\sin(\phi/2)},$$

which is the desired relation. This equation holds only for θ so chosen that the light ray passes symmetrically through the prism. For this condition the deviation angle ψ is a minimum; if θ is either increased or decreased, a larger deviation will be produced. ◀

41–2 Huygens' Principle

A theory of light would not be accepted if it were not able to predict the well-established laws of reflection and refraction. These laws can be derived from Maxwell's equations, but mathematical complexity prevents us from

doing so. Fortunately, these and several other laws of optics can be derived on the basis of a simpler but less comprehensive theory of light, put forward by the Dutch physicist Christian Huygens in 1678. This theory simply assumes that light is a wave rather than, say, a stream of particles. It says nothing about the nature of the wave and, in particular—since Maxwell's theory of electromagnetism appeared only after the lapse of a century—gives no hint of the electromagnetic character of light. Huygens did not know whether light was a transverse wave or a longitudinal one; he did not know the wavelengths of visible light; he had little knowledge of the speed of light. Nevertheless, his theory was a useful guide to experiment for many years and remains useful today for pedagogic and certain other practical purposes. We must not expect it to yield the same wealth of detailed information that Maxwell's more complete electromagnetic theory does.

Huygens' theory is based on a geometrical construction, called *Huygens' principle*, that allows us to tell where a given wavefront will be at any time in the future if we know its present position; it is: *All points on a wavefront can be considered as point sources for the production of spherical secondary wavelets. After a time t the new position of the wavefront will be the surface of tangency to these secondary wavelets.*

We illustrate this by a trivial example: Given a wavefront (*ab* in Fig. 41–7) in a plane wave in free space, where will the wavefront be a time t later? Following Huygens' principle, we let several points on this plane (see dots) serve as centers for secondary spherical wavelets. In a time t the radius of these spherical waves is ct, where c is the speed of light in free space. The plane of tangency to these spheres at time t is represented by dc. As we

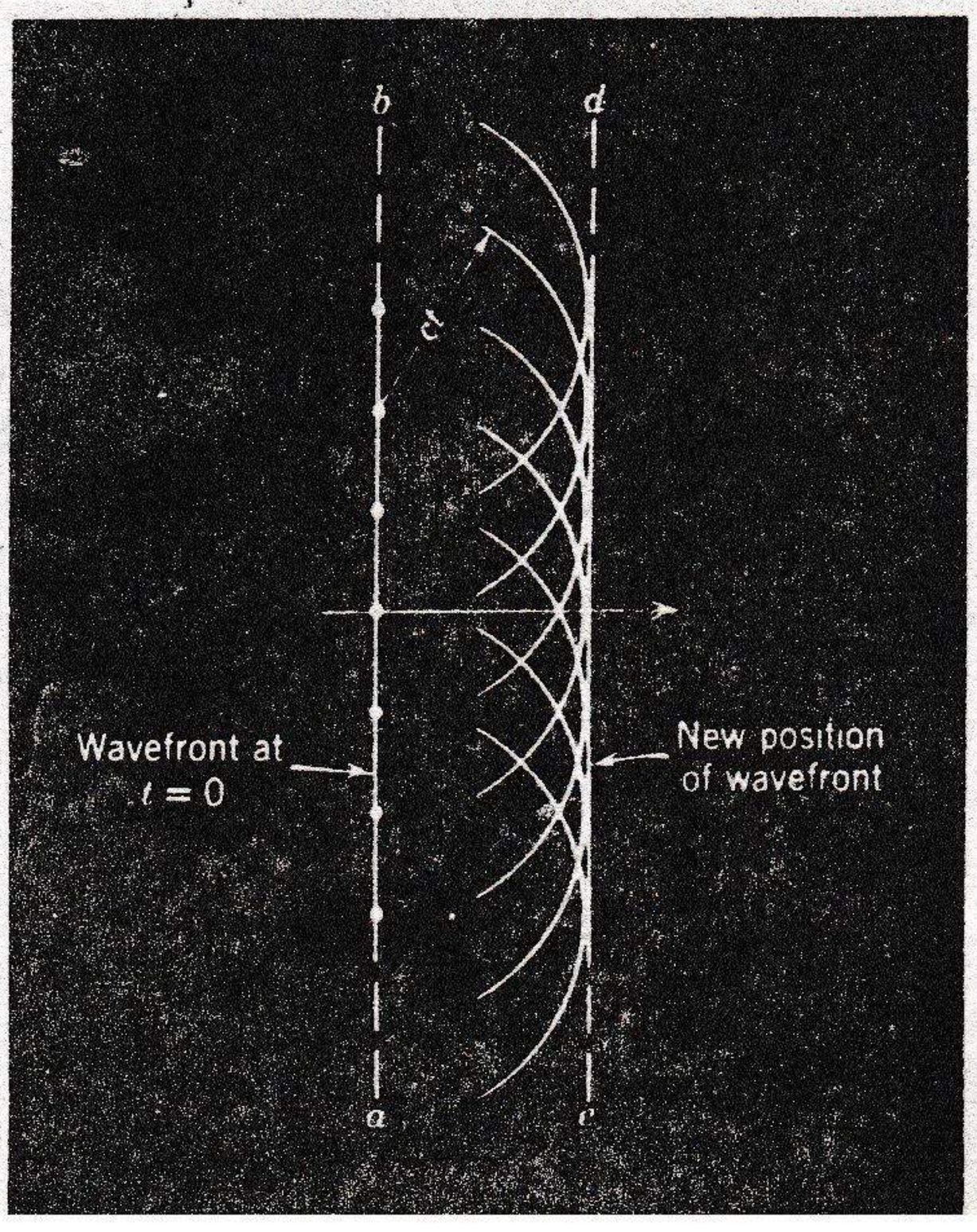


Fig. 41–7 The propagation of a plane wave in free space is described by the Huygens construction. Note that the ray (horizontal arrow) representing the wave is perpendicular to the wavefronts.

expect, it is parallel to plane *ab* and a perpendicular distance *ct* from it. Thus plane wavefronts are propagated as planes and with speed *c*. Note that the Huygens method involves a three-dimensional construction and that Fig. 41–7 is the intersection of this construction with the plane of the page.

We might expect that, contrary to observation, a wave should be radiated backward as well as forward from the dots in Fig. 41–7. This result is avoided by assuming that the *intensity* of the spherical wavelets is not uniform in all directions but varies continuously from a maximum in the forward direction to a minimum of zero in the back direction. This is suggested by the shading of the circular arcs in Fig. 41–7. Huygens' method can be applied quantitatively to *all* wave phenomena; see Problem 1. The method was put on a firm mathematical footing by Augustin Fresnel (1788–1827).

41–3 Huygens' Principle and the Law of Reflection

Figure 41–8*a* shows three wavefronts in a plane wave falling on mirror *MM'*. For convenience they are chosen to be one wavelength apart. Note that θ_1, the angle between the wavefronts and the mirror, is the same as the angle between the incident ray and the normal to the mirror. In other

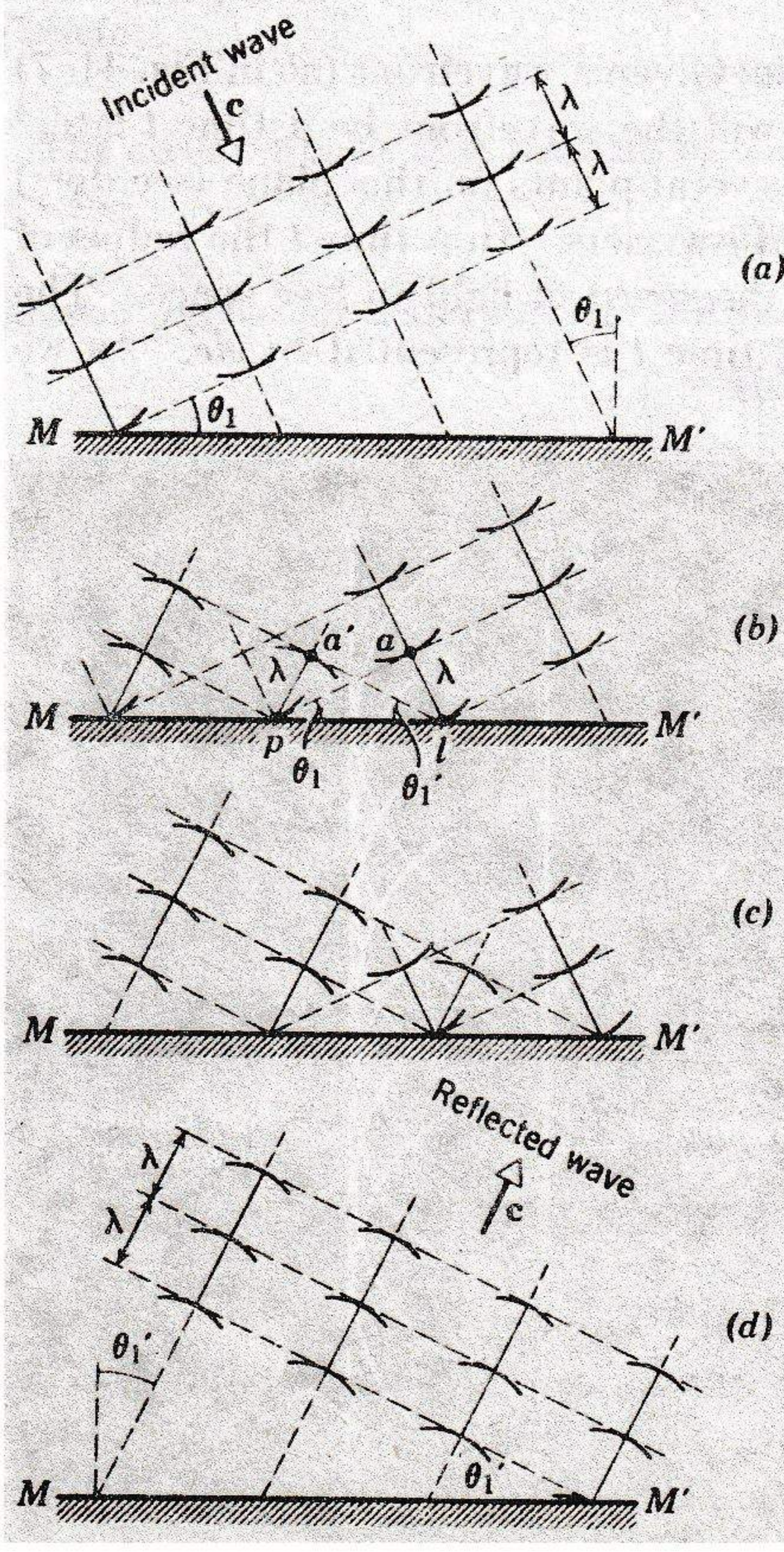


Fig. 41–8 The reflection of a plane wave as described by the Huygens construction.

words, θ_1 is the *angle of incidence*. The three wavefronts are related to each other by the Huygens construction, as in Fig. 41–7.

In Fig. 41–8*b* a Huygens wavelet centered on point *a* will expand to include point *l* after a time λ/c. Light from point *p* in this same wavefront cannot move beyond the mirror but must expand upward as a spherical Huygens wavelet. Setting a compass to radius λ and swinging an arc about *p* provides a semicircle to which the reflected wavefront must be tangent. Since point *l* must lie on the new wavefront, this tangent must pass through *l*. Note that the angle θ_1' between the wavefront and the mirror is the same as the angle between the reflected ray and the normal to the mirror. In other words, θ_1' is the *angle of reflection*.

Consider right triangles *alp* and *a'lp*. They have side *lp* in common and side $al\ (=\lambda)$ is equal to side $a'p$. The two right triangles are thus congruent and we may conclude that

$$\theta_1 = \theta_1',$$

as required by the law of reflection. If the student recalls that the Huygens construction is three-dimensional and that the arcs shown represent segments of spherical surfaces, he will be able to convince himself that the reflected ray lies in the plane formed by the incident ray and the normal to the mirror, that is, the plane of Fig. 41–8. This is also a requirement of the law of reflection; see p. 1014.

41–4 Huygens' Principle and the Law of Refraction

Figure 41–9 shows four stages in the refraction of three successive wavefronts in a plane wave falling on an interface between air (medium 1) and glass (medium 2). For convenience, we assume that the incident wavefronts are separated by λ_1, the wavelength as measured in medium 1. Let the speed of light in air be v_1 and that in glass be v_2. We assume that

$$v_2 < v_1. \tag{41–5}$$

This assumption about the speeds is vital to the derivation that follows. It was not possible to test it experimentally because of technical difficulties until 1850, when the assumption was shown by Foucault to be correct.

The wavefronts in Fig. 41–9*a* are related to each other by the Huygens construction of Fig. 41–7. As in Fig. 41–8, θ_1 is the angle of incidence. In Fig. 41–9*b* consider the time $(=\lambda_1/v_1)$ during which a Huygens wavelet from point *e* moves to include point *c*. Light from point *h*, traveling through glass at a reduced speed (recall the assumption of Eq. 41–5) will move a shorter distance

$$\lambda_2 = \lambda_1 \frac{v_2}{v_1} \tag{41–6}$$

during this time. The refracted wavefront must be tangent to an arc of this radius centered on *h*. Since *c* lies on the new wavefront, the tangent must pass through this point, as shown. Note that θ_2, the angle between the refracted wavefront and the air-glass interface, is the same as the angle

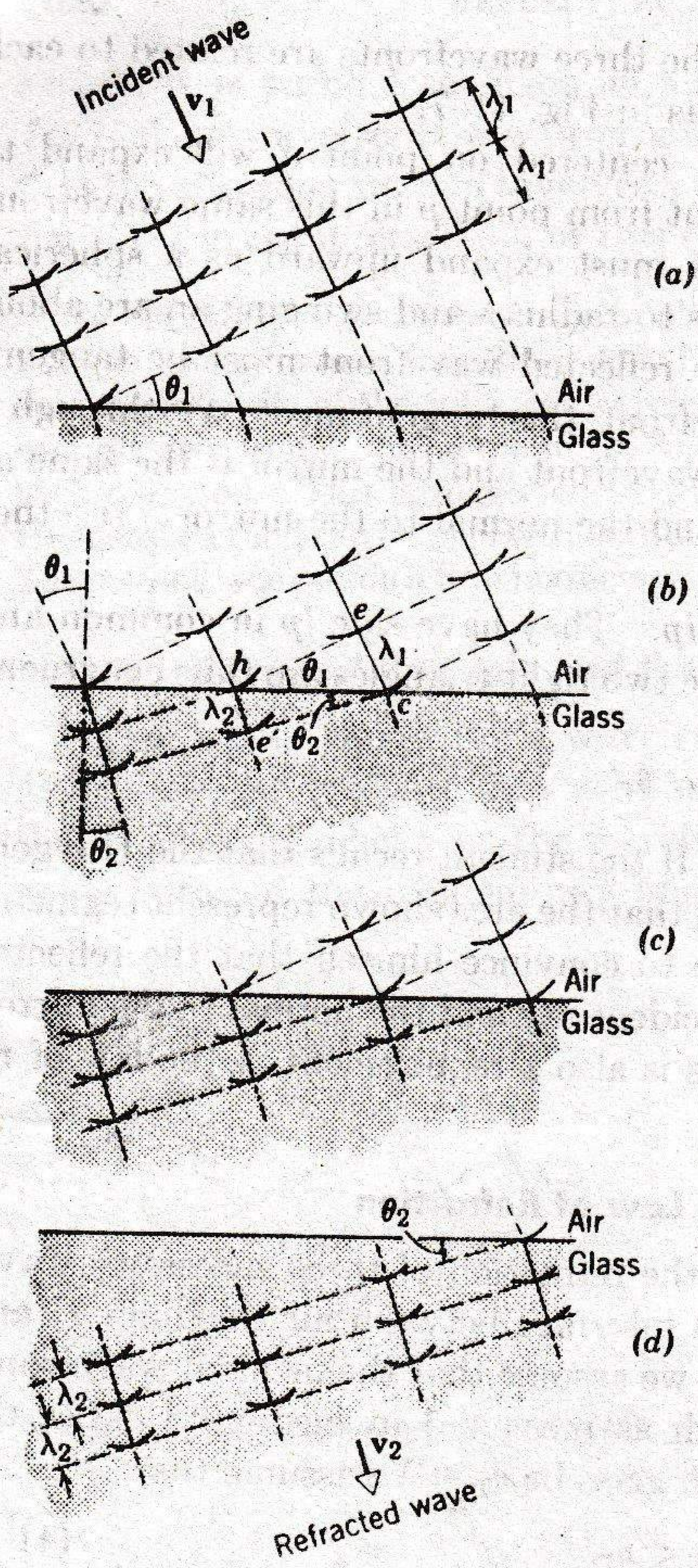


Fig. 41–9 The refraction of a plane wave as described by the Huygens construction; the reflected wave is omitted for simplicity. Note the change in wavelength on refraction.

between the refracted ray and the normal to this interface. In other words, θ_2 is the *angle of refraction*. Note, too, that the wavelength in glass (λ_2) is less than the wavelength in air (λ_1).

For the right triangles *hce* and *hce′* we may write

$$\sin \theta_1 = \frac{\lambda_1}{hc} \qquad \text{(for } hce\text{)}$$

and

$$\sin \theta_2 = \frac{\lambda_2}{hc} \qquad \text{(for } hce'\text{)}.$$

Dividing and using Eq. 41–6 yields

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{\lambda_1}{\lambda_2} = \frac{v_1}{v_2} = \text{a constant.} \qquad (41\text{–}7)$$

The law of refraction, as stated in Eq. 41–2, is

$$\frac{\sin \theta_1}{\sin \theta_2} = n_{21}, \tag{41–2}$$

so that n_{21} is now revealed as the ratio of the speeds of light in the two media, or

$$n_{21} = \frac{v_1}{v_2}. \tag{41–8}$$

We may rewrite Eq. 41–7 as

$$\left(\frac{c}{v_1}\right) \sin \theta_1 = \left(\frac{c}{v_2}\right) \sin \theta_2, \tag{41–9}$$

in which c is the speed of light in free space. The quantities (c/v_1) and (c/v_2) (see Eq. 41–8) are the indices of refraction of medium 1 and of medium 2, respectively, with respect to a vacuum. Introducing the symbols n_1 and n_2 for these quantities allows us to write the law of refraction as

$$n_1 \sin \theta_1 = n_2 \sin \theta_2. \tag{41–10}$$

If we assume that the medium above the glass in Fig. 41–9 is a vacuum rather than air, the speed v_1 becomes c and the wavelength, called λ_1 in Fig. 41–9, assumes a value λ that is characteristic of the wave in free space. Equation 41–6 may thus be written

$$\lambda_2 = \lambda \frac{v_2}{c} = \frac{\lambda}{n_2}. \tag{41–11}$$

This shows specifically that the wavelength of light in a material medium is less than the wavelength of the same wave in a vacuum. Figure 41–9 shows clearly the difference in wavelength in the two media.

The application of Huygens' principle to refraction requires that if a light ray is bent toward the normal in passing from air to an optically dense medium then the speed of light in that optically dense medium (glass, say) must be *less* than that in air; see Eq. 41–5. This requirement holds for all wave theories of light. For the early particle theory of light put forward by Newton, refraction can be explained only if the speed of light in the medium in which light is bent toward the normal (the optically dense medium) is *greater* than that in air. The dense medium was thought to exert attractive forces on the light "corpuscles" as they neared the surface, speeding them up and changing their direction to cause them to make a smaller angle with the normal. Figure 41–10 shows a figure from a 1637 work of René Descartes, in which he makes an analogy between the refraction of light and the motion of a tennis ball on entering a medium in which it moves more slowly.

An experimental comparison of the speed of light in water and in air is decisive, therefore, between the wave and corpuscular theories of light. Such a measurement was first carried out by Foucault in 1850; he showed conclusively that *light travels more slowly in water than in air*, thus ruling out the corpuscular theory of Newton.

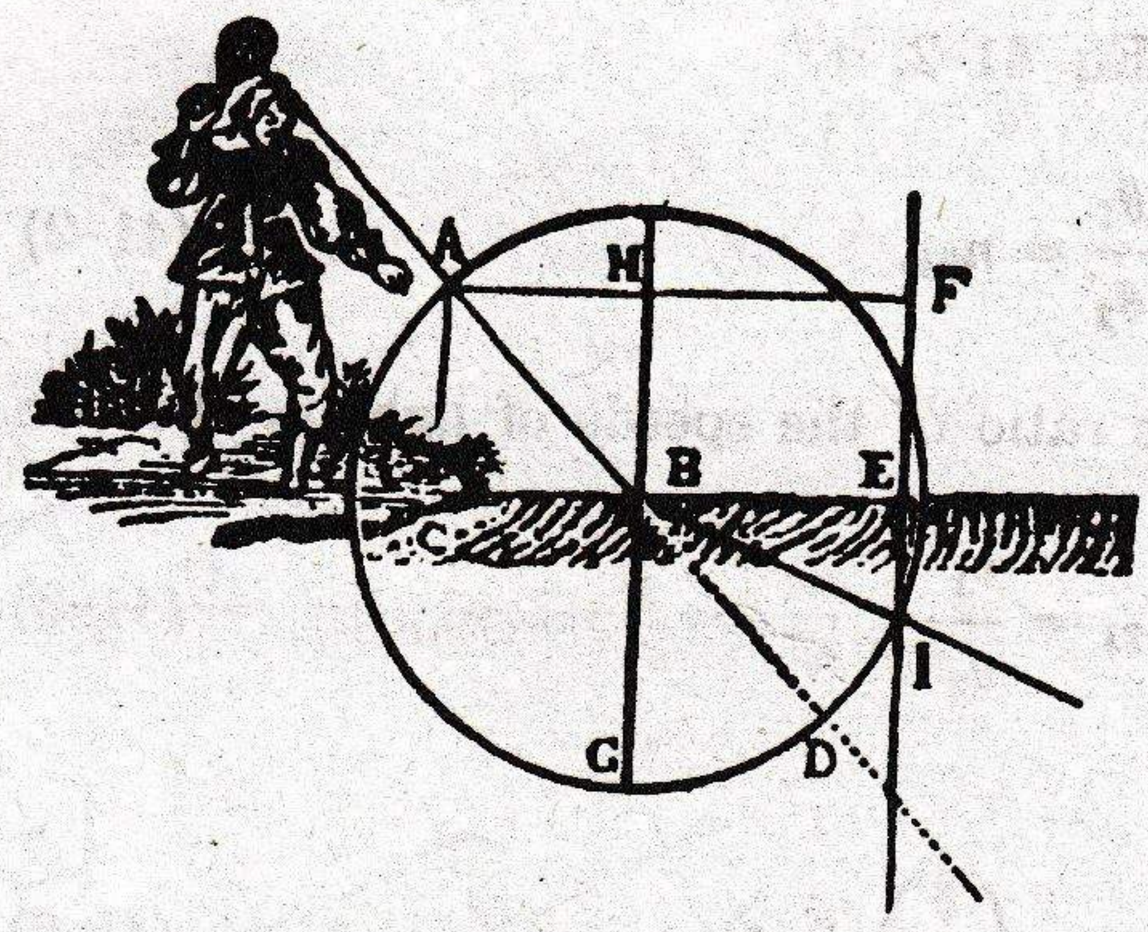


Fig. 41–10 According to the early (incorrect) corpuscular theory of light, ABI is the path of a ray on entering a medium in which its speed is *less*. The (correct) wave theory predicts that for the ray shown the speed in the lower medium (below BE) must be *greater*; from Descartes' *La Dioptrique* (1637).

41–5 Total Internal Reflection

Let light rays in an optically dense medium (glass, say) fall on a surface on the other side of which is a less optically dense medium (air, say); see Fig. 41–11. As the angle of incidence θ is increased, a situation is reached (see ray e) at which the refracted ray points along the surface, the angle of refraction being 90°. For angles of incidence larger.than this *critical angle* θ_c no refracted ray exists, giving rise to a phenomenon called *total internal reflection*.

The critical angle is found by putting $\theta_2 = 90°$ in the law of refraction (see Eq. 41–10):

$$n_1 \sin \theta_c = n_2 \sin 90°,$$

or

$$\sin \theta_c = \frac{n_2}{n_1}. \tag{41–12}$$

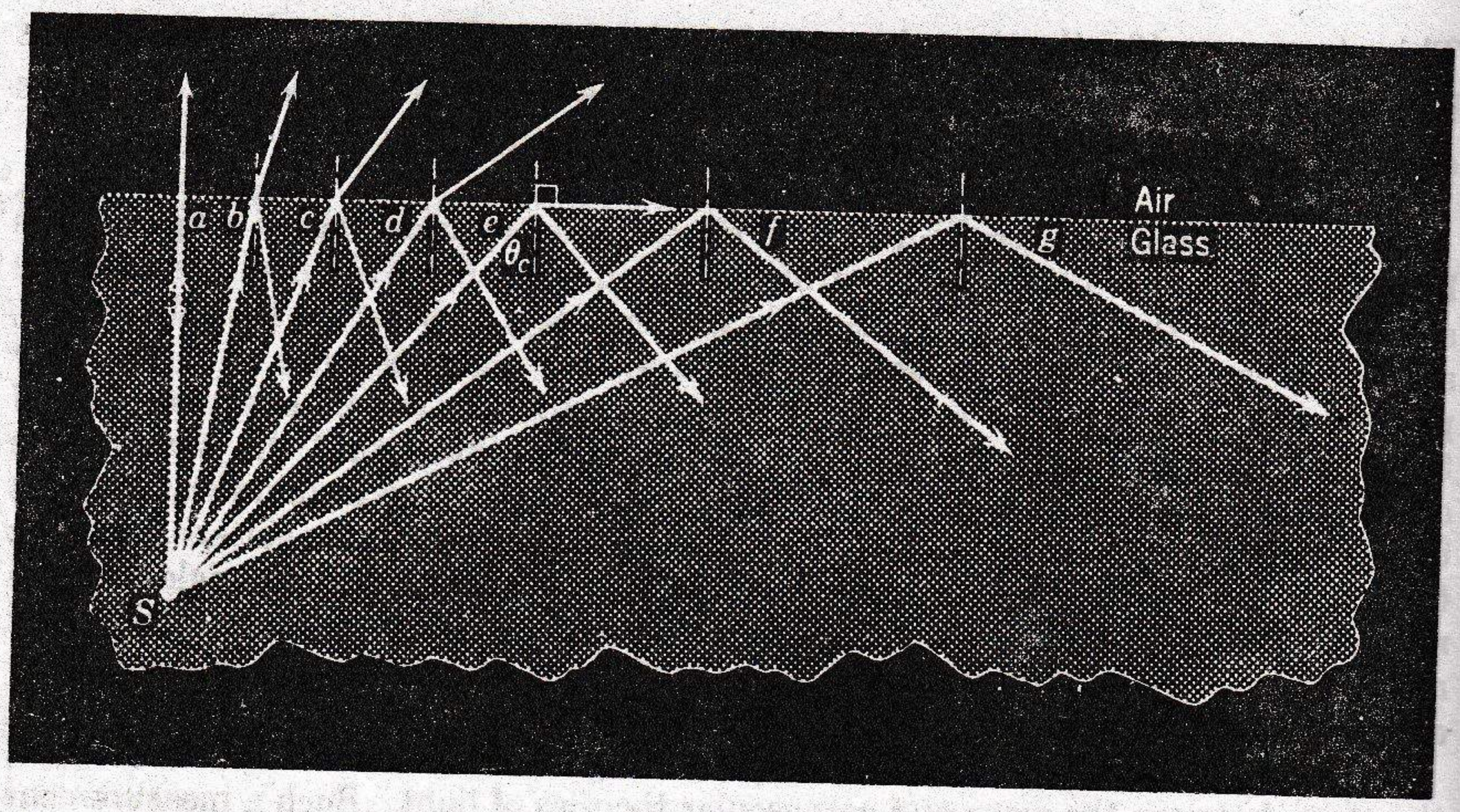


Fig. 41–11 Showing the total internal reflection of light from a source S; the *critical angle* is θ_c.

For glass and air $\sin\theta_c = (1.00/1.50) = 0.667$, which yields $\theta_c = 41.8°$. Total internal reflection does not occur when light originates in the medium of lower index of refraction.

Light can be "piped" from one point to another with little loss by allowing it to enter one end of a rod of transparent plastic. The light will undergo total internal

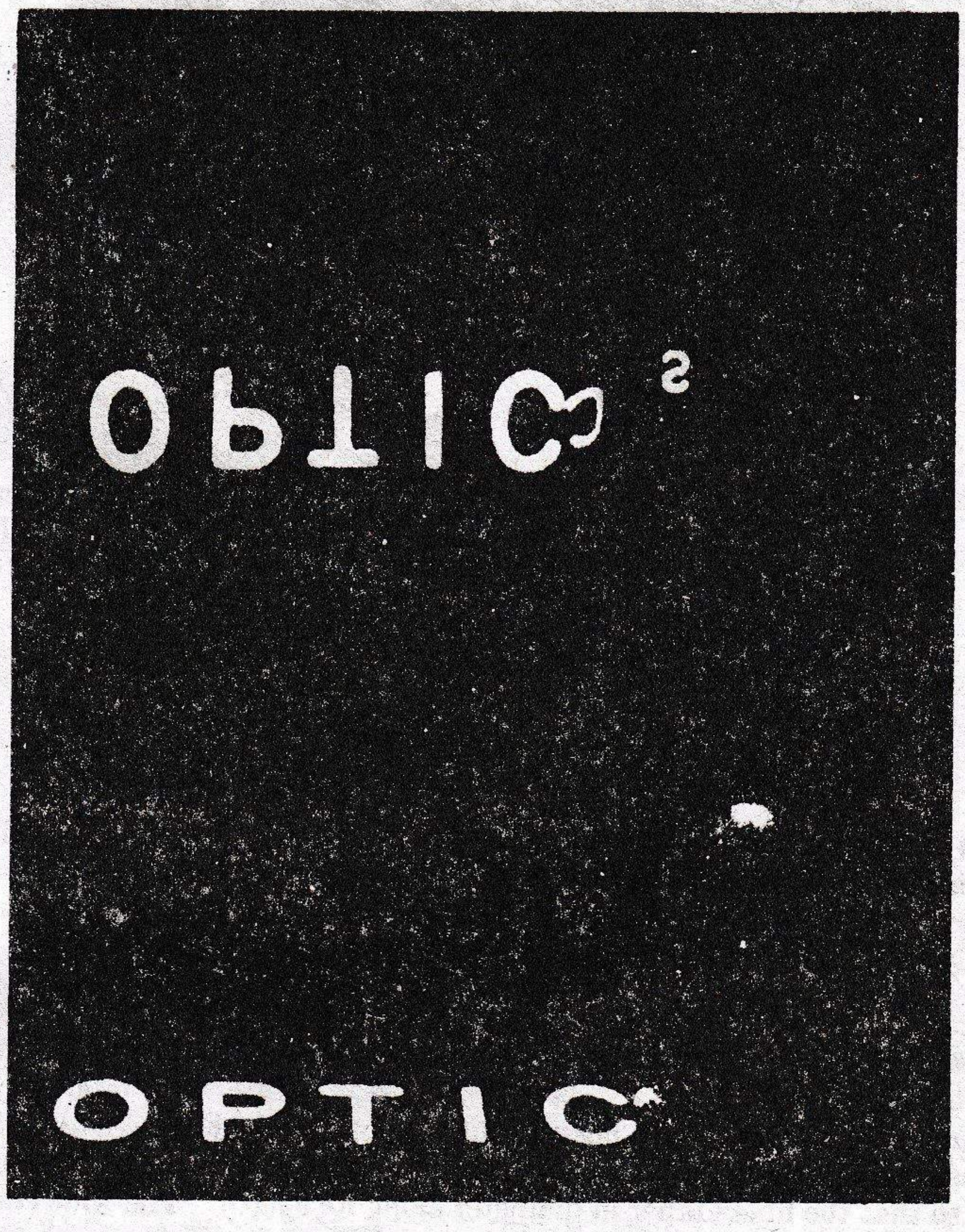


Fig. 41–12 A bundle of tapered fibers (*below*) is placed over the letter *S*. Above, with the aid of a mirror, we see that the image, reduced in size, is transmitted to the top of the bundle by total internal reflection in the individual fibers. (Courtesy Dr. N. S. Kapany, Optics Technology, Inc.)

reflection at the boundary of the rod and will follow its contour, emerging at its far end. Images may be transferred from one location to another, using a bundle of fine glass fibers, each fiber transmitting a small fraction of the image.* Such bundles can be made in which a (flexible) seven-foot length delivers half the energy entering it at the far end. In a fiber in such a bundle a typical ray may undergo 48,000 reflections. Most of the energy loss is due to absorption within the glass, reflection being almost

* See "Fiber Optics," by N. S. Kapany, *Scientific American*, November 1960.

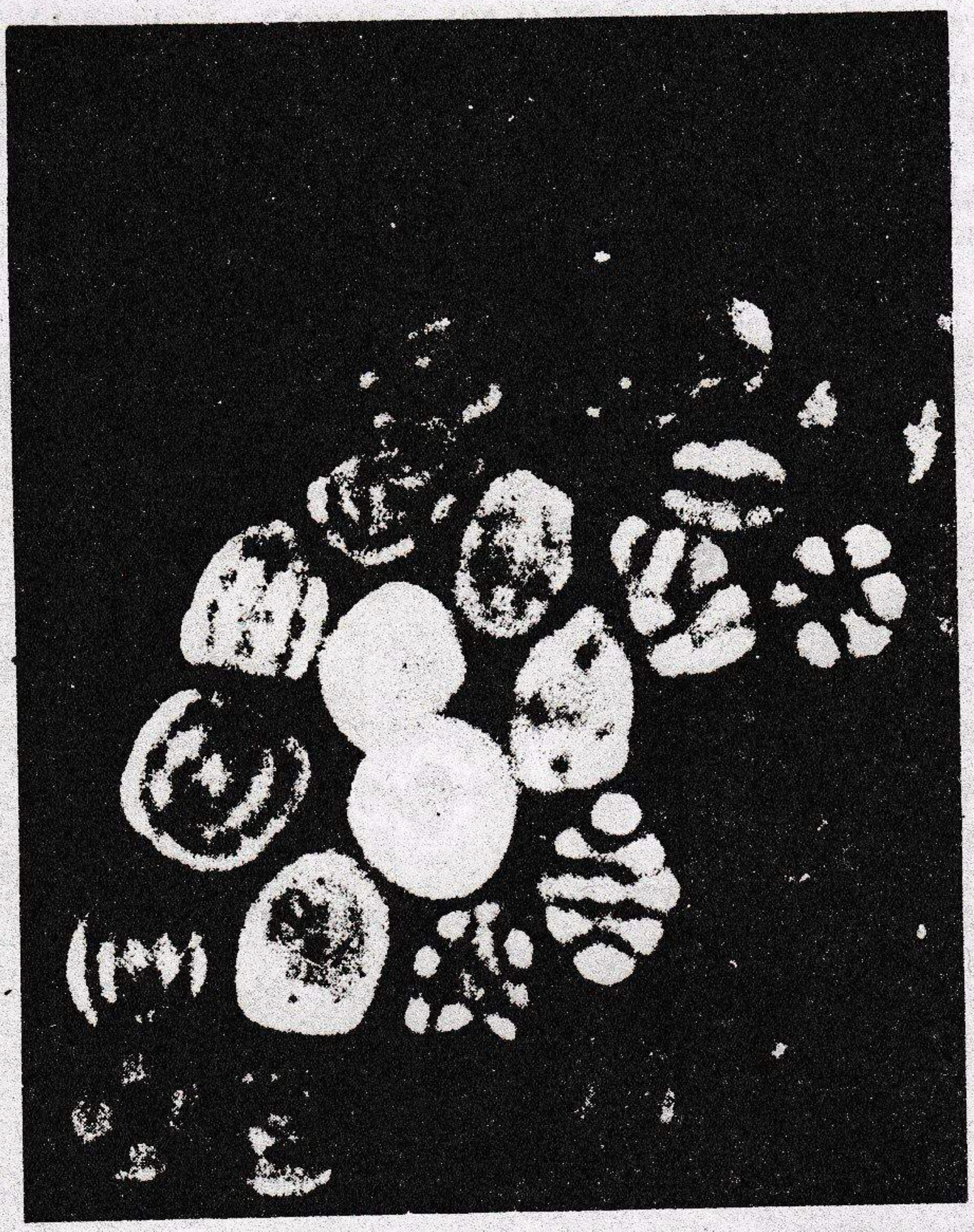


Fig. 41–13 A photomicrograph of light emerging from the end of a bundle of fibers. The fiber diameters approach the wavelength of light so that each fiber acts like an "optical waveguide." We have here convincing visual evidence that light is an electromagnetic wave. (Courtesy of Dr. N. S. Kapany, Optics Technology, Inc.)

truly total. Measurable amounts of light can be transmitted through single fibers 150 feet long.

Fiber optics techniques make possible many useful optical devices for transmitting and transforming luminous images. Figure 41–12 shows a short fiber bundle constructed so that the fibers taper in diameter along its length. The wide end is shown placed over the letter *S* in the printed word "OPTICS." We see, with the aid of a mirror placed above the bundle, that a letter *S*, reduced in size, has been transmitted by total internal reflection through the bundle to its narrow end.

Figure 41–13 is an enlarged view of a cross section of such a fiber bundle in which the diameters of the individual fibers are made so small that they are of the order of magnitude of the wavelength of light. This condition violates the spirit of our assumption of p. 1016, namely that the transverse dimensions of reflecting and refracting surfaces would be large compared to the wavelength of light. Consequently, a description of the reflection and refraction of light in terms of rays, as in Figs. 41–1*b* and 41–11, is not possible. Figure 41–13 is readily interpreted, however, on the basis of the electromagnetic wave theory of light and provides convincing pictorial supporting evidence for that theory. The fibers behave like waveguides *

* A dielectric rod can serve as a waveguide in the same way that a hollow metal pipe can, and it has similar properties.

(see Section 39–3), and the patterns of darkness and light represent the distribution of the **E** and **B** vectors for various modes of oscillation of the electromagnetic waves traveling down the "guides."

▶ **Example 4.** Figure 41–14*a* shows a triangular prism of glass, a ray incident normal to one face being totally reflected. If θ_1 is 45°, what can you conclude about the index of refraction n of the glass?

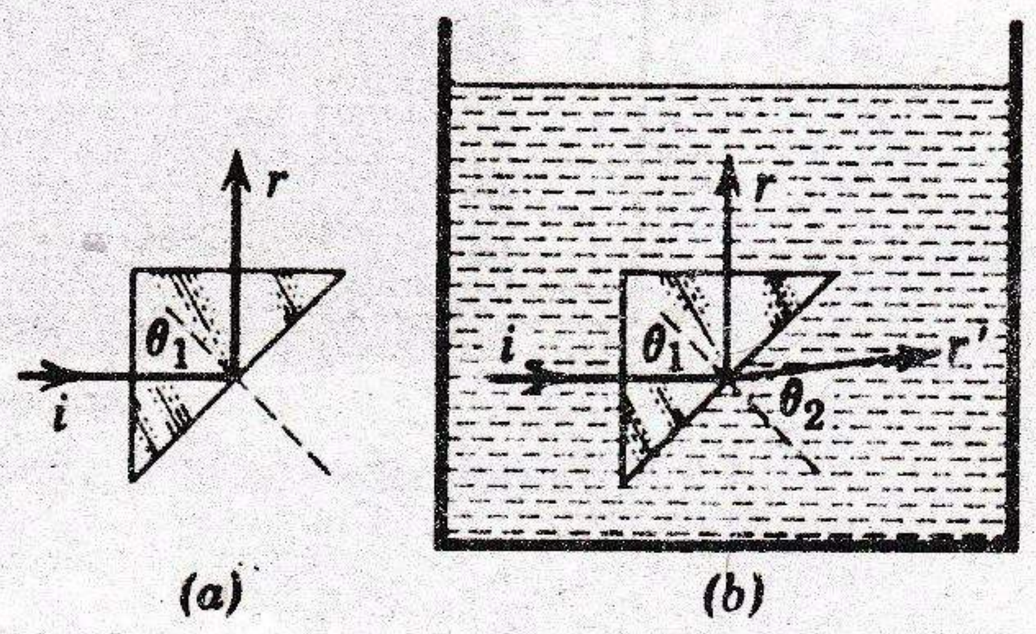


Fig. 41–14 Example 4.

The angle θ_1 must be equal to or greater than the critical angle θ_c where θ_c is given by Eq. 41–12:

$$\sin \theta_c = \frac{n_2}{n_1} = \frac{1}{n},$$

in which, for all practical purposes, the index of refraction of air ($= n_2$) is set equal to unity. Suppose that the index of refraction of the glass is such that total internal reflection just occurs, that is, that $\theta_c = 45°$. This would mean

$$n = \frac{1}{\sin 45°} = 1.41.$$

Thus the index of refraction of the glass must be *equal to or larger than* 1.41. If it were less, total internal reflection would not occur.

Example 5. What happens if the prism in Example 4 (assume that $n = 1.50$) is immersed in water ($n = 1.33$)? See Fig. 41–14*b*.

The new critical angle, given by Eq. 41–12, is

$$\sin \theta_c = \frac{n_2}{n_1} = \frac{1.33}{1.50} = 0.887,$$

which corresponds to $\theta_c = 62.5°$. The actual angle of incidence ($= 45°$) is less than this so that we do *not* have total internal reflection.

There is a reflected ray, with an angle of reflection of 45°, as Fig. 41–14*b* shows. There is also a refracted ray, with an angle of refraction given by

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

$$(1.50)(\sin 45°) = (1.33) \sin \theta_2,$$

which yields $\theta_2 = 52.9°$. ◀

Maxwell's equations permit us to calculate how the incident energy is divided between the reflected and the refracted beams. Figure 41–15 shows the theoretical

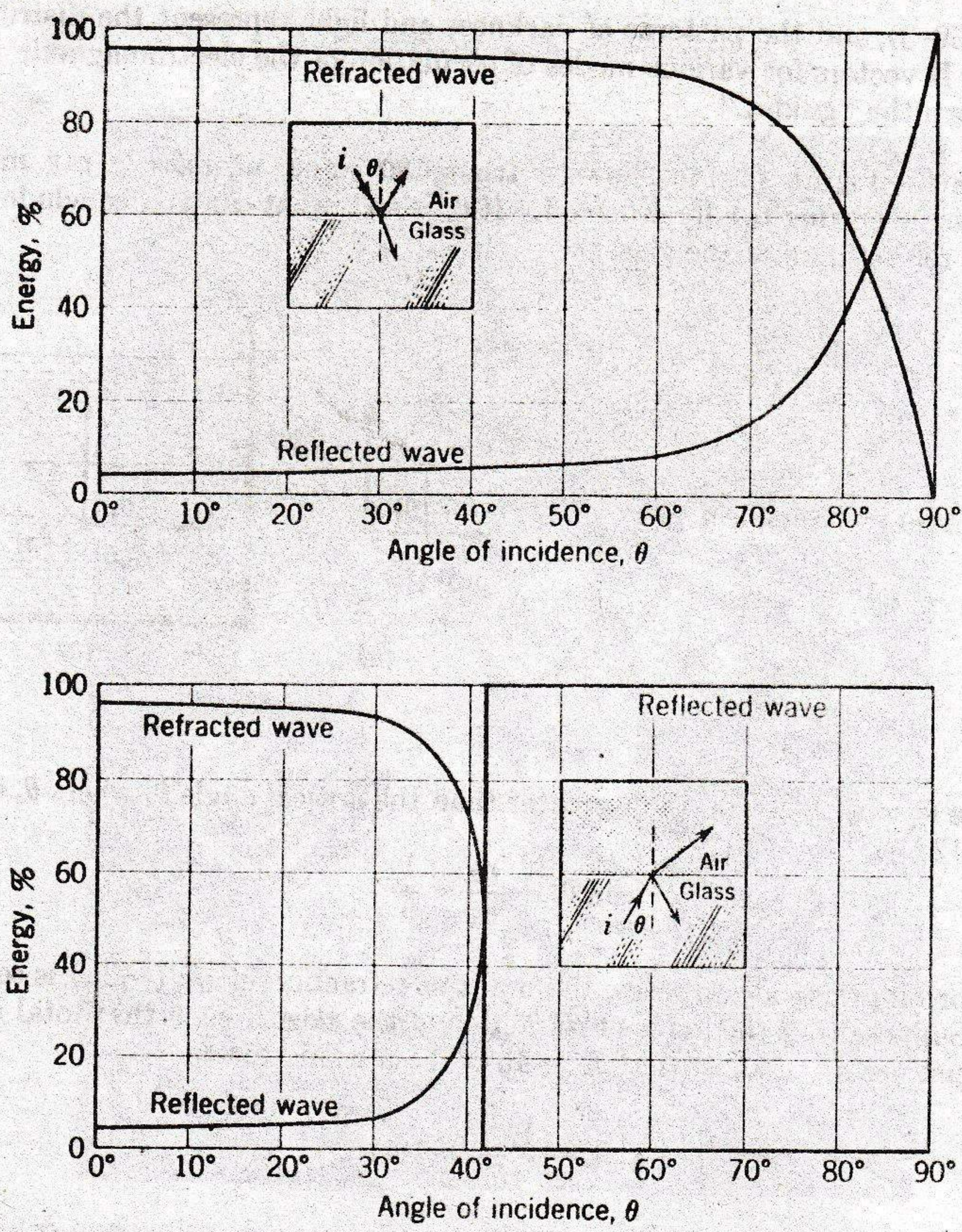


Fig. 41–15 (*a*) The per cent of the energy reflected and refracted when an incident wave in air falls on glass ($n = 1.50$). (*b*) The same for the incident wave in glass, showing total internal reflection.

prediction for (*a*) a light beam in air falling on a glass-air interface and (*b*) a light beam in glass falling on such an interface. Figure 41–15*a* shows that for angles of incidence up to about 50°, less than 10% of the light energy is reflected. At grazing incidence, however (that is, angles of incidence near 90°), the surface becomes an excellent reflector. We are all familiar with the high reflecting power of a wet road for light from automobile headlights that strikes near grazing incidence.

Figure 41–15*b* shows clearly that at a certain critical angle (41.8° in this case; see Eq. 41–12) *all* the light is reflected. For angles of incidence appreciably below this value, about 4% of the energy is reflected.

