

INTRODUCTION

The treatment of light as wave motion allows for a region of approximation in which the wavelength is considered to be negligible compared with the dimensions of the relevant components of the optical system. This region of approximation is called *geometrical optics*. When the wave character of the light may not be so ignored, the field is known as *physical optics*. Thus, geometrical optics forms a special case of physical optics in a way that may be summarized as follows:

limit {physical optics} = {geometrical optics}

Since the wavelength of light—around 500 nm—is very small compared to ordinary objects, early unrefined observations of the behavior of a light beam passing through apertures or around obstacles in its path could be handled by geometrical optics. Recall that the appearance of *distinct shadows* influenced Newton to assert that the apparent rectilinear propagation of light was due to a stream of light corpuscles rather than a wave motion. Wave motion characterized by longer wavelengths, such as those in water waves and sound waves, was known to give distinct bending around obstacles. Newton's model of light propagation, therefore, seemed not to allow for the existence of a wave motion with very small wavelengths. There was in fact already evidence of some degree of bending, even for light waves, in the time of Isaac Newton. The Jesuit Francesco Grimaldi had noticed the fine structure in the edge of a shadow, a structure not explainable in terms of the rectilinear propagation of light. This bending of light waves around the edges of an obstruction came to be called *diffraction*. Within the approximation represented by geometrical optics, light is understood to travel out from its source along straight lines, or *rays*. The ray is then simply the path along which light energy is transmitted from one point to another in an optical system. The ray is a useful construct, although abstract in the sense that a light beam, in practice, cannot be narrowed down indefinitely to approach a straight line. A pencil-like laser beam is perhaps the best actual approximation to a ray of light. (When an aperture through which the beam is passed is made small enough, however, even a laser beam begins to spread out in a characteristic diffraction pattern.) When a light ray traverses an optical system consisting of several homogeneous media in sequence, the optical path is a sequence of straight-line segments. Discontinuities in the line segments occur each time the light is reflected or refracted. The laws of geometrical optics that describe the subsequent direction of the rays are the Law of Reflection and the Law of Refraction.

Law of Reflection

When a ray of light is reflected at an *interface* dividing two optical media, the reflected ray remains within the *plane of incidence*, and the angle of reflection θ_r equals the angle of incidence θ_i . The *plane of incidence* is the plane containing the incident ray and the surface normal at the point of incidence.

Law of Refraction (Snell's Law)

When a ray of light is refracted at an interface dividing two transparent media, the transmitted ray remains within the plane of incidence and the sine of the angle of refraction θ_t is directly proportional to the sine of the angle of incidence θ_i . These two laws are summarized in Figure 1, which depicts the general case in which an incident ray is partially reflected and partially transmitted at a plane interface separating two transparent media.

1 HUYGENS' PRINCIPLE

The Dutch physicist Christian Huygens envisioned light as a series of pulses emitted from each point of a luminous body and propagated in relay fashion



Figure 1 Reflection and refraction at an *interface* between two optical media. Incident, reflected, and refracted rays are shown in the *plane of incidence*.

by the particles of the ether, an elastic medium filling all space. Consistent with his conception, Huygens imagined each point of a propagating disturbance as capable of originating new pulses that contributed to the disturbance an instant later. To show how his model of light propagation implied the laws of geometrical optics, he enunciated a fruitful principle that can be stated as follows: Each point on the leading surface of a wave disturbance the wavefront-may be regarded as a secondary source of spherical waves (or wavelets), which themselves progress with the speed of light in the medium and whose envelope at a later time constitutes the new wavefront. Simple applications of the principle are shown in Figure 2 for a plane and spherical wave. In each case, AB forms the initial wave disturbance or wavefront, and A'B' is the new wavefront at a time t later. The radius of each wavelet is, accordingly, vt, where v is the speed of light in the medium. Notice that the new wavefront is tangent to each wavelet at a single point. According to Huygens, the remainder of each wavelet is to be disregarded in the application of the principle. Indeed, were the remainder of the wavelet considered to be effective in propagating the light disturbance, Huygens could not have derived the law of rectilinear propagation from his principle. To see this more clearly, refer to Figure 3, which shows a spherical wave disturbance originating at O and incident upon an aperture with an opening SS'. According to the notion



Figure 2 Illustration of Huygens' principle for (a) plane and (b) spherical waves.