41–6 Fermat's Principle

In 1650 Pierre Fermat discovered a remarkable principle which we often express today in these terms: *A light ray traveling from one point to another will follow a path such that, compared with nearby paths, the time required is either a minimum or a maximum or will remain unchanged (that is, it will be stationary).*

The laws of reflection and refraction can readily be derived from this principle.

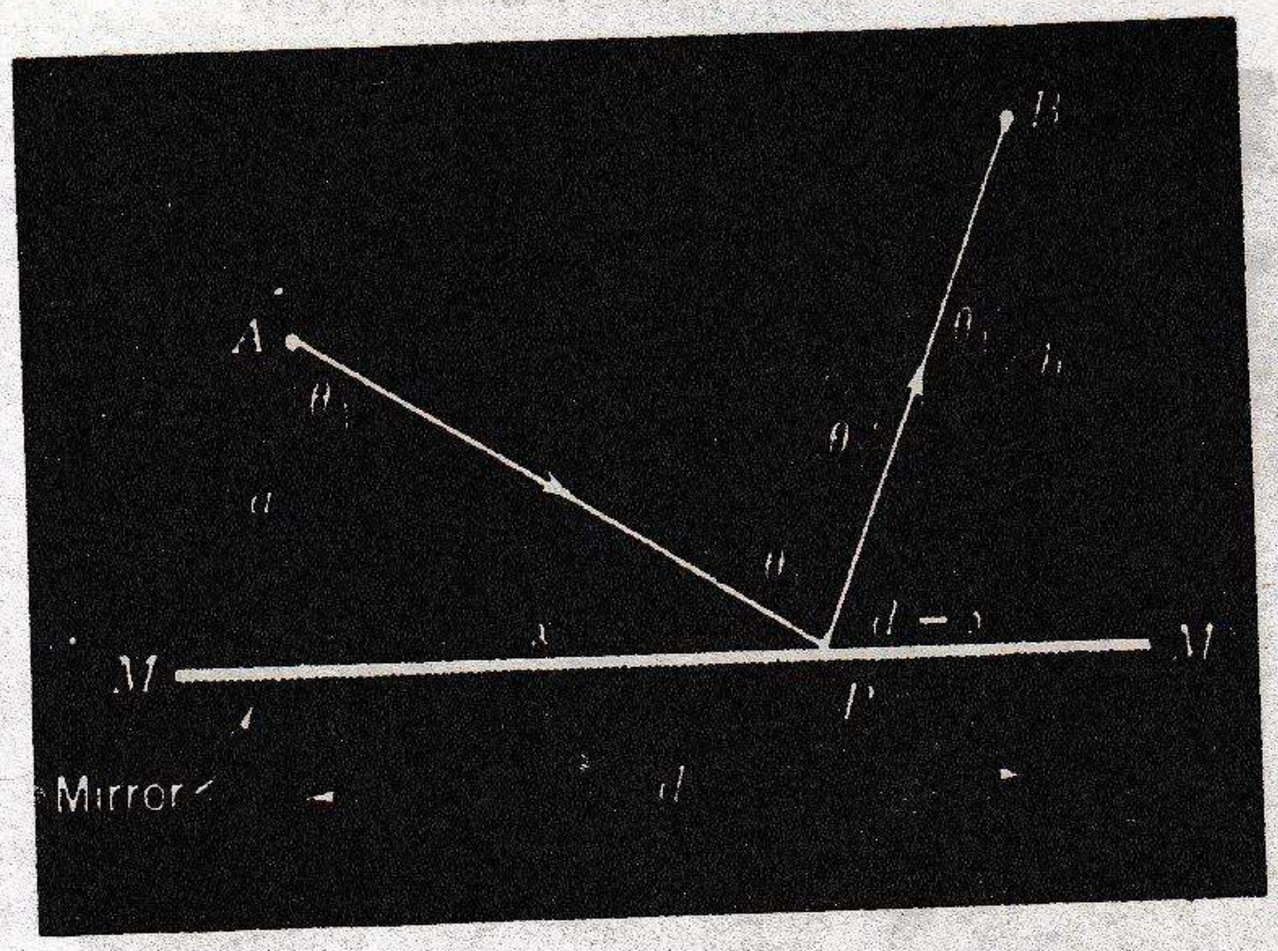


Fig. 41–16 A ray from A passes through B after reflection at P.

Figure 41–16 shows two fixed points A and B and a ray APB connecting them.* The total length l of this ray is

$$l = \sqrt{a^2 + x^2} + \sqrt{b^2 + (d - x)^2},$$

where x locates the point P at which the ray touches the mirror.

According to Fermat's principle, P will have a position such that the time of travel of the light must be a minimum (or a maximum or must remain unchanged). Expressed in another way, the total length l of the ray must be a minimum (or a maximum or must remain unchanged). In either case, the methods of the calculus require that dl/dx be zero. Taking this derivative yields

$$\frac{dl}{dx} = (\tfrac{1}{2})(a^2 + x^2)^{-1/2}(2x) + \tfrac{1}{2}[b^2 + (d - x)^2]^{-1/2}(2)(d - x)(-1) = 0,$$

which can be rewritten as

$$\frac{x}{\sqrt{a^2 + x^2}} = \frac{d - x}{\sqrt{b^2 + (d - x)^2}}.$$

Comparison with Fig. 41–16 shows that this can be written as

$$\sin \theta_1 = \sin \theta_1',$$

or

$$\theta_1 = \theta_1',$$

which is the law of reflection.

To prove the law of refraction from Fermat's principle, consider Fig. 41–17, which shows two points A and B in two different media and a ray APB connecting them. The time t is given by

$$t = \frac{l_1}{v_1} + \frac{l_2}{v_2}.$$

Using the relation $n = c/v$ this can be written as

$$t = \frac{n_1 l_1 + n_2 l_2}{c} = \frac{l}{c}.$$

The quantity l $(= n_1 l_1 + n_2 l_2)$ is called the *optical path length* of the ray. Equation 41–11 (written as $\lambda = n\lambda_n$) shows that the optical path length is equal to the length

* We assume that ray APB lies in the plane of the figure; see Problem 20.

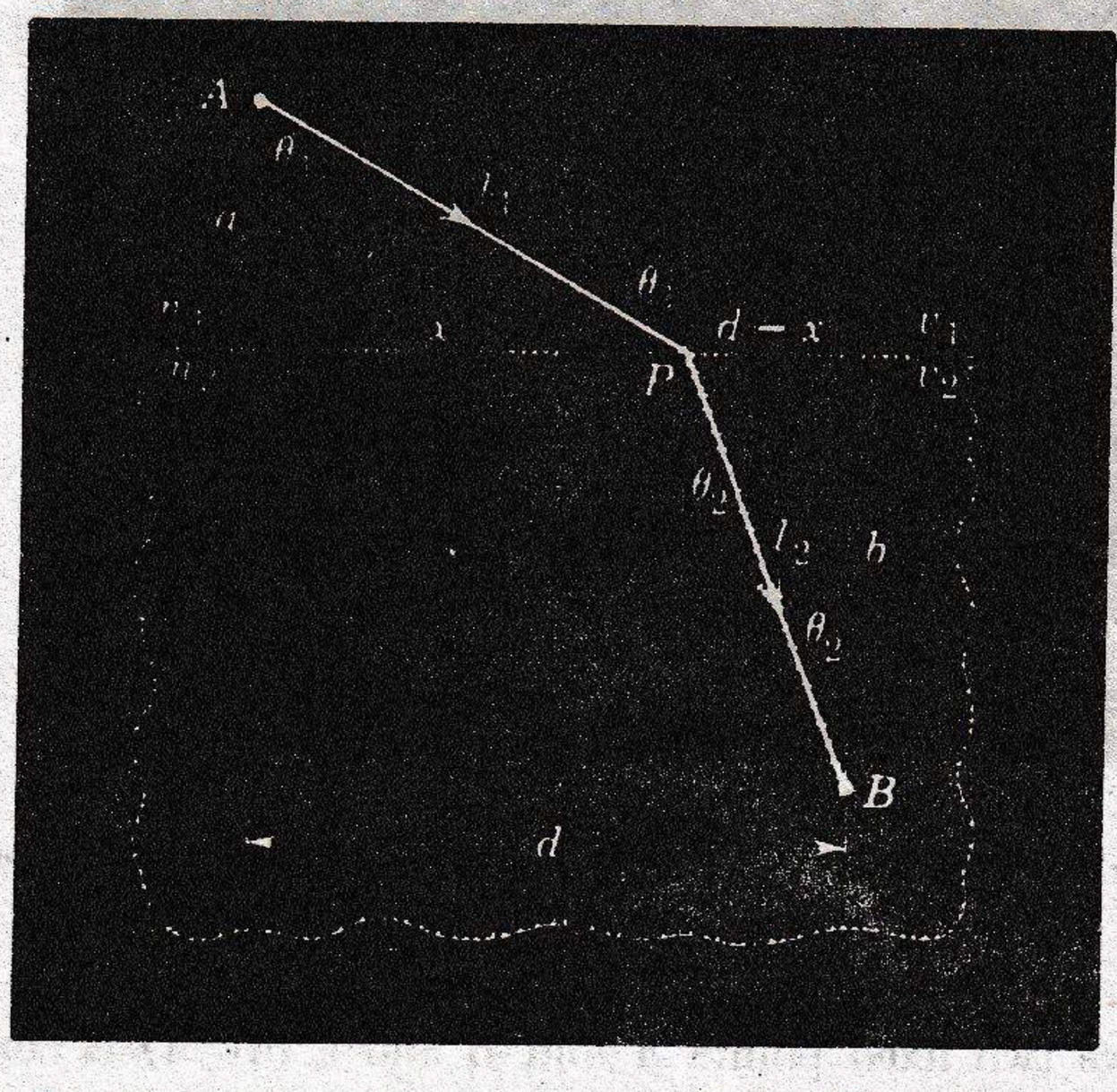


Fig. 41–17 A ray from A passes through B after refraction at P.

that this same number of waves would have if the medium were a vacuum. The optical path length must not be confused with the geometrical path length, which is $l_1 + l_2$.

Fermat's principle requires that l be a minimum (or a maximum or must remain unchanged) which, in turn, requires that x be so chosen that $dl/dx = 0$. The optical path length is

$$l = n_1 l_1 + n_2 l_2 = n_1 \sqrt{a^2 + x^2} + n_2 \sqrt{b^2 + (d - x)^2}.$$

Differentiating yields

$$\frac{dl}{dx} = n_1(\tfrac{1}{2})(a^2 + x^2)^{-1/2}(2x) + n_2(\tfrac{1}{2})[b^2 + (d - x)^2]^{-1/2}(2)(d - x)(-1) = 0,$$

which can be rewritten as

$$n_1 \frac{x}{\sqrt{a^2 + x^2}} = n_2 \frac{d - x}{\sqrt{b^2 + (d - x)^2}}.$$

Comparison with Fig. 41–17 shows that this, in turn, can be written as

$$n_1 \sin \theta_1 = n_2 \sin \theta_2,$$

which is the law of refraction.

In each of the examples of this section the time required, or, what is equivalent, the optical path length, proves to be a *minimum*. Problem 22 describes a case in which it may be a maximum, a minimum, or stationary.

QUESTIONS

1. Discuss the propagation of spherical and of cylindrical waves, using Huygens' principle.
2. Does Huygens' principle apply to sound waves in air?
3. If Huygens' principle predicts the laws of reflection and refraction, why is it necessary or desirable to view light as an electromagnetic wave, with all its attendant complexity?

4. Would you expect sound waves to obey the laws of reflection and of refraction obeyed by light waves?

5. A street light, viewed by reflection across a body of water in which there are ripples, appears very elongated. Explain.

6. The light beam in Fig. 41–1*a* is broadened on entering the water. Explain.

7. By what per cent does the speed of blue light in fused quartz differ from that of red light?

8. Can (*a*) reflection phenomena or (*b*) refraction phenomena be used to determine the wavelength of light?

9. How did Foucault's measurement of the speed of light in water decide between the wave and the particle theories of light?

10. Why does a diamond "sparkle" more than a glass imitation cut to the same shape?

11. Is it plausible that the wavelength of light should change in passing from air into glass but that its frequency should not? Explain.

12. How can one determine the indices of refraction of the media in Table 41–1 relative to water, given the data in that table?

13. You are given a cube of glass. How can you find the speed of light (from a sodium light source) in this cube?

14. Describe and explain what a fish sees as he looks in various directions above his "horizon."

15. Design a periscope, taking advantage of total internal reflection. What are the advantages compared with silvered mirrors?

16. What characteristics must a material have in order to serve as an efficient "light pipe"?

17. Can the optical path length between two points ever be less than the geometrical path length between those points?

PROBLEMS

1. One end of a stick is dragged through water at a speed v which is greater than the speed u of water waves. Applying Huygens' construction to the water waves, show that a conical wavefront is set up and that its half-angle α is given by

$$\sin \alpha = u/v.$$

This is familiar as the bow wave of a ship or the shock wave caused by an object moving through air with a speed exceeding that of sound, as in Fig. 20–12.

2. Prove that if a mirror is rotated through an angle α, the reflected beam is rotated through an angle 2α. Is this result reasonable for $\alpha = 45°$?

3. What is the speed in fused quartz of light of wavelength 5500 A; see Fig. 41–2?

4. The wavelength of yellow sodium light in air is 5890 A. (*a*) What is its frequency? (*b*) What is its wavelength in glass whose index of refraction is 1.52? (*c*) From the results of (*a*) and (*b*) find its speed in this glass.

5. The speed of yellow sodium light in a certain liquid is measured to be 1.92×10^8 meters/sec. What is the index of refraction of this liquid, with respect to air, for sodium light?

6. Suppose that the speed of light in air has been measured with an uncertainty of, say, 1 km/sec. In calculating the speed in vacuum, suppose that it is not certain whether n for air is 1.00029 or 1.00030. (*a*) How much extra uncertainty is introduced into the calculated value for c? (*b*) Estimate how accurately n should be known for this purpose.

7. In Fig. 41–6 show by graphical ray tracing, using a protractor, that if θ for the incident ray is *either* increased *or* decreased, the deviation angle ψ is increased. The symmetrical situation shown in this figure is called the *position of minimum deviation*.

8. A ray of light is incident normally on the face ab of a glass prism ($n = 1.52$), as shown in Fig. 41–18. (a) Assuming that the prism is immersed in air, find the largest value for the angle ϕ so that the ray is totally reflected at face ac. (b) Find ϕ if the prism is immersed in water.

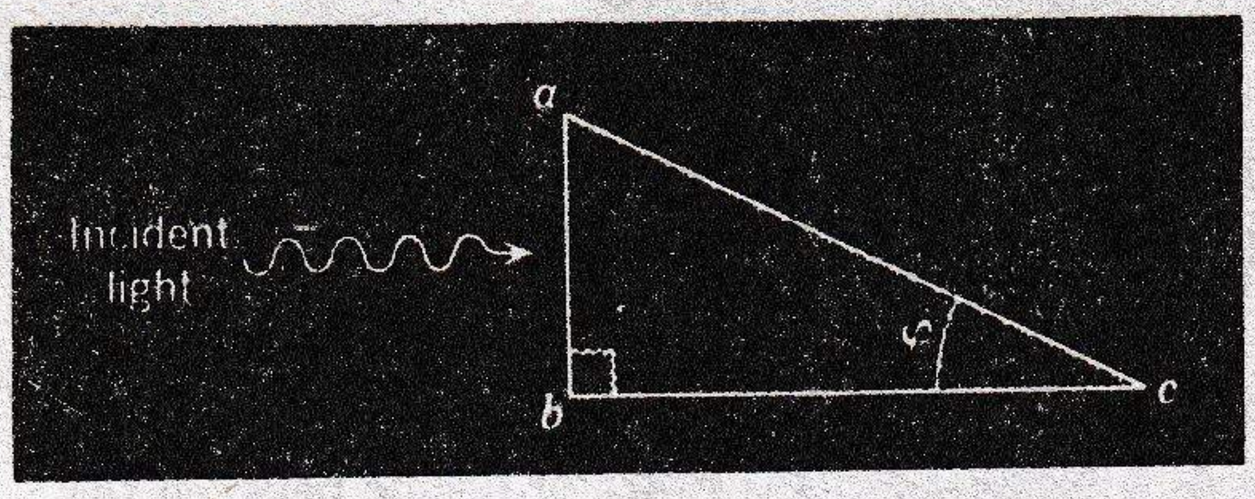


Fig. 41–18

9. Show that for a thin prism (ϕ small) and light not far from normal incidence (θ_1 small) the deviation angle is independent of the angle of incidence and is equal to $(n - 1)\,\phi$ (see Fig. 41–6).

10. A glass prism with an apex angle of 60° has $n = 1.60$. (a) What is the smallest angle of incidence for which a ray can enter one face of the prism and emerge from the other? (b) What angle of incidence would be required for the ray to pass through the prism symmetrically, as in Fig. 41–6?

11. A 60° prism is made of fused quartz. A ray of light falls on one face, making an angle of 45° with the normal. Trace the ray through the prism graphically with some care, showing the paths traversed by rays representing (a) blue light, (b) yellow-green light, and (c) red light. See Figs. 40–2 and 41–2.

12. Ptolemy, who lived at Alexandria toward the end of the first century A.D., gave the following measured values for the angle of incidence θ_1 and the angle of refraction θ_2 for a light beam passing from air to water:

θ_1	θ_2	θ_1	θ_2
10°	7°45′	50°	35°0′
20°	15°30′	60°	40°30′
30°	22°30′	70°	45°30′
40°	29°0′	80°	50°0′

Are these data consistent with Snell's law; if so, what index of refraction results? These data are interesting as the oldest recorded physical measurements.

13. Prove that a ray of light incident on tne surface of a sheet of plate glass of thickness t emerges from the opposite face parallel to its initial direction but displaced sideways, as in Fig. 41–19. Show that, for small angles of incidence θ, this displacement is given by

$$x = t\theta\,\frac{n - 1}{n}$$

where n is the index of refraction and θ is measured in radians.

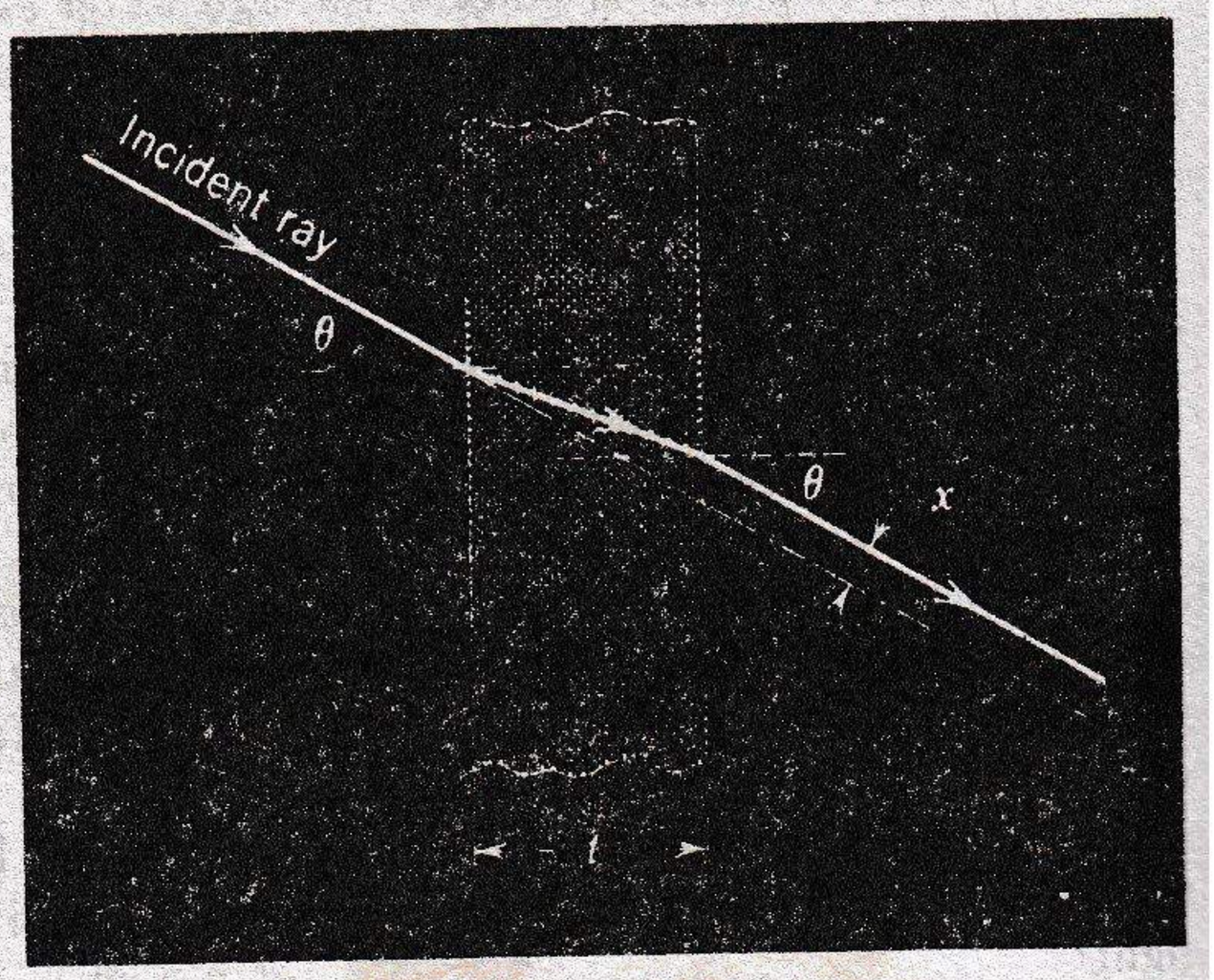


Fig. 41–19

14. A plane wave of white light traveling in fused quartz strikes a plane surface of the quartz, making an angle of incidence θ. Is it possible for the internally reflected beam to appear (a) bluish or (b) reddish? Roughly what value of θ must be used? (Hint: White light will appear bluish if wavelengths corresponding to red are removed from the spectrum.)

15. A point source of light is placed a distance h below the surface of a large deep lake. (a) Show that the fraction f of the light energy that escapes directly from the water surface

is independent of h and is given by

$$f = \frac{1}{2} - \frac{1}{2n}\sqrt{n^2 - 1}$$

where n is the index of refraction of water. (Note: Absorption within the water and reflection at the surface—except where it is total—have been neglected.) (*b*) Evaluate this fraction for $n = 1.33$.

16. Figure 41–20 shows a *constant-deviation prism*. Although made of one piece of glass, it is equivalent to two 30°–60°–90° prisms and one 45°–45°–90° prism. White light is incident in the direction i. θ_1 is changed by rotating the prism so that, in turn, light of any desired wavelength may be made to follow the path shown, emerging at r. Show that, if $\sin\theta_1 = \frac{1}{2}n$, then $\theta_2 = \theta_1$ and beams i and r are at right angles.

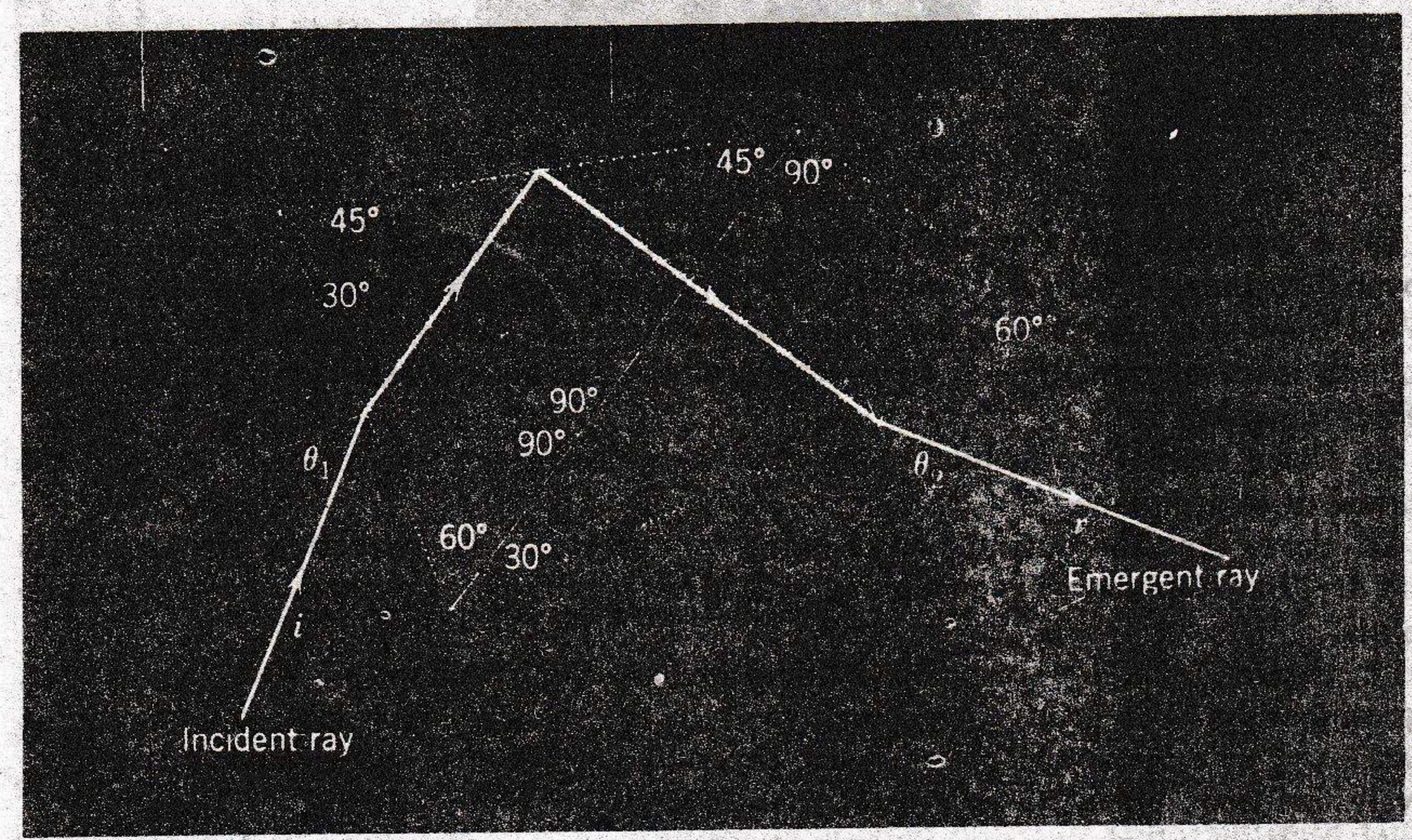


Fig. 41–20

17. A drop of liquid may be placed on a semicircular slab of glass as in Fig. 41–21. Show how to determine the index of refraction of the liquid by observing total internal reflection. The index of refraction of the glass is unknown and must also be determined. Is the range of indices of refraction that can be measured in this way restricted in any sense?

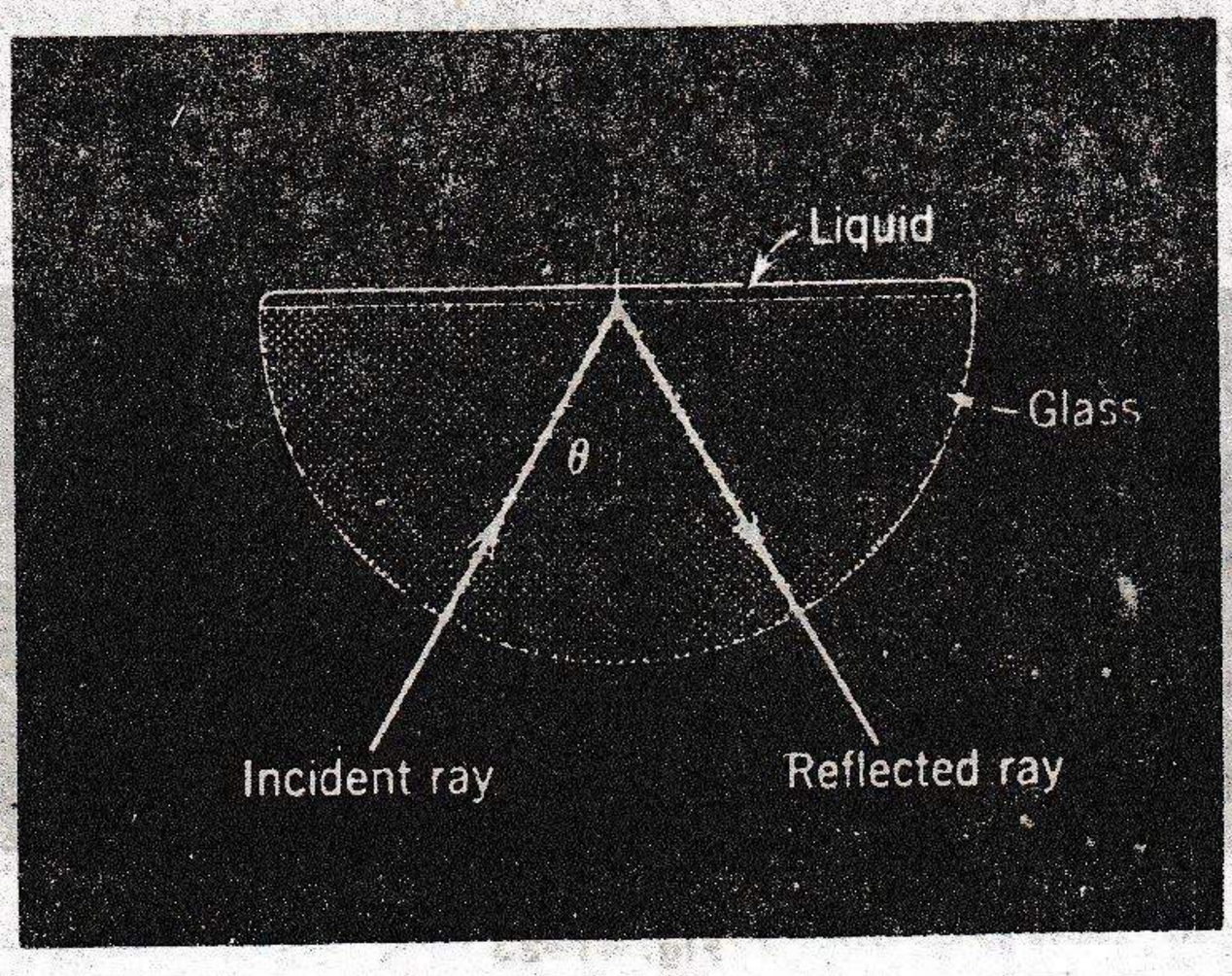


Fig. 41–21

18. A light ray falls on a square glass slab as in Fig. 41–22. What must the index of refraction of the glass be if total internal reflection occurs at the vertical face?

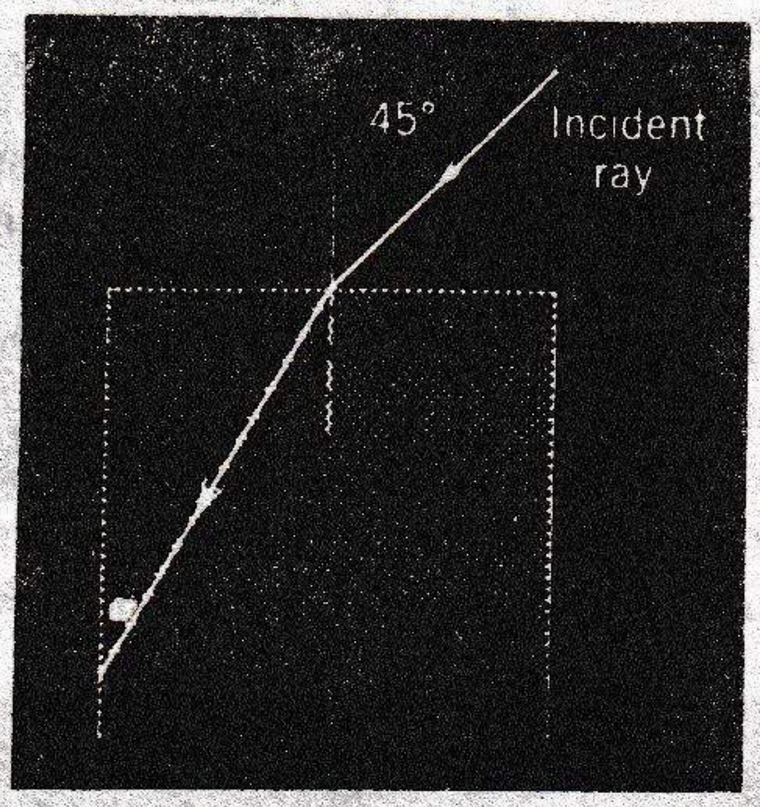


Fig. 41–22

19. A glass cube has a small spot at its center. What parts of the cube face must be covered to prevent the spot from being seen, no matter what the direction of viewing? What fraction of the cube surface must be so covered? Assume a cube edge of 1.0 cm and an index of refraction of 1.50. (Neglect the subsequent behavior of an internally reflected ray.)

20. Using Fermat's principle, prove that the reflected ray, the incident ray, and the normal lie in one plane.

21. Prove that the optical path lengths for reflection and refraction in Figs. 41–16 and 41–17 are minima when compared with other nearby paths connecting the same two points.

22. Figure 41–23 shows two points A and B connected by a light ray AvB. Show that, by comparison with nearby rays, such as AcB, the ray AvB represents a minimum, a stationary, or a maximum optical path, depending on whether the distance a is respectively less than, equal to, or greater than, the quantity

$$R\,\frac{n+1}{n-1}.$$

R is the radius of curvature of the spherical surface and n is the index of refraction of the medium to the right of this surface. (Hint: Express the optical path length l of the arbitrary path AcB in terms of the angle ϕ in Fig. 41–23. Use approximations that are appropriate when ϕ is small, since AcB must be a *nearby* path. Take the first and second derivatives of l with respect to ϕ.)

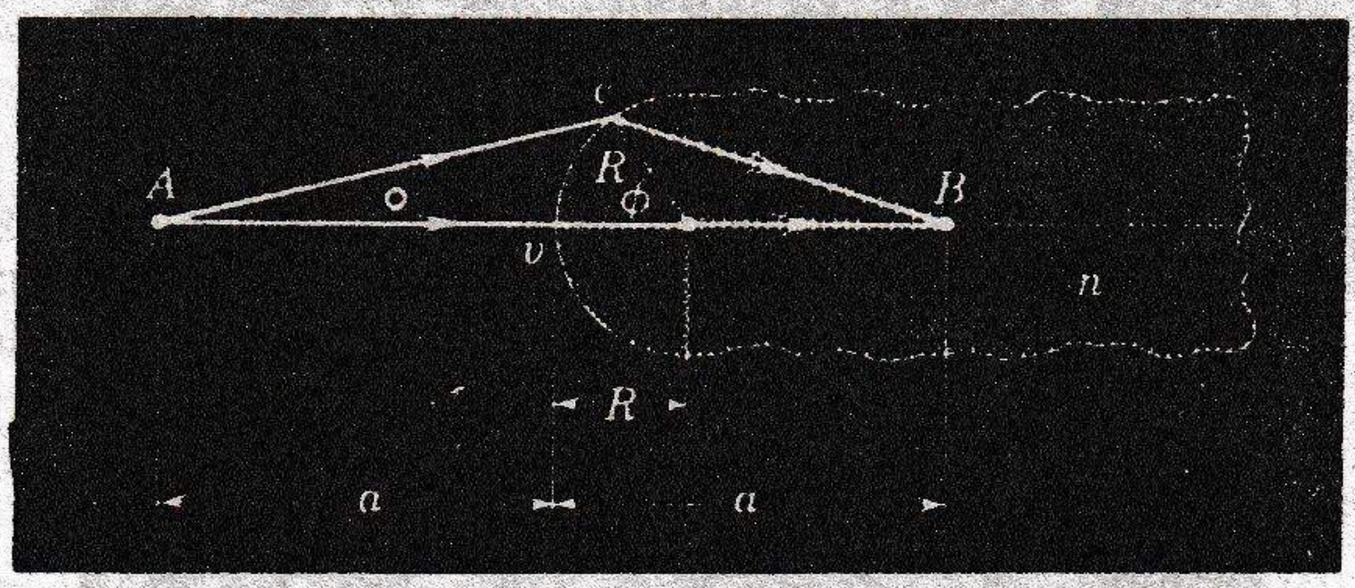


Fig. 41–23

Reflection and Refraction—Spherical Waves and Spherical Surfaces

CHAPTER 42

42–1 Geometrical Optics and Wave Optics

In Chapter 41 we described the reflection and refraction of plane waves at plane surfaces. In this chapter we consider the more general case of spherical waves falling on spherical reflecting and refracting surfaces. All of the results of Chapter 41 will emerge as special cases of the results of this chapter, since a plane can be viewed as a spherical surface with an infinite radius of curvature.

Both in Chapter 41 and in this chapter we make extensive use of *rays*. Although a ray is a convenient construction, it proves impossible to isolate one physically. Figure 42–1*a* shows schematically a plane wave of wavelength λ falling on a slit of width $a = 5\lambda$. We find that the light flares out into the geometrical shadow of the slit, a phenomenon called *diffraction*. Figures 42–1*b* ($a = 3\lambda$) and 42–1*c* ($a = \lambda$) show that diffraction becomes more pronounced as $a/\lambda \rightarrow 0$ and that attempts to isolate a single ray from the incident plane wave are futile.

Figure 42–2 shows water waves in a shallow *ripple tank*, produced by tapping the water surface periodically and automatically with the edge of a flat stick. We see that the plane wave so generated flares out by diffraction when it encounters a gap in a barrier placed across it. Diffraction is characteristic of waves of all types. We can hear around corners, for example, because of the diffraction of sound waves.

The diffraction of waves at a slit (or at an obstacle such as a wire) is expected from Huygens' principle. Consider the portion of the wavefront that arrives at the position of the slit in Fig. 42–1. Every point on it can be

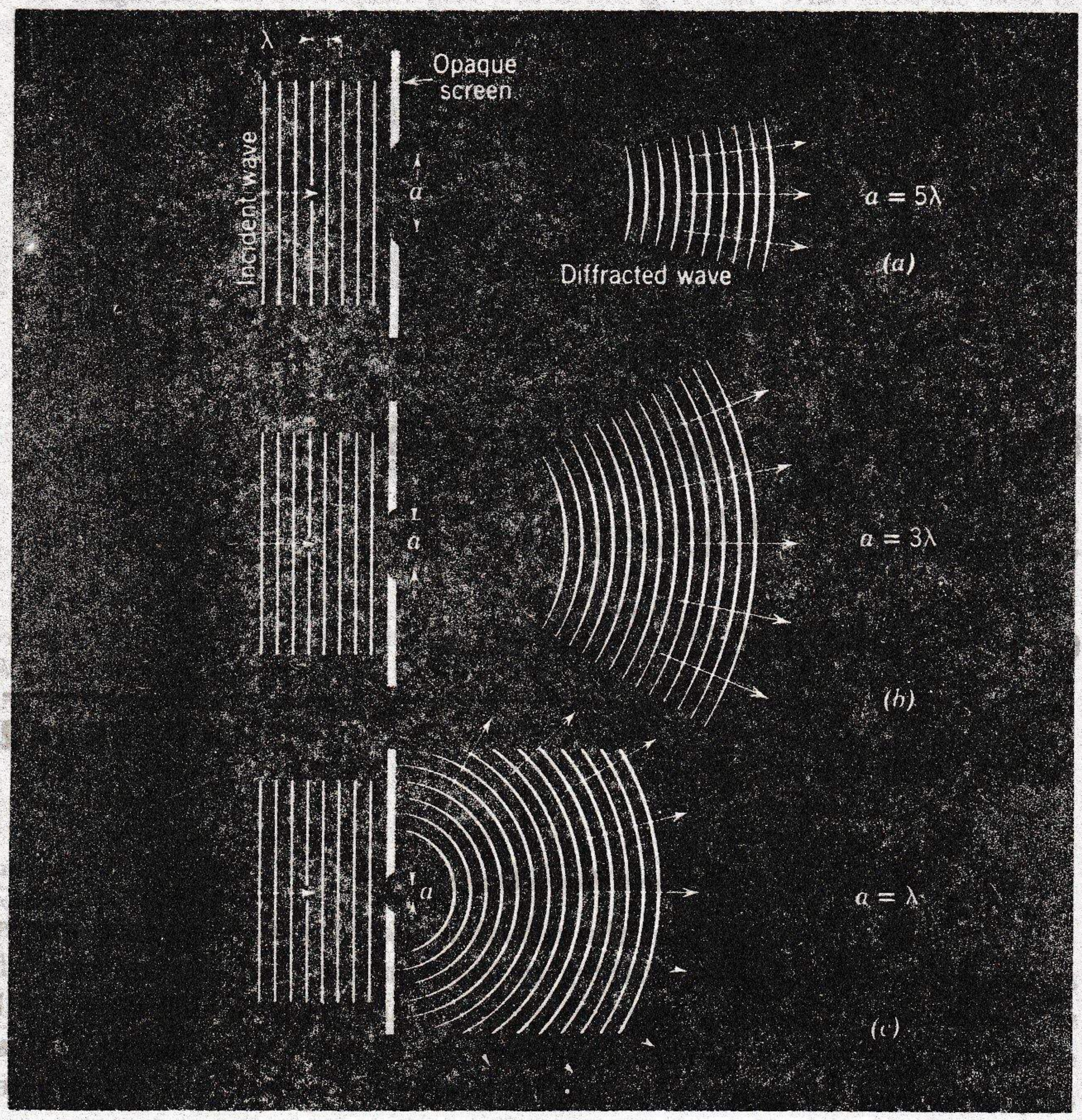


Fig. 42–1 An attempt to isolate a ray by reducing the slit width a fails because of diffraction, which becomes more pronounced as a/λ approaches zero.

viewed as the site of an expanding spherical Huygens' wavelet. The "bending" of light into the region of the geometrical shadow is associated with the blocking off of Huygens' wavelets from those parts of the incident wavefront that lie behind the slit edges.

Figure 42–3 was made by allowing parallel light to fall on a slit placed 50 cm in front of a photographic plate. In Fig. 42–3*a* the slit width was about 6×10^{-3} mm. The central band of light is much wider than this, showing that light has "flared out" into the geometric shadow of the slit. In addition, many secondary maxima, omitted from Fig. 42–1 for simplicity, appear. Figure 42–3*b* shows what happens when the slit width is *reduced* by a factor of two. The central maximum becomes *wider*, in agreement with Fig. 42–1. Figure 42–3*c* shows the effect of reducing the slit width by an additional factor of 7, to 4×10^{-4} mm. The central maximum is now much wider and the secondary maxima, whose intensities relative to the central

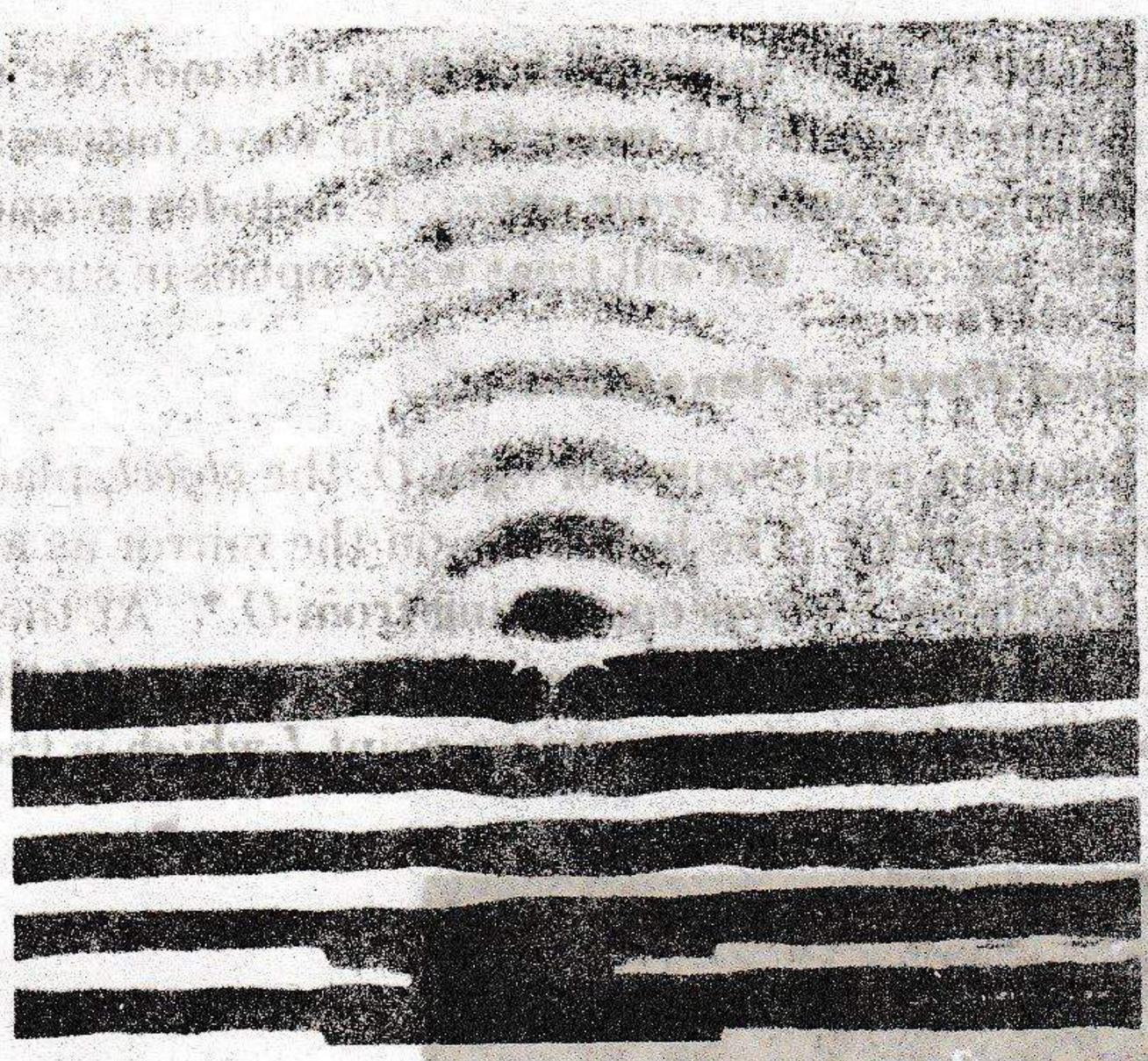


Fig. 42–2 Diffraction of water waves at a slit in a ripple tank. Note that the slit width is about the same size as the wavelength. (Courtesy of Educational Services Incorporated.)

maximum have been deliberately overemphasized by long exposure, are very evident.

Diffraction can be ignored if the ratio a/λ is large enough, a being a measure of the smallest sideways dimension of the slit or obstacle. If $a \gg \lambda$, light appears to travel in straight lines which can be represented by rays that obey the laws of reflection (Eq. 41–1) and refraction (Eq. 41–2). In Chapter 41 this condition, called *geometrical optics*, prevailed, the lateral dimensions of all mirrors, prisms, etc., being much greater than the wavelength. We assume in this chapter also that the conditions for geometrical optics are satisfied.

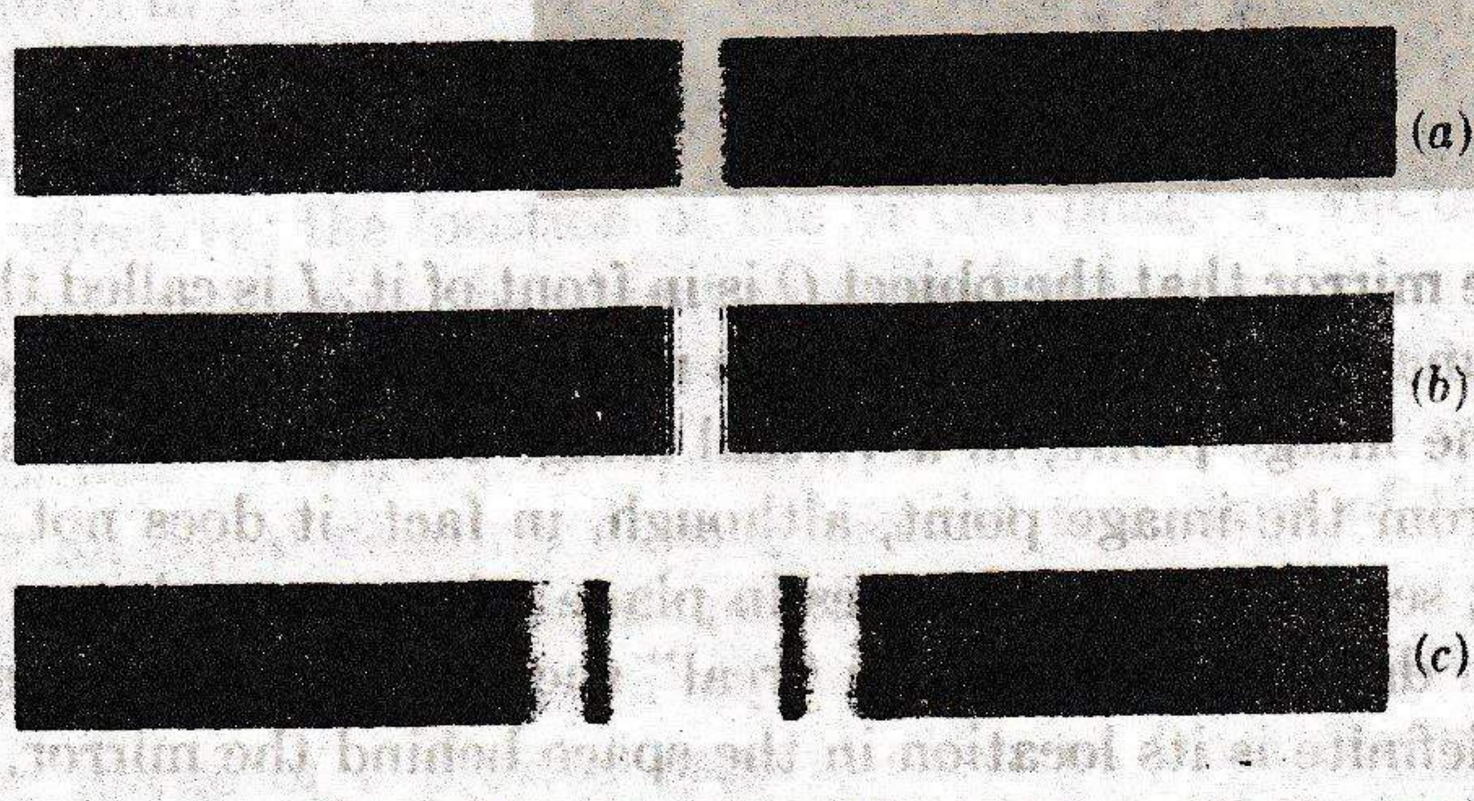


Fig. 42–3 (*a*) The intensity of light diffracted from a slit of width $a \cong 6 \times 10^{-3}$ mm and falling on a screen 50 cm beyond. (*b*) The slit width is reduced by a factor of two. (*c*) The slit width is further reduced by an additional factor of seven. Note that secondary maxima, made prominent in this case by deliberate overexposure, appear on either side of the central maximum. These secondary maxima have been omitted from Fig. 42–1 for simplicity.

If the requirement for geometrical optics is not met, we cannot describe the behavior of light by rays but must take its wave nature specifically into account. This subject is called *wave optics;* it includes geometrical optics as an important limiting case. We will treat wave optics in succeeding chapters.

42–2 Spherical Waves—Plane Mirror

Figure 42–4 shows a point source of light O, the *object*, placed a distance o in front of a plane mirror. The light falls on the mirror as a spherical wave represented in the figure by rays emanating from O.* At the point at which each ray strikes the mirror we construct a reflected ray. If the reflected rays are extended backward, they intersect in a point I which is the same distance

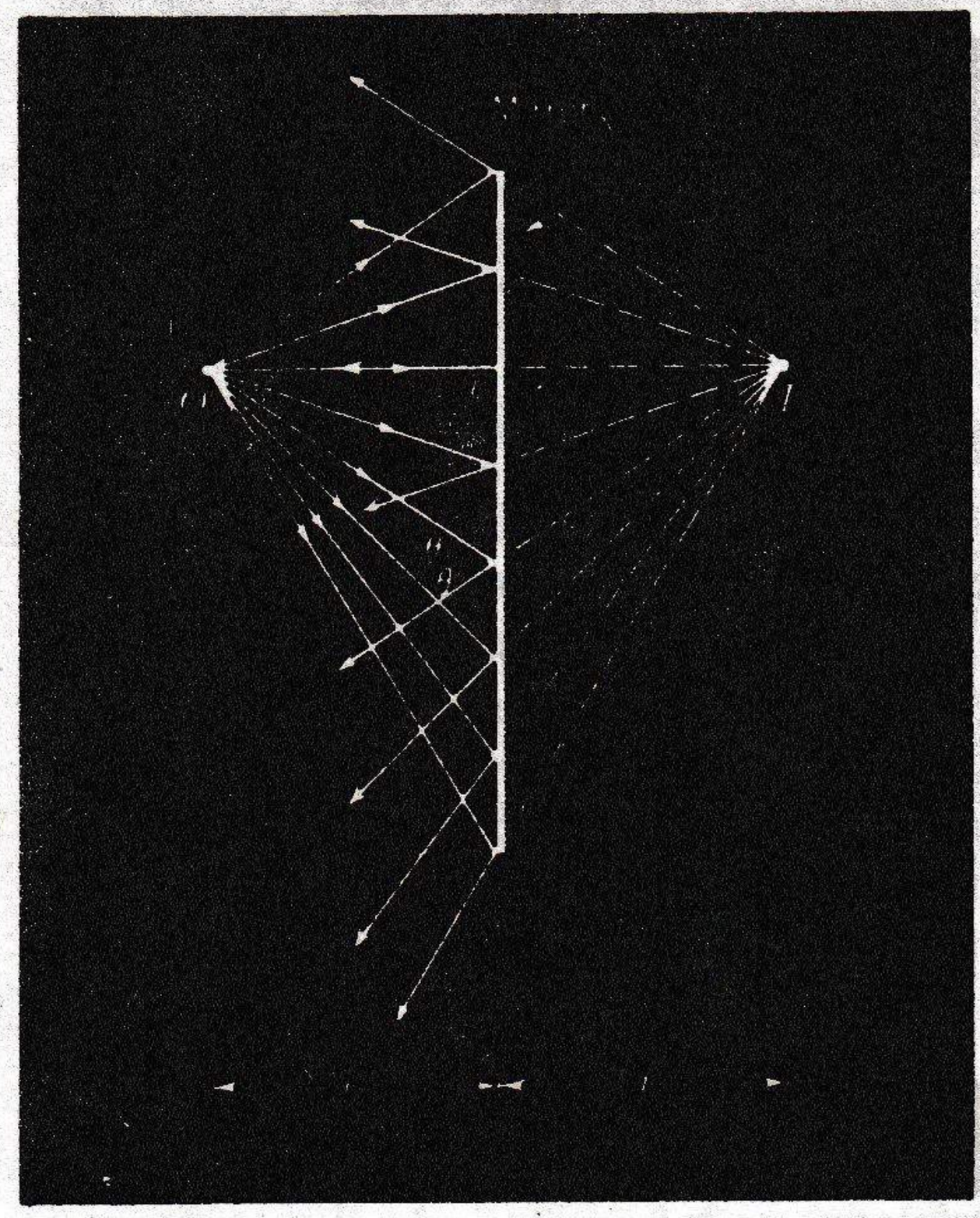


Fig. 42–4 A point object O forms a virtual image I in a plane mirror. The rays appear to emanate from I, but actually light energy does not pass through this point.

behind the mirror that the object O is in front of it; I is called the *image* of O.

Images may be *real* or *virtual*. In a real image light energy actually passes through the image point; in a virtual image the light behaves as though it diverges from the image point, although, in fact, it does not pass through this point; see Fig. 42–4. Images in plane mirrors are always virtual. We know from daily experience how "real" such a virtual image appears to be and how definite is its location in the space behind the mirror, even though this space may, in fact, be occupied by a brick wall.

Figure 42–5 shows two rays from Fig. 42–4. One strikes the mirror at v, along a perpendicular line. The other strikes it at an arbitrary point a, making an angle of incidence θ with the normal at that point. Elementary

* In our discussion of reflection from mirrors in Chapter 41 (see Fig. 41–8) we assumed an incident *plane* wave; the incident rays are parallel to each other in that case.

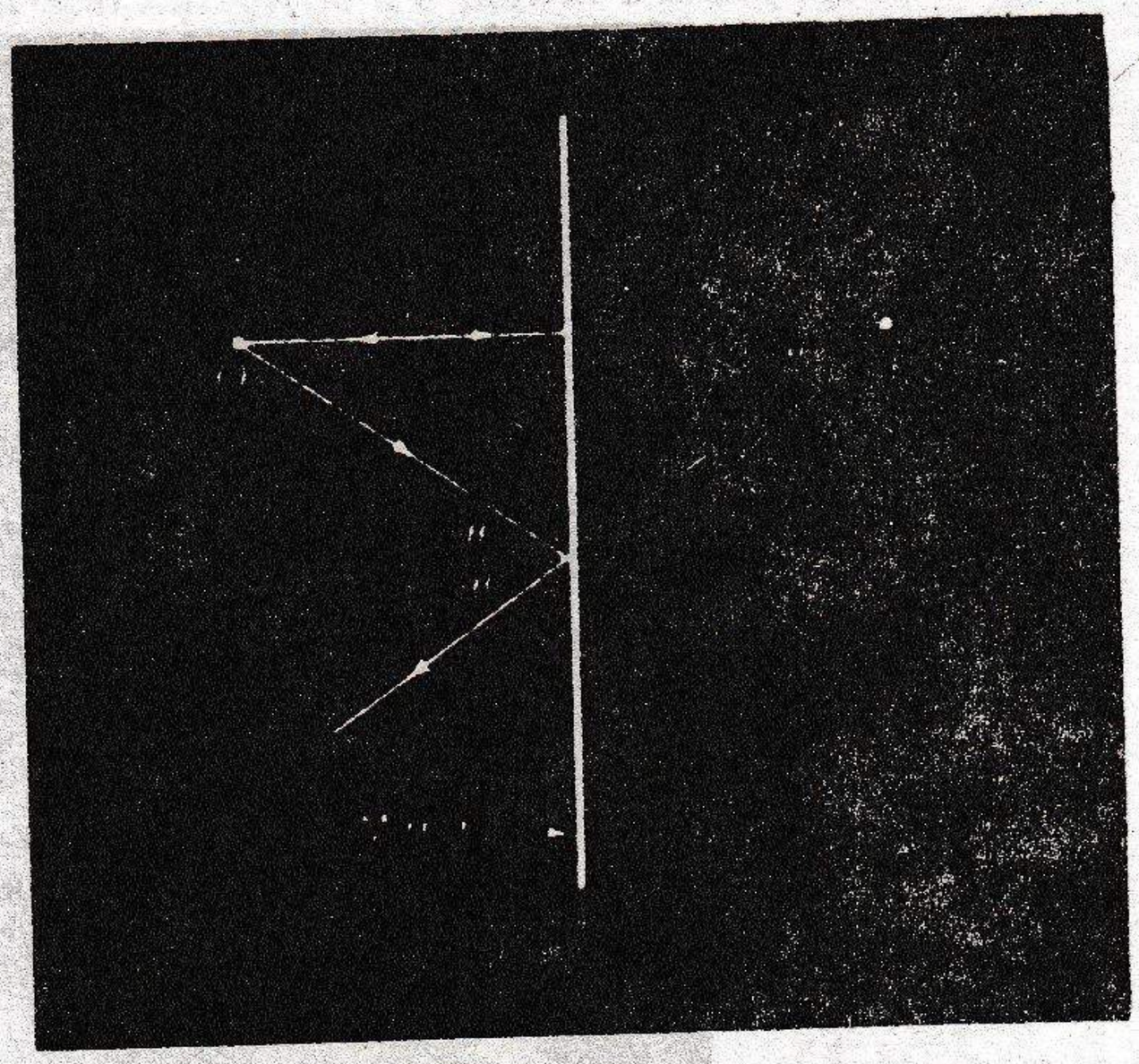


Fig. 42–5 Two rays from Fig. 42–4; ray *Oa* makes an arbitrary angle θ with the normal.

geometry shows that the angles *aOv* and *aIv* are also equal to θ. Thus the right triangles *aOva* and *aIva* are congruent and

$$o = -i, \tag{42–1}$$

in which the minus sign is arbitrarily introduced to show that I and O are on opposite sides of the mirror. Equation 42–1 does not involve θ, which means that *all* rays striking the mirror pass through I when extended backward, as we have seen above. Beyond assuming that the mirror is truly plane and that the conditions for geometrical optics hold, we have made no approximations in deriving Eq. 42–1. A point object produces a point image in a plane mirror, with $o = -i$, no matter how large the angle θ in Fig. 42–5.

Because of the finite diameter of the pupil of the eye, only rays that lie fairly close together can enter the eye after reflection at a mirror. For the eye position shown in Fig. 42–6 only a small patch of the mirror near point *a* is effective in forming the image; the rest of the mirror may be covered up or removed. If the eye is moved to another location, a different patch of the mirror will be effective; the location of the virtual image I will remain unchanged, however, as long as the object remains fixed.

If the object is an extended source such as the head of a person, a virtual image is also formed. From Eq. 42–1, every point of the source has an image point that lies an equal distance directly behind the plane of the mirror. Thus the image reproduces the object point by point.

Images in plane mirrors differ from objects in that left and right are interchanged. The image of a printed page is different from the page itself. Similarly, if a top is made to spin clockwise, the image, viewed in a vertical mirror, will seem to spin counterclockwise. Figure 42–7 shows an image of a left hand, constructed by using point-by-point application of Eq. 42–1; the image has the symmetry of a right hand.*

* See "The Overthrow of Parity" by Philip Morrison, *Scientific American*, April, 1957, for a discussion of the distinction in nature between right and left.

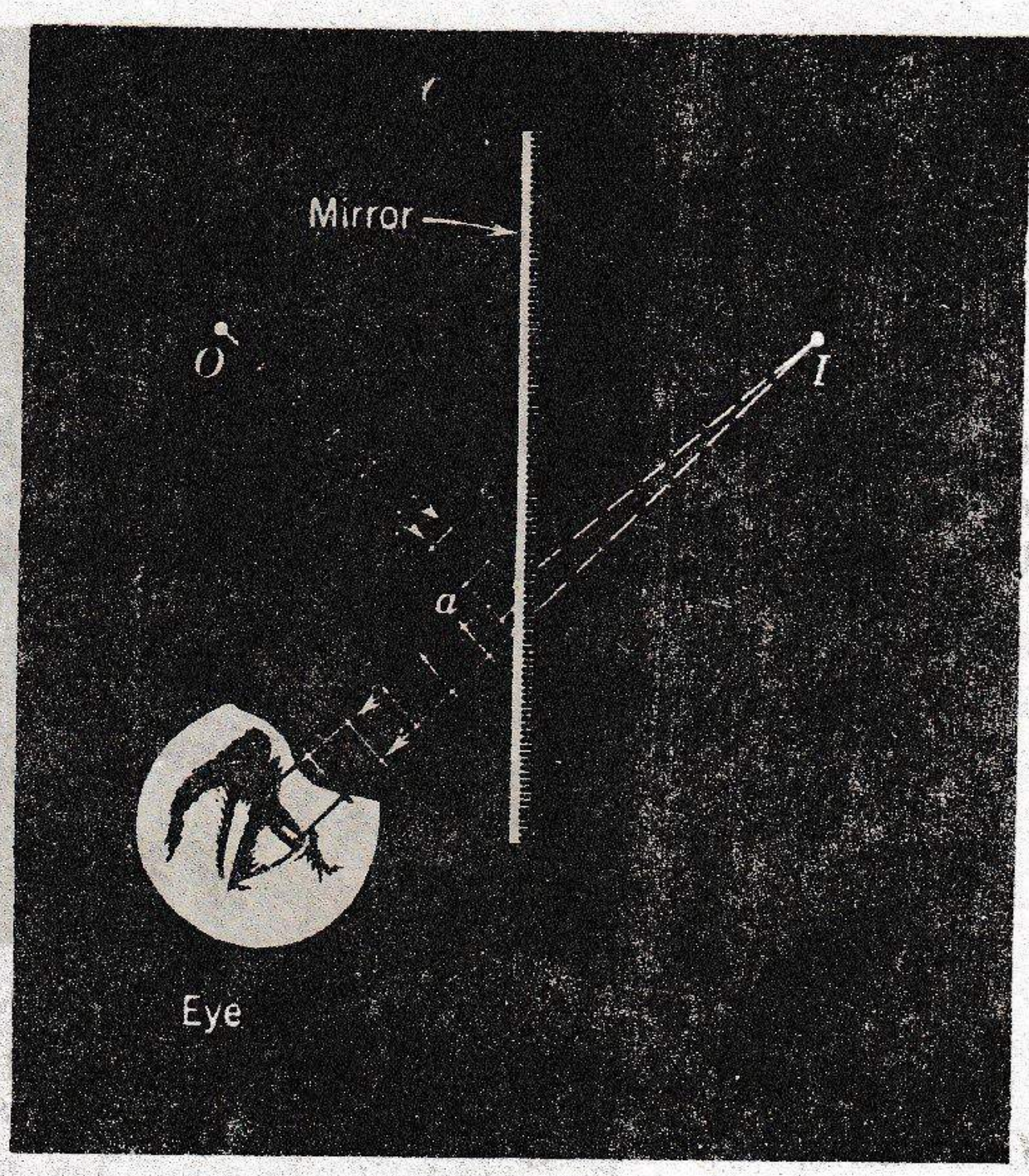


Fig. 42–6 A pencil of rays from O enters the eye after reflection at the mirror. Only a small portion of the mirror near a is effective. The small arcs represent portions of spherical wavefronts.

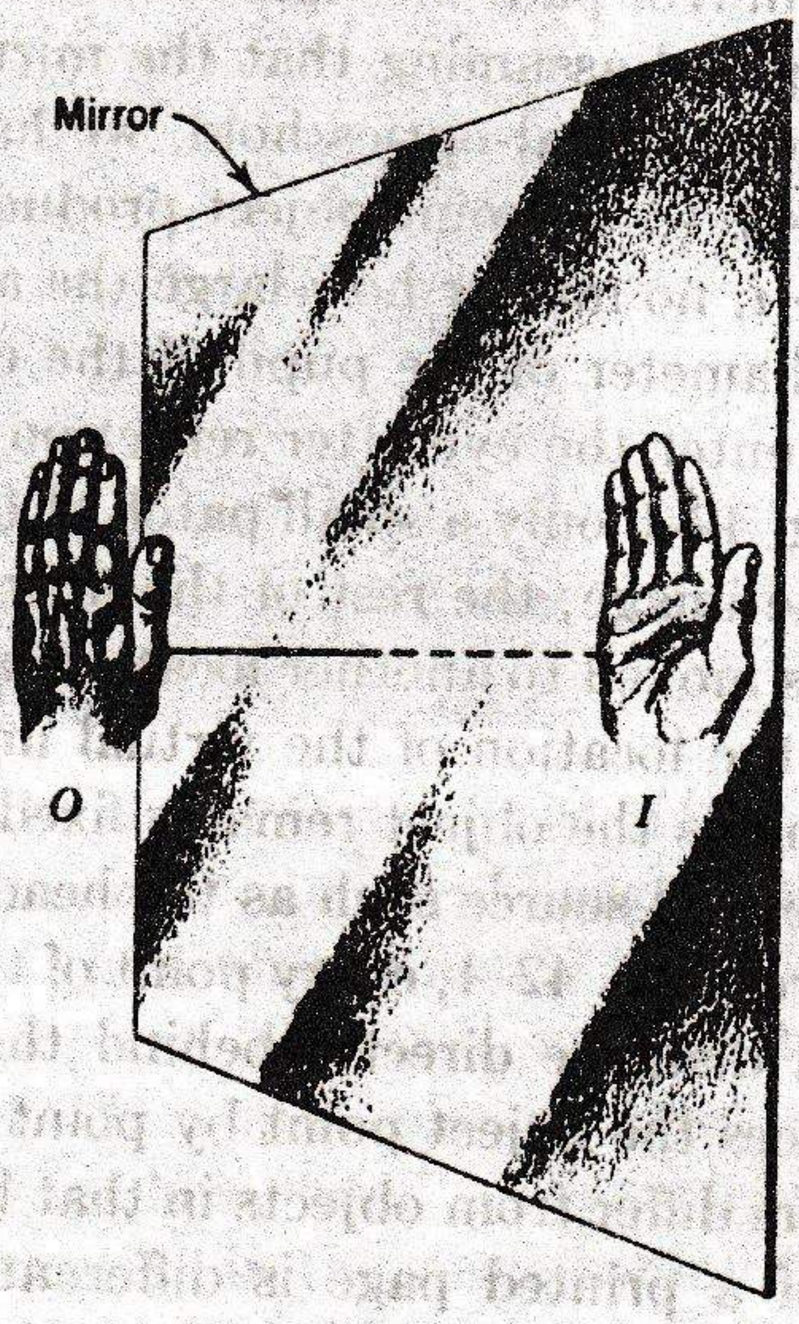


Fig. 42–7 A plane mirror reverses right and left.

▶ **Example 1.** How tall must a vertical mirror be if a person 6 ft high is to be able to see his entire length? Assume that his eyes are 4 in. below the top of his head.

Figure 42–8 shows the paths followed by light rays leaving the top of the man's head and the tips of his toes. These rays, chosen so that they will enter the eye e after reflection, strike the vertical mirror at points a and b, respectively. The mirror need occupy only the region between these two points. Calculation shows that b is 2 ft, 10 in. and a is 5 ft, 10 in. above the floor. The length of the mirror is thus 3 ft, or half the height of the person. Note that this height is independent of the distance between the person and the mirror. Mirrors that extend below point b show reflections of the floor between the person and the mirror.

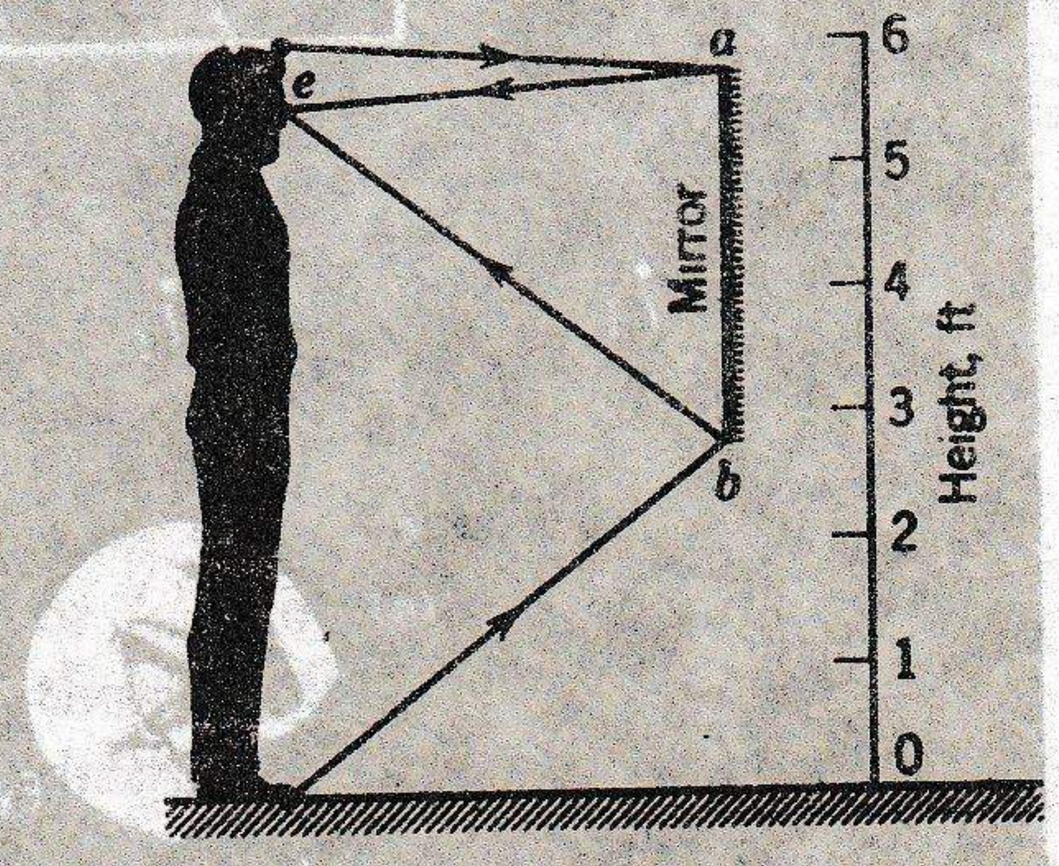


Fig. 42–8 Example 1. A person can view his full-length image in a mirror that is only half his height.

Example 2. Two plane mirrors are placed at right angles, and a point object O is located on the perpendicular bisector, as shown in Fig. 42–9*a*. Locate the images. Images I_1 and I_2 are formed in mirrors ab and cd, respectively. There is also a third image; it may be considered to be the image of I_1 in mirror cd or the image of I_2 in mirror ab. The three images and the object O lie on a circle whose center is on the line of intersection of the mirrors and whose plane is at right angles to that line.

In viewing I_3 the light entering the observer's eye is reflected *twice* after leaving the source. Figure 42–9*b* shows a typical bundle of rays. In viewing I_1 or I_2, the light is reflected only once, as in Fig. 42–6. ◀

42–3 Spherical Waves—Spherical Mirror

In Fig. 42–10 a spherical light wave from a point object O falls on a concave spherical mirror whose radius of curvature is r.* A line through O and the center of curvature C makes a convenient reference axis.

A ray from O that makes an arbitrary angle α with this axis intersects the axis at I after reflection from the mirror at a. A ray that leaves O along the axis will be reflected back along itself at v and will also pass through I. Thus, for these two rays at least, I is the image of O; it is a *real* image because light energy actually passes through I. Let us find the location of I.

A useful theorem is that the exterior angle of a triangle is equal to the sum of the two opposite interior angles. Applying this to triangles $OaCO$ and

---

* A spherical shell, viewed from inside, is everywhere *concave*; viewed from outside it is everywhere *convex*. In this chapter concave and convex will always be judged from the point of view of an observer sighting along the direction of the incident light.

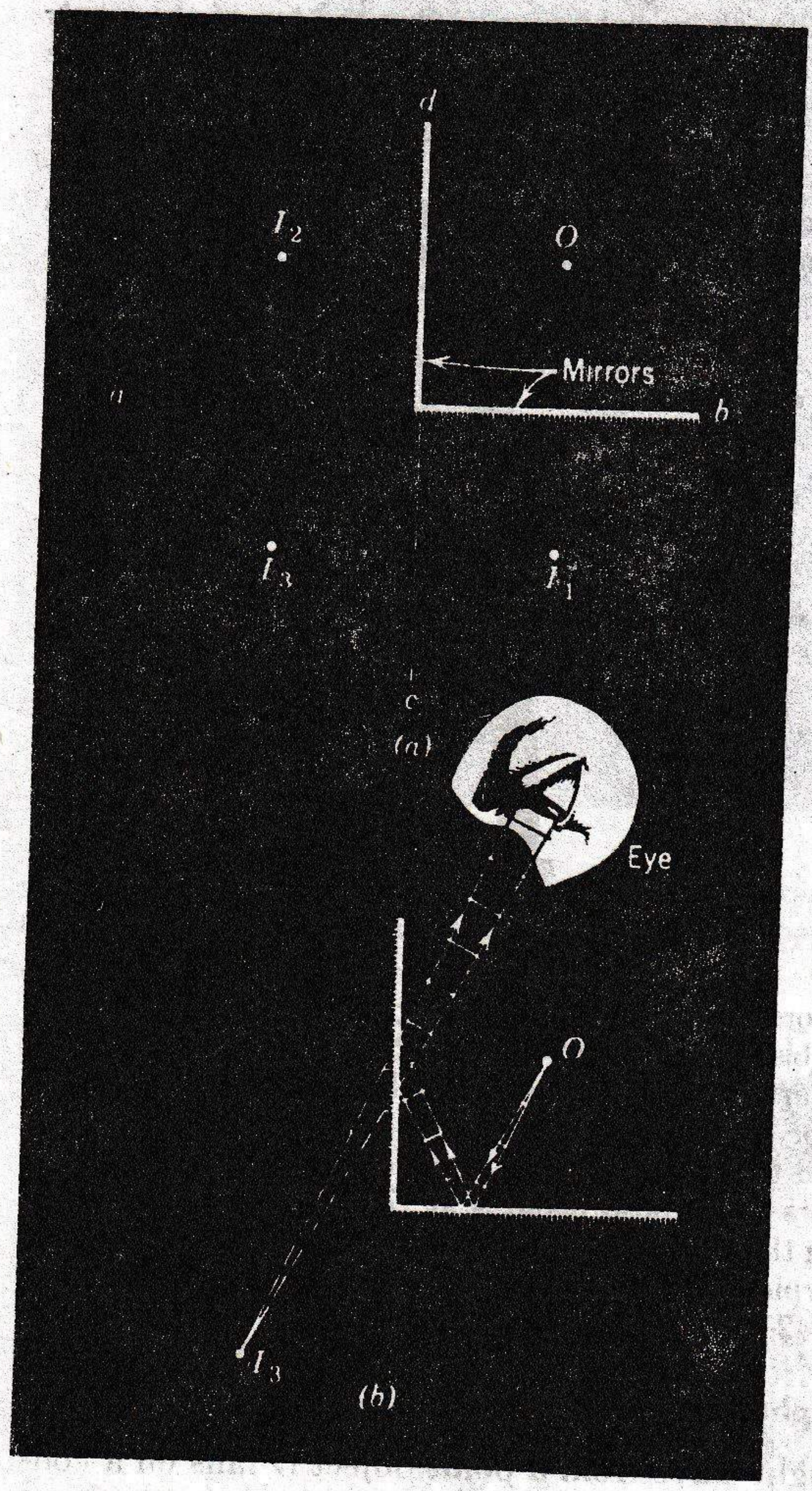


Fig. 42–9 Example 2. (*a*) Object *O* has three virtual images. (*b*) A typical bundle of rays used to view I_3.

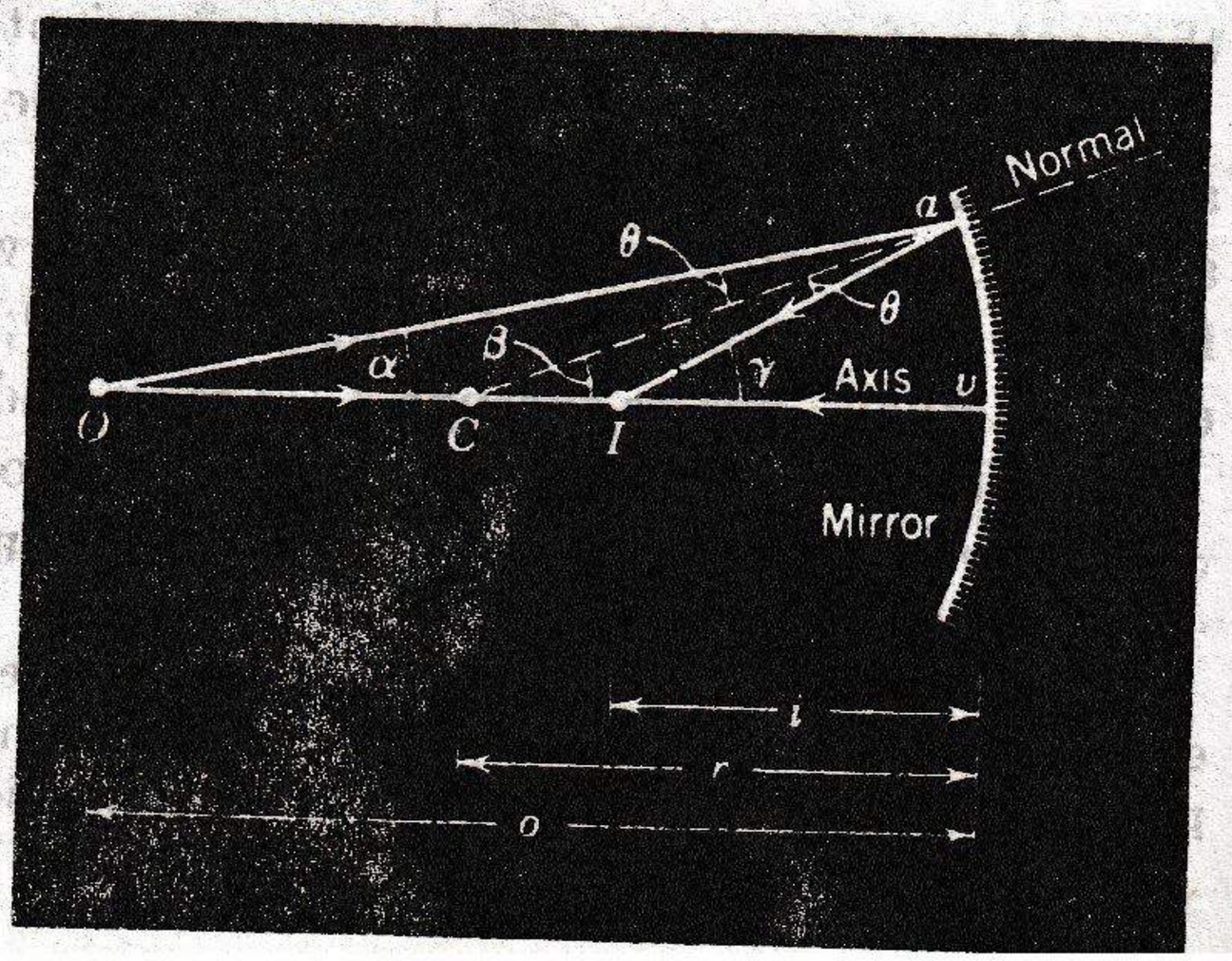


Fig. 42–10 Two rays from *O* converge after reflection in a spherical concave mirror, forming a real image at *I*.

OaIO in Fig. 42–10 yields

$$\beta = \alpha + \theta$$

and

$$\gamma = \alpha + 2\theta.$$

Eliminating θ between these equations leads to

$$\alpha + \gamma = 2\beta. \tag{42–2}$$

In radian measure we can write angles α, β, and γ as

$$\alpha \cong \frac{av}{vO} = \frac{av}{o}$$

$$\beta = \frac{av}{vC} = \frac{av}{r} \tag{42–3}$$

$$\gamma \cong \frac{av}{vI} = \frac{av}{i}.$$

Note that only the equation for β is exact, for the reason that the center of curvature of arc av is at C and not at O or I. However, the equations for α and for γ are approximately correct if these angles are sufficiently small. *In all that follows we assume that the rays diverging from the object make only a small angle α with the axis of the mirror.* Such rays, which lie close to the mirror axis, are called *paraxial rays.* We did not find it necessary to make such an assumption for plane mirrors. Substituting these equations into Eq. 42–2 and canceling av yields

$$\frac{1}{o} + \frac{1}{i} = \frac{2}{r}, \tag{42–4}$$

in which o is the *object distance* and i is the *image distance.* Both these distances are measured from the *vertex* of the mirror, which is the point v at which the axis intercepts the mirror.

Significantly, Eq. 42–4 does not contain α (or β, γ, or θ), so that it holds for all rays that strike the mirror provided that they are sufficiently paraxial. In an actual case the rays can be made as paraxial as one likes by putting a circular diaphragm in front of the mirror, centered about the vertex v; this will impose a certain maximum value of α.

As α in Fig. 42–10 is permitted to become larger, it will become less true that a point object will form a point image; the image will become extended and fuzzy. No sharp criterion for deciding whether a given ray is paraxial can be laid down. If the maximum permitted value of α is reduced, the rays will become more paraxial and the image will become sharper. Unfortunately, the image will also become fainter because less total light energy will be reflected from the mirror. A compromise must often be made between image brightness and image quality.

As for plane mirrors, the image (real or virtual) in a spherical mirror can be seen only if the eye is located so that light rays from the object can enter

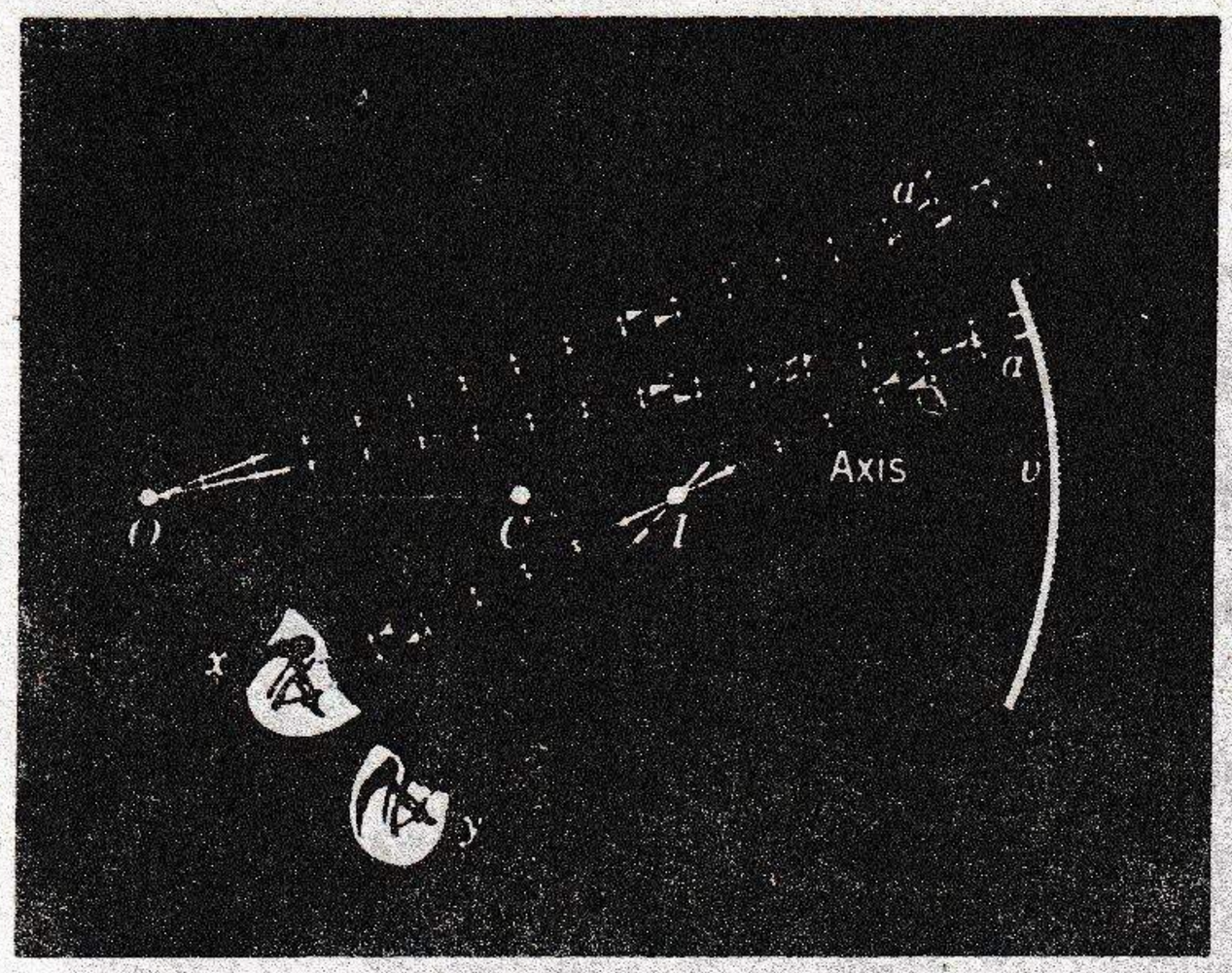


Fig. 42–11 The eye must be properly located to see image *I*.

it after reflection. In Fig. 42–11 a bundle of light rays is shown entering the eye in position x; only the small patch of mirror near a is effective for this eye position. If the observer moves his eye to position y, the image will vanish for him because the mirror does not exist near point a'.

Although Eq. 42–4 was derived for the special case in which the object is located beyond the center of curvature, it is generally true, no matter where the object is located. It is also true for convex mirrors, as in Fig. 42–12.

In applying Eq. 42–4 we must be careful to follow a consistent convention of signs for o, i, and r. As the basis for the sign conventions to be used in this book, we start from this statement:

> In Fig. 42–10, in which light *diverges* from a *real* object, falls on a *concave* mirror, and *converges* after reflection to form a *real* image, the quantities o, i, and r in Eq. 42–4 are given positive numerical values.

Figure 42–10 was used to derive Eq. 42–4 and the student should associate them in his mind as an aid in getting the signs correct.

Let us fix our minds on the side of the mirror from which incident light comes. Because mirrors are opaque, the light, after reflection, must remain on this side, and if an image is formed here it will be a *real* image. There-

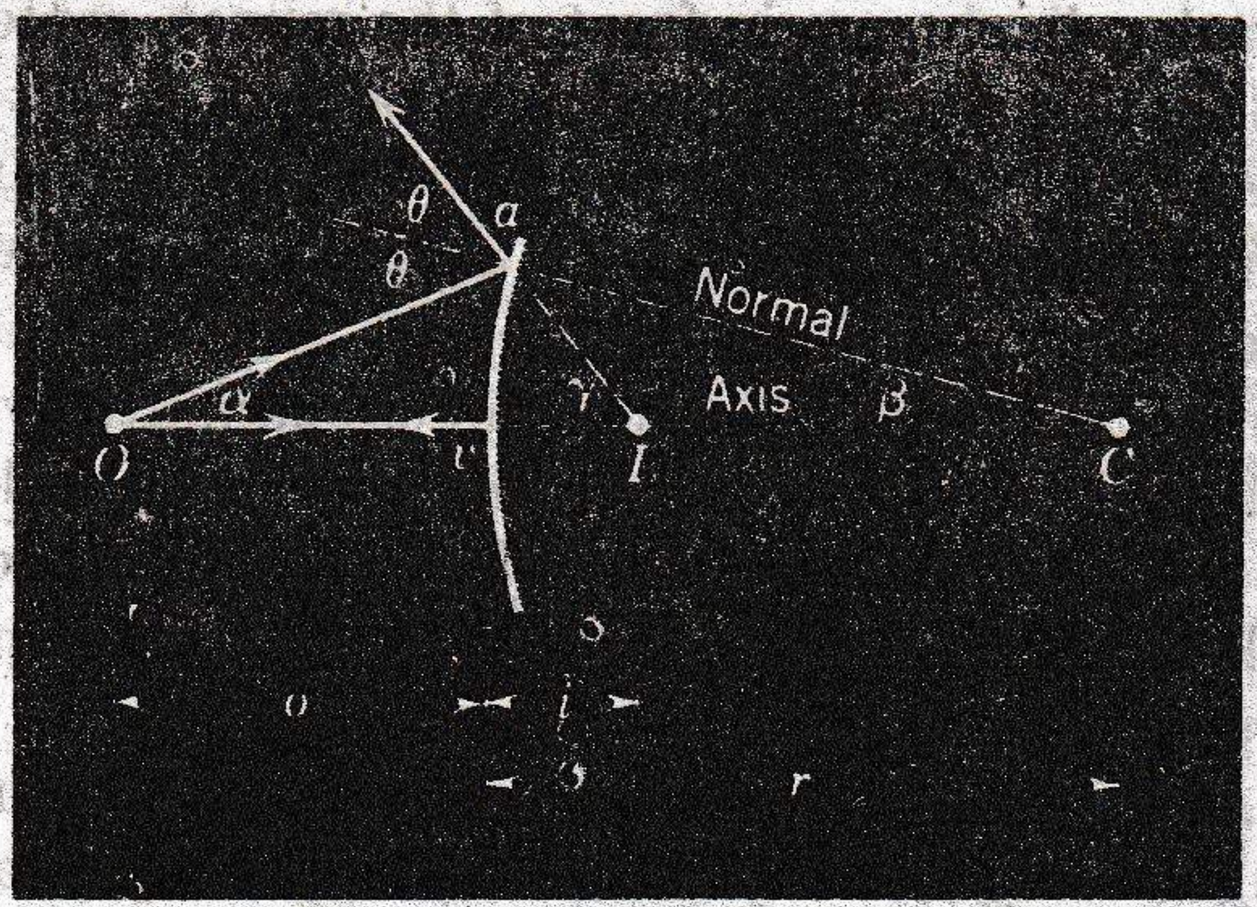


Fig. 42–12 Two rays from O diverge after reflection in a spherical convex mirror, forming a virtual image at I, the point from which they appear to originate. Compare Fig. 42–10.

fore, we call the side of the mirror from which the light comes the *R-side* (for *real image*). We call the back of the mirror the *V-side* (for *virtual image*), because images formed on this side of the mirror must be virtual, no light energy being present on this side.*

In the indented statement above we associated real images with positive image distances. This suggests our first sign convention:

1. The image distance i is positive if the image (real) lies on the R-side of the mirror, as in Fig. 42–10; i is negative if the image (virtual) lies on the V-side, as in Fig. 42–12.