Figure 3 Huygens' construction for an obstructed wavefront.

of rectilinear propagation, the lines OA and OB form the sharp edges of the shadow to the right of the aperture. Some of the wavelets that originate from points of the wavefront (arc SS'), however, overlap into the region of shadow. According to Huygens, however, these are ignored and the new wavefront ends abruptly at points P and P', precisely where the extreme wavelets originating at points S and S' are tangent to the new wavefront. In so disregarding the effectiveness of the overlapping wavelets, Huygens avoided the possibility of diffraction of the light into the region of geometric shadow. Huygens also ignored the wavefront formed by the back half of the wavelets, since these wavefronts implied a light disturbance traveling in the opposite direction. Despite weaknesses in this model, remedied later by Fresnel and others, Huygens was able to apply his principle to prove the laws of both reflection and refraction, as we show in what follows.

Figure 4a illustrates the Huygens construction for a narrow, parallel beam of light to prove the law of reflection. Huygens' principle must be modified slightly to accommodate the case in which a wavefront, such as AC, encounters a plane interface, such as XY, at an angle. Here the angle of incidence of the rays AD, BE, and CF relative to the perpendicular PD is θ_i . Since points along the plane wavefront do not arrive at the interface simultaneously, allowance is made for these differences in constructing the wavelets that determine the reflected wavefront. If the interface XY were not present, the Huygens construction would produce the wavefront GI at the instant ray CF reached the interface at I. The intrusion of the reflecting surface, however, means that during the same time interval required for ray CF to progress from F to I, ray BE has progressed from E to J and then a distance equivalent to JH after reflection. Thus, a wavelet of radius JN = JH centered at J is drawn above the reflecting surface. Similarly, a wavelet of radius DG is drawn centered at D to represent the propagation after reflection of the lower part of the beam. The new wavefront, which must now be tangent to these wavelets at points M and N, and include the point I, is shown as KI in the figure. A representative reflected ray is DL, shown perpendicular to the reflected wavefront. The normal PD drawn for this ray is used to define angles of incidence and reflection for the beam. The construction makes clear the equivalence between the angles of incidence and reflection, as outlined in Figure 4a.

Similarly, in Figure 4b, a Huygens construction is shown that illustrates the law of refraction. Here we must take into account a different speed of light in the upper and lower media. If the speed of light in vacuum is c, we express the speed in the upper medium by the ratio c/n_i , where n_i is a constant that characterizes the medium and is referred to as the *refractive index*. Similarly, the speed of light in the lower medium is c/n_t . The points D, E, and F on the incident wavefront arrive at points D, J, and I of the plane interface XY at different times. In the absence of the refracting surface, the wavefront GI is formed at the instant ray CF reaches I. During the progress of ray CFfrom F to I in time t, however, the ray AD has entered the lower medium, where its speed is, let us say, slower. Thus, if the distance DG is $v_i t$, a wavelet of radius $v_i t$ is constructed with center at D. The radius DM can also be expressed as

$$DM = v_t t = v_t \left(\frac{DG}{v_i}\right) = \left(\frac{n_i}{n_t}\right) DG$$

Similarly, a wavelet of radius $(n_i/n_t) JH$ is drawn centered at J. The new wavefront KI includes point I on the interface and is tangent to the two wavelets at points M and N, as shown. The geometric relationship between the angles θ_i and θ_t , formed by the representative incident ray AD and refracted ray DL, is Snell's law, as outlined in Figure 4b. Snell's law of refraction may be expressed as

$$n_i \sin \theta_i = n_t \sin \theta_t \tag{1}$$



Figure 4 (a) Huygens' construction to prove the law of reflection. (b) Huygens' construction to prove the law of refraction.

2 FERMAT'S PRINCIPLE

The laws of geometrical optics can also be derived, perhaps more elegantly, from a different fundamental hypothesis. The root idea had been introduced by Hero of Alexandria, who lived in the second century B.C. According to Hero, when light is propagated between two points, it takes the shortest path. For propagation between two points in the same uniform medium, the path is clearly the straight line joining the two points. When light from the first point A, Figure 5, reaches the second point B after reflection from a plane surface, however, the same principle predicts the law of reflection, as follows. Figure 5



Figure 5 Construction to prove the law of reflection from Hero's principle.

shows three possible paths from A to B, including the correct one, ADB. Consider, however, the arbitrary path ACB. If point A' is constructed on the perpendicular AO such that AO = OA', the right triangles AOC and A'OC are equal. Thus, AC = A'C and the distance traveled by the ray of light from A to B via C is the same as the distance from A' to B via C. The shortest distance from A' to B is obviously the straight line A'DB, so the path ADB is the correct choice taken by the actual light ray. Elementary geometry shows that for this path, $\theta_i = \theta_r$. Note also that to maintain A'DB as a single straight line, the reflected ray must remain within the plane of incidence, that is, the plane of the page.