If the mirror in Fig. 42–10, which is concave as viewed from the direction of the incident light, is made convex, the rays will *diverge* after reflection and will form a *virtual* image, as Fig. 42–12 shows. Thus the indented statement above suggests our second sign convention:

2. The radius of curvature r is positive if the center of curvature of the mirror lies on the R-side, as in Fig. 42–10; r is negative if the center of curvature lies on the V-side, as in Fig. 42–12.

The student should not commit these sign conventions to memory but should deduce them in each case from the basic statement on p. 1044, using Fig. 42–10 as a mnemonic aid.

For all cases in this book the object distance o is to be taken as positive. In systems of two or more mirrors (or combinations of mirrors and refracting surfaces) it is possible to arrange that *converging* light falls on the mirror. In such cases the object is called *virtual* and the object distance o is negative; we limit our discussion here to real objects.

▶ **Example 3.** A convex mirror has a radius of curvature of 20 cm. If a point source is placed 14 cm away from the mirror, as in Fig. 42–12, where is the image?

A rough graphical construction, applying the law of reflection at a in the figure, shows that the image will be on the V-side of the mirror and thus will be virtual. We may verify this quantitatively and analytically from Eq. 42–4, noting that r is negative here because the center of curvature of the mirror is on its V-side. We have

$$\frac{1}{o} + \frac{1}{i} = \frac{2}{r}$$

or

$$\frac{1}{+14\text{ cm}} + \frac{1}{i} = \frac{2}{-20\text{ cm}},$$

which yields $i = -5.8$ cm, in agreement with the graphical prediction. The negative sign for i reminds us that the image is on the V-side of the mirror and thus is virtual. ◀

When *parallel* light falls on a mirror (Fig. 42–13), the image point (real or virtual) is called the *focal point* F of the mirror. The focal length f is the distance between F and the vertex. If we put $o \to \infty$ in Eq. 42–4, thus insuring parallel incident light, we have

$$i = \tfrac{1}{2}r = f.$$

* This nomenclature may seem unnecessarily cumbersome at this point. We adopt it because of the later extension of these ideas to refracting surfaces and lenses.

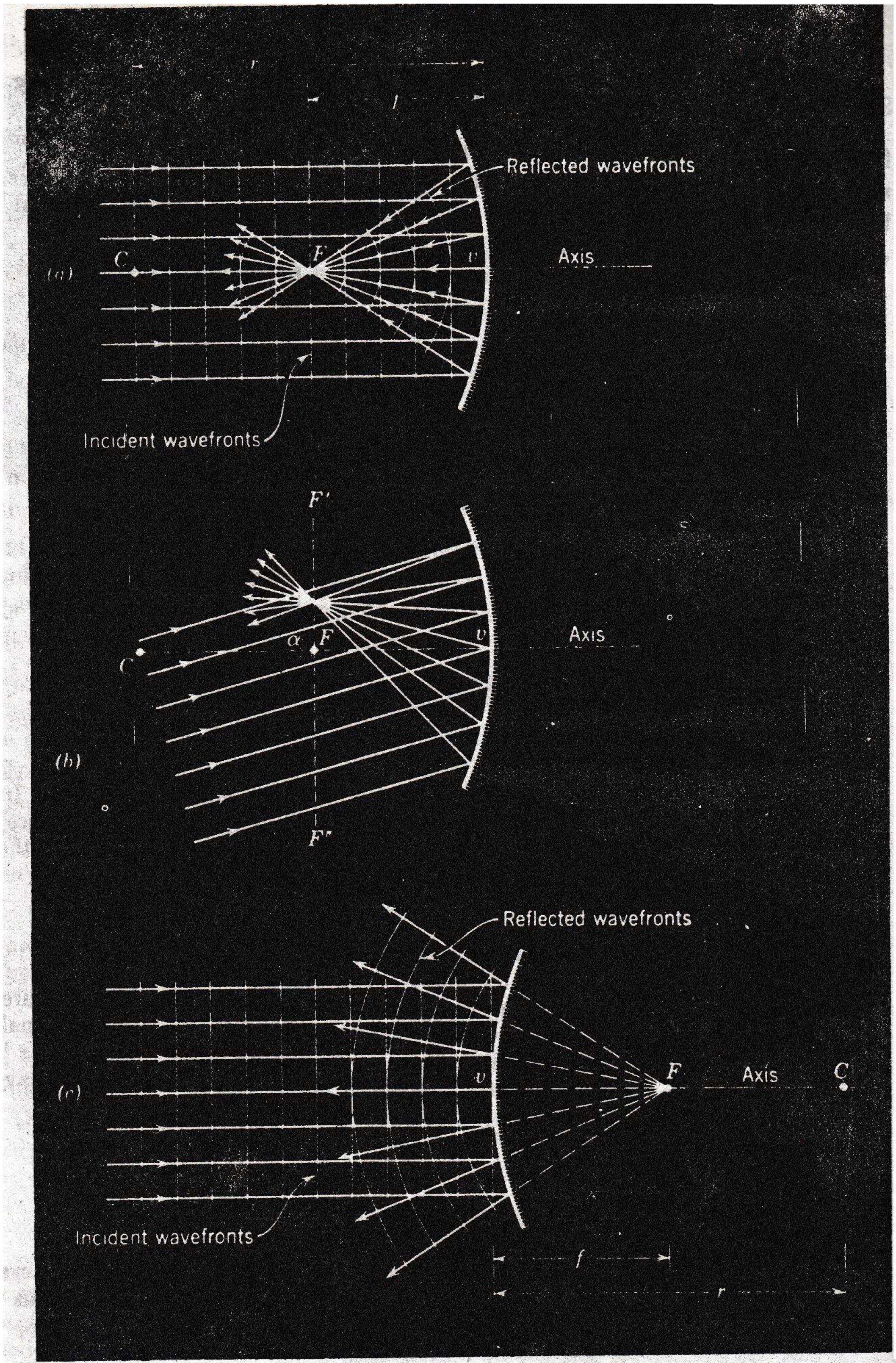


Fig. 42–13 (*a*) The focal point for a concave spherical mirror, showing both the rays and the wavefronts. F and C lie on the R-side, the focal point is real, and the focal length f of the mirror is positive (as is r). (*b*) The same, except that the incident light makes an angle α with the mirror axis; the rays are focused at a point in the *focal plane* $F'F''$. (*c*) Same as (*a*) except that the mirror is convex; F and C lie on the V-side of the mirror. The focal point is virtual and the focal length f is negative (as is r).

Equation 42–4 can then be rewritten

$$\frac{1}{o} + \frac{1}{i} = \frac{1}{f}, \tag{42–5}$$

where f, like r, is taken as positive for mirrors whose centers of curvature are on the R-side (that is, for *concave*, or *converging* mirrors; see Fig. 42–13*a*), and negative for those whose centers of curvature are on the V-side (that is, for *convex*, or *diverging* mirrors; see Fig. 42–13*c*). Figure 42–13*b* shows an incident plane wave that makes a small angle α with the mirror axis. The rays are focused at a point in the *focal plane* of the mirror. This is a plane at right angles to the mirror axis at the focal point.

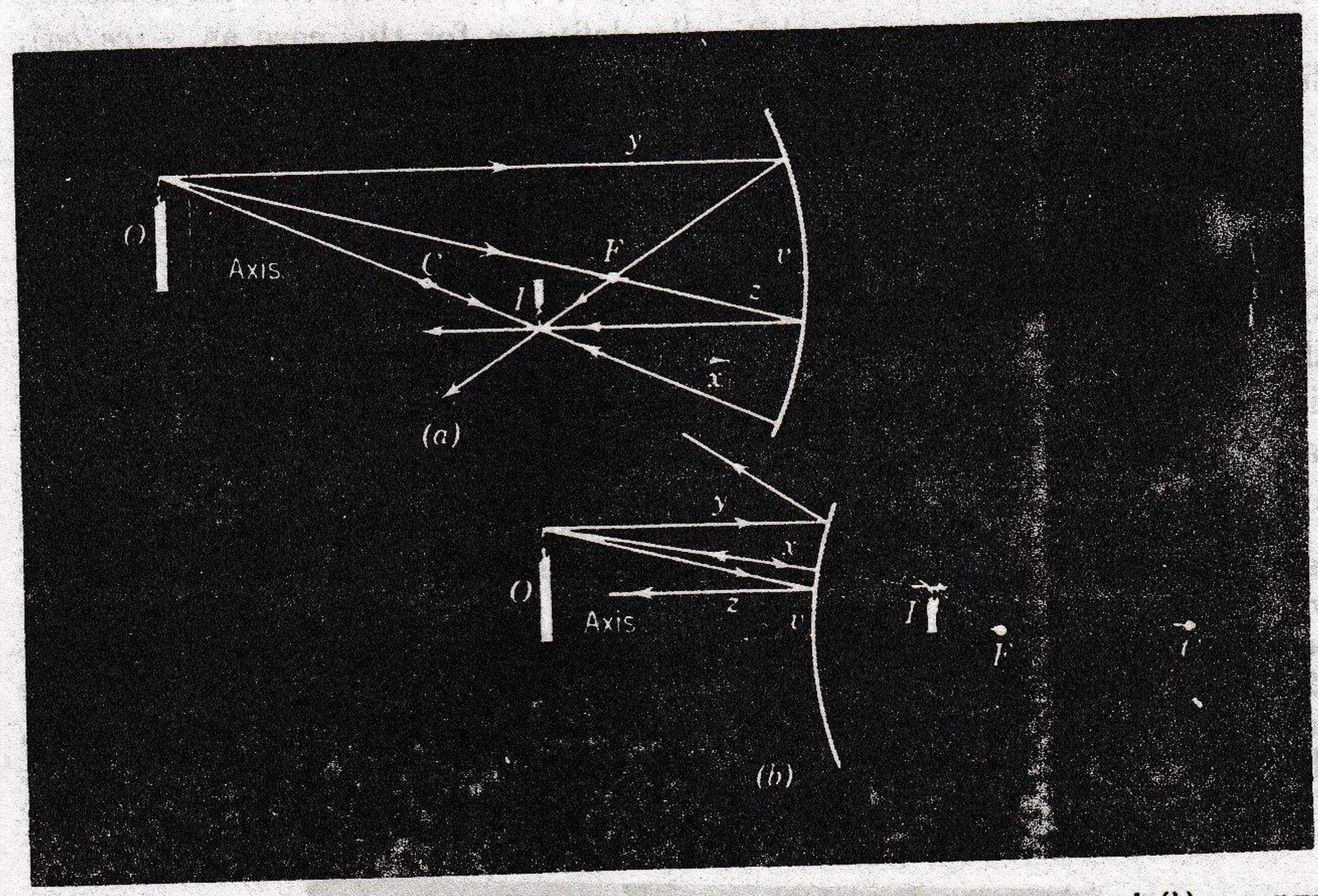


Fig. 42–14 The image of an extended object in (*a*) a concave mirror and (*b*) a convex mirror is located graphically. Any two of the three special rays shown are sufficient.

We now consider objects that are not points. Figure 42–14 shows a candle in front of (*a*) a concave mirror and (*b*) a convex mirror. We choose to draw the mirror axis through the foot of the candle and, of course, through the center of curvature. The image of any off-axis point, such as the tip of the candle, can be found graphically, using the following facts:

1. A ray that strikes the mirror after passing (either directly or upon being extended) through the center of curvature C returns along itself (ray x in Fig. 42–13). Such rays strike the mirror at right angles.
2. A ray that strikes the mirror parallel to its axis passes (or will pass when extended) through the focal point (ray y).

3. A ray that strikes the mirror after passing (either directly or upon being extended) through the focal point emerges parallel to the axis (ray *z*).

Figure 42–15 shows a ray (*dve*) that originates on the tip of the object candle of Fig. 42–14*a*, is reflected from the mirror at point *v*, and passes through the tip of the image candle. The law of reflection demands that this ray make equal angles θ with the mirror axis as shown. For the two similar right triangles in the figure we can write

$$\frac{ce}{bd} = \frac{vc}{vb}.$$

The quantity on the left (apart from a question of sign) is the *lateral magnification m* of the mirror. Since we want to represent an *inverted* image by a *negative* magnification, we arbitrarily define *m* for this case as $-(ce/bd)$. Since $vc = i$ and $vb = o$, we have at once

$$m = -\frac{i}{o}. \tag{42–6}$$

This equation gives the magnification for spherical and plane mirrors under all circumstances. For a plane mirror, $o = -i$ and the predicted magnification is +1 which, in agreement with experience, indicates an *erect* image the same size as the object.

Images in spherical mirrors suffer from several "defects" that arise because the assumption of paraxial rays is never completely justified. In general, a point source will not produce a point image; see Problem 5. Apart from this, distortion arises because the magnification varies somewhat with distance from the mirror axis, Eq. 42–6 being strictly correct only for paraxial rays. Superimposed on these defects are *diffraction effects* which come about because the basic assumption of geometrical optics, that light travels in straight lines, must always be considered an approximation.

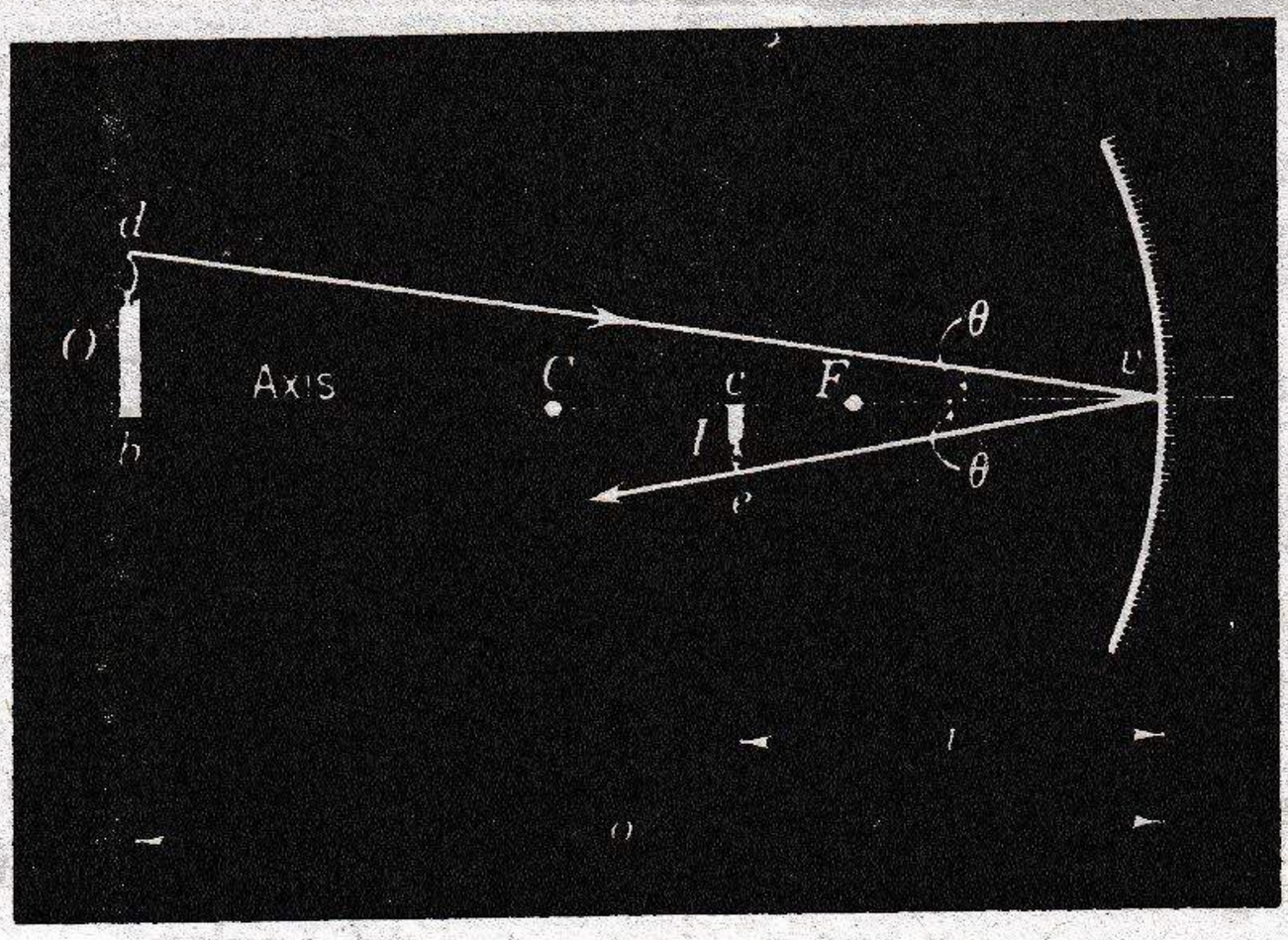


Fig. 42–15 A particular ray for the arrangement of Fig. 42–14, used to show that the *lateral magnification m* is given by $-i/o$.

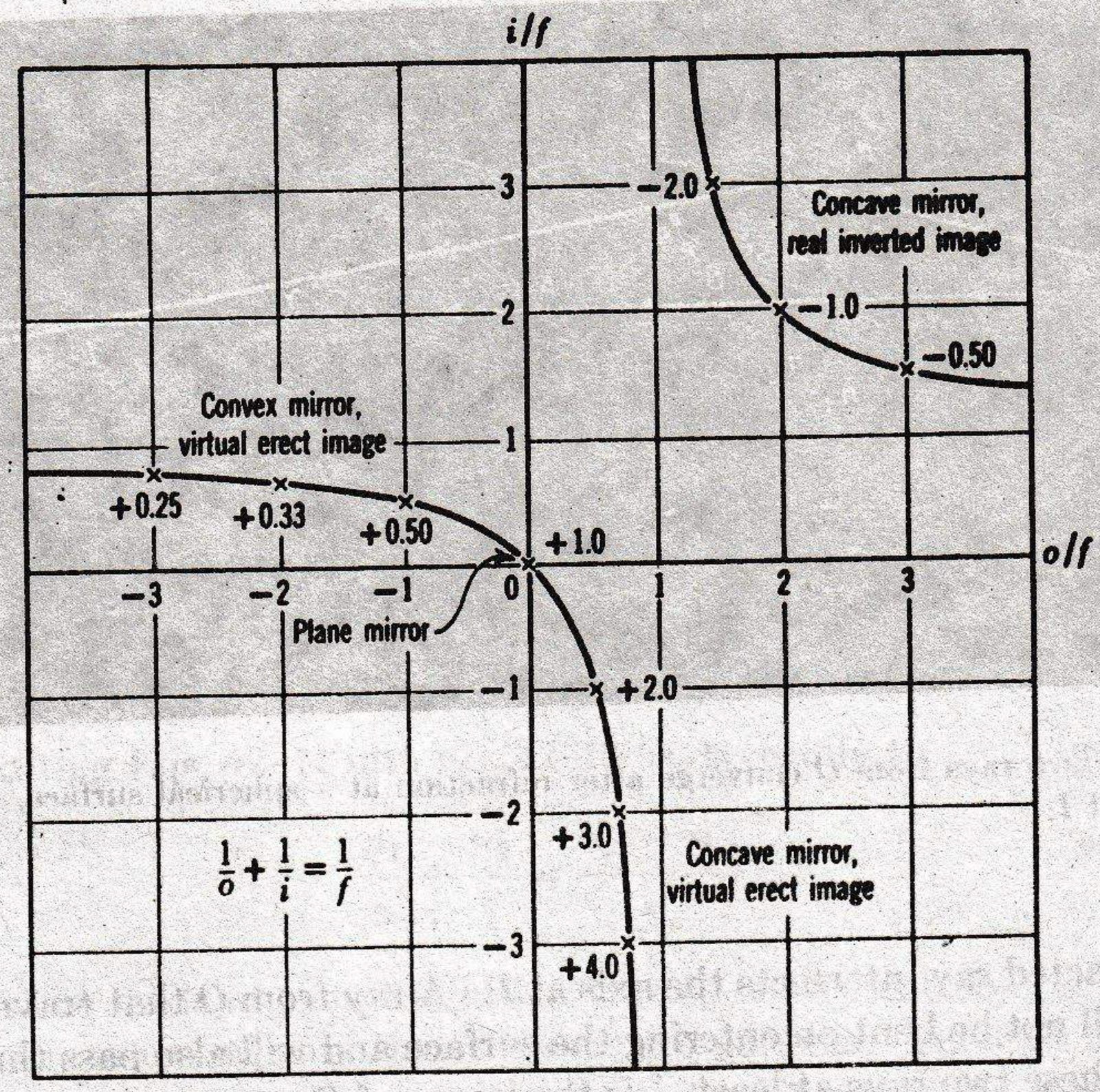


Fig. 42–16 A graphical representation of the mirror formula (Eq. 42–5). The numbers on the curves give the lateral magnifications (Eq. 42–6).

Figure 42–16 summarizes the predictions of the mirror equation (Eq. 42–5) and of the lateral magnification equation (Eq. 42–6). For convex mirrors the image is always virtual. For concave mirrors the image is virtual if the object is inside the focal point ($o/f < 1$) and real otherwise ($o/f > 1$). As $o \to \infty$ for both types of mirror, the image approaches the focal point ($i/f \to 1$). A plane mirror ($m = +1$) is represented in the figure by a point at the origin. It is appropriate that this point form the central point of the lower-left branch of Fig. 42–16 because both convex and concave mirrors approach plane mirrors as r is increased in magnitude, that is, as $f \to \infty$ (or as $o/f \to 0$ in Fig. 42–16, assuming a fixed object distance).

42–4 Spherical Refracting Surface

Figure 42–17 shows a point source O near a convex spherical *refracting* surface of radius of curvature r. The surface separates two media whose indices of refraction differ, *that of the medium in which the incident light falls on the surface being* n_1 and that on the other side of the surface being n_2.

From O we draw a line through the center of curvature C of the refracting surface, thus establishing a convenient axis which intercepts the surface at vertex v. From O we draw a ray that makes a small but arbitrary angle α with the axis and strikes the refracting surface at a, being refracted according to

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

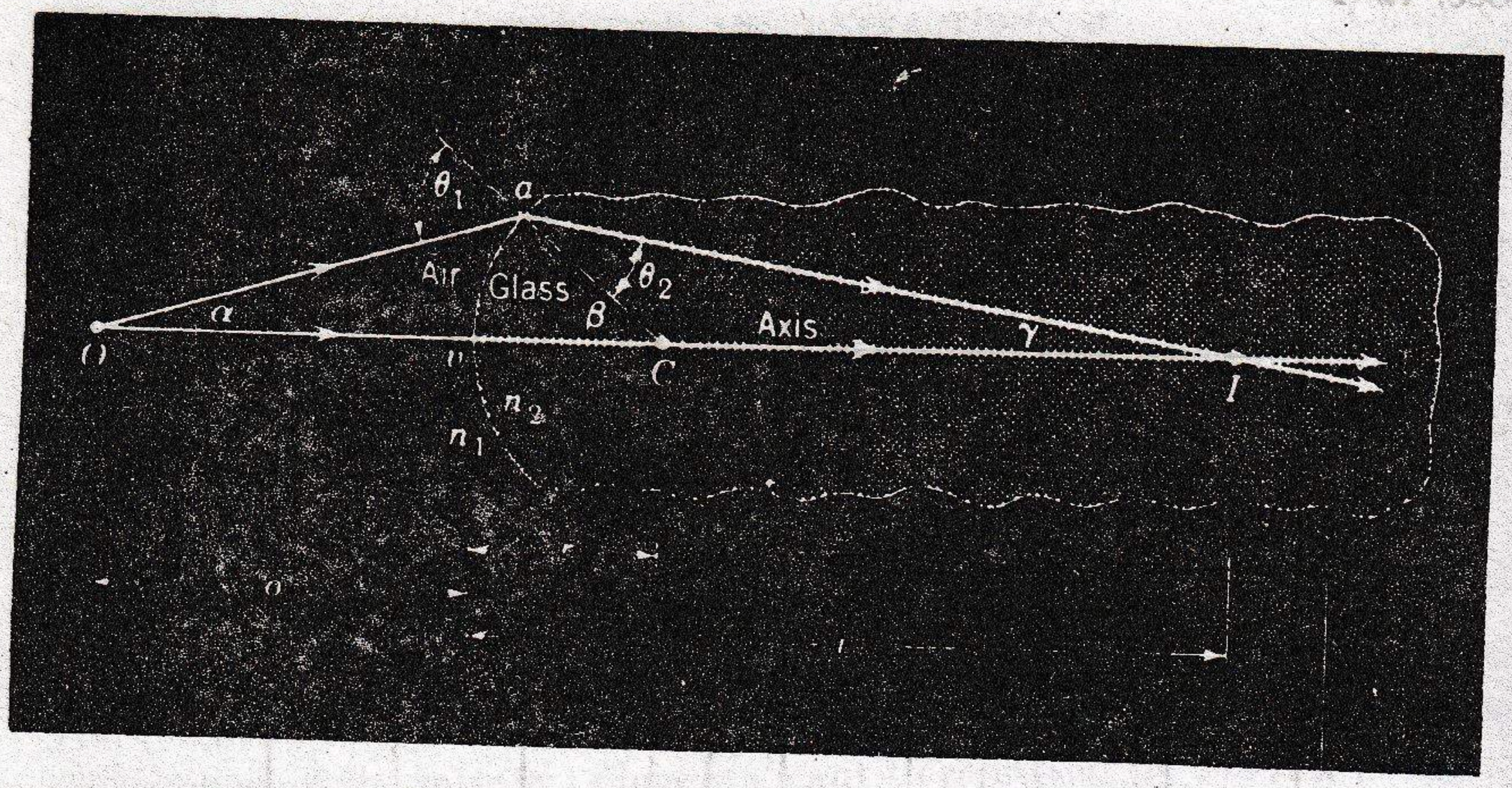


Fig. 42–17 Two rays from O converge after refraction at a spherical surface, forming a real image at I.

The refracted ray intersects the axis at I. A ray from O that travels along the axis will not be bent on entering the surface and will also pass through I. Thus, for these two rays at least, I is the image of O.

As in the derivation of the mirror equation, we use the theorem that the exterior angle of a triangle is equal to the sum of the two opposite interior angles. Applying this to triangles $COaC$ and $ICaI$ yields

$$\theta_1 = \alpha + \beta \tag{42–7}$$

and

$$\beta = \theta_2 + \gamma. \tag{42–8}$$

As α is made small, angles β, γ, θ_1, and θ_2 in Fig. 42–17 also become small. *We at once assume that* α, *hence all these angles, are arbitrarily small,* this paraxial ray assumption was also made for spherical mirrors. Replacing the sines of the angles by the angles themselves—since the angles are required to be small—permits us to write the law of refraction as

$$n_1\theta_1 \cong n_2\theta_2. \tag{42–9}$$

Combining Eqs. 42–8 and 42–9 leads to

$$\beta = \frac{n_1}{n_2}\theta_1 + \gamma.$$

Eliminating θ_1 between this equation and Eq. 42–7 leads, after rearrangement, to

$$n_1\alpha + n_2\gamma = (n_2 - n_1)\beta. \tag{42–10}$$

In radian measure the angles α, β, and γ in Fig. 42–17 are

$$\alpha \cong \frac{av}{o}$$

$$\beta = \frac{av}{r} \qquad (42\text{–}11)$$

$$\gamma \cong \frac{av}{i}.$$

Only the second of these equations is exact. The other two are approximate because I and O are *not* the centers of circles of which av is an arc. However, for paraxial rays (α small enough) the inaccuracies in Eq. 42–11 can be made as small as desired.

Substituting Eqs. 42–11 into Eq. 42–10 leads readily to

$$\frac{n_1}{o} + \frac{n_2}{i} = \frac{n_2 - n_1}{r}. \qquad (42\text{–}12)$$

This equation holds whenever light is refracted from point objects at spherical surfaces, assuming only that the rays are paraxial. As with the mirror formula, care must be taken to use Eq. 42–12 with consistent signs for o, i, and r. Once again we establish our sign conventions by physical reasoning from a particular case, that of Fig. 42–17:

> In Fig. 42–17, in which light *diverges* from a *real* object, falls on a *convex* refracting surface, and *converges* after refraction to form a *real* image, the quantities o, i, and r in Eq. 42–12 have positive numerical values.

Figure 42–17 was used to derive Eq. 42–12, and the student should associate them in his mind as an aid in getting the signs correct. This basic statement is quite similar to that which we made for mirrors on p. 1044.

We fix our attention on the side of the refracting surface from which the incident light falls on the surface. In contrast to mirrors, the light energy *passes through* a refracting surface to the other side, and if a real image is formed it must appear on the far side, which we call the R-side. The side from which the incident light comes is called the V-side because virtual images must appear here. Figure 42–18 suggests this important distinction between reflection and refraction.

In the indented statement above we associated real images with positive-image distances. Thus we are led to the sign convention:

1. The image distance i is positive if the image (real) is on the R-side of the refracting surface, as in Fig. 42–17; i is negative if the image (virtual) lies on the V-side, as in Fig. 42–19.

The refracting surface in Fig. 42–17 is convex. If it is made concave (still assuming that $n_2 > n_1$), the rays will *diverge* after refraction and form

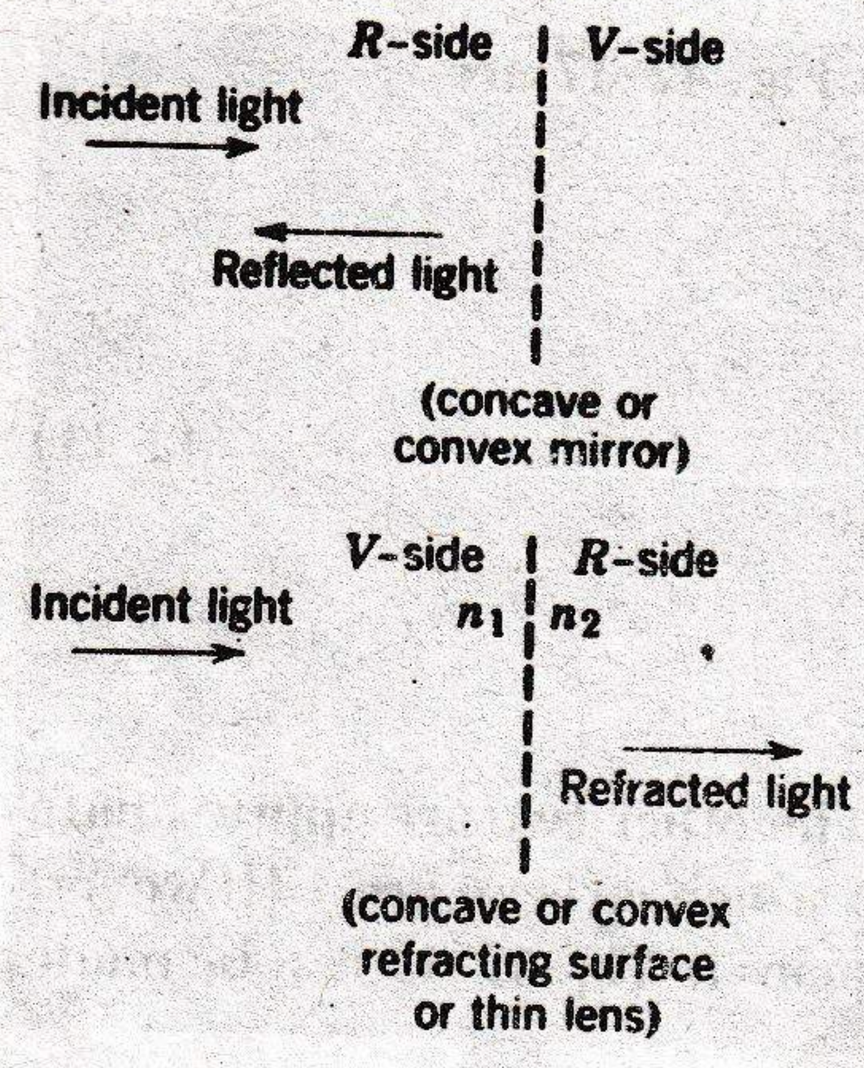


Fig. 42–18 Real images are formed on the same side as the incident light for mirrors but on the opposite side for refracting surfaces and thin lenses. This is so because the incident light is reflected back by mirrors but is transmitted through by refracting surfaces.

a *virtual* image, as Fig. 42–19 shows. Thus we are led to our second sign convention:

2. The radius of curvature r is positive if the center of curvature of the refracting surface lies on the R-side, as in Fig. 42–17; r is negative if the center of curvature is on the V-side, as in Fig. 42–19.

The sign conventions for refracting surfaces are the same as for mirrors (p. 1045), the fundamental difference between the two situations being absorbed in the definitions of R-side and V-side in Fig. 42–18. This difference is easily remembered on physical grounds.

For all cases in this book the object distance o is to be taken as positive. In systems of two or more refracting surfaces (or combinations of refracting surfaces and mirrors) it is possible to arrange that *converging* light falls on the refracting surface. In such cases the object is called *virtual* and the object distance o is negative; we limit our discussions here to real objects.

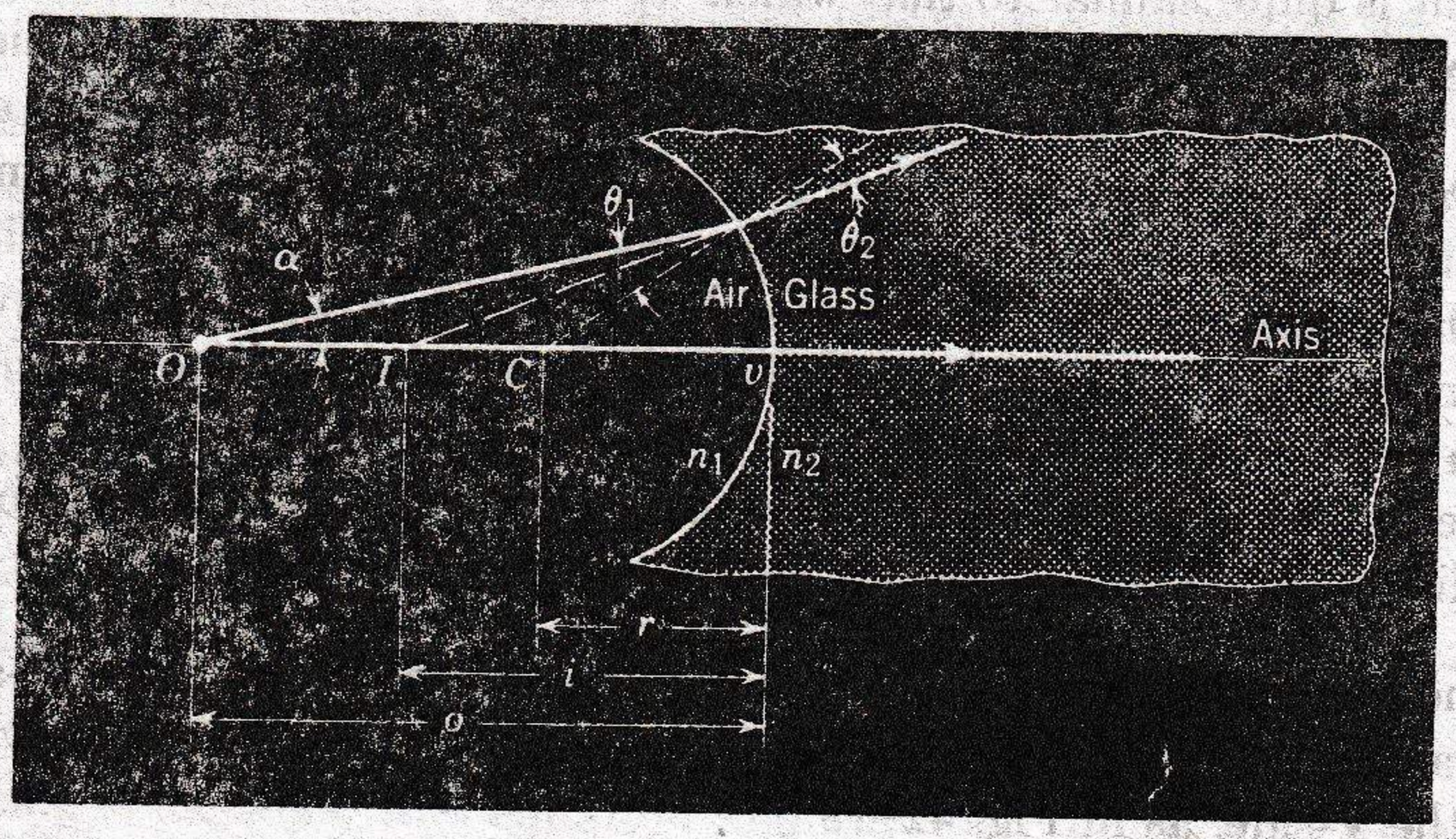


Fig. 42–19 Two rays from O diverge after refraction at a spherical surface, forming a virtual image at I.

▶ **Example 4.** Locate the image for the geometry shown in Fig. 42–17, assuming the radius of curvature to be 10 cm, n_2 to be 2.0, and n_1 to be 1.0. Let the object be 20 cm to the left of v.

From Eq. 42–12,

$$\frac{n_1}{o} + \frac{n_2}{i} = \frac{n_2 - n_1}{r},$$

we have

$$\frac{1.0}{+20\text{ cm}} + \frac{2.0}{i} = \frac{2.0 - 1.0}{+10\text{ cm}}.$$

Note that r is positive because the center of curvature of the surface lies on the R-side. This relation yields $i = +40$ cm in agreement with the graphical construction. The light energy actually passes through I so that the image is real, as indicated by the positive sign for i.

Example 5. An object is immersed in a medium with $n_1 = 2.0$, being 15 cm from the spherical surface whose radius of curvature is -10 cm, as in Fig. 42–20; r is negative because C lies on the V-side. Locate the image.

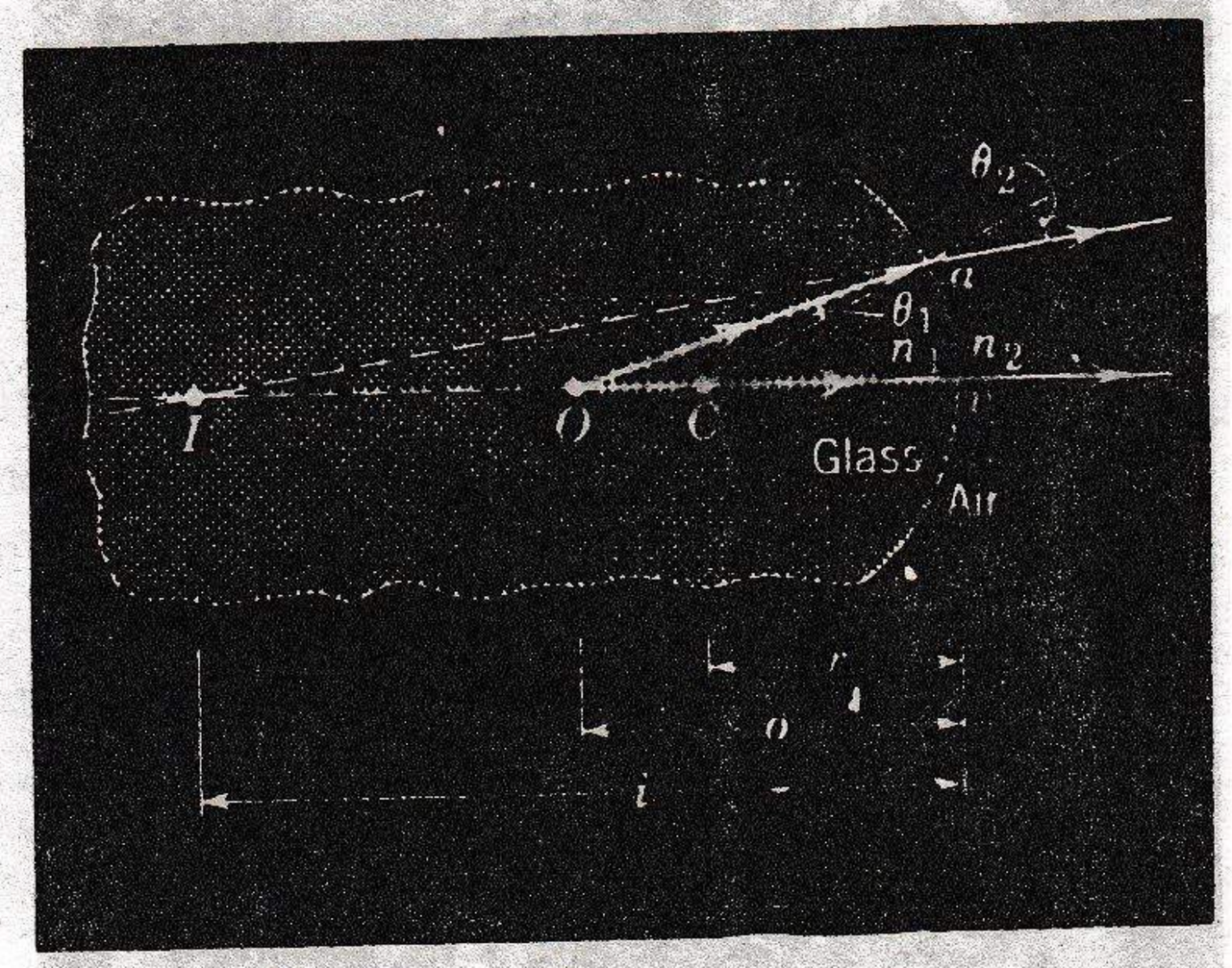


Fig. 42–20 Two rays from O appear to originate from I (virtual image) after refraction at a spherical surface.

Figure 42–20 shows a ray traced through the surface by applying the law of refraction at point a. A second ray from O along the axis emerges undeflected at v. The image I is found by extending these two rays backward; it is virtual.

From Eq. 42–12,

$$\frac{n_1}{o} + \frac{n_2}{i} = \frac{n_2 - n_1}{r},$$

we have

$$\frac{2.0}{+15\text{ cm}} + \frac{1.0}{i} = \frac{1.0 - 2.0}{-10\text{ cm}},$$

which yields $i = -30$ cm, in agreement with Fig. 42–20 and with the sign conventions. Note that n_1 always refers to the medium on the side of the surface from which the light comes.

Example 6. What is the relationship between i and o if the refracting surface is plane?

A plane surface has an infinite radius of curvature. Putting $r \to \infty$ in Eq. 42–12 leads to

$$i = -o\frac{n_2}{n_1}.$$

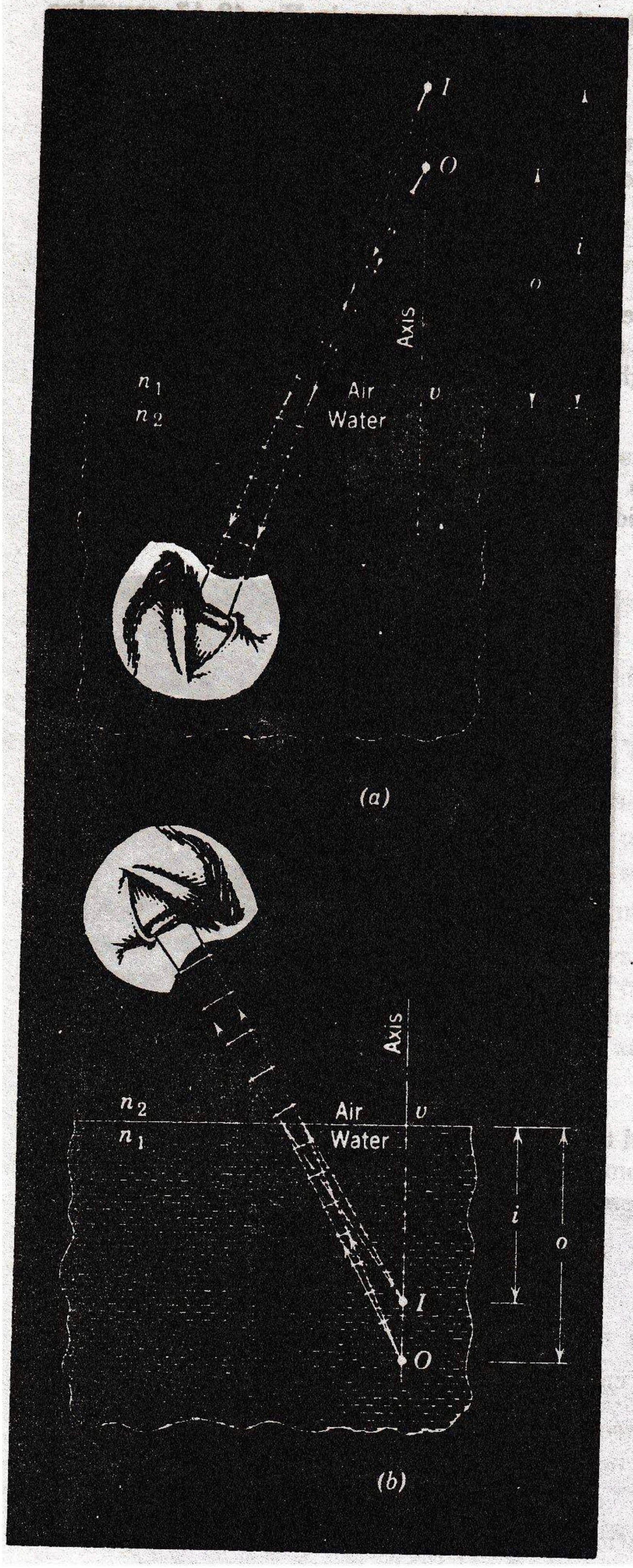


Fig. 42–21 Refraction at a plane surface at near-normal incidence, showing a pencil of rays and the corresponding wavefronts entering the pupil. (*a*) Source in air and (*b*) source in water.

Figure 42–21 illustrates the situation graphically (*a*) for an object in air as seen from below water and (*b*) for an object in water with air above. This shows that a diver, looking upward at, say, an overhanging tree branch, will think it higher than it is by the factor 1.33/1.00. Similarly, an observer in air will think that objects on the bottom of a water tank are closer to the surface than they actually are, in the ratio 1.00/1.33. These considerations, being based on Eq. 42–12, hold only for paraxial rays, which means that the incident rays can make only a small angle with the normal; this angle has been exaggerated in the figure for clarity. Note again that n_1 is always identified with the medium that lies on the side of the surface containing the incident light. ◀

42–5 Thin Lenses

In most refraction situations there is more than one refracting surface. This is true even for a spectacle lens, the light passing from air into glass and then from glass into air. In microscopes, telescopes, cameras, etc., there are often many more than two surfaces.

Figure 42–22*a* shows a thick glass "lens" of length l whose surfaces are ground to radii r' and r''. A point object O' is placed near the left surface as shown. A ray leaving O' along the axis is not deflected on entering or leaving the lens because it falls on each surface along a normal.

A second ray leaving O', at an arbitrary angle α with the axis, strikes the surface at point a', is refracted, and strikes the second surface at point a''. The ray is again refracted and crosses the axis at I'', which, being the intersection of two rays from O'', is the image of point O', formed after refraction at two surfaces.

Figure 42–22*b* shows the first surface, which forms a virtual image of O' at I'. To locate I', we use Eq. 42–12,

$$\frac{n_1}{o} + \frac{n_2}{i} = \frac{n_2 - n_1}{r}.$$

Putting $n_1 = 1.0$ and $n_2 = n$ and bearing in mind that the image distance is negative (that is, $i = -i'$ in Fig. 42–22*b*), we obtain

$$\frac{1}{o'} - \frac{n}{i'} = \frac{n - 1}{r'}. \qquad (42\text{–}13)$$

In this equation i' will be a positive number because we have arbitrarily introduced the minus sign appropriate to a virtual image.

Figure 42–22*c* shows the second surface. Unless an observer at point a'' were aware of the existence of the first surface, he would think that the light striking that point originated at point I' in Fig. 42–22*b* and that the region to the left of the surface was filled with glass. Thus the (virtual) image I' formed by the first surface serves as a real object O'' for the second surface. The distance of this object from the second surface is

$$o'' = i' + l. \qquad (42\text{–}14)$$

In applying Eq. 42–12 to the second surface, we insert $n_1 = n$ and $n_2 = 1.0$ because the object behaves as if it were imbedded in glass. If we use

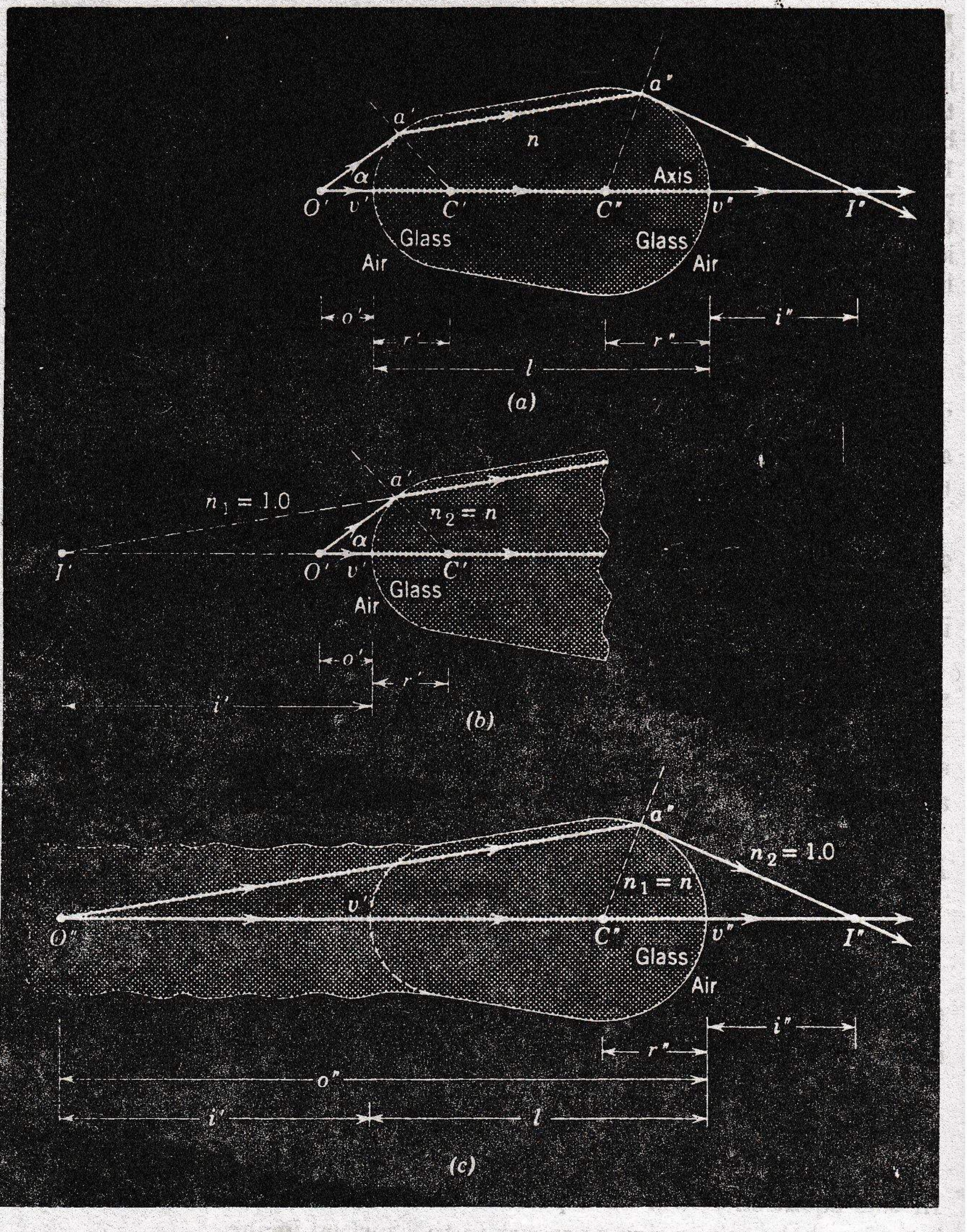


Fig. 42–22 (*a*) Two rays from O' intersect at I'' (real image) after refraction at two spherical surfaces. (*b*) The first surface and (*c*) the second surface shown separately. The quantities α and n have been exaggerated for clarity.

Eq. 42–14, Eq. 42–12 becomes

$$\frac{n}{i' + l} + \frac{1}{i''} = \frac{1 - n}{r''}. \tag{42–15}$$

Let us now assume that the thickness l of the "lens" in Fig. 42–22 is so small that it can be neglected in comparison with other linear quantities in

this figure (such as o', i', o'', i'', r', and r''). In all that follows we make this *thin-lens approximation*. Putting $l = 0$ in Eq. 42–15 leads to

$$\frac{n}{i'} + \frac{1}{i''} = -\frac{n-1}{r''}. \tag{42–16}$$

Adding Eqs. 42–13 and 42–16 leads to

$$\frac{1}{o'} + \frac{1}{i''} = (n-1)\left(\frac{1}{r'} - \frac{1}{r''}\right).$$

Finally, calling the original object distance simply o and the final image distance simply i leads to

$$\frac{1}{o} + \frac{1}{i} = (n-1)\left(\frac{1}{r'} - \frac{1}{r''}\right). \tag{42–17}$$

This equation holds only for paraxial rays and only if the lens is so thin that it essentially makes no difference from which surface of the lens the quantities o and i are measured. In Eq. 42–17 r' refers to the first surface struck by the light as it traverses the lens and r'' to the second surface.

The sign conventions for Eq. 42–17 are the same as those for mirrors and for single refracting surfaces. Because the lens is assumed to be thin, we refer to the R-side and the V-side of the lens itself (see Fig. 42–18) rather than those of its separate surfaces. The sign conventions then are the following:

1. The image distance i is positive if the image (real) lies on the R-side of the lens, as in Fig. 42–23a; i is negative if the image (virtual) lies on the V-side of the lens, as in Fig. 42–23b.
2. The radii of curvature r' and r'' are positive if their respective centers

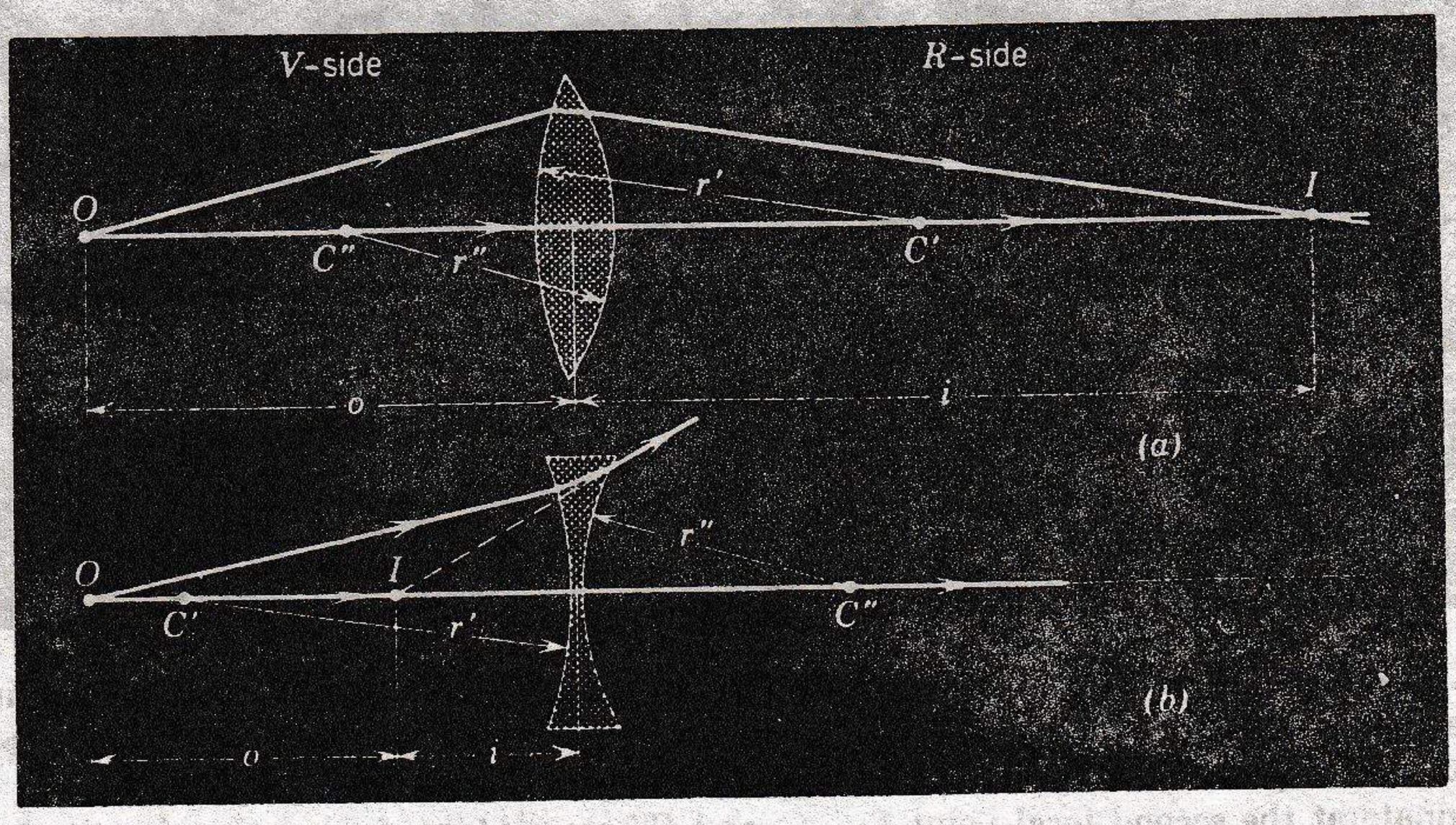


Fig. 42–23 Illustrations used to establish sign conventions for thin lenses.

of curvature lie on the R-side of the lens. They are negative if their centers lie on the V-side. In Fig. 42–23*a*, r' is positive and r'' is negative; in Fig. 42–23*b*, r' is negative and r'' is positive.

Figure 42–24*a* and *c* shows parallel light from a distant object falling on a thin lens. The image location is called the *second focal point* F_2 of the lens. The distance from F_2 to the lens is called the *focal length* f. The *first focal point* for a thin lens (F_1 in figure) is the object position for which the image

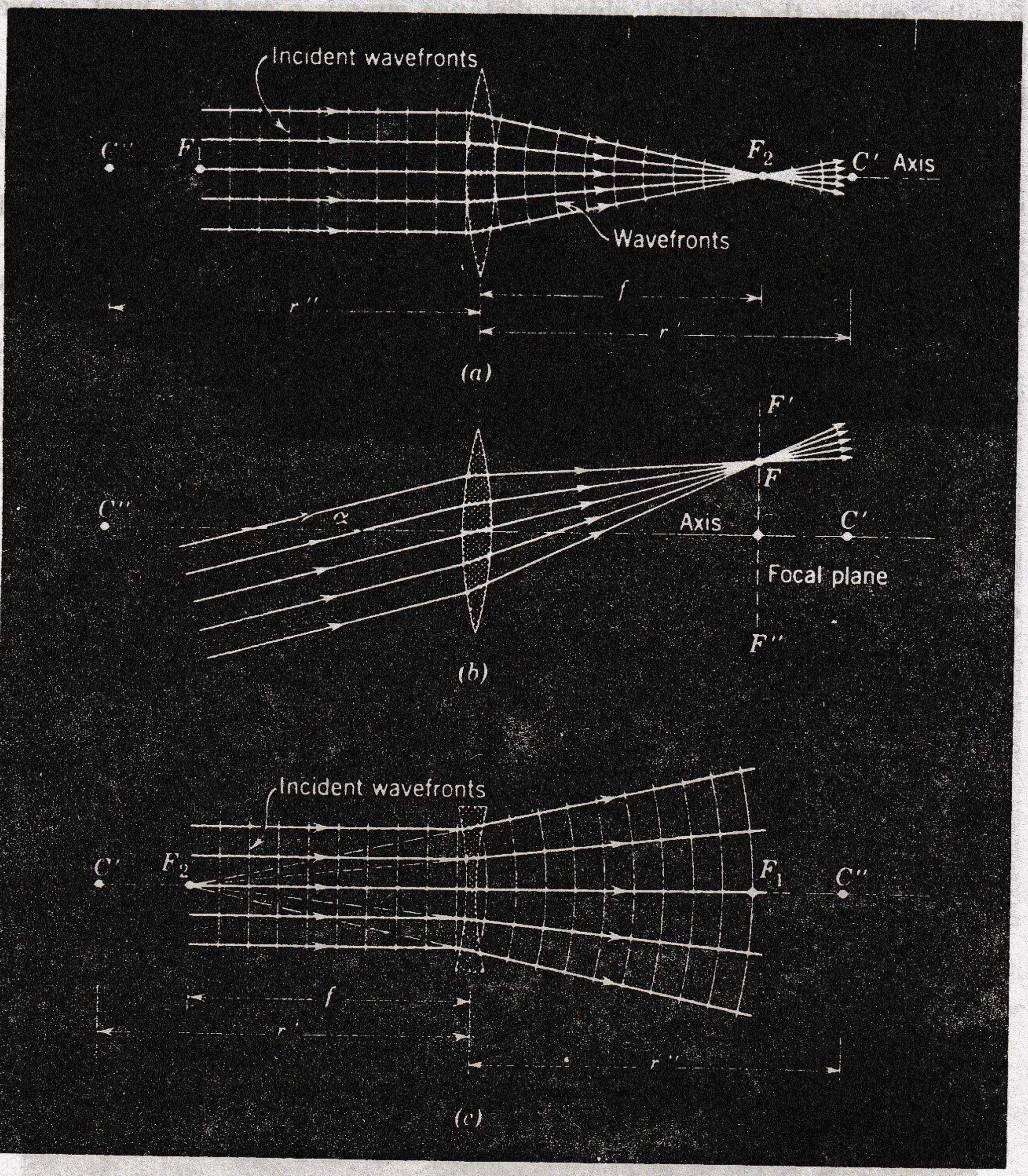


Fig. 42–24 (*a*) Parallel light passes through the second focal poiut F_2 of a converging lens. (*b*) The incident light makes an angle α with the lens axis, the rays being focused in the focal plane $F'F''$ (*c*) Parallel light, passing through a diverging lens, seems to originate at the second focal point F_2. C' and C'' are centers of curvature for the lens surfaces; F_1 is the first focal point.

is at infinity. For thin lenses the first and second focal points are on opposite sides of the lens and are equidistant from it.

The focal length can be computed from Eq. 42–17 by inserting $o \to \infty$ and $i = f$. This yields

$$\frac{1}{f} = (n - 1)\left(\frac{1}{r'} - \frac{1}{r''}\right). \tag{42–18}$$

This relation is called the *lens maker's equation* because it allows us to compute the focal length of a lens in terms of the radii of curvature and the index of refraction of the material. Combining Eqs. 42–17 and 42–18 allows us to write the thin-lens equation as

$$\frac{1}{o} + \frac{1}{i} = \frac{1}{f}. \tag{42–19}$$

Figure 42–24*b* shows parallel incident rays that make a small angle α with the lens axis; they are brought to a focus in the *focal plane F′F″*, as shown. This is a plane normal to the lens axis at the focal point.

In Fig. 42–24*a* we note that all rays in the figure contain the same number of wavelengths; in other words, they have the same *optical path lengths;* see Section 41–6. The optical path lengths are the same because the wavefronts are surfaces over which the wave disturbance has the same constant value and because all the rays shown pass through the same number of wavefronts.

▶ **Example 7.** The lenses of Fig. 42–24 have radii of curvature of magnitude 40 cm and are made of glass with $n = 1.65$. Compute their focal lengths.

Since C' lies on the R-side of the lens in Fig. 42–24*a*, r' is positive ($= +40$ cm). Since C'' lies on the V-side, r'' is negative ($= -40$ cm). Substituting in Eq. 42–18 yields

$$\frac{1}{f} = (n - 1)\left(\frac{1}{r'} - \frac{1}{r''}\right) = (1.65 - 1)\left(\frac{1}{+40\text{ cm}} - \frac{1}{-40\text{ cm}}\right),$$

or

$$f = +31\text{ cm}.$$

A positive focal length indicates that in agreement with Fig. 42–24*a* the focal point F_2 is on the R-side of the lens and parallel incident light converges after refraction to form a real image.

In Fig. 42–24*c* C' lies on the V-side of the lens so that r' is negative ($= -40$ cm). Since r'' is positive ($= +40$ cm), Eq. 42–17 yields

$$f = -31\text{ cm}.$$

A negative focal length indicates that in agreement with Fig. 42–24*c* the focal point F_2 is on the V-side of the lens and incident light diverges after refraction to form a virtual image. ◀

The location of the image of an extended object such as a candle (Fig. 42–25) can be found graphically by using the following three facts:

1. A ray parallel to the axis and falling on the lens passes, either directly or when extended, through the second focal point (ray x in Fig. 42–25).
2. A ray falling on a lens after passing, either directly or when extended,

through the first focal point will emerge from the lens parallel to the axis (ray y).

3. A ray falling on the lens at its center will pass through undeflected. There is no deflection because the lens, near its center, behaves like a thin piece of glass with parallel sides. The direction of the light rays is not changed and the sideways displacement can be neglected because the lens thickness has been assumed to be negligible (ray z; see also Problem 13, Chapter 41).

Figure 42–26, which represents part of Fig. 42–25a, shows a ray passing from the tip of the object through the center of curvature to the tip of the image. For the similar triangles abc and dec we may write

$$\frac{de}{ab} = \frac{dc}{ac}.$$

The right side of this equation is i/o and the left side is $-m$, where m is the

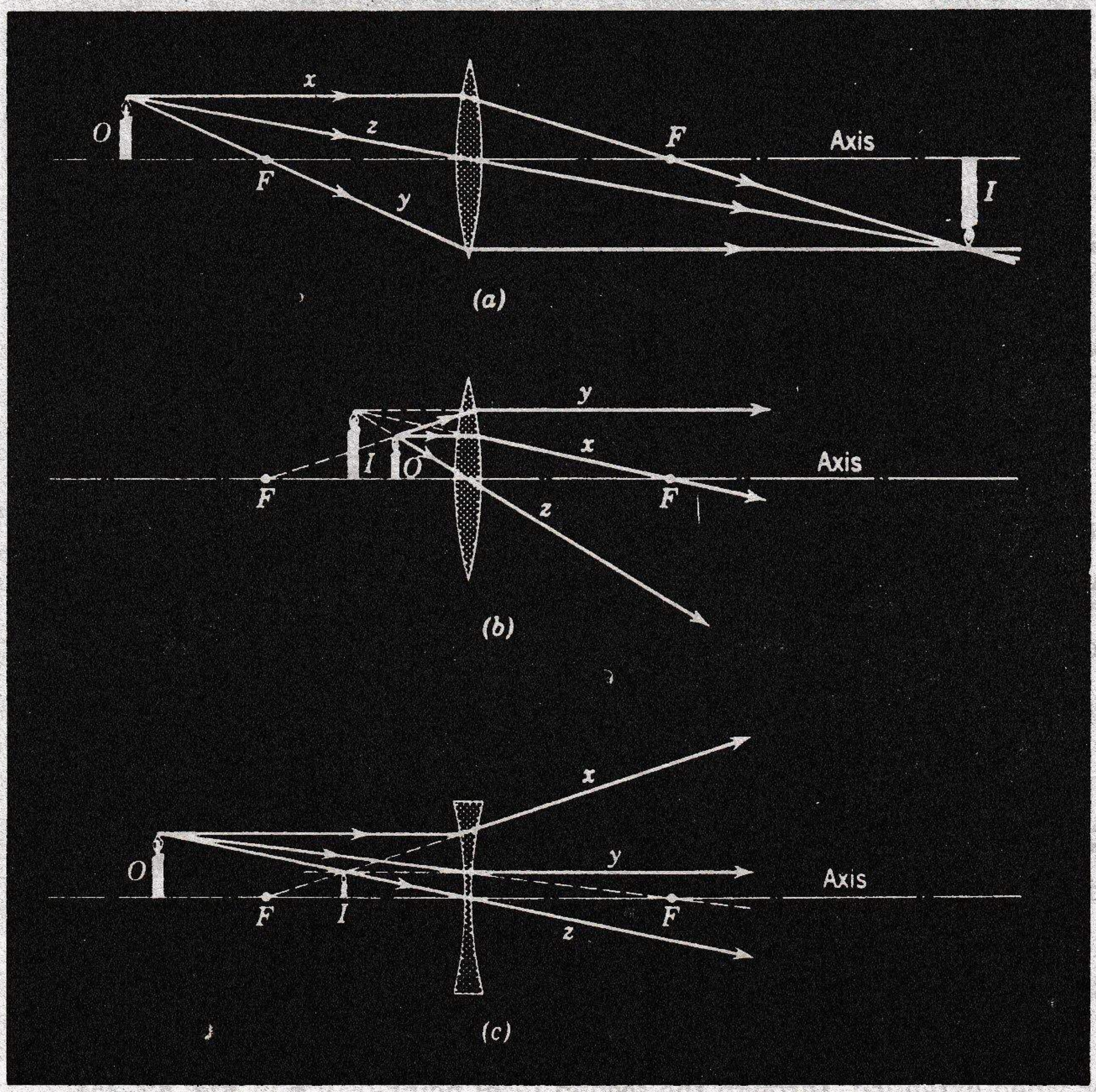


Fig. 42–25 Showing the graphical location of images for three thin lenses.

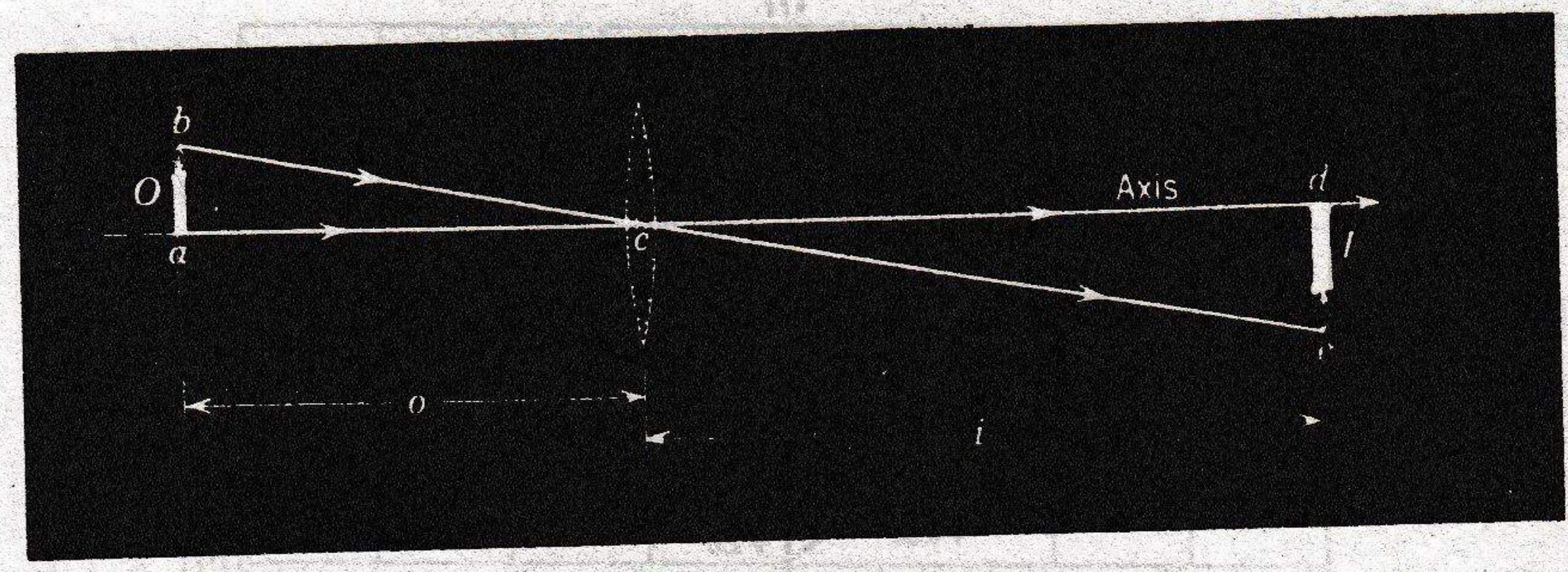


Fig. 42–26 Two rays for the situation of Fig. 42–25*a*.

lateral magnification. The minus sign is required because we wish m to be negative for an inverted image. This yields

$$m = -\frac{i}{o}, \tag{42–20}$$

which holds for all types of thin lenses and for all object distances.

▶ **Example 8.** A converging thin lens has a focal length of +24 cm. An object is placed 9.0 cm from the lens as in Fig. 42–25*b*; describe the image.

From Eq. 42–19,

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

we have

$$\frac{1}{+9.0\text{ cm}} + \frac{1}{i} = \frac{1}{+24\text{ cm}},$$

which yields $i = -14.4$ cm, in agreement with the figure. The minus sign means that the image is on the V-side of the lens and is thus virtual.

The lateral magnification is given by

$$m = -\frac{i}{o} = -\frac{-14.4\text{ cm}}{+9.0\text{ cm}} = +1.6,$$

again in agreement with the figure. The plus signifies an erect image.

Images formed by lenses suffer from defects similar to those discussed for mirrors on p. 1048. There are effects connected with the failure of a point object to form a point image, with the variation of magnification with distance from the lens axis, and with diffraction. For lenses, but not for mirrors, there are also *chromatic aberrations* associated with the fact that the refracting properties of the lens vary with wavelength because the index of refraction of the lens material does. If a point object on the lens axis emits white light, the image, neglecting other lens defects, will be a series of colored points spread out along the axis. We have all seen the colored images produced by inexpensive lenses. A great deal of ingenious optical engineering goes into the design of lenses (more commonly lens systems) in which the various lens defects are minimized. The lens surfaces are normally not spherical.

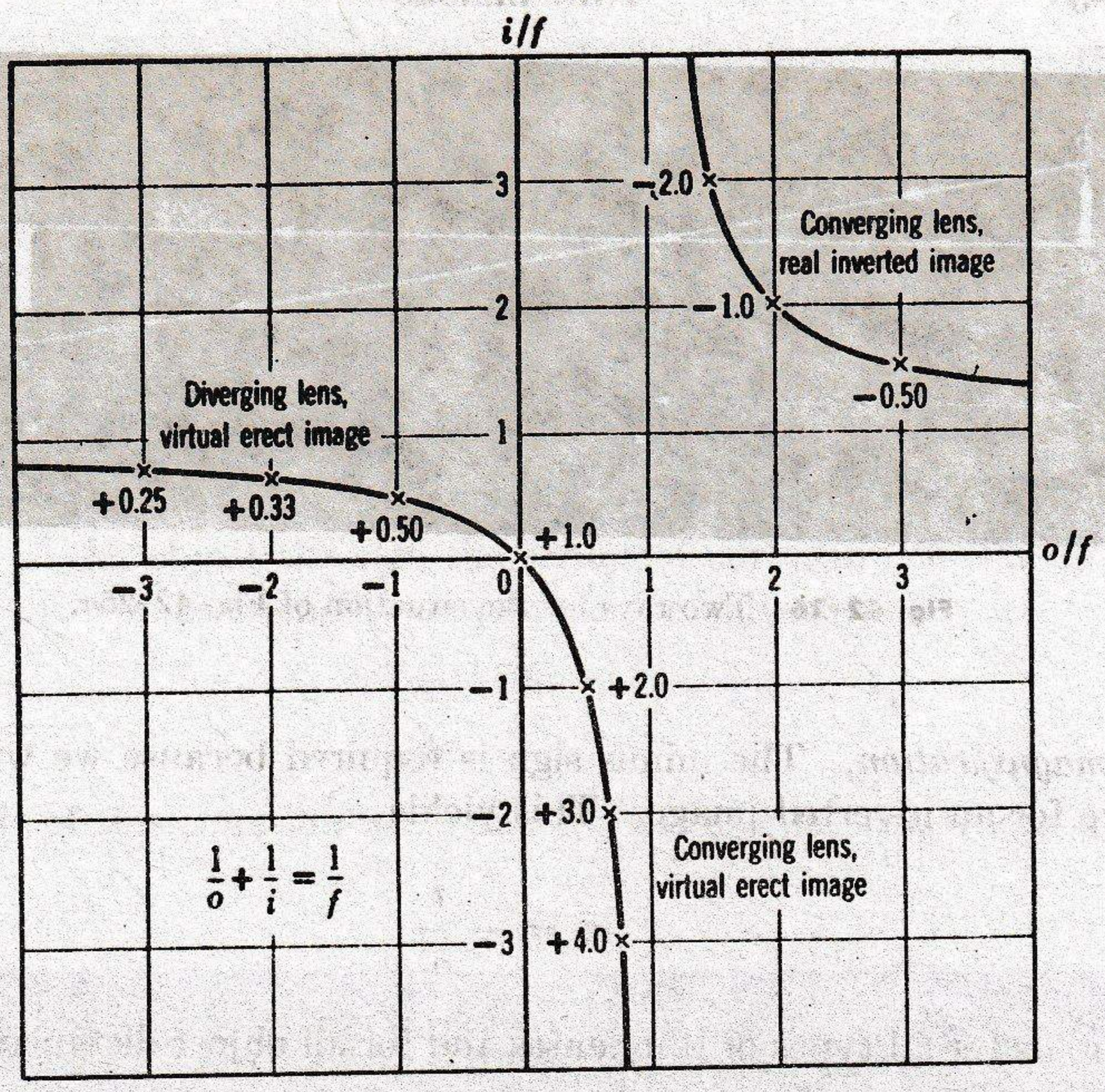


Fig. 42–27 A graphical representation of the thin-lens formula (Eq. 42–19). The numbers on the curves give the lateral magnifications (Eq. 42–20). Compare Fig. 42–16.

Figure 42–27 is a graphical representation of the object-image relationships for thin lenses; see Eqs. 42–19 and 42–20. The fact that it is identical in form with Fig. 42–16, which holds for mirrors, is not surprising because the basic equations for mirrors and thin lenses (Eqs. 42–19 and 42–5; Eqs. 42–20 and 42–6) are identical. A converging lens of a given focal length is quantitatively similar to a concave mirror of the same focal length, the difference being that the lens forms real images on the opposite side from the (real) object and the mirror forms them on the same side.

QUESTIONS

1. If a mirror reverses right and left, why doesn't it reverse up and down?
2. Is it possible to photograph a virtual image?
3. What approximations were made in deriving the mirror equation (Eq. 42–4):

$$\frac{1}{o} + \frac{1}{i} = \frac{2}{r}?$$

4. Can you think of a simple test or observation to prove that the law of reflection is the same for all wavelengths, under conditions in which geometrical optics prevails?
5. Under what conditions will a spherical mirror, which may be concave or convex, form (*a*) a real image, (*b*) an inverted image, and (*c*) an image smaller than the object?
6. An unsymmetrical thin lens forms an image of a point object on its axis. Is the image location changed if the lens is reversed?
7. Why has a lens two focal points and a mirror only one?

8. Under what conditions will a thin lens, which may be converging or diverging, form (*a*) a real image, (*b*) an inverted image, and (*c*) an image smaller than the object?

9. Does the apparent depth of an object below water depend on the angle of view of the observer in air? Explain and illustrate with ray diagrams.

10. A skin diver wants to use an air-filled plastic bag as a converging lens for underwater use. Sketch a suitable cross section for the bag.

11. What approximations were made in deriving the thin lens equation (Eq. 42–19):

$$\frac{1}{o} + \frac{1}{i} = \frac{1}{f}?$$

12. Under what conditions will a thin lens have a lateral magnification (*a*) of -1 and (*b*) of $+1$?

13. How does the focal length of a glass lens for blue light compare with that for red light, assuming the lens is (*a*) diverging and (*b*) converging?

14. Does the focal length of a lens depend on the medium in which the lens is immersed? Is it possible for a given lens to act as a converging lens in one medium and a diverging lens in another medium?

15. Are the following statements true for a glass lens in air? (*a*) A lens that is thicker at the center than at the edges is a converging lens for parallel light. (*b*) A lens that is thicker at the edges than at the center is a diverging lens for parallel light. Explain and illustrate, using wavefronts.

16. Under what conditions would the lateral magnification, ($m = -i/o$) for lenses and mirrors become infinite? Is there any practical significance to such a condition?

17. Light rays are reversible. Discuss the situation in terms of objects and images if all rays in Figs. 42–10, 42–14, 42–17, 42–19, 42–23, and 42–25 are reversed in direction.

18. What significance can be given to the origin of coordinates in the graphical representation of the thin-lens formula (Fig. 42–27)?

19. In connection with Fig. 42–24*a*, we pointed out that all rays originating on the same wavefront in the incident wave have the same optical path length to the image point. Discuss this in connection with Fermat's principle (Section 41–6).

PROBLEMS

1. Solve Example 2 if the angle between the mirrors is (*a*) 45°, (*b*) 60°, (*c*) 120°, the object always being placed on the bisector of the mirrors.

2. Two plane mirrors make an angle of 90° with each other. What is the largest number of images of an object placed between them that can be seen by a properly placed eye? The object need *not* lie on the mirror bisector.

3. A small object is 10 cm in front of a plane mirror. If you stand behind the object, 30 cm from the mirror, and look at its image, for what distance must you focus your eyes?

4. A small object *O* is placed one-third of the way between two parallel plane mirrors as in Fig. 42–28. Trace appropriate bundles of rays for viewing the four images that lie closest to the object.

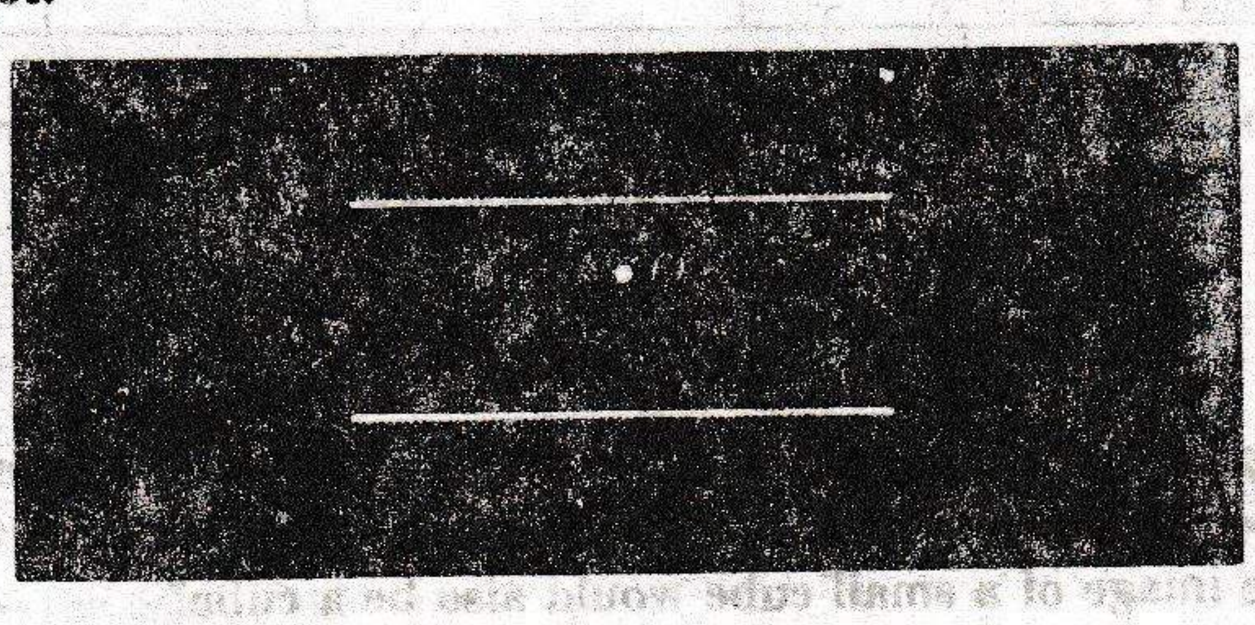


Fig. 42–28

5. Redraw Fig. 42–29 on a large sheet of paper and trace carefully the reflected rays, using the law of reflection. Is a point focus formed? Discuss.

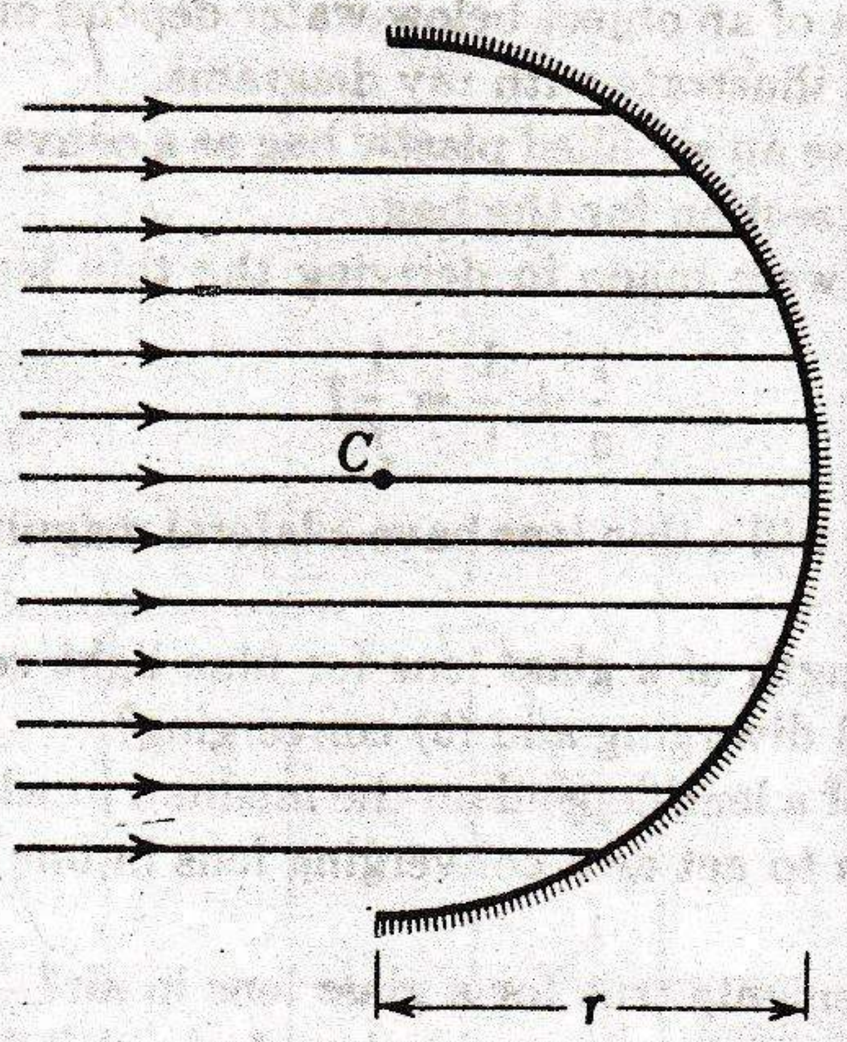


Fig. 42–29

6. Fill in this table, each column of which refers to a spherical mirror. Check your results by graphical analysis. Distances are in centimeters; if a number has no plus or minus sign in front of it, it may have either sign.

	a	*b*	*c*	*d*	*e*	*f*	*g*	*h*
Type	concave						convex	
f	20		+20			20		
r					−40		40	
i					−10		4	
o	+10	+10	+30	+60				+24
m		+1		−0.5		+0.10		0.50
Real image?		no						
Erect image?								no

7. A short linear object of length l lies on the axis of a spherical mirror, a distance o from the mirror. (*a*) Show that its image will have a length l' where

$$l' = l\left(\frac{f}{o-f}\right)^2.$$

(*b*) Show that the *longitudinal magnification* $m'(=l'/l)$ is equal to m^2 where m is the lateral magnification discussed in Section 42–3. (*c*) Is there any condition such that, neglecting all mirror defects, the image of a small cube would also be a cube?

8. As an example of the effect of relaxing the paraxial ray assumption find, analytically,

the apparent depth of a swimming pool having water ($n = 1.33$) of actual depth 8.0 ft (*a*) when viewed at normal incidence and (*b*) when viewed at an angle of 30° from the normal.

9. A layer of water ($n = 1.33$) 2.0 cm thick floats on carbon tetrachloride ($n = 1.46$) 4.0 cm thick. How far below the water surface, viewed at normal incidence, does the bottom of the tank seem to be?

10. Fill out the following table, each column of which refers to a spherical surface separating two media with different indices of refraction. Distances are measured in centimeters.

	a	*b*	*c*	*d*	*e*	*f*	*g*	*h*
n_1	1.0	1.0	1.0	1.0	1.5	1.5	1.5	1.5
n_2	1.5	1.5	1.5		1.0	1.0	1.0	
o	+10	+10		+20	+10		+70	+100
i		−13	+600	−20	−6	−7.5		+600
r	+30		+30	−20		−30	+30	−30
Real image?								

Draw a figure for each situation and construct the appropriate rays graphically. Assume a point object.

11. Fill in this table, each column of which refers to a thin lens, to the extent possible. Check your results by graphical analysis. Distances are in centimeters; if a number (except in row *n*) has no plus sign or minus sign in front of it, it may have either sign.

	a	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*
Type	converging								
f	10	+10	10	10					
r′					+30	−30	−30		
r″					−30	+30	−60		
i									
o	+20	+5	+5	+5	+10	+10	+10	+10	+10
n					1.5	1.5	1.5		
m			>1	<1				0.5	0.5
Real image?									yes
Erect image?								yes	

Draw a figure for each situation and construct the appropriate rays graphically. Assume a finite object.

12. Define and locate the first and second focal points (see p. 966) for a single spherical refracting surface such as that of Fig. 42–17.

13. A luminous object and a screen are a fixed distance D apart. (*a*) Show that a converging lens of focal length f will form a real image on the screen for two positions that are separated by

$$d = \sqrt{D(D - 4f)}.$$

(*b*) Show that the ratio of the two image sizes for these two positions is

$$\left(\frac{D - d}{D + d}\right)^2.$$



Fig. 42–30

14. Using the lens maker's equation (Eq. 42–18), decide which of the thin lenses in Fig. 42–30 is converging and which diverging for parallel incident light.

15. Two thin lenses of focal length f_1 and f_2 are in contact. Show that they are equivalent to a single thin lens with a focal length given by

$$f = \frac{f_1 f_2}{f_1 + f_2}.$$

16. Show that the distance between an object and its real image formed by a thin converging lens is always greater than four times the focal length of the lens.

17. A double-convex lens is to be made of glass with an index of refraction of 1.50. One surface is to have twice the radius of curvature of the other and the focal length is to be 6.0 cm. What are the radii?

18. The formula

$$\frac{1}{o} + \frac{1}{i} = \frac{1}{f}$$

is called the *Gaussian form* of the thin lens formula. Another form of this formula, the *Newtonian* form, is obtained by considering the distance x from the object to the first focal point and the distance x' from the second focal point to the image. Show that

$$xx' = f^2.$$

19. A parallel incident beam falls on a solid glass sphere at normal incidence. Locate the image in terms of the index of refraction n and the sphere radius r.

20. An erect object is placed a distance in front of a converging lens equal to twice the focal length f_1 of the lens. On the other side of the lens is a converging mirror of focal length f_2 separated from the lens by a distance $2(f_1 + f_2)$. (*a*) Find the location, nature and relative size of the final image. (*b*) Draw the appropriate ray diagram. See Fig. 42–31.