The French mathematician Pierre de Fermat generalized Hero's principle to prove the law of refraction. If the terminal point B lies below the surface of a second medium, as in Figure 6, the correct path is definitely not the shortest path or straight line AB, for that would make the angle of refraction equal to the angle of incidence, in violation of the empirically established law of refraction. Appealing to the "economy of nature," Fermat supposed instead that the ray of light traveled the path of least *time* from A to B, a generalization that included Hero's principle as a special case. If light travels more slowly in the second medium, as assumed in Figure 6, light bends at the interface so as to take a path that favors a shorter time in the second medium, thereby minimizing the overall transit time from A to B. Mathematically, we are required to minimize the total time,

$$t = \frac{AO}{v_i} + \frac{OB}{v_t}$$

where v_i and v_t are the velocities of light in the incident and transmitting media, respectively. Employing the Pythagorean theorem and the distances defined in Figure 6, we have $AO = \sqrt{a^2 + x^2}$ and $OB = \sqrt{b^2 + (c - x)^2}$, so that

$$t = \frac{\sqrt{a^2 + x^2}}{v_i} + \frac{\sqrt{b^2 + (c - x)^2}}{v_t}$$

Since other choices of path change the position of point *O* and therefore the distance *x*, we can minimize the time by setting dt/dx = 0:

$$\frac{dt}{dx} = \frac{x}{v_i \sqrt{a^2 + x^2}} - \frac{c - x}{v_i \sqrt{b^2 + (c - x)^2}} = 0$$

Again from Figure 6, in the two right triangles containing AO and OB, respectively, the angles of incidence and refraction can be conveniently



Figure 6 Construction to prove the law of refraction from Fermat's principle.

introduced into the preceding condition, since $\sin \theta_i = \frac{x}{\sqrt{a^2 + x^2}}$ and $\sin \theta_t = \frac{c - x}{\sqrt{b^2 + (c - x)^2}}$, giving

$$\frac{dt}{dx} = \frac{\sin \theta_i}{v_i} - \frac{\sin \theta_t}{v_t} = 0$$

Simplifying the equation set equal to zero, we obtain at once $v_t \sin \theta_i = v_i \sin \theta_t$. Introducing the refractive indices of the media through the relation v = c/n, we arrive at Snell's law:

$$n_i \sin \theta_i = n_t \sin \theta_t$$

Fermat's principle, like that of Huygens, required refinement to achieve more general applicability. Situations exist where the actual path taken by a light ray may represent a maximum time or even one of many possible paths, all requiring equal time. As an example of the latter case, consider light propagating from one focus to the other inside an ellipsoidal mirror, along any of an infinite number of possible paths. Since the ellipse is the locus of all points whose combined distances from the two foci is a constant, all paths are indeed of equal time. A more precise statement of Fermat's principle, which requires merely an extremum relative to neighboring paths, may be given as follows: The actual path taken by a light ray in its propagation between two given points in an optical system is such as to make its optical path equal, in the first approximation, to other paths closely adjacent to the actual one.

With this formulation, Fermat's principle falls in the class of problems called *variational calculus*, a technique that determines the form of a function that minimizes a definite integral. In optics, the definite integral is the integral of the time required for the transit of a light ray from starting to finishing points.¹

3 PRINCIPLE OF REVERSIBILITY

Refer again to the cases of reflection and refraction pictured in Figures 5 and 6. If the roles of points A and B are interchanged, so that B is the source of light rays, Fermat's principle of least time must predict the same path as determined for the original direction of light propagation. In general, then, any actual ray of light in an optical system, if reversed in direction, will retrace the same path backward. This principle of *reversibility* will be found very useful in various applications to be dealt with later.