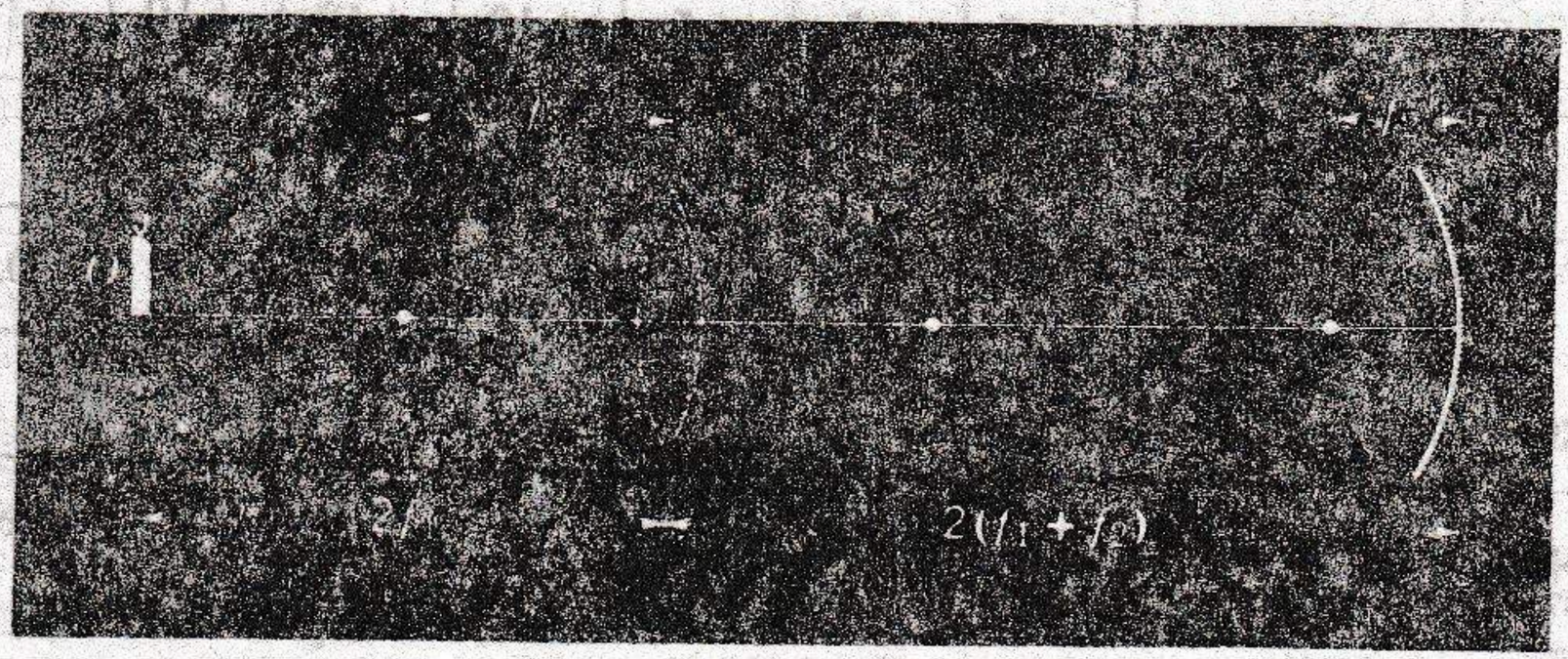


Fig. 42–31

21. An object is placed at a center of curvature of a double concave lens, both of whose radii of curvature are the same. (*a*) According to the convention used in this book, what are the signs of the two radii of curvature? (*b*) Find the location of the image in terms of the radius of curvature r and the index of refraction, n, of the glass. (*c*) Describe the nature of the image. (*d*) Verify your result with a ray diagram.

22. Two lenses with their focal points are shown in Fig. 42–32. With the given object, find the image formed by lens 1 and, using this as an object for lens 2, find the final image. Discuss the nature of the two images. Verify your conclusions by copying the figure to scale and drawing a ray diagram.

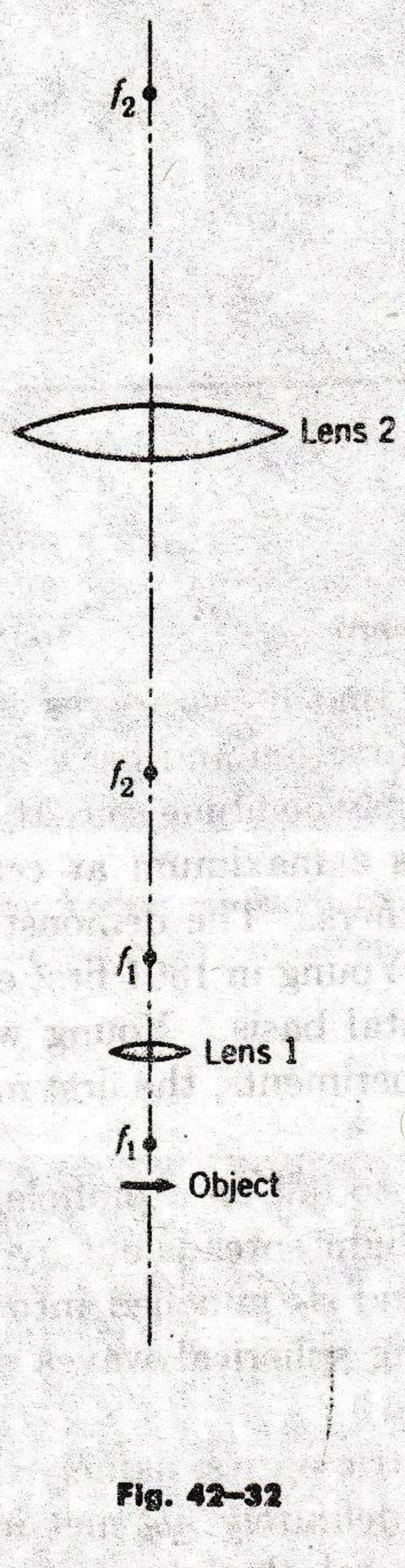


Fig. 42–32

23. An object is 20 cm to the left of a lens with a focal length of +10 cm. A second lens of focal length +12.5 cm is 30 cm to the right of the first lens. (*a*) Using the image formed by the first lens as the object for the second, find the location and relative size of the final image. (*b*) Verify your conclusions by drawing the lens system to scale and constructing a ray diagram. (*c*) Describe the final image.

Interference

CHAPTER 43

43–1 Young's Experiment

In Section 19–7 we saw that if two waves of the same frequency travel in approximately the same direction and have a phase difference that remains constant with time, they may combine so that their energy is not distributed uniformly in space but is a maximum at certain points and a minimum (possibly even zero) at others. The demonstration of such *interference* effects for light by Thomas Young in 1801 first established the wave theory of light on a firm experimental basis. Young was able to deduce the wavelength of light from his experiments, the first measurement of this important quantity.

Young allowed sunlight to fall on a pinhole S_0 punched in a screen A in Fig. 43–1. The emerging light spreads out by diffraction (see Section 42–1) and falls on pinholes S_1 and S_2 punched into screen B. Again diffraction occurs and two overlapping spherical waves expand into the space to the right of screen B.

The condition for geometric optics, namely that $a \gg \lambda$ where a is the diameter of the pinholes, is definitely *not* met in this experiment. The pinholes do not cast geometrical shadows but act as sources of expanding Huygens' wavelets. We are dealing here (and in succeeding chapters) with *wave optics* rather than with geometrical optics.

Figure 43–2, taken from an 1803 paper of Young, shows the region between screens B and C. The blackening represents the minima of the wave disturbance; the white space between represents the maxima. If you hold the page with your eye close to the left edge and look at a grazing angle along the figure, you will see that along lines marked by x's there is cancella-

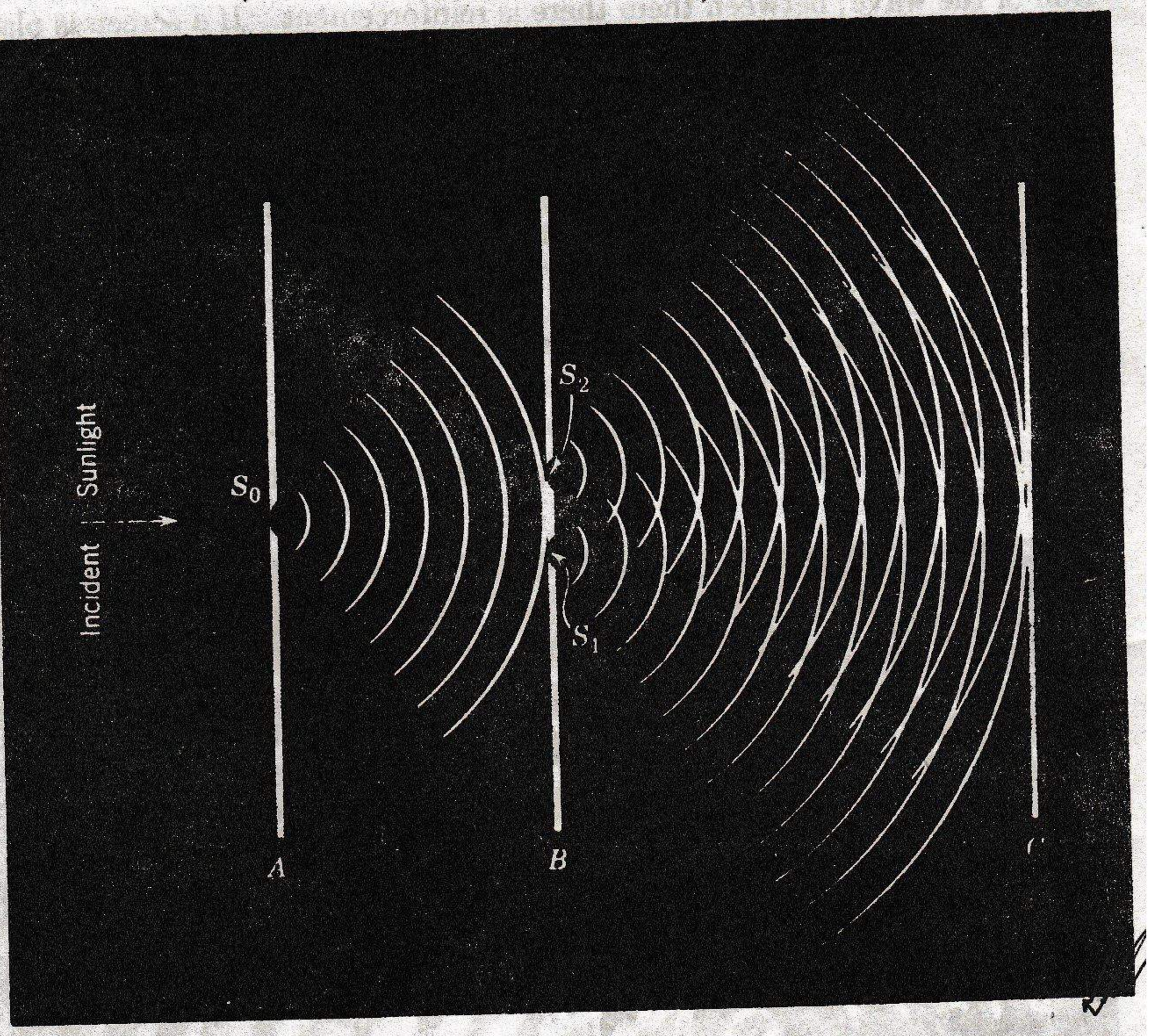


Fig. 43–1 Showing how Thomas Young produced an interference pattern by allowing diffracted waves from pinholes S_1 and S_2 to overlap on screen C.

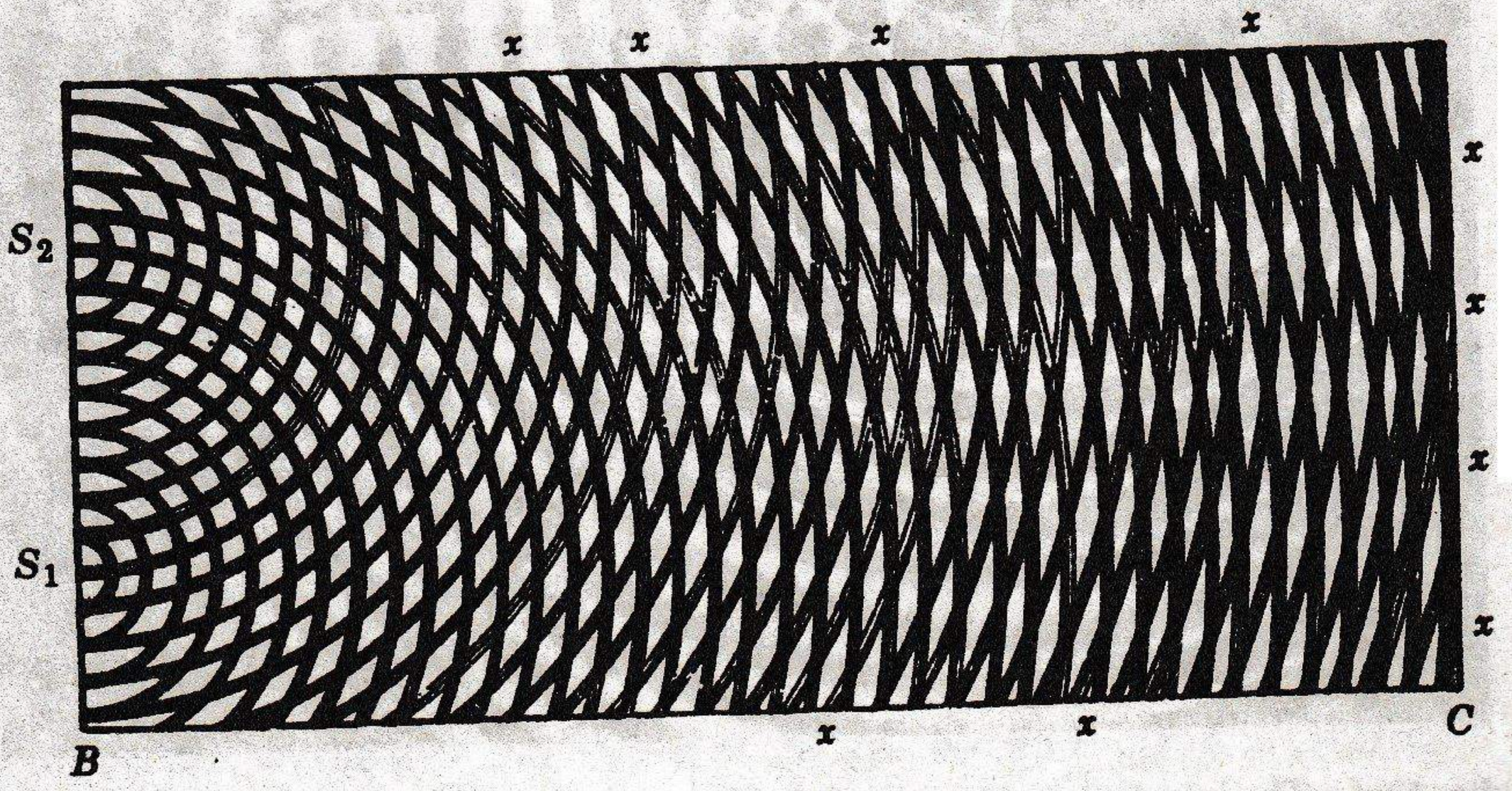


Fig. 43–2 Thomas Young's original drawing showing interference effects in overlapping waves. Place the eye near the left edge and sight at a grazing angle along the figure. (From Thomas Young, *Phil. Transactions*, 1803.)

tion of the wave; between them there is reinforcement. If a screen is placed anywhere across the superimposed waves, we expect to find alternate bright and dark spots on it. Figure 43–3 shows a photograph of such *interference*



Fig. 43–3 Interference fringes for monochromatic light, made with an arrangement like that of Fig. 43–1, using long narrow slits rather than pinholes.

fringes; in keeping with modern technique, long narrow slits rather than pinholes were used in preparing this figure.

Interference is not limited to light waves but is a characteristic of all wave phenomena. Figure 43–4, for example, shows the interference pattern of



Fig. 43–4 The interference of water waves in a ripple tank. There is destructive interference along the lines marked "Line of nodes" and constructive interference between these lines. (Courtesy Physical Science Study Committee.)

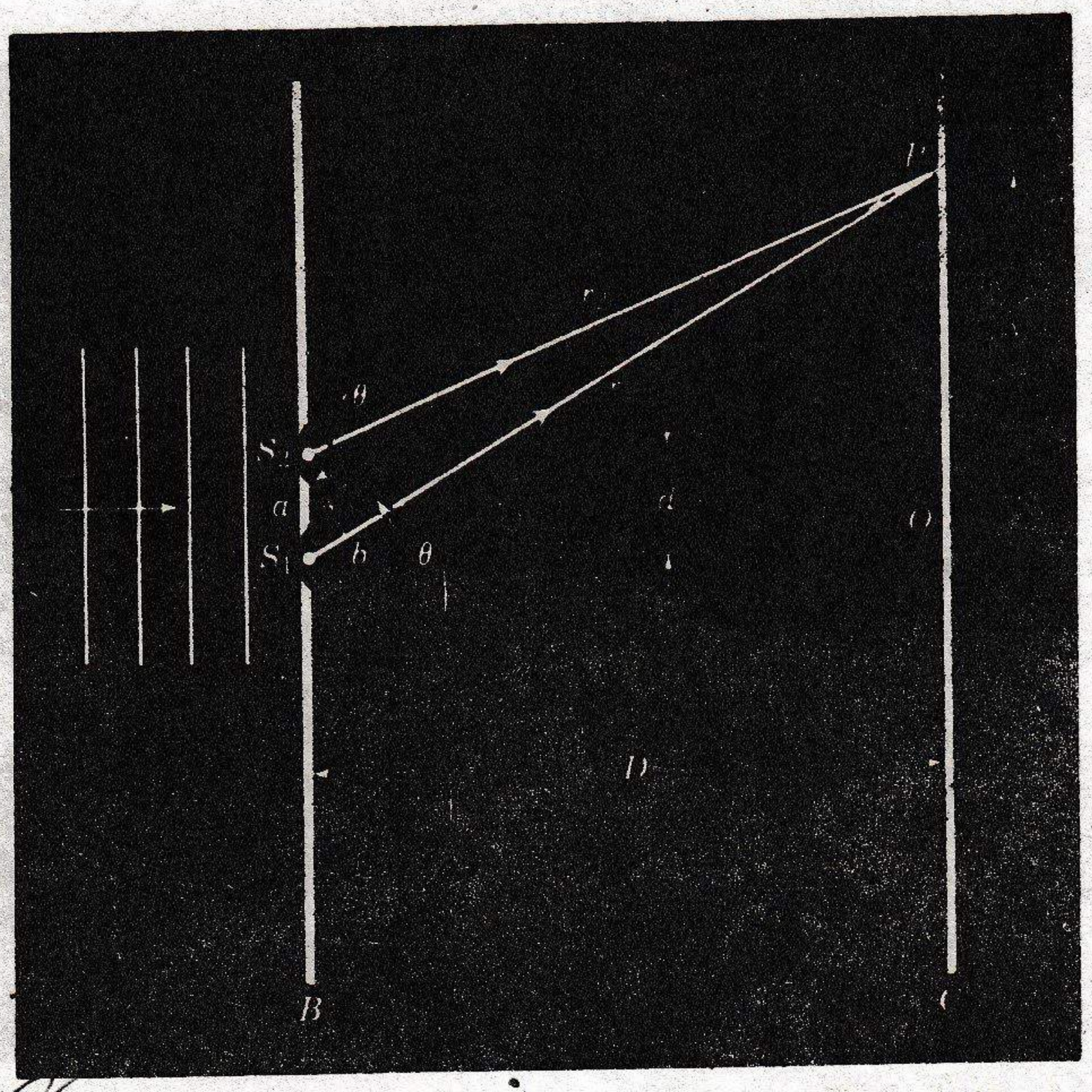


Fig. 43–5 Rays from S_1 and S_2 combine at P. The light falling on screen B has been taken as parallel. Actually, $D \gg d$, the figure being distorted for clarity.

water waves in a shallow *ripple tank*. The waves are generated by two vibrators that tap the water surface in synchronism, producing two expanding spherical waves.

Let us analyze Young's experiment quantitatively, assuming that the incident light consists of a single wavelength only. In Fig. 43–5 P is an arbitrary point on the screen, a distance r_1 and r_2 from the narrow slits S_1 and S_2, respectively. Let us draw a line from S_2 to b in such a way that the lines PS_2 and Pb are equal. If d, the slit spacing, is much smaller than the distance D between the two screens (the ratio d/D in the figure has been exaggerated for clarity), S_2b is then almost perpendicular to both r_1 and r_2. This means that angle S_1S_2b is almost equal to angle PaO, both angles being marked θ in the figure. This is equivalent to saying that the lines r_1 and r_2 may be taken as parallel.

We often put a lens in front of the two slits, as in Fig. 43–6, the screen C being in the focal plane of the lens. Under these conditions light focused at P must have struck the lens parallel to the line Px, drawn from P through the center of the (thin) lens. Under these conditions rays r_1 and r_2 are strictly parallel even though the requirement $D \gg d$ is not met. The lens L may in practice be the lens and cornea of the eye, screen C being the retina.

The two rays arriving at P in Figs. 43–5 or 43–6 from S_1 and S_2 are in phase at the source slits, both being derived from the same wavefront in the incident plane wave. Because the rays have different optical path lengths,

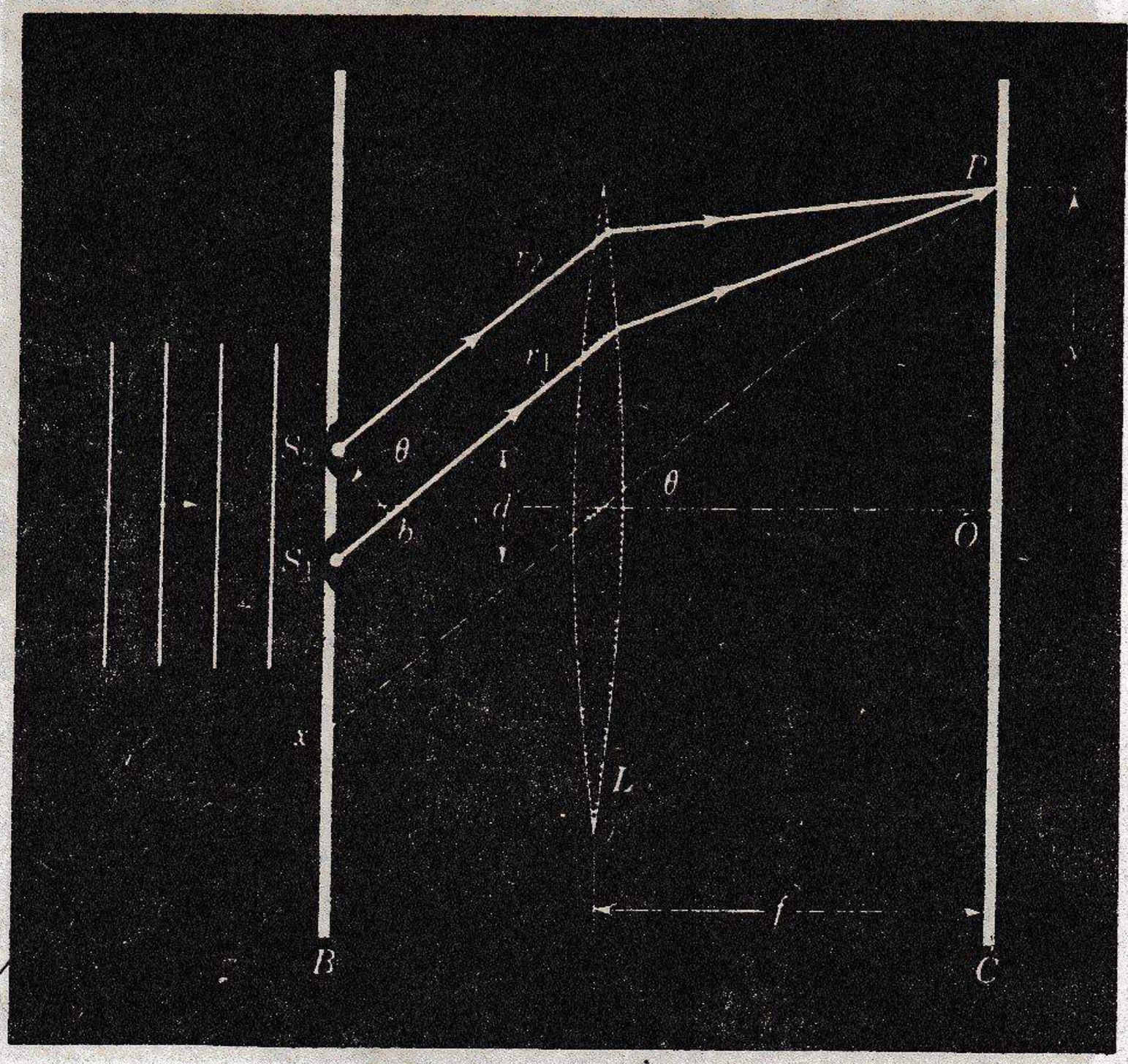


Fig. 43–6 A lens is normally used to produce interference fringes; compare with Fig. 43–5. The figure is again distorted for clarity in that $f \gg d$ in practice.

they arrive at P with a phase difference. The number of wavelengths contained in S_1b, which is the path difference, determines the nature of the interference at P.

To have a *maximum* at P, S_1b $(= d \sin\theta)$ must contain an integral number of wavelengths, or

$$S_1b = m\lambda \qquad m = 0, 1, 2, \ldots,$$

which can be written as

$$d \sin\theta = m\lambda \qquad m = 0, 1, 2, \ldots \text{ (maxima)}. \tag{43–1}$$

Note that each maximum above O in Figs. 43–5 and 43–6 has a symmetrically located maximum below O. There is a central maximum described by $m = 0$.

For a *minimum* at P, S_1b $(= d \sin\theta)$ must contain a half-integral number of wavelengths, or

$$d \sin\theta = (m + \tfrac{1}{2})\lambda \qquad m = 0, 1, 2, \ldots \text{ (minima)}. \tag{43–2}$$

If a lens is used as in Fig. 43–6, it may seem that a phase difference should develop between the rays beyond the plane represented by S_2b, the path lengths between this plane and P being clearly different. In Section 42–5, however, we saw that for such parallel rays focused by a lens the *optical path lengths* are identical. Two rays with the same optical path lengths contain the same number of wavelengths, so that no phase difference will result because of the light passing through the lens.

Example 1. The double-slit arrangement in Fig. 43–5 is illuminated with light from a mercury vapor lamp so filtered that only the strong green line ($\lambda = 5460$ A) is effective. The slits are 0.10 mm apart, and the screen on which the interference pattern appears is 20 cm away. What is the angular position of the first minimum? Of the tenth maximum?

At the first minimum we put $m = 0$ in Eq. 43–2, or

$$\sin\theta = \frac{(m + \frac{1}{2})\lambda}{d} = \frac{(\frac{1}{2})(546 \times 10^{-9}\ \text{meter})}{0.10 \times 10^{-3}\ \text{meter}} = 0.0027.$$

This value for $\sin\theta$ is so small that we can take it to be the value of θ, expressed in radians; expressed in degrees it is 0.16°.

At the tenth maximum (not counting the central maximum) we must put $m = 10$ in Eq. 43–1. Doing so and calculating as before leads to an angular position of 3.8°. For these conditions we see that the angular spread of the first dozen or so fringes is small.

Example 2. In Example 1 what is the linear distance on screen C between adjacent maxima?

If θ is small enough, we can use the approximation

$$\sin\theta \cong \tan\theta \cong \theta.$$

From Fig. 43–5 we see that

$$\tan\theta = \frac{y}{D}.$$

Substituting this into Eq. 43–1 for $\sin\theta$ leads to

$$y = m\frac{\lambda D}{d} \qquad m = 0, 1, 2, \ldots \text{ (maxima)}.$$

The positions of any two adjacent maxima are given by

$$y_m = m\frac{\lambda D}{d}$$

and

$$y_{m+1} = (m + 1)\frac{\lambda D}{d}.$$

Their separation Δy is found by subtracting:

$$\Delta y = y_{m+1} - y_m = \frac{\lambda D}{d}$$

$$= \frac{(546 \times 10^{-9}\ \text{meter})(20 \times 10^{-2}\ \text{meter})}{0.10 \times 10^{-3}\ \text{meter}} = 1.09\ \text{mm}.$$

As long as θ in Figs. 43–5 and 43–6 is small, the separation of the interference fringes is independent of m; that is, the fringes are evenly spaced. Note that if the incident light contains more than one wavelength the separate interference patterns, which will have different fringe spacings, will be superimposed. ◀

Equation 43–1 can be used to determine the wavelength of light; to quote Thomas Young:

> From a comparison of various experiments, it appears that the breadth of the undulations [that is, the wavelength] constituting the extreme red light must be supposed to be, in air, about one 36 thousandth of an inch, and those of the extreme violet about one 60 thousandth; the mean of the whole spectrum, with respect to the intensity of light, being one 45 thousandth.

Young's value for the average effective wavelength present in sunlight (1/45,000 in.) can be written as 5700 A, which agrees rather well with the wavelength at which the eye sensitivity is a maximum, 5550 A (see Fig. 40–2). It must not be supposed that Young's work was received without criticism. One of his contemporaries, evidently a firm believer in the corpuscular theory of light, wrote:

> We wish to raise our feeble voice against innovations that can have no other effect than to check the progress of science, and renew all those wild phantoms of the imagination which Bacon and Newton put to flight from her temple. This paper contains nothing that deserves the name of either experiment or discovery.

Needless to say, posterity has decided in favor of Young.

43–2 Coherence

Analysis of the derivation of Eqs. 43–1 and 43–2 shows that a fundamental requirement for the existence of well-defined interference fringes on screen C in Fig. 43–1 is that the light waves that travel from S_1 and S_2 to any point P on this screen must have a sharply defined phase difference ϕ that remains constant with time. If this condition is satisfied, a stable, well-defined fringe pattern will appear. At certain points P, ϕ will be given, independent of time, by $n\pi$ where $n = 1, 3, 5, \ldots$ so that the resultant intensity will be strictly zero and will remain so throughout the time of observation. At other points ϕ will be given by $n\pi$ where $n = 0, 2, 4 \ldots$ and the resultant intensity will be a maximum. Under these conditions the two beams emerging from slit S_1 and S_2 are said to be completely *coherent*.

Let the source in Fig. 43–1 be removed and let slits S_1 and S_2 be replaced by two completely independent light sources, such as two fine incandescent wires placed side by side in a glass envelope. No interference fringes will appear on screen C but only a relatively uniform illumination. We can interpret this if we make the reasonable assumption that for completely independent light sources the phase difference between the two beams arriving at P will vary with time in a random way. At a certain instant conditions may be right for cancellation and a short time later (perhaps 10^{-8} sec) they may be right for re-enforcement. This same random phase behavior holds for all points on screen C with the result that this screen is uniformly illuminated. The intensity at any point is equal to the sum of the intensities that each source S_1 and S_2 produces separately at that point. Under these conditions the two beams emerging from S_1 and S_2 are said to be completely *incoherent*.

Note that for completely coherent light beams one (1) combines the amplitudes vectorially, taking the (constant) phase difference properly into account, and then (2) squares this resultant amplitude to obtain a quantity proportional to the resultant intensity. For completely incoherent light beams, on the other hand, one (1) squares the individual amplitudes to obtain quantities proportional to the individual intensities and then (2) adds the individual intensities to obtain the resultant intensity. This procedure is in agreement with the experimental fact that for completely independent light sources the resultant intensity at every point is always greater than the intensity produced at that point by either light source acting alone.

It remains to investigate under what experimental conditions coherent or incoherent beams may be produced and to give an explanation for coherence in terms of the mode of production of the radiation. Consider first a parallel beam of microwave radiation emerging from an antenna connected by a coaxial cable to an oscillator based on an electromagnetic resonant cavity. The cavity oscillations (see Section 38–6) are completely periodic with time and produce, at the antenna, a completely periodic variation of **E** and **B** with time. The radiated wave at large enough distances from the antenna is well represented by Fig. 39–11. Note that (1) the wave has essentially infinite extent in time, including both future times ($t > 0$, say) and past times ($t < 0$); see Fig. 43–7*a*. At any point, as the wave passes by, the wave disturbance (i.e., **E** or **B**) varies with time in a perfectly periodic way. (2) The wavefronts at points far removed from the antenna are parallel planes of essentially infinite extent at right angles to the propagation direction. At any instant of time the wave disturbance varies with distance along the propagation direction in a perfectly periodic way.

Two beams generated from a single traveling wave like that of Fig. 39–11 will be completely coherent. One way to generate two such beams is to put an opaque screen containing two slits in the path of the beam. The waves emerging from the slits will always have a constant phase difference at any point in the region in which they overlap and interference fringes will be produced. Coherent radio beams can also be readily established, as can coherent elastic waves in solids, liquids and gases. The two prongs of the vibrating tapper in Fig. 43–4, for example, generate two coherent waves in the water of the ripple tank.

The technique of producing two beams from a single beam (and thus from a single source) tests specifically whether the wavefronts in the single parallel beam are truly planes, that is, whether all points in a plane at right angles to the direction of propagation have the same phase at any given instant. By dividing the beam in another way it is possible to test whether the beam is truly periodic over a large number of cycles of oscillation. This can be done, as we show in detail in Section 43–6, by inserting in the beam at 45° to it a thin sheet of material possessing the property that two beams are produced, one (which will be at right angles to the incident beam) by reflection and a second (which will be in the direction of the incident beam) by transmission. In the visible region such a sheet, called a *half-silvered mirror*, may be formed from a glass plate by depositing on it an appropriately thin film of silver. By appropriate use of mirrors (see Section 43–7) these two sub-beams can be recombined into a single beam traveling in a chosen direction. If the beams travel different distances before they are recombined, we are comparing in the combined beam a sample of the original beam with another sample an arbitrarily large number of cycles away. If the original beam is truly periodic in space and time, the two sub-beams will be completely coherent and interference fringes will be produced when they are recombined.

If we turn from microwave sources to common sources of visible light, such as incandescent wires or an electric discharge passing through a gas, we become aware of a fundamental difference. In both of these sources the fundamental light emission processes occur in individual atoms and these atoms do not act together in a cooperative (i.e., *coherent*) way. The act of

light emission by a single atom takes, in a typical case, about 10^{-8} sec and the emitted light is properly described as a *wavetrain* (Fig. 43–7*b*) rather than as a wave (Fig. 43–7*a*). For emission times such as these the wavetrains are a few meters long.

Interference effects from ordinary light sources may be produced by putting a very narrow slit (S_0 in Fig. 43–1) directly in front of the source. This insures that the wavetrains that strike slits S_1 and S_2 in screen B in this figure originate from the same small region of the source. The diffracted beams emerging from S_1 and S_2 thus represent the same population of wavetrains and are coherent with respect to each other. If the phase of the light emitted from S_0 changes, this change is transmitted simultaneously to S_1 and S_2. Thus, at any point on screen C, a constant phase difference is maintained between the beams from these two slits and a stationary interference pattern occurs.

If the width of slit S_0 in Fig. 43–1 is gradually increased it will be observed experimentally that the maxima of the interference fringes become reduced in intensity and that the intensity in the fringe minima is no longer strictly zero. In other words, the fringes become less distinct. If S_0 is opened extremely wide, the lowering of the maximum intensity and the raising of the minimum intensity will be such that the fringes disappear, leaving only a uniform illumination. Under these conditions we say that the beams from S_1 and S_2 pass continuously from a condition of complete coherence to one of complete incoherence. When not at either of these two limits the beams are said to be partially coherent.

Partial coherence can also be demonstrated in two beams that are produced, as described on p. 1090 by inserting a "half-silvered mirror" in a beam at an angle of 45° to the direction of propagation. The two beams so produced, by reflection and transmission, traverse paths of different lengths before they are recombined. If the path difference is small compared to the average length of a wavetrain, the interference fringes will be sharply defined and will go essentially to zero at their minima. If the path difference is deliberately made longer, the fringes will become less distinct, and finally, when the path difference is larger than the average length of a wavetrain, the fringes will disappear altogether. Thus it is possible once again to progress smoothly in an experimental arrangement from complete coherence, through partial coherence, to complete incoherence.

The lack of coherence of the light from ordinary sources such as glowing wires is due to the fact that the emitting atoms do not act cooperatively

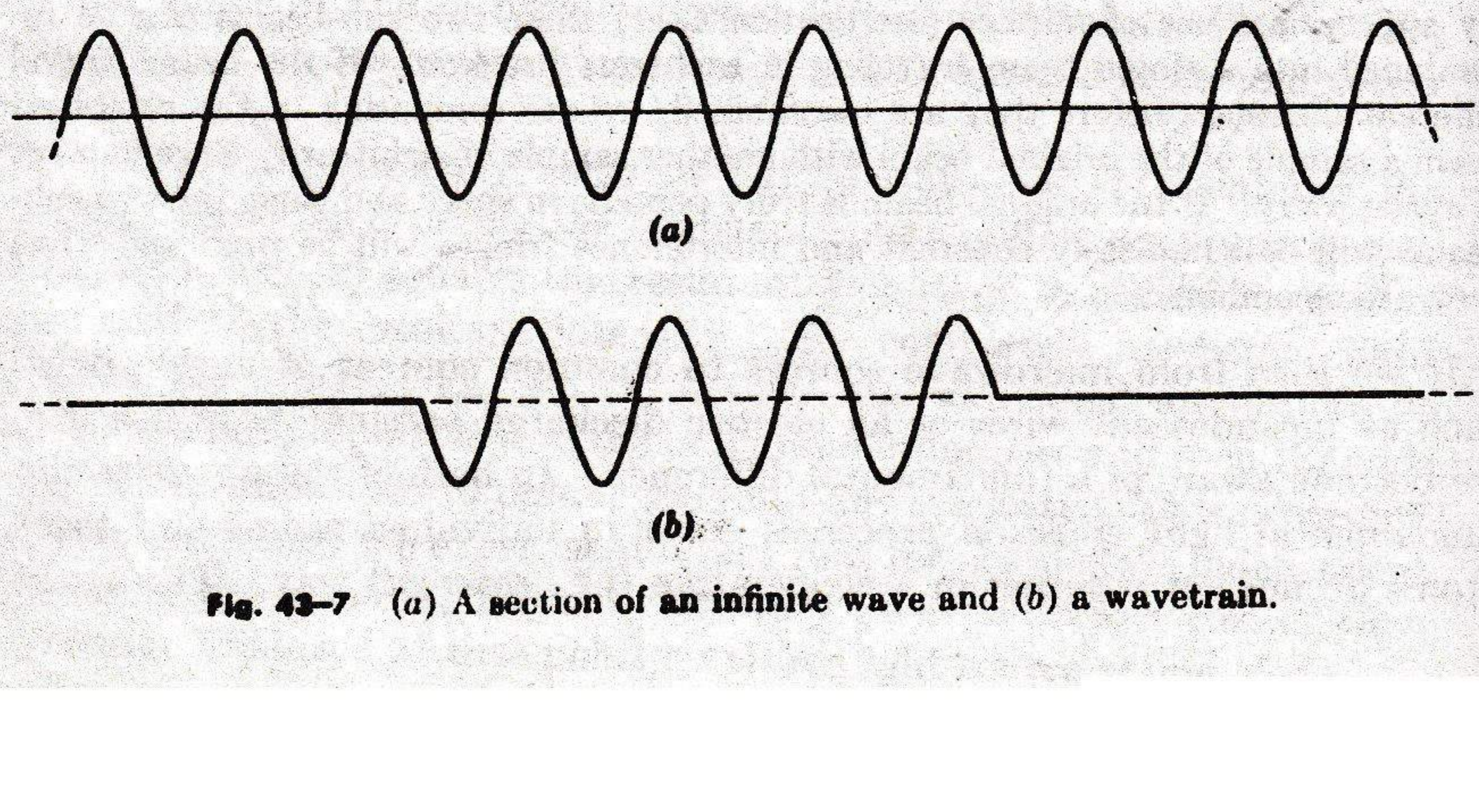


Fig. 43–7 (*a*) A section of an infinite wave and (*b*) a wavetrain.

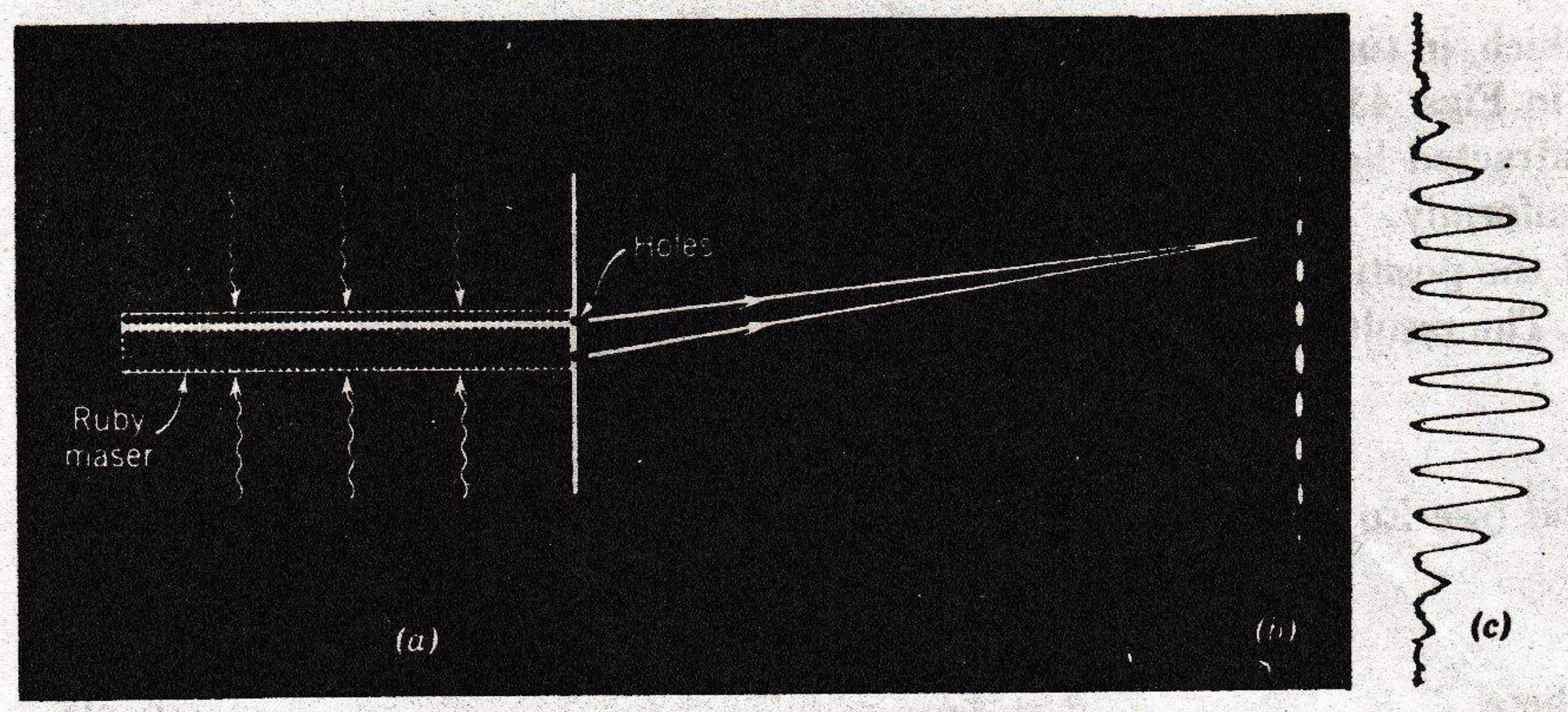


Fig. 43–8 A screen with two small holes is placed against the end of an optical maser. (*a*) The light passing through the holes forms an interference pattern on a strip of photographic film placed at (*b*). The fact that the pattern is formed shows that the light emitted from the maser is coherent across the beam cross section. At (*b*) is shown the image as formed on a photographic strip exposed in this manner. At (*c*) is an intensity plot of the film made by measuring the degree to which the film has been exposed. (Courtesy of D. F. Nelson and R. J. Collins, Bell Telephone Laboratories.)

(i.e., *coherently*). Since 1960 it has proved possible to construct sources of visible light in which the atoms *do* act cooperatively and in which the emitted light is highly coherent. Such devices are called *optical masers* or *lasers*; * their light output is extremely monochromatic, intense, and highly collimated. The coherence of the emitted light can be demonstrated by placing a screen with two holes in it in the beam emerging from the laser. An interference pattern results, as Fig. 43–8 shows. These methods permit, for the first time, a degree of control of visible light approaching that possible for radio and for microwaves. The practical applications of lasers, which include the amplification of weak light signals, the use of light beams as highly efficient carriers of information from point to point (see problem 13), and the production of high temperatures by intense local heating, remain to be fully exploited.

43–3 Intensity in Young's Experiment

Let us assume that the electric field components of the two waves in Fig. 43–5 vary with time at point P as

$$E_1 = E_0 \sin \omega t \tag{43–3}$$

and

$$E_2 = E_0 \sin (\omega t + \phi) \tag{43–4}$$

where ω $(= 2\pi\nu)$ is the angular frequency of the waves and ϕ is the phase difference between them. Note that ϕ depends on the location of point P,

* Laser, like maser, is a coined word, meaning "*l*ight *a*mplification through *s*timulated *e*mission of *r*adiation." The interested student is referred to an article called "Optical Masers" by Arthur L. Schawlow, *Scientific American*, June 1961.

which, in turn, for a fixed geometrical arrangement, is described by the angle θ in Figs. 43–5 and 43–6. We assume that the slits are so narrow that the diffracted light from each slit illuminates the central portion of the screen uniformly. This means that near the center of the screen E_0 is independent of the position of P, that is, of the value of θ.

The resultant wave disturbance * at P is found from

$$E = E_1 + E_2 \qquad (43\text{–}5)$$

and (see Eq. 19–17) proves to be

$$E = E_\theta \sin(\omega t + \beta), \qquad (43\text{–}6a)$$

where $$\beta = \tfrac{1}{2}\phi \qquad (43\text{–}6b)$$

and $$E_\theta = 2E_0 \cos\beta = E_m \cos\beta. \qquad (43\text{–}6c)$$

E_m, the maximum possible amplitude for E_θ, is equal to twice the amplitude of the combining waves ($= 2E_0$), corresponding to complete reinforcement. The student should verify Eq. 43–6 carefully. The amplitude E_θ of the resultant wave disturbance, which determines the intensity of the interference fringes, will turn out to depend strongly on the value of θ, that is, on the location of point P in Figs. 43–5 and 43–6.

In Section 19–6 we showed that the intensity of a wave I, measured perhaps in watts/meter^2, is proportional to the square of its amplitude. For the resultant wave then, ignoring the proportionality constant,

$$I_\theta \propto E_\theta^2. \qquad (43\text{–}7)$$

This relationship seems reasonable if we recall (Eq. 30–27) that the energy density in an electric field is proportional to the *square* of the electric field strength. This is true for rapidly varying electric fields, such as those in a light wave, as well as for static fields.

The ratio of the intensities of two light waves is the ratio of the squares of the amplitudes of their electric fields. If I_θ is the intensity of the resultant wave at P and I_0 is the intensity that a single wave acting alone would produce, then

$$\frac{I_\theta}{I_0} = \left(\frac{E_\theta}{E_0}\right)^2. \qquad (43\text{–}8)$$

Combining with Eq. 43–6c leads to

$$I_\theta = 4I_0 \cos^2\beta = I_m \cos^2\beta. \qquad (43\text{–}9)$$

---

* The electric field **E** in the light wave rather than the magnetic field **B** is normally identified with the "wave disturbance" because the effects of **B** on the human eye and on various light detectors are exceedingly small. Radiation pressure (Section 40–2) is one such effect. Note too that, although Eq. 43–5 should be a vector equation, in most cases of interest the **E** vectors in the two interfering waves are closely parallel so that an algebraic equation suffices.

Note that the intensity of the resultant wave at any point P varies from zero [for a point at which $\phi\ (= 2\beta) = \pi$, say] to I_m, which is four times the intensity I_0 of each individual wave [for a point at which $\phi\ (= 2\beta) = 0$, say]. Let us compute I_θ as a function of the angle θ in Figs. 43–5 or 43–6.

The phase difference ϕ in Eq. 43–4 is associated with the path difference S_1b in Fig. 43–5 or 43–6. If S_1b is $\frac{1}{2}\lambda$, ϕ will be π; if S_1b is λ, ϕ will be 2π, etc. This suggests that

$$\frac{\text{phase difference}}{2\pi} = \frac{\text{path difference}}{\lambda},$$

$$\phi = \frac{2\pi}{\lambda}(d \sin\theta),$$

or, finally, from Eq. 43–6*b*,

$$\beta = \tfrac{1}{2}\phi = \frac{\pi d}{\lambda}\sin\theta. \qquad (43\text{–}10)$$

This expression for β can be substituted into Eq. 43–9 for I_θ, yielding the latter quantity as a function of θ. For convenience we collect here the expressions for the amplitude and the intensity in double-slit interference.

[Eq. 43–6*c*]	$E_\theta = E_m \cos\beta$	interference	(43–11*a*)
[Eq. 43–9]	$I_\theta = I_m \cos^2\beta$	from narrow	(43–11*b*)
[Eq. 43–10]	$\beta\ (= \frac{1}{2}\phi) = \frac{\pi d}{\lambda}\sin\theta$	slits (that is, $a \ll \lambda$)	(43–11*c*)

To find the positions of the intensity maxima, we put

$$\beta = m\pi \qquad m = 0, 1, 2, \ldots$$

in Eq. 43–11*b*. From Eq. 43–11*c* this reduces to

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

which is the equation derived in Section 43–1 (Eq. 43–1). To find the intensity minima we write

$$\frac{\pi d \sin\theta}{\lambda} = (m + \tfrac{1}{2})\pi \qquad m = 0, 1, 2, \ldots \text{ (minima)},$$

which reduces to the previously derived Eq. 43–2.

Figure 43–9 shows the intensity pattern for double-slit interference. The horizontal solid line is I_0; this describes the (uniform) intensity pattern on the screen if one of the slits is covered up. If the two sources were incoherent the intensity would be uniform over the screen and would be $2I_0$; see the horizontal dashed line in Fig. 43–9. For coherent sources we expect the

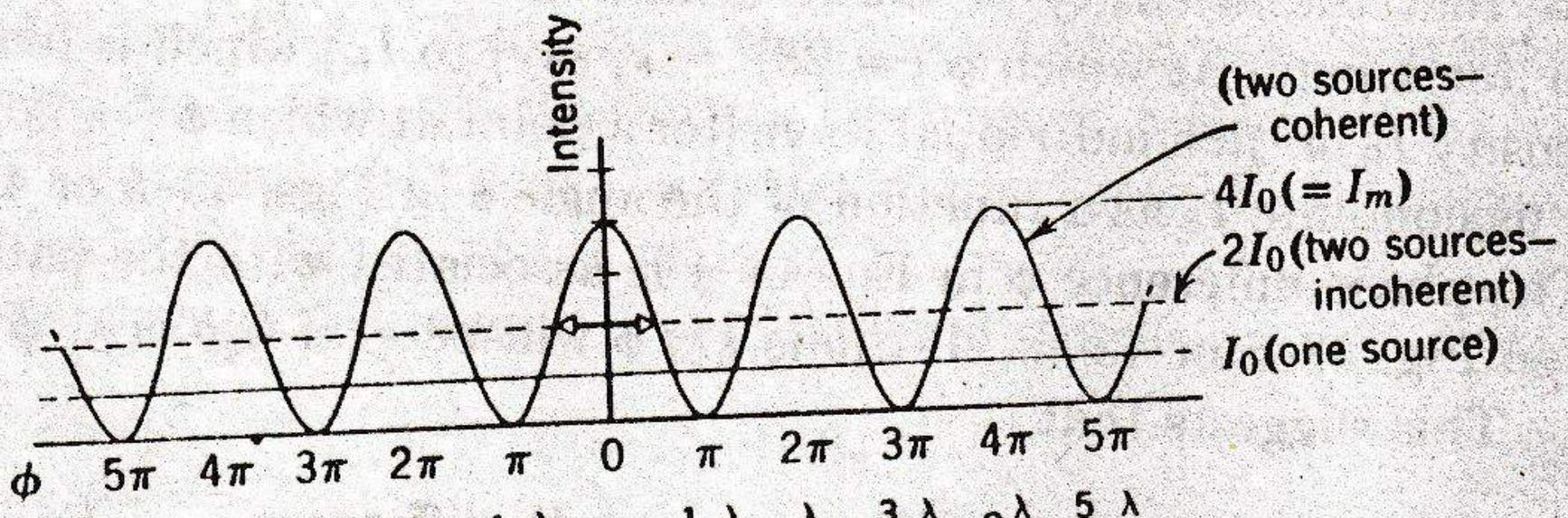


Fig. 43–9 The intensity pattern for double-slit interference. The heavy arrow in the central peak represents the half-width of the peak. This figure is constructed on the assumption that the two interfering waves each illuminate the central portion of the screen uniformly, that is, I_0 is independent of position as shown.

energy to be merely redistributed over the screen, since energy is neither created nor destroyed by the interference process. Thus the *average* intensity in the interference pattern should be $2I_0$, as for incoherent sources. This follows at once if, in Eq. 43–11*b*, we substitute one-half for the cosine-squared term and if we recall that $I_m = 4I_0$. We have seen several times that the average value of the square of a sine or a cosine term over one or more half-cycles is one-half.

43–4 Adding Wave Disturbances

In Section 43–3 we combined two time-varying wave disturbances, namely

$$E_1 = E_0 \sin \omega t \tag{43–3}$$

and

$$E_2 = E_0 \sin (\omega t + \phi), \tag{43–4}$$

which have the same angular frequency ω and amplitude E_0 but which have a phase difference ϕ between them. In this case the result (Eqs. 43–11*a* and *c*) is easily obtained algebraically.

In later chapters we will want to add larger numbers of wave disturbances, often an infinite number, with infinitesimal individual amplitudes. Since analytic methods become more difficult in such cases we describe a graphical method, illustrating it by rederiving Eq. 43–11*a*.

A sinusoidal wave disturbance such as that represented by Eq. 43–3 can be represented graphically, using a rotating vector. In Fig. 43–10*a* a vector of magnitude E_0 is allowed to rotate about the origin in a counterclockwise direction with an angular frequency ω. Following electrical engineering practice we call such a rotating vector a *phasor*. The alternating wave disturbance E_1 (Eq. 43–3) is represented by the projection of this phasor on the vertical axis.

A second wave disturbance E_2, which has the same amplitude E_0 but a

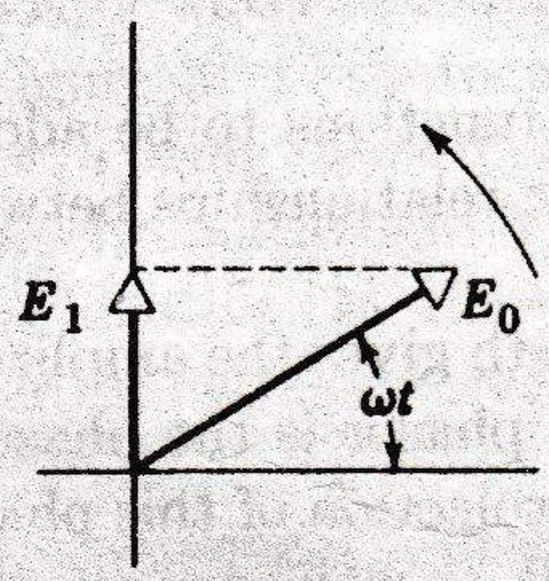
phase difference ϕ with respect to E_1,

$$E_2 = E_0 \sin (\omega t + \phi), \tag{43–4}$$

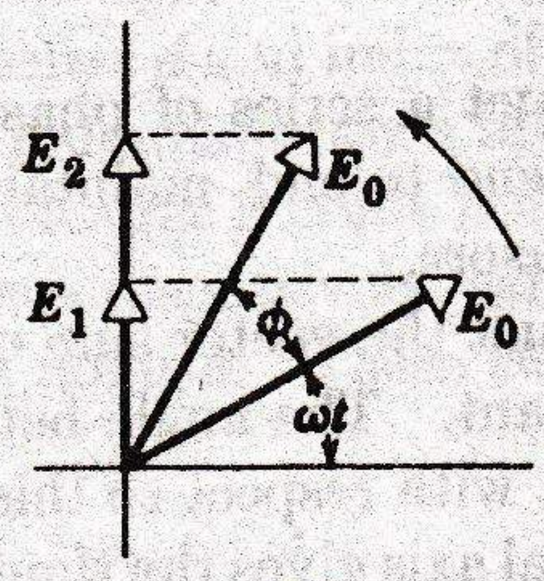
can be represented graphically (Fig. 43–10*b*) as the projection on the vertical axis of a second phasor of magnitude E_0 which makes an angle ϕ with the first phasor. As this figure shows, the sum E of E_1 and E_2 is the sum of the projections of the two phasors on the vertical axis. This is revealed more clearly if the phasors are redrawn, as in Fig. 43–10*c*, placing the foot of one arrow at the head of the other, maintaining the proper phase difference, and letting the whole assembly rotate counterclockwise about the origin.

In Fig. 43–10*c* E can also be regarded as the projection on the vertical axis of a phasor of length E_0, which is the vector sum of the two phasors of magnitude E_0. Note that the (algebraic) sum of the projections of the two phasors is equal to the projection of the (vector) sum of the two phasors.

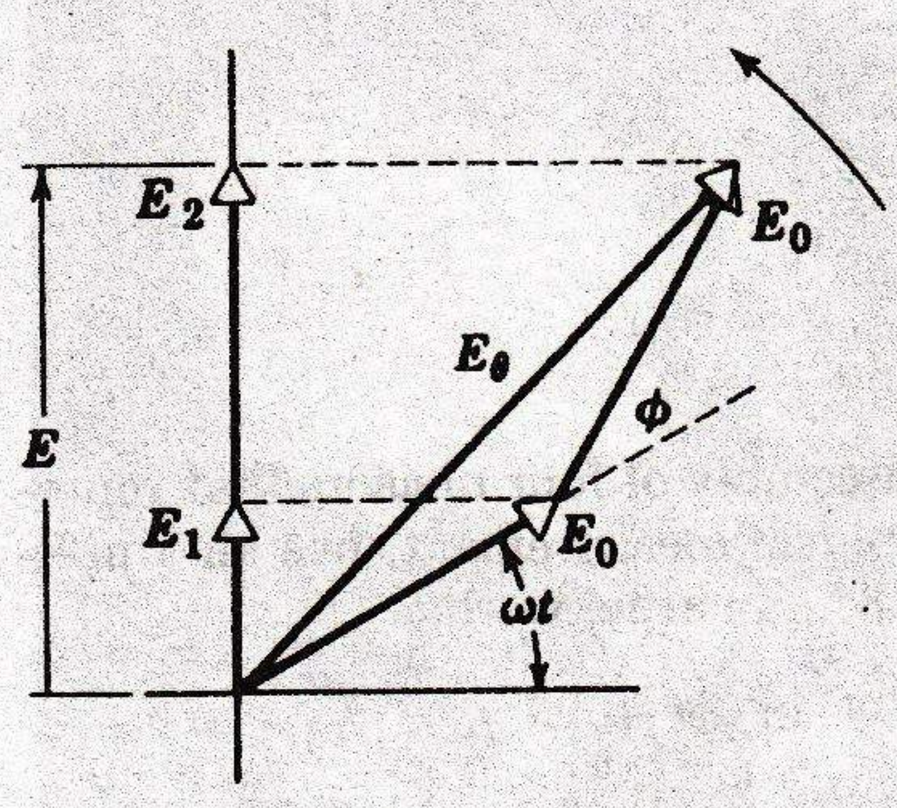
In most problems in optics we are concerned only with the *amplitude* E_0 of the resultant wave disturbance and not with its time variation. This is because the eye and other common measuring instruments respond to the resultant intensity of the light (that is, to the square of the amplitude) and cannot detect the rapid time variations that characterize visible light. For sodium light, for example (λ = 5890 A), the frequency ν ($= \omega/2\pi$) is 5.1×10^{14} cycles/sec. Often, then, we need not consider the rotation of the phasors but can confine our attention to finding the magnitude of the resultant phasor.



(a)



(b)



(c)

Fig. 43–10 (*a*) A wave disturbance E_1 is represented by a rotating vector or *phasor*. (*b*) Two wave disturbances E_1 and E_2, with a phase difference ϕ between them, are so represented. These two phasors can represent the two wave disturbances in the double-slit problem; see Eqs. 43–3 and 43–4. (*c*) Another way of drawing (*b*).

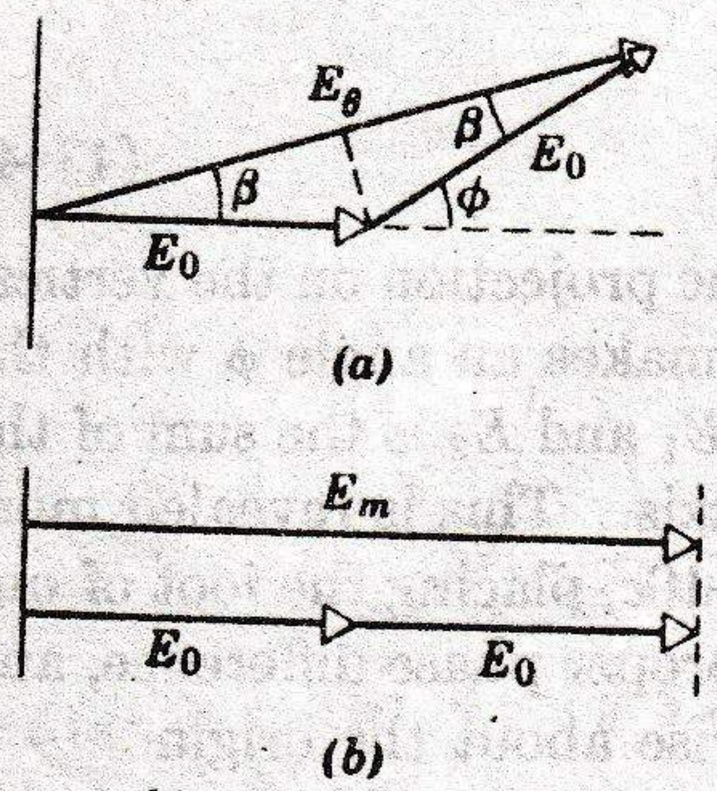


Fig. 43–11 (*a*) A construction to find the amplitude E_θ of two wave disturbances of amplitude E_0 and phase difference ϕ. (*b*) The maximum possible amplitude for these wave disturbances occurs for $\phi = 0$ and has the value $E_m = 2E_0$.

Figure 43–11*a* shows the phasors for double-slit interference at time $t = 0$; compare Fig. 43–10*c*. We see that

$$E_\theta = 2E_0 \cos \beta = E_m \cos \beta,$$

in which, from the theorem that the exterior angle of a triangle (ϕ) is equal to the sum of its two opposite interior angles ($\beta + \beta$),

$$\beta = \tfrac{1}{2}\phi.$$

This is exactly the result arrived at earlier algebraically; compare Eqs. 43–11*a* and *c*.

In a more general case we might want to find the resultant of a number (>2) of sinusoidally varying wave disturbances. The general procedure is the following:

1. Construct a series of phasors representing the functions to be added. Draw them end to end, maintaining the proper phase relationships between adjacent phasors.
2. Construct the vector sum of this array. Its length gives the amplitude of the resultant. The angle between it and the first phasor is the phase of the resultant with respect to this first phasor. The projection of this phasor on the vertical axis gives the time variation of the resultant wave disturbance.

▶ **Example 3.** Find graphically the resultant $E(t)$ of the following wave disturbances:

$$E_1 = 10 \sin \omega t$$
$$E_2 = 10 \sin (\omega t + 15°)$$
$$E_3 = 10 \sin (\omega t + 30°)$$
$$E_4 = 10 \sin (\omega t + 45°).$$

Figure 43–12 in which E_0 equals 10, shows the assembly of four phasors that represents these functions. Their vector sum, by graphical measurement, has an amplitude E_R of 38 and a phase ϕ_0 with respect to E_1 of 23°. In other words

$$E(t) = E_1 + E_2 + E_3 + E_4 = 38 \sin (\omega t + 23°).$$

Check this result by trigonometric calculation. ◀

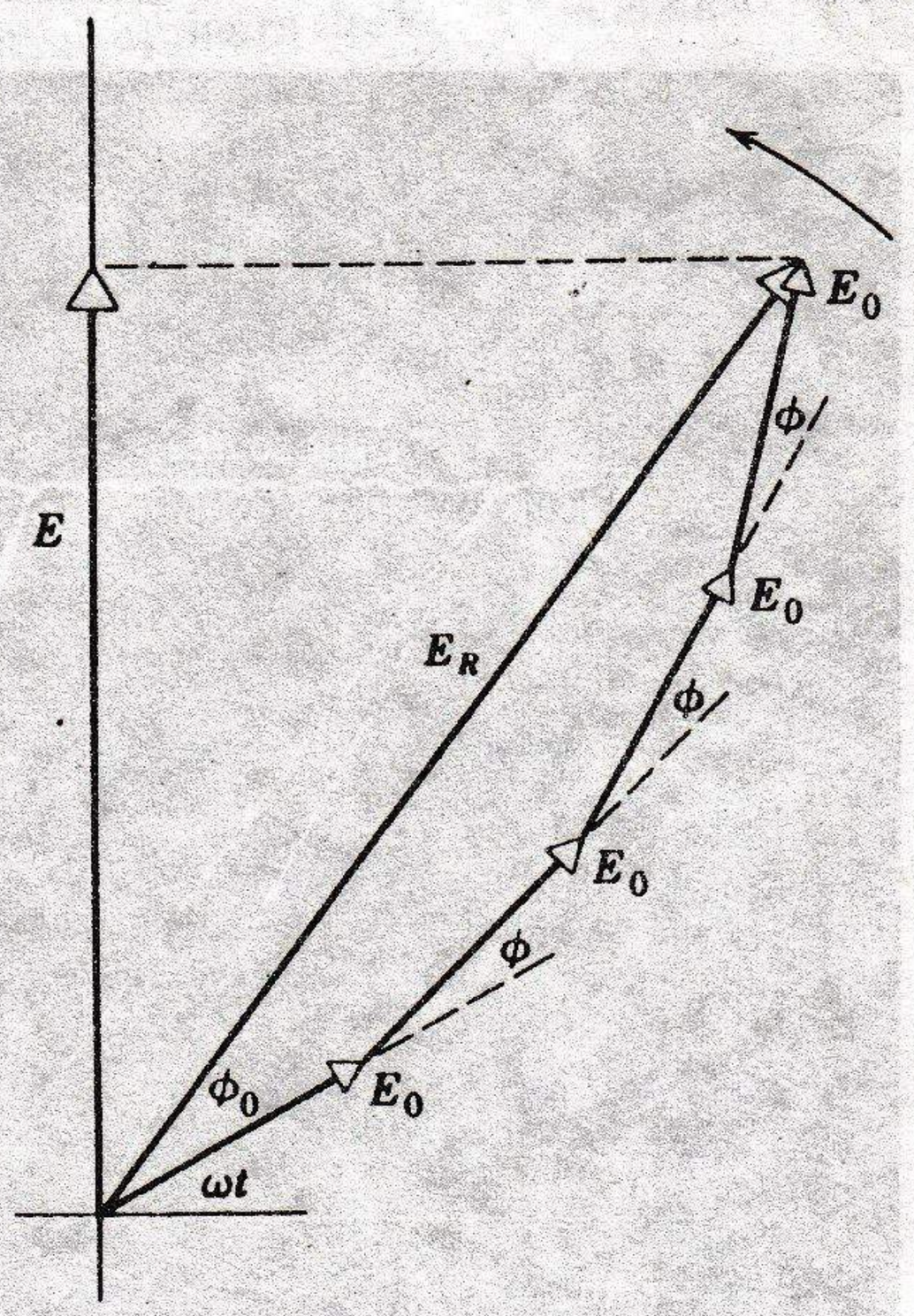


Fig. 43–12 Example 3. Four wave disturbances are added graphically, using the method of phasors.

43–5 Interference from Thin Films

The colors of soap bubbles, oil slicks, and other thin films are the result of interference. Figure 43–13 shows interference effects in a thin vertical film of soapy water illuminated by monochromatic light.

Figure 43–14 shows a film of uniform thickness d and index of refraction n, the eye being focused on spot a. The film is illuminated by a broad source of monochromatic light S. There exists on this source a point P such that two rays, identified by the single and double arrows, respectively, can leave P and enter the eye as shown, after passing through point a. These two rays follow

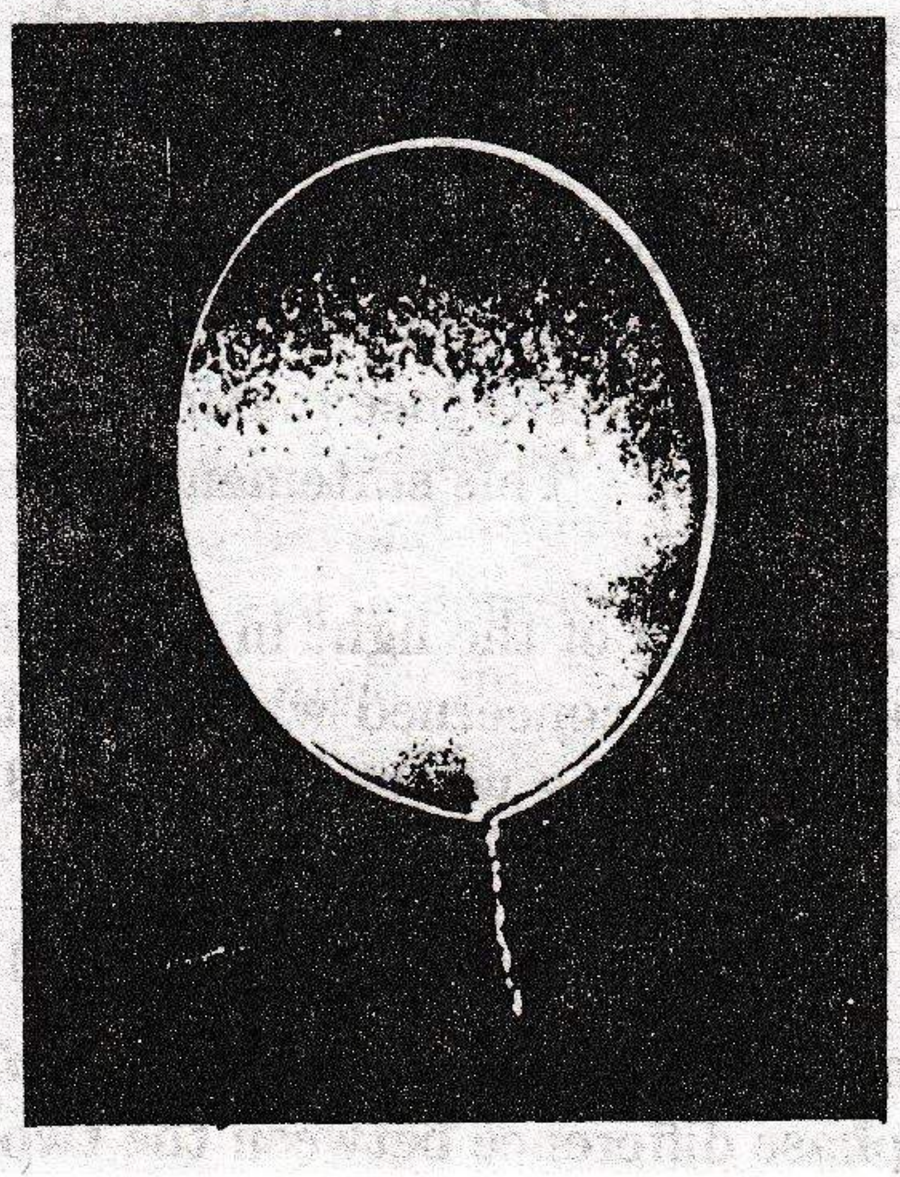


Fig. 43–13 A soapy water film on a wire loop, viewed by reflected light. The black segment at the top is not a tear. It arises because the film, by drainage, is so thin here that there is destructive interference between the light reflected from its front surface and that reflected from its back surface. Recall that these two waves differ in phase by 180°

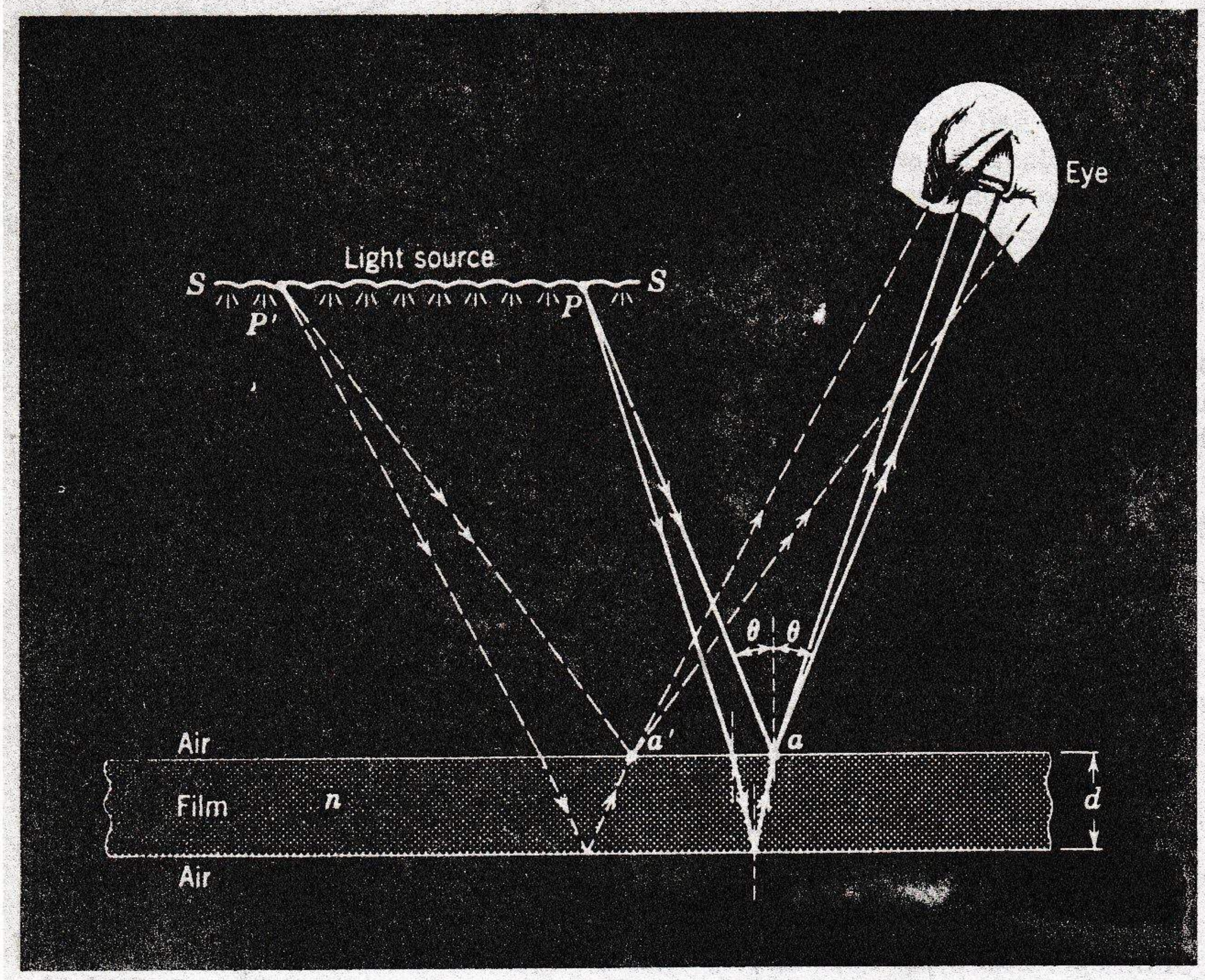


Fig. 43–14 Interference by reflection from a thin film, assuming an extended source S.

different paths in going from P to the eye, one being reflected from the upper surface of the film, the other from the lower surface. Whether point a appears bright or dark depends on the nature of the interference between the two waves that diverge from a. These waves are coherent because they both originate from the same point P on the light source.

If the eye looks at another part of the film, say a', the light that enters the eye must originate from a different point P' of the extended source, as suggested by the dashed lines in Fig. 43–14.

For near-normal incidence ($\theta \cong 0$ in Fig. 43–14) the geometrical path difference for the two rays from P will be close to $2d$. We might expect the resultant wave reflected from the film near a to be an interference maximum if the distance $2d$ is an integral number of wavelengths. This statement must be modified for two reasons.

First, the wavelength must refer to the wavelength of the light in the film λ_n and not to its wavelength in air λ; that is, we are concerned with optical path lengths rather than geometrical path lengths. The wavelengths λ and λ_n (see Eq. 41–11) are related by

$$\lambda_n = \lambda/n. \tag{43–12}$$

To bring out the second point, let us assume that the film is so thin that $2d$ is very much less than a wavelength. The phase difference between the two

waves would be close to zero on our assumption, and we would expect such a film to appear bright on reflection. However, it appears dark. This is clear from Fig. 43–13, in which the action of gravity produces a wedge-shaped film, extremely thin at its top edge. As drainage continues, the dark area increases in size. To explain this and many similar phenomena, we assume that one or the other of the two rays of Fig. 43–14 suffers an abrupt phase change of π ($= 180°$) associated either with reflection at the air-film interface or transmission through it. As it turns out, the ray reflected from the upper surface suffers this phase change. The other ray is not changed abruptly in phase, either on transmission through the upper surface or on reflection at the lower surface.

In Section 19–9 we discussed phase changes on reflection for transverse waves in strings. To extend these ideas, consider the composite string of Fig. 43–15, which consists of two parts with different masses per unit length, stretched to a given tension. If a pulse moves to the right in Fig. 43–15*a*, approaching the junction, there will be a reflected and a transmitted pulse,

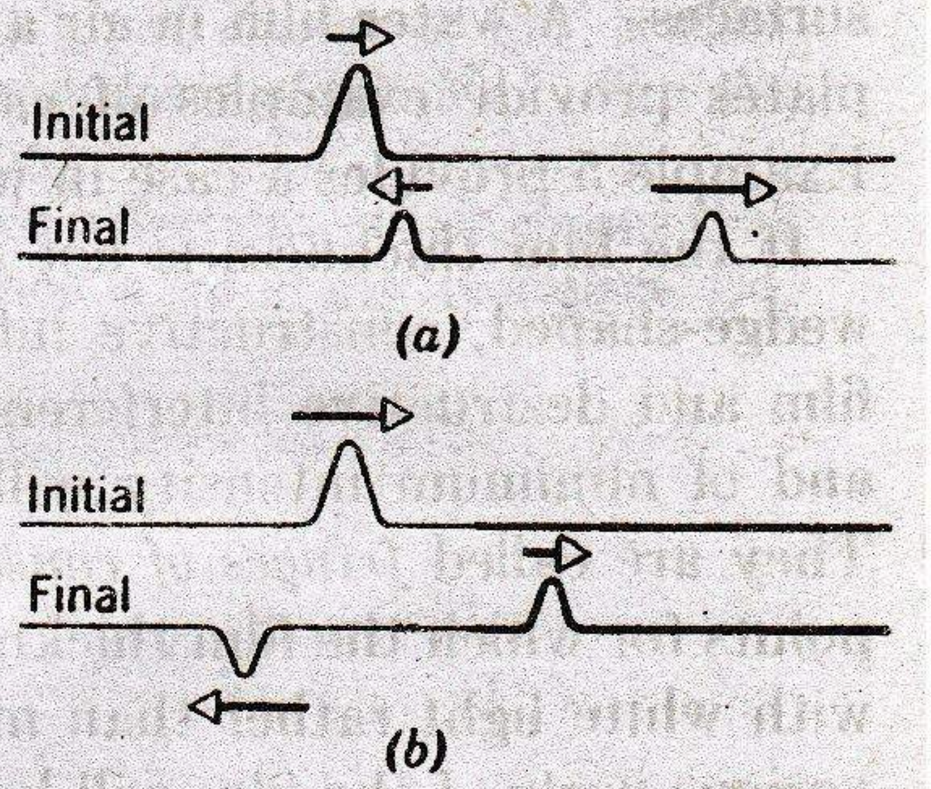


Fig. 43–15 Phase changes on reflection at a junction between two stretched composite strings. (*a*) Incident pulse in heavy string and (*b*) incident pulse in light string.

the reflected pulse being *in phase* with the incident pulse. In Fig. 43–15*b* the situation is reversed, the incident pulse now being in the less massive string. In this case the reflected pulse will differ in phase from the incident pulse by π ($= 180°$). In each case the transmitted pulse will be in phase with the incident pulse.

Figure 43–15*a* suggests a light wave in glass, say, approaching a surface beyond which there is a less optically dense medium (one of lower index of refraction) such as air. Figure 43–15*b* suggests a light wave in air approaching glass. To sum up the optical situation, when reflection occurs from an interface beyond which the medium has a *lower* index of refraction, the reflected wave undergoes *no phase change;* when the medium beyond the interface has a *higher* index, there is a phase change of π.* The transmitted wave does not experience a change of phase in either case.

---

* These statements, which can be proved rigorously from Maxwell's equations (see also Section 43–6), must be modified for light falling on a less dense medium at an angle such that total internal reflection occurs. They must also be modified for reflection from metallic surfaces.

We are now able to take into account both factors that determine the nature of the interference, namely, differences in optical path length and phase changes on reflection. For the two rays of Fig. 43–14 to combine to give a *maximum* intensity, assuming normal incidence, we must have

$$2d = (m + \tfrac{1}{2})\lambda_n \qquad m = 0, 1, 2, \ldots.$$

The term $\frac{1}{2}\lambda_n$ is introduced because of the phase change on reflection, a phase change of 180° being equivalent to half a wavelength. Substituting λ/n for λ_n yields finally

$$2dn = (m + \tfrac{1}{2})\lambda \qquad m = 0, 1, 2, \ldots \text{ (maxima)}. \qquad (43\text{–}13)$$

The condition for a *minimum* intensity is

$$2dn = m\lambda \qquad m = 0, 1, 2, \ldots \text{ (minima)}. \qquad (43\text{–}14)$$

These equations hold when the index of refraction of the film is either greater or less than the indices of the media on *each* side of the film. Only in these cases will there be a relative phase change of 180° for reflections at the two surfaces. A water film in air and an air film in the space between two glass plates provide examples of cases to which Eqs. 43–13 and 43–14 apply. Example 5 provides a case in which they do not apply.

If the film thickness is not uniform, as in Fig. 43–13, where the film is wedge-shaped, constructive interference will occur in certain parts of the film and destructive interference will occur in others. Lines of maximum and of minimum intensity will appear—these are the interference fringes. They are called *fringes of constant thickness*, each fringe being the locus of points for which the film thickness d is a constant. If the film is illuminated with white light rather than monochromatic light, the light reflected from various parts of the film will be modified by the various constructive or destructive interferences that occur. This accounts for the brilliant colors of soap bubbles and oil slicks.

Only if the film is "thin," which implies that d is no more than a few wavelengths of light, will fringes of the type described, that is, fringes that appear localized on the film and associated with a variable film thickness, be possible. For very thick films (say $d \cong 1$ cm), the path difference between the two rays of Fig. 43–14 will be many wavelengths and the phase difference at a given point on the film will change rapidly as we move even a small distance away from a. For "thin" films, however, the phase difference at a also holds for reasonably nearby points; there is a characteristic "patch brightness" for any point on the film, as Fig. 43–13 shows. Interference fringes can be produced for thick films; they are not localized on the film but are at infinity. See Section 43–7.

▶ **Example 4.** A water film ($n = 1.33$) in air is 3200 A thick. If it is illuminated with white light at normal incidence, what color will it appear to be in reflected light?

By solving Eq. 43–13 for λ,

$$\lambda = \frac{2dn}{m + \frac{1}{2}} = \frac{(2)(3200\text{ A})(1.33)}{m + \frac{1}{2}} = \frac{8500\text{ A}}{m + \frac{1}{2}} \qquad \text{(maxima)}.$$

From Eq. 43–14 the minima are given by

$$\lambda = \frac{8500\ \text{A}}{m} \qquad \text{(minima)}.$$

Maxima and minima occur for the following wavelengths:

m	0 (max)	1 (min)	1 (max)	2 (min)	2 (max)
λ, A	17000	8500	5700	4250	3400

Only the maximum corresponding to $m = 1$ lies in the visible region (see Fig. 40–2); light of this wavelength appears yellow-green. If white light is used to illuminate the film, the yellow-green component will be enhanced when viewed by reflection.

Example 5. *Nonreflecting glass.* Lenses are often coated with thin films of transparent substances like MgF_2 ($n = 1.38$) in order to reduce the reflection from the glass surface, using interference. How thick a coating is needed to produce a minimum reflection at the center of the visible spectrum (5500 A)?

We assume that the light strikes the lens at near-normal incidence (θ is exaggerated for clarity in Fig. 43–16), and we seek destructive interference between rays r and r_1. Equation 43–14 does not apply because in this case a phase change of 180° is associated with *each* ray, for at *both* the upper and lower surfaces of the MgF_2 film the reflection is from a medium of greater index of refraction.

There is no net change in phase produced by the two reflections, which means that the optical path difference for destructive interference is $(m + \frac{1}{2})\lambda$ (compare Eq. 43–13), leading to

$$2dn = (m + \tfrac{1}{2})\lambda \qquad m = 0, 1, 2, \ldots \text{ (minima)}.$$

Solving for d and putting $m = 0$ yields

$$d = \frac{(m + \frac{1}{2})\lambda}{2n} = \frac{\lambda}{4n} = \frac{5500\ \text{A}}{(4)(1.38)} = 1000\ \text{A}.$$

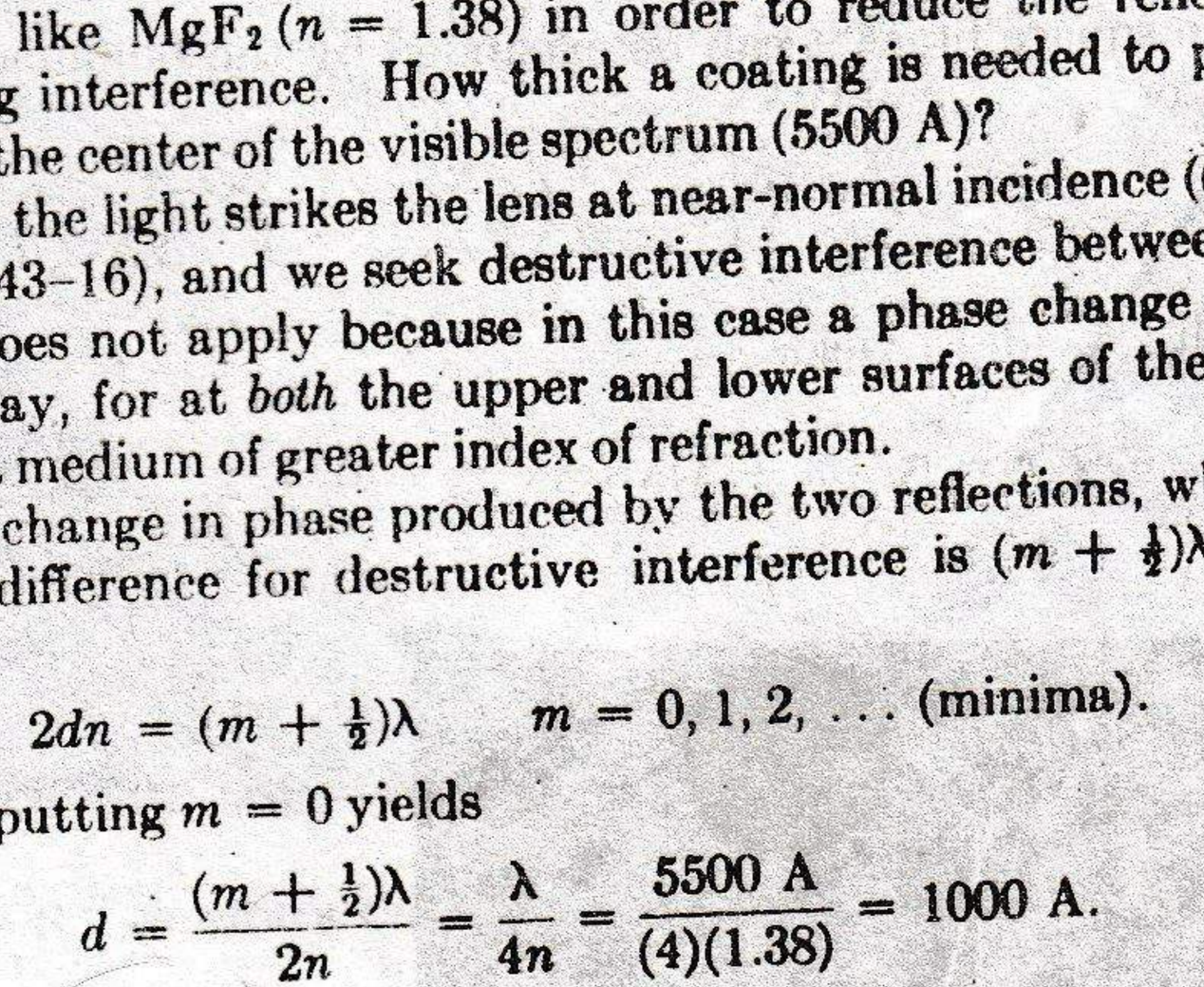


Fig. 43–16 Example 5. Unwanted reflections from glass can be reduced by coating the glass with a thin transparent film.

Example 6. *Newton's rings.* Figure 43–17 shows a lens of radius of curvature R resting on an accurately plane glass plate and illuminated from above by light of wavelength λ. Figure 43–18 shows that circular interference fringes (Newton's rings) appear, associated with the variable thickness air film between the lens and the plate. Find the radii of the circular interference maxima.

Here it is the ray from the *bottom* of the (air) film rather than from the top that undergoes a phase change of 180°, for it is the one reflected from a medium of higher refractive index. The condition for a maximum remains unchanged, however (Eq. 43–13), and is

$$2d = (m + \tfrac{1}{2})\lambda \qquad m = 0, 1, 2, \ldots, \tag{43-15}$$

the index of refraction of the air film being assumed to be unity. From Fig. 43–17 we can write

$$d = R - \sqrt{R^2 - r^2} = R - R\left[1 - \left(\frac{r}{R}\right)^2\right]^{1/2}.$$

If $r/R \ll 1$, the square bracket can be expanded by the binomial theorem, keeping only two terms, or

$$d = R - R\left[1 - \frac{1}{2}\left(\frac{r}{R}\right)^2 + \cdots\right] \cong \frac{r^2}{2R}.$$

Combining with Eq. 43–15 yields

$$r = \sqrt{(m + \tfrac{1}{2})\lambda R} \qquad m = 0, 1, 2, \ldots,$$

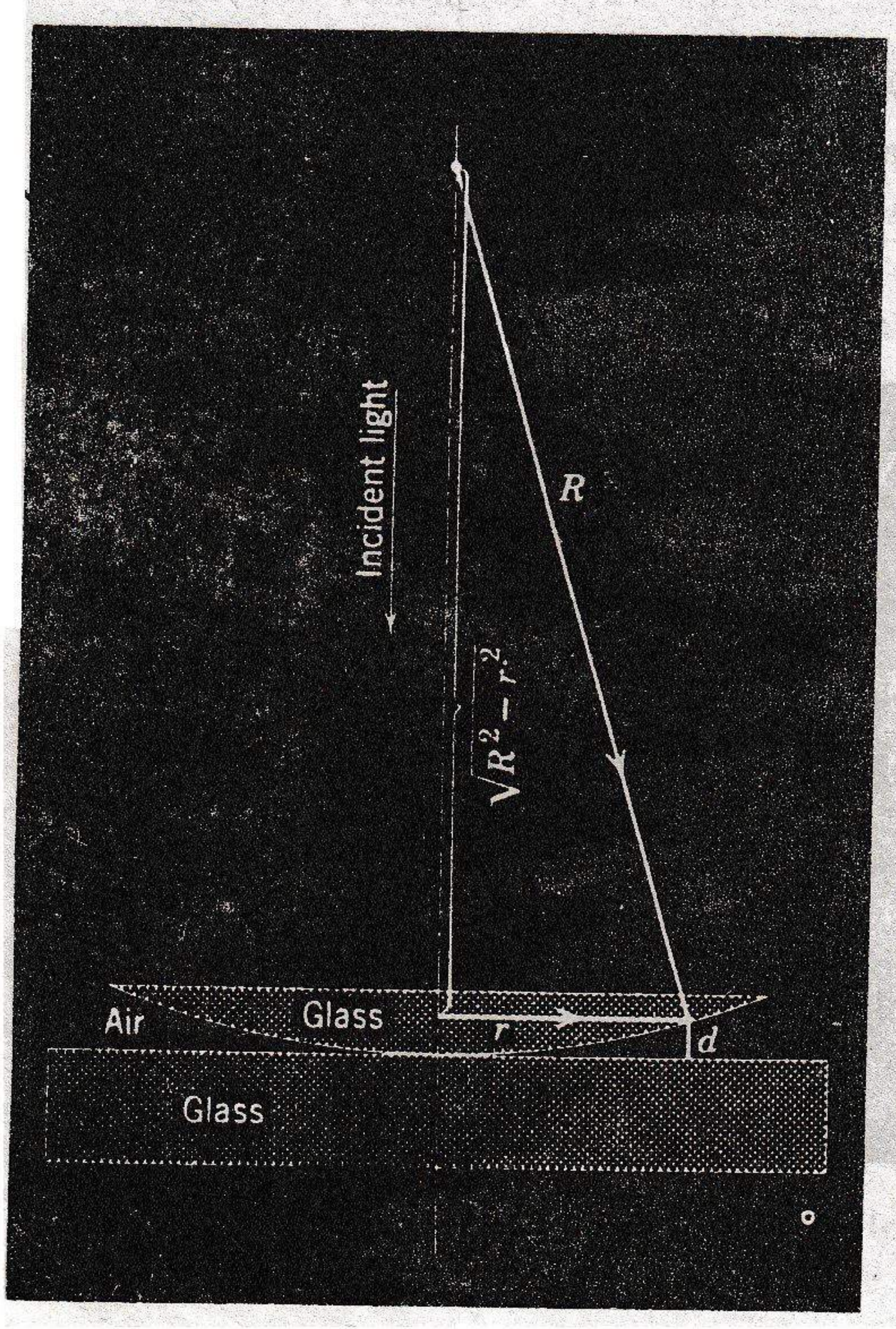


Fig. 43–17 Example 6. Apparatus for observing Newton's rings.

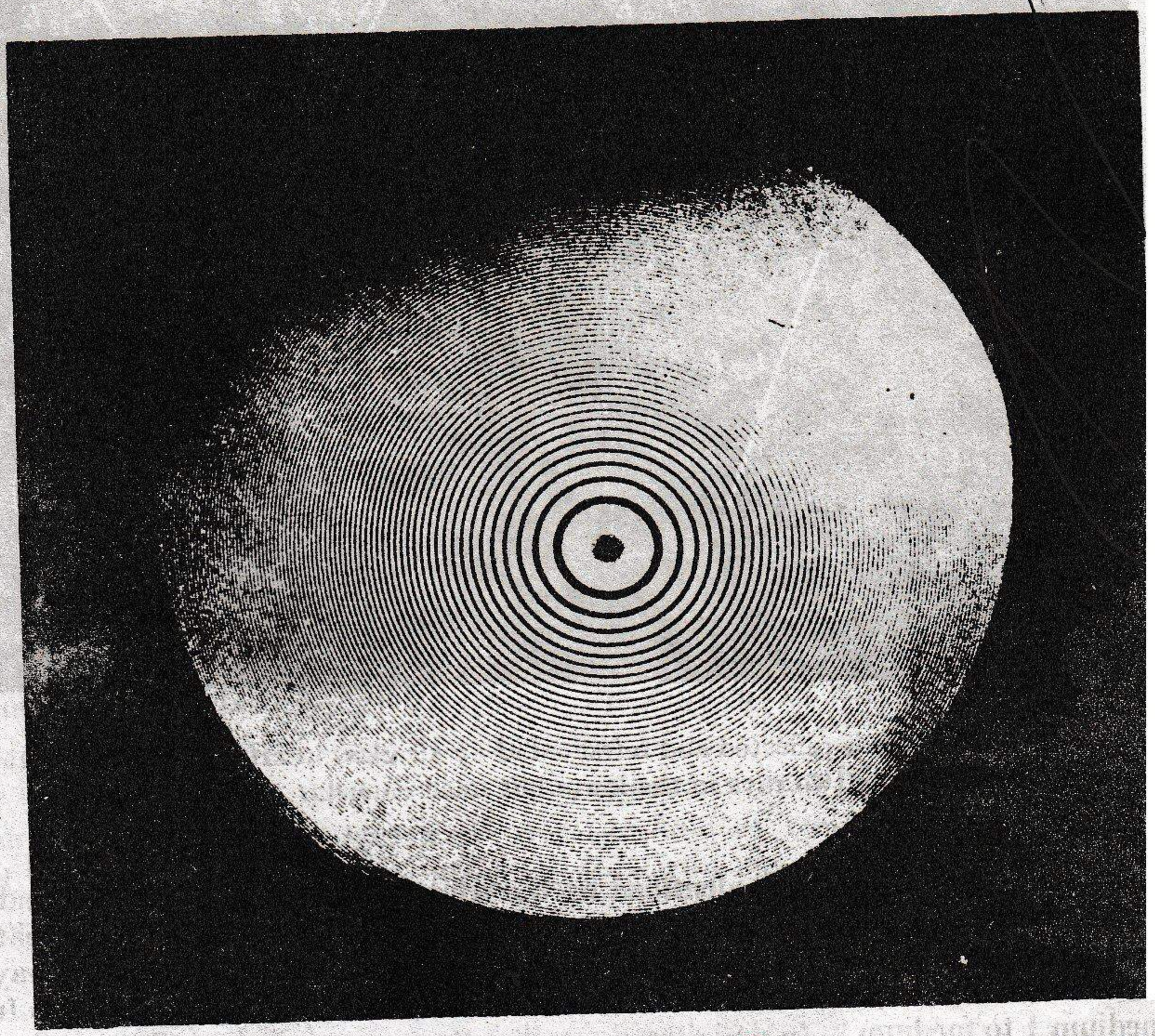


Fig. 43–18 Example 6. Newton's rings. (Courtesy of Bausch and Lomb Optical Co.)

which gives the radii of the bright rings. If white light is used, each spectrum component will produce its own set of circular fringes, the sets all overlapping. ◀

43–6 Phase Changes on Reflection

G. G. Stokes (1819–1903) used the principle of optical reversibility to investigate the reflection of light at an interface between two media. The principle states that if there is no absorption of light a light ray that is reflected or refracted will retrace its original path if its direction is reversed. This reminds us that any mechanical system can run backward as well as forward, provided there is no absorption of energy because of friction, etc.

Figure 43–19*a* shows a wave of amplitude E reflected and refracted at a surface separating media 1 and 2, where $n_2 > n_1$. The amplitude of the reflected wave is $r_{12}E$, in which r_{12} is an *amplitude reflection coefficient*. The amplitude of the refracted wave is $t_{12}E$, where t_{12} is an *amplitude transmission coefficient*.

We consider only the possibility of phase changes of 0 or 180°. If $r_{12} = +0.5$, for example, we have a reduction in amplitude on reflection by one-half and no change in phase. For $r_{12} = -0.5$ we have a phase change of 180° because

$$E \sin (\omega t + 180°) = -E \sin \omega t.$$

Figure 43–19*b* suggests that if we reverse these two rays they should combine to produce the original ray reversed in direction. Ray $r_{12}E$, identified by the single arrows in the figure, is reflected and refracted, producing the rays of amplitudes

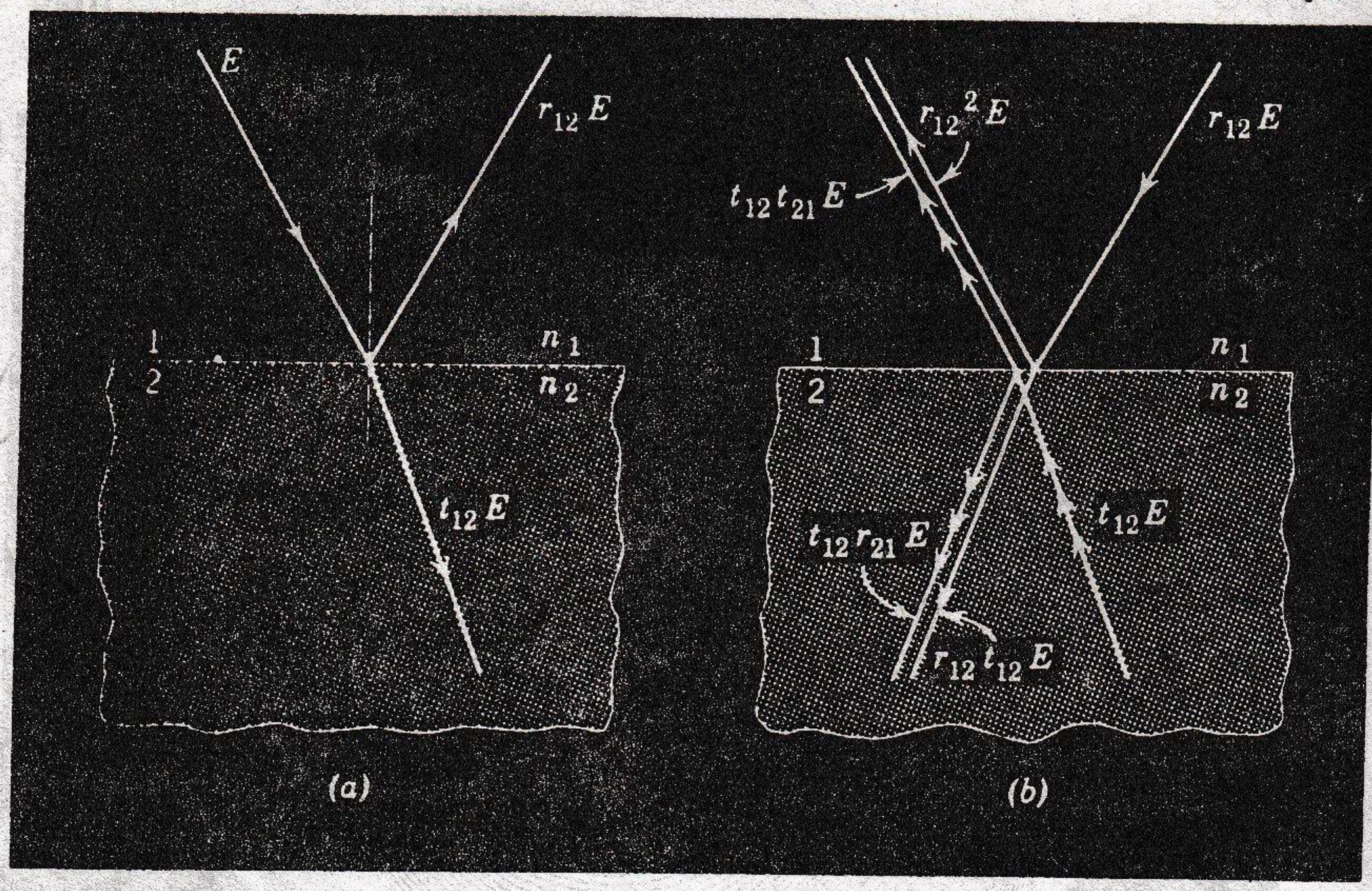


Fig. 43–19 (*a*) A ray is reflected and refracted at an air-glass interface. (*b*) The optically reversed situation; the two rays in the lower left must cancel.

r_{12}^2E and $r_{12}t_{12}E$. Ray $t_{12}E$, identified by the triple arrows, is also reflected and refracted, producing the rays of amplitudes $t_{12}t_{21}E$ and $t_{12}r_{21}E$ as shown. Note that r_{12} describes a ray in medium 1 reflected from medium 2, and r_{21} describes a ray in medium 2 reflected from medium 1. Similarly, t_{12} describes a ray that passes from medium 1 to medium 2; t_{21} describes a ray that passes from medium 2 to medium 1.

The two rays in the upper left of Fig. 43–19*b* must be equivalent to the incident ray of Fig. 43–19*a*, reversed; the two rays in the lower left of Fig. 43–19*b* must cancel. This second requirement leads to

$$r_{12}t_{12}E + t_{12}r_{21}E = 0,$$

or

$$r_{12} = -r_{21}.$$

This result tells us that if we compare a wave reflected from medium one with one reflected from medium 2, they behave differently in that one or the other undergoes a phase change of 180°. We must rely on experiment to show that, as we pointed out earlier, the ray reflected from the more optically dense medium is the one that experiences the phase change of 180°.

43–7 Michelson's Interferometer

An interferometer is a device that can be used to measure lengths or changes in length with great accuracy by means of interference fringes. We describe the form originally built by Michelson in 1881.

Consider light that leaves point P on extended source S (Fig. 43–20) and falls on half-silvered mirror M. This mirror has a silver coating just thick enough to transmit half the incident light and to reflect half; in the figure we have assumed that this mirror, for convenience, possesses negligible thickness. At M the light divides into two waves. One proceeds by transmission toward mirror M_1; the other proceeds by reflection toward M_2. The waves

are reflected at each of these mirrors and are sent back along their directions of incidence, each wave eventually entering the eye E. Since the waves are coherent, being derived from the same point on the source, they will interfere.

If the mirrors M_1 and M_2 are exactly perpendicular to each other, the effect is that of light from an extended source S falling on a uniformly thick slab of air, between glass, whose thickness is equal to $d_2 - d_1$. Interference fringes appear, caused by small changes in the angle of incidence of the light from different points on the extended source as it strikes the equivalent air film. For *thick* films a path difference of one wavelength can be brought about by a very small change in the angle of incidence.

If M_2 is moved backward or forward, the effect is to change the thickness of the equivalent air film. Suppose that the center of the (circular) fringe pattern appears bright and that M_2 is moved just enough to cause the first bright circular fringe to move to the center of the pattern. The path of the light beam striking M_2 has been changed by one wavelength. This means (because the light passes twice through the equivalent air film) that the mirror must have moved one-half a wavelength.

The interferometer is used to measure changes in length by counting the number of interference fringes that pass the field of view as mirror M_2 is moved. Length measurements made in this way can be accurate if large numbers of fringes are counted.

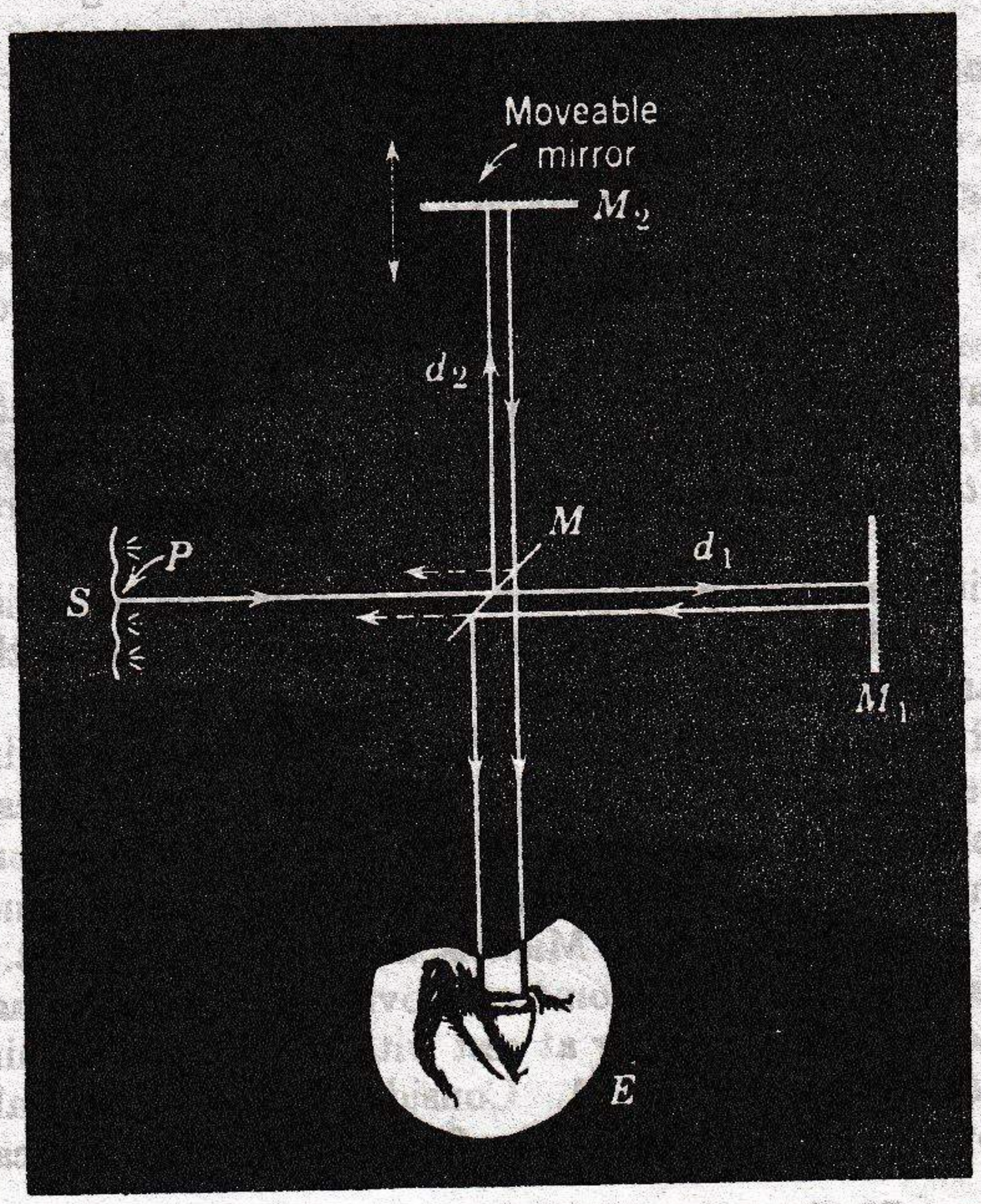


Fig. 43–20 Michelson's interferometer, showing the path of a particular ray originating at point P of an extended source S.

Michelson measured the length of the standard meter, kept in Paris, in terms of the wavelength of certain monochromatic red light emitted from a light source containing cadmium. He showed that the standard meter was equivalent to 1,553,163.5 wavelengths of the red cadmium light.

Physicists have long speculated on the advantages of discarding the standard meter bar as the basic standard of length and of *defining* the meter in terms of the wavelength of some carefully chosen monochromatic radiation. This would make the primary length standard readily available in laboratories all over the world. It would improve the accuracy of length measurements, since one would no longer need to compare an unknown object with a standard object (the meter bar), using interferometer techniques, but could measure the unknown object *directly* and in an absolute sense, using these techniques. There is the additional advantage that if the standard meter bar were destroyed it could never be replaced, whereas light sources and interferometers will (presumably) always be available.

In 1961 such an atomic standard of length was adopted by international agreement. Quoting from an article * describing the event:

> The wavelength of the orange-red light of krypton-86 has replaced the platinum iridium bar as the world standard of length. Formerly the wavelength of this light was defined as a function of the length of the meter bar. Now the meter is defined as a multiple (1,650,763.73) of the wavelength of the light.

The light from krypton-86 was used in preference to that from cadmium or other sources because it produces sharper interference fringes in the interferometer over the long optical paths sometimes used in length measurement.

43–8 Michelson's Interferometer and Light Propagation

In Section 40–4 we presented Einstein's hypothesis, now well verified, that in free space light is propagated with the same speed c no matter what the relative velocity of the source and the observer may be. We pointed out that this hypothesis contradicted the views of nineteenth-century physicists regarding wave propagation. It was difficult for these physicists, trained as they were in the classical physics of the time, to believe that a wave could be propagated without a medium. If such a medium could be established, the speed c of light would naturally be construed as the speed *with respect to that medium*, just as the speed of sound always refers to a medium such as air.

Although no medium for light propagation was obvious, the physicists postulated one, called the *ether* † and hypothesized that its properties were such that it was undetectable by ordinary means such as weighing.

In 1881 (24 years before Einstein's hypothesis) A. A. Michelson set himself the task of forcing the ether, assuming that it existed, to submit to direct physical verification. In particular, Michelson, later joined by E. W. Morley, tried to measure the speed u with which the earth moves through the ether. Michelson's interferometer was their instrument of choice for this now-famous Michelson-Morley experiment.

The earth together with the interferometer moving with velocity u through the ether is equivalent to the interferometer at rest with the ether streaming through it with velocity $-u$, as shown in Fig. 43–21. Consider a wave moving along the path MM_1M and one moving along MM_2M. The first corresponds classically to a man

---

* *Scientific American*, p.75, December 1960.

† More fully, the *luminiferous* (or light-carrying) *ether*.

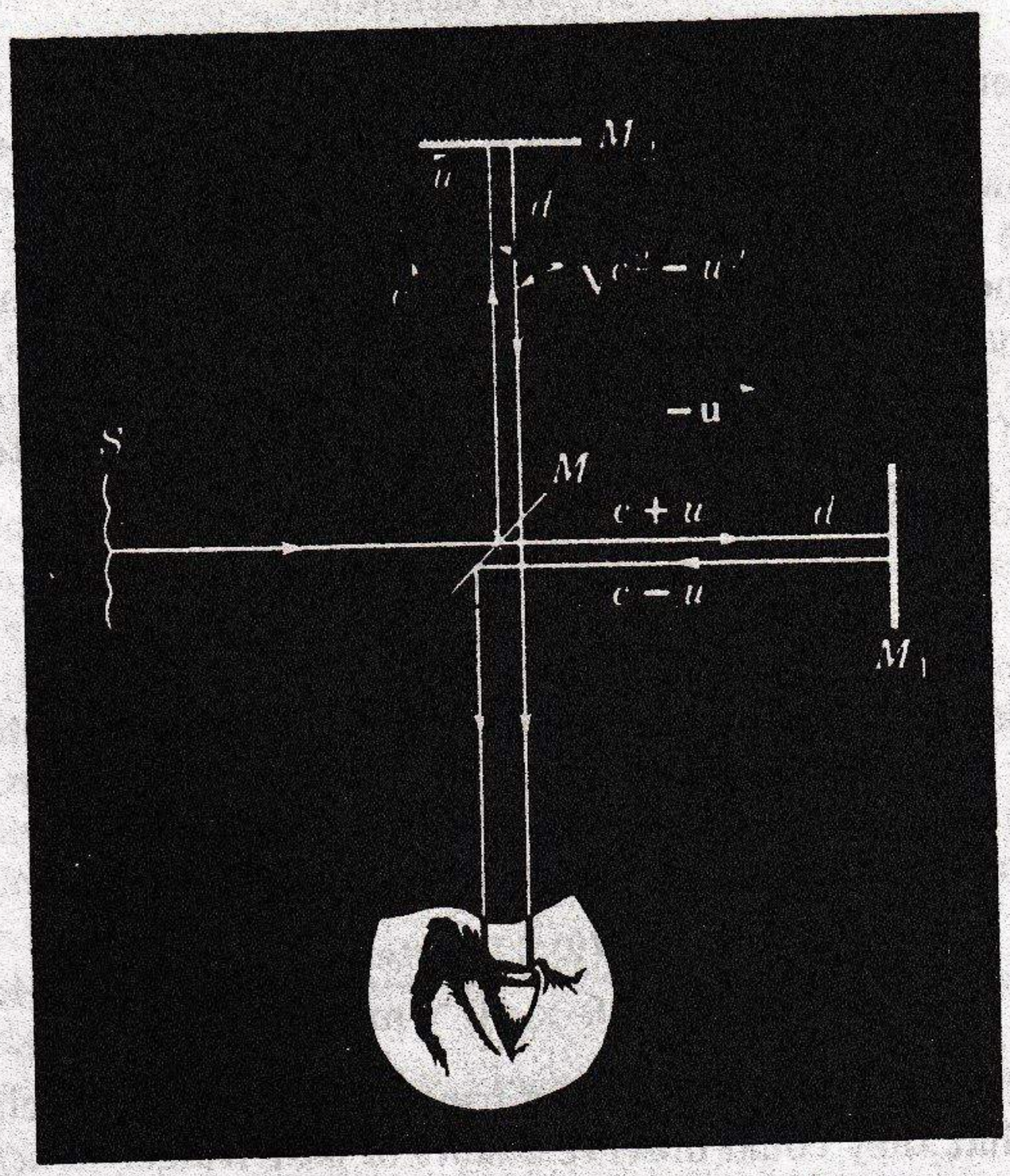


Fig. 43–21 The "ether" is streaming with velocity —u through Michelson's interferometer. The wave speeds shown are on the basis of the (incorrect) ether hypothesis.

rowing a boat a distance d downstream and the same distance upstream; the second corresponds to rowing a boat a distance d across a stream and back.

On the ether hypothesis the speed of light on the path MM_1 is $c + u$; on the return path M_1M it is $c - u$. The time required for the complete trip is

$$t_1 = \frac{d}{c+u} + \frac{d}{c-u} = d\frac{2c}{c^2-u^2} = \frac{2d}{c}\frac{1}{1-(u/c)^2}.$$

The speed of light, *on the ether hypothesis*, for path MM_2 is $\sqrt{c^2-u^2}$, as Fig. 43–21 suggests. This same speed holds for the return path M_2M, so that the time required for this complete path is

$$t_2 = \frac{2d}{\sqrt{c^2-u^2}} = \frac{2d}{c}\frac{1}{\sqrt{1-(u/c)^2}}.$$

The difference of time for the two paths is

$$\Delta t = t_1 - t_2$$
$$= \frac{2d}{c}\left\{\left[1-\left(\frac{u}{c}\right)^2\right]^{-1} - \left[1-\left(\frac{u}{c}\right)^2\right]^{-\frac{1}{2}}\right\}.$$

Assuming $u/c \ll 1$, we can expand the quantities in the square brackets by using the binomial theorem, retaining only the first two terms. This leads to

$$\Delta t = \frac{2d}{c}\left\{\left[1+\left(\frac{u}{c}\right)^2+\cdots\right] - \left[1+\frac{1}{2}\left(\frac{u}{c}\right)^2+\cdots\right]\right\}$$
$$= \frac{2d}{c}\left\{\frac{1}{2}\left(\frac{u}{c}\right)^2\right\} = \frac{du^2}{c^3}. \qquad (43\text{–}16)$$

Now let the entire interferometer be rotated through 90°. This will interchange the roles of the two light paths, MM_1M now being the "cross-stream" path and MM_2M the "down- and upstream" path. The time difference between the two waves entering the eye is also reversed; this changes the phase difference between the combining waves and alters the positions of the interference maxima. *The experiment consists of looking for a shift of the interference fringes as the apparatus is rotated.*

The *change* in time difference is $2\Delta t$, which corresponds to a fringe shift of $2\Delta t/T$ where $T\ (= \lambda/c)$ is the period of vibration of the light. The expected maximum shift in the number of fringes on a 90° rotation (see Eq. 43–16) is

$$\Delta N = \frac{2\Delta t}{T} = \frac{2\Delta t c}{\lambda} = \frac{2d}{\lambda}\left(\frac{u}{c}\right)^2. \tag{43–17}$$

In the Michelson-Morley interferometer let $d = 11$ meters (obtained by multiple reflection in the interferometer) and $\lambda = 5.9 \times 10^{-7}$ meter. If u is assumed to be roughly the orbital speed of the earth, then $u/c \cong 10^{-4}$. The expected maximum fringe shift when the interferometer is rotated through 90° is then

$$\Delta N = \frac{2d}{\lambda}\left(\frac{u}{c}\right)^2 = \frac{(2)(11 \text{ meters})}{5.9 \times 10^{-7} \text{ meter}}(10^{-4})^2 = 0.4.$$

Even though a shift of only about 0.4 of a fringe was expected, Michelson and Morley were confident that they could observe a shift of 0.01 fringe. *They found from their experiment, however, that there was no observable fringe shift!*

The analogy between a light wave in the supposed ether and a boat moving in water, which seemed so evident in 1881, is simply incorrect. The derivation based on this analogy is incorrect for light waves. When the analysis is carried through on Einstein's hypothesis, the observed negative result is clearly predicted, the speed of light being c for all paths. The motion of the earth around the sun and the rotation of the interferometer have, in Einstein's view, no effect whatever on the speed of the light waves in the interferometer.

It should be made clear that although Einstein's hypothesis is completely consistent with the negative result of the Michelson-Morley experiment this experiment standing alone cannot serve as a proof for Einstein's hypothesis. Einstein said that no number of experiments, however large, could prove him right but that a single experiment could prove him wrong. Our present-day belief in Einstein's hypothesis rests on consistent agreement in a large number of experiments designed to test it. The "single experiment" that might prove Einstein wrong has never been found.

QUESTIONS

1. Is Young's experiment an interference experiment or a diffraction experiment, or both?

2. Do interference effects occur for sound waves? Recall that sound is a longitudinal wave and that light is a transverse wave.

3. In Young's double-slit interference experiment, using a monochromatic laboratory light source, why is screen A in Fig. 43–1 necessary? What would happen if one gradually enlarged the hole in this screen?

4. Describe the pattern of light intensity on screen C in Fig. 43–5 if one slit is covered with a red filter and the other with a blue filter, the incident light being white.

5. Is coherence important in reflection and refraction?

6. Define carefully, and distinguish between, the angles θ and ϕ that appear in Eq. 43–10.

7. If one slit in Fig. 43–5 is covered, what change occurs in the intensity of light in the center of the screen?

8. What changes occur in the pattern of interference fringes if the apparatus of Fig. 43–5 is placed under water?

9. What are the requirements for a maximum intensity when viewing a thin film by *transmitted* light?

10. In a Newton's rings experiment, is the central spot, as seen by reflection, dark or light? Explain.

11. Why must the film of Fig. 43–14 be "thin" for us to see an interference pattern of the type described?

12. Why do coated lenses (see Example 5) look purple by reflected light?

13. A person wets his eyeglasses to clean them. As the water evaporates he notices that for a short time the glasses become markedly more nonreflecting. Explain.

14. A lens is coated to reduce reflection, as in Example 5. What happens to the energy that had previously been reflected? Is it absorbed by the coating?

15. Very small changes in the angle of incidence do not change the interference conditions much for "thin" films but they do change them for "thick" films. Why?

16. The directional characteristics of a certain radar antenna as a receiver of radiation are known. What can be said about its directional characteristics as a transmitter?

17. A person in a dark room, looking through a small window, can see a second person standing outside in bright sunlight. The second person cannot see the first person. Is this a failure of the principle of optical reversibility? Assume no absorption of light.

18. Why is it necessary to rotate the interferometer in the Michelson-Morley experiment?

19. How is the negative result of the Michelson-Morley experiment interpreted according to Einstein's theory of relativity?

20. If interference between light waves of different frequencies is possible, one should observe light beats, just as one obtains sound beats from two sources of sound with slightly different frequencies. Discuss how one might experimentally look for this possibility.

21. In Young's double-slit experiment suppose that screen A in Fig. 43–1 contained *two* very narrow parallel slits instead of one. (*a*) Show that if the spacing between these slits is properly chosen the interference fringes can be made to disappear. (*b*) Under these conditions, would you call the beams emerging from slits S_1 and S_2 in screen B coherent? They do not produce interference fringes. (*c*) Discuss what would happen to the interference fringes in the case of a single slit in screen A if the slit width were gradually increased.

PROBLEMS

1. Design a double-slit arrangement that will produce interference fringes 1° apart on a distant screen. Assume sodium light ($\lambda = 5890$ A).

2. Two point sources in Fig. 43–22 emit coherent waves. Show that curves, such as that given, over which the phase difference for rays r_1 and r_2 is a constant are hyperbolas. Extend the analysis to three dimensions. (Hint: A constant phase difference implies a constant difference in length between r_1 and r_2.)

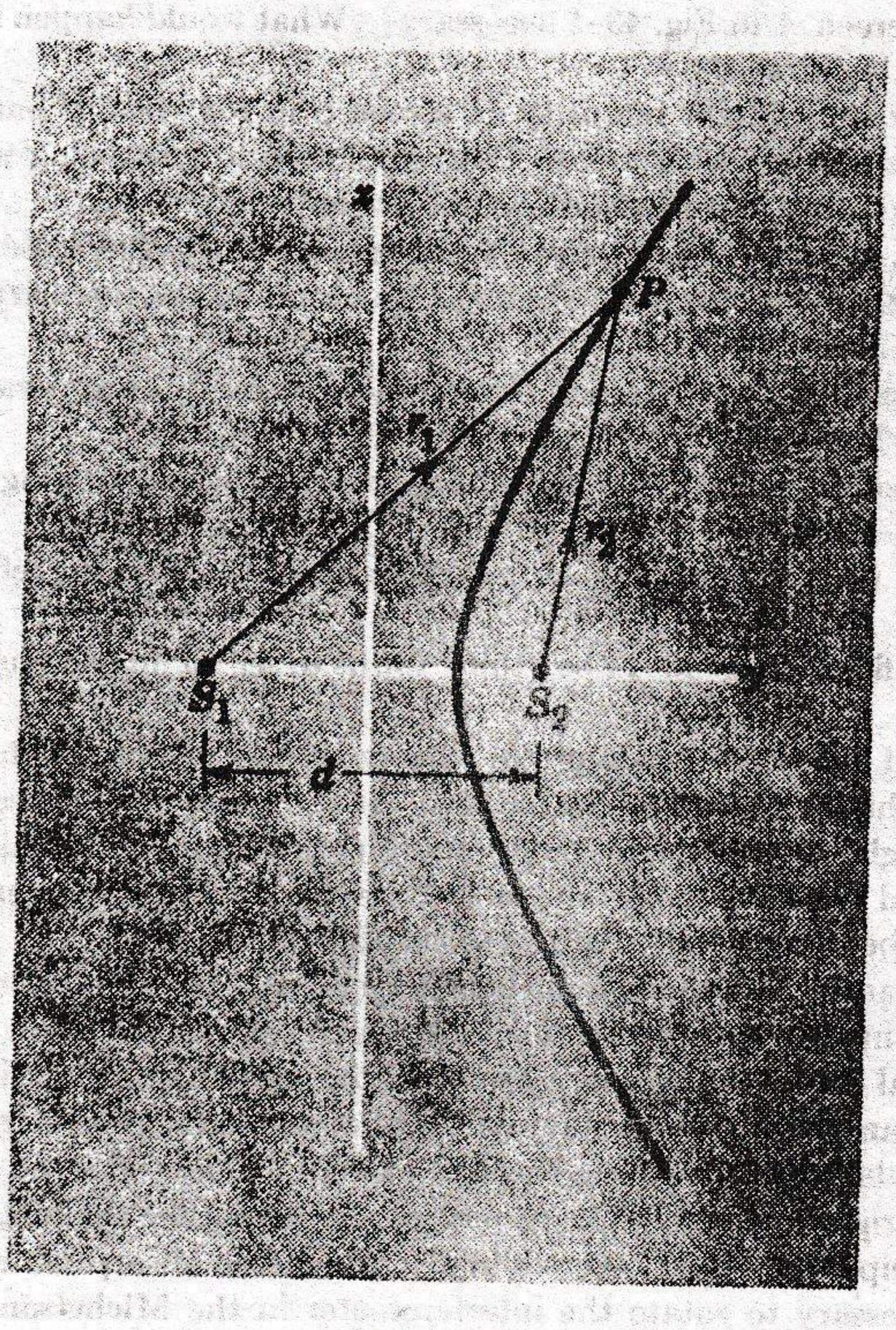


Fig. 43–22

3. A double-slit arrangement produces interference fringes for sodium light ($\lambda = 5890$ A) that are 0.20° apart. What is the angular fringe separation if the entire arrangement is immersed in water?

4. A double-slit arrangement produces interference fringes for sodium light ($\lambda = 5890$ A) that are 0.20° apart. For what wavelength would the angular separation be 10% greater?

5. A thin flake of mica ($n = 1.58$) is used to cover one slit of a double-slit arrangement. The central point on the screen is occupied by what used to be the seventh bright fringe. If $\lambda = 5500$ A, what is the thickness of the mica?

6. Sodium light ($\lambda = 5890$ A) falls on a double slit of separation $d = 2.0$ mm. D in Fig. 43-5 is 4 cm. What per cent error is made in locating the tenth bright fringe if it is *not* assumed that $D \gg d$?

7. Sodium light ($\lambda = 5890$ A) falls on a double slit of separation $d = 0.20$ mm. A thin lens ($f = +1.0$ meter) is placed near the slit as in Fig. 43–6. What is the linear fringe separation on a screen placed in the focal plane of the lens?

8. S_1 and S_2 in Fig. 43–23 are effective point sources of radiation, excited by the same oscillator. They are coherent and in phase with each other. Placed 4.0 meters apart, they emit equal amounts of power in the form of 1.0-meter wavelength electromagnetic waves. (*a*) Find the positions of the first (that is, the nearest), the second, and the third maxima of the received signal, as the detector is moved out along Ox. (*b*) Is the intensity at the nearest minimum equal to zero? Justify your answer.

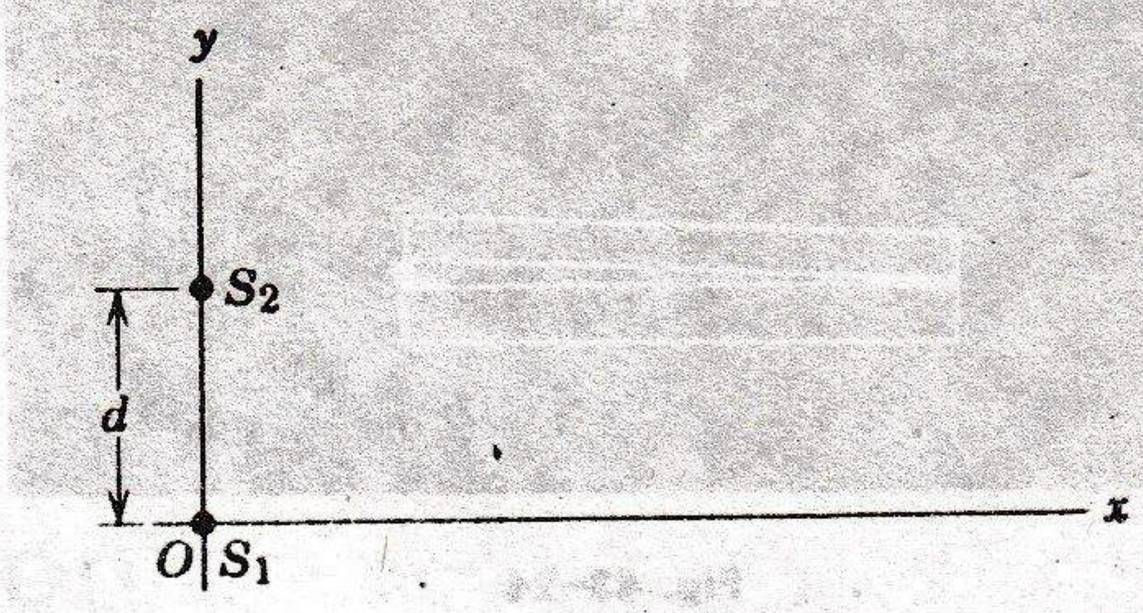


Fig. 43–23

9. One of the slits of a double-slit system is wider than the other, so that the amplitude of the light reaching the central part of the screen from one slit, acting alone, is twice that from the other slit, acting alone. Derive an expression for I_θ in terms of θ, corresponding to Eqs. 43–11*b* and *c*.

10. Show that the half-width $\Delta\theta$ of the double-slit interference fringes (see arrow in Fig. 43–9) is given by

$$\Delta\theta = \frac{\lambda}{2d}$$

if θ is small enough so that $\sin\theta \cong \theta$.

11. Find the sum of the following quantities (*a*) by the vector method and (*b*) analytically:

$$y_1 = 10 \sin \omega t$$

$$y_2 = 8 \sin (\omega t + 30°).$$

12. Add the following quantities graphically, using the vector method:

$$y_1 = 10 \sin \omega t$$

$$y_2 = 15 \sin (\omega t + 30°)$$

$$y_3 = 5 \sin (\omega t - 45°).$$

13. The frequency width of a standard television channel is 4×10^6 cycles/sec. If it were possible to transmit television signals in the visible portion of the electromagnetic spectrum, possibly by using optical maser or laser techniques, how many such standard channels would be available for allocation in the visible spectrum? Assume that this spectrum extends from 4000 to 7000 A.

14. In a Newton's rings experiment the radius of curvature R of the lens is 5.0 meters and its diameter is 2.0 cm. (*a*) How many rings are produced? (*b*) How many rings would be seen if the arrangement were immersed in water ($n = 1.33$)? Assume that $\lambda = 5890$ A.

15. The diameter of the tenth bright ring in a Newton's rings apparatus changes from 1.40 to 1.27 cm as a liquid is introduced between the lens and the plate. Find the index of refraction of the liquid.

16. A broad source of light ($\lambda = 6800$ A) illuminates normally two glass plates 12 cm long that touch at one end and are separated by a wire 0.048 mm in diameter at the other (Fig. 43–24). How many bright fringes appear over the 12-cm distance?

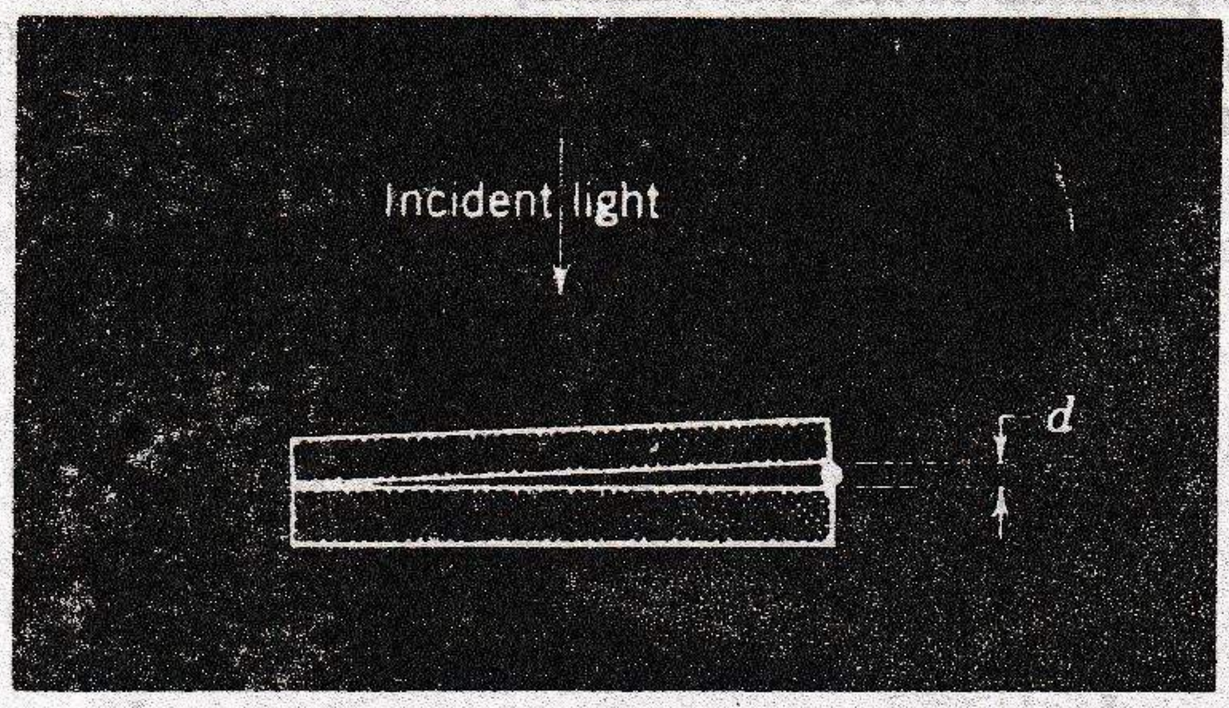


Fig. 43–24

17. A thin film 4×10^{-5} cm thick is illuminated by white light normal to its surface. Its index of refraction is 1.5. What wavelengths within the visible spectrum will be intensified in the reflected beam?

18. White light reflected at perpendicular incidence from a soap film has, in the visible spectrum, an interference maximum at 6000 A and a minimum at 4500 A, with no minimum in between. If $n = 1.33$ for the film, what is the film thickness, assumed uniform?

19. A plane wave of monochromatic light falls normally on a uniformly thin film of oil which covers a glass plate. The wavelength of the source can be varied continuously. Complete destructive interference of the reflected light is observed for wavelengths of 5000 and 7000 A and for no other wavelengths in between. If the index of refraction of the oil is 1.30 and that of the glass is 1.50, find the thickness of the oil film.

20. In Example 5 assume that there is zero reflection for light of wavelength 5500 A at normal incidence. Calculate the factor by which the reflection is diminished by the coating at 4500 and at 6500 A.

21. A plane monochromatic light wave in air falls at normal incidence on a thin film of oil which covers a glass plate. The wavelength of the source may be varied continuously. Complete destructive interference in the reflected beam is observed for wavelengths of 5000 and 7000 A and for no other wavelength in between. The index of refraction of glass is 1.50. Show that the index of refraction of the oil must be less than 1.50.

22. White light reflected at perpendicular incidence from a soap bubble has, in the visible spectrum, a single interference maximum (at $\lambda = 6000$ A) and a single minimum at the violet end of the spectrum. If $n = 1.33$ for the film, calculate its thickness.

23. If mirror M_2 in Michelson's interferometer is moved through 0.233 mm, 792 fringes are counted. What is the wavelength of the light?

24. A thin film with $n = 1.40$ for light of wavelength 5890 A is placed in one arm of a Michelson interferometer. If a shift of 7.0 fringes occurs, what is the film thickness?

25. (*a*) What is the wavelength of the orange-red line of krypton-86 in angstrom units? Take 1 A $= 10^{-10}$ meter exactly. (*b*) Does this question really make sense, in view of the fact that the meter is defined in terms of this wavelength? Explain.

26. A Michelson interferometer is used with a sodium discharge tube as a light source. The yellow sodium light consists of two wavelengths, 5890 and 5896 A. It is observed that the interference pattern disappears and reappears periodically as one moves mirror M_2 in Fig. 43–20. (*a*) Explain this effect. (*b*) Calculate the change in path difference between two successive reappearances of the interference pattern.