4 REFLECTION IN PLANE MIRRORS

Before discussing the formation of images in a general way, we discuss the simplest—and experientially, the most accessible—case of images formed by plane mirrors. In this context it is important to distinguish between *specular reflection* from a perfectly smooth surface and *diffuse reflection* from a granular or rough surface. In the former case, all rays of a parallel beam incident on the surface obey the law of reflection from a plane surface and therefore reflect as a parallel beam; in the latter case, though the law of reflection is obeyed locally for each ray, the microscopically granular surface results in

¹It is of interest to note here that a similar principle, called *Hamilton's principle of least action* in mechanics, that calls for a minimum of the definite integral of the Lagrangian function (the kinetic energy minus the potential energy), represents an alternative formulation of the laws of mechanics and indeed implies Newton's laws of mechanics themselves.

rays reflected in various directions and thus a diffuse scattering of the originally parallel rays of light. Every plane surface will produce some such scattering, since a perfectly smooth surface can only be approximated in practice. The treatment that follows assumes the case of specular reflection.

Consider the specular reflection of a single light ray *OP* from the *xy*plane in Figure 7a. By the law of reflection, the reflected ray *PQ* remains within the plane of incidence, making equal angles with the normal at *P*. If the path *OPQ* is resolved into its *x*-, *y*-, and *z*-components, it is clear that the direction of ray *OP* is altered by the reflection only along the *z*-direction, and then in such a way that its *z*-component is simply reversed. If the direction of the incident ray is described by its unit vector, $\hat{\mathbf{r}}_1 = (x, y, z)$, then the reflection causes

$$\hat{\mathbf{r}}_1 = (x, y, z) \longrightarrow \hat{\mathbf{r}}_2 = (x, y, -z)$$

It follows that if a ray is incident from such a direction as to reflect sequentially from all three rectangular coordinate planes, as in the "corner reflector" of Figure 7b,

$$\hat{\mathbf{r}}_1 = (x, y, z) \longrightarrow \hat{\mathbf{r}}_2 = (-x, -y, -z)$$

and the ray returns precisely parallel to the line of its original approach. A network of such corner reflectors ensures the exact return of a beam of light—a headlight beam from highway reflectors, for example, or a laser beam from a mirror on the moon.

Image formation in a plane mirror is illustrated in Figure 8a. A point object S sends rays toward a plane mirror, which reflect as shown. The law of reflection ensures that pairs of triangles like SNP and S'NP are equal, so all reflected rays appear to originate at the *image point S'*, which lies along the normal line SN, and at such a depth that the *image distance* S'N equals the object distance SN. The eye sees a point image at S' in exactly the same way it would see a real point object placed there. Since none of the actual rays of light lies below the mirror surface, the image is said to be a virtual image. The image S' cannot be projected on a screen as in the case of a real image. All points of an extended object, such as the arrow in Figure 8b, are imaged by a plane mirror in similar fashion: Each object point has its image point along its normal to the mirror surface and as far below the reflecting surface as the object point lies above the surface. Note that the image position does not depend on the position of the eye. Further, the construction of Figure 8b makes clear that the image size is identical with the object size, giving a mag*nification* of unity. In addition, the transverse orientation of object and image are the same. A right-handed object, however, appears left-handed in its image. In Figure 8c, where the mirror does not lie directly below the object, the mirror plane may be extended to determine the position of the image as seen by an eye positioned to receive reflected rays originating at the object. Figure 8d illustrates multiple images of a point object O formed by two perpendicular mirrors. Images I_1 and I_2 result from single reflections in the two mirrors, but a third image I_3 results from sequential reflections from both mirrors.

5 REFRACTION THROUGH PLANE SURFACES

Consider light ray (1) in Figure 9a, incident at angle θ_1 at a plane interface separating two transparent media characterized, in order, by refractive indices n_1 and n_2 . Let the angle of refraction be the angle θ_2 . Snell's law, which now takes the form

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \tag{2}$$



Figure 7 Geometry of a ray reflected from a plane.







Figure 8 Image formation in a plane mirror.

requires an angle of refraction such that refracted rays bend away from the normal, as shown in Figure 9a, for rays 1 and 2, when $n_2 < n_1$. For $n_2 > n_1$, on the other hand, the refracted ray bends toward the normal. The law also requires that ray 3, incident normal to the surface ($\theta_1 = 0$), be transmitted without change of direction ($\theta_2 = 0$), regardless of the ratio of refractive indices.

In Figure 9a, the three rays shown originate at a source point S below an interface and emerge into an upper medium of lower refractive index, as in the case of light emerging from water $(n_1 = 1.33)$ into air $(n_2 = 1.00)$. A unique image point is not determined by these rays because they have no common intersection or virtual image point below the surface from which they appear to originate after refraction, as shown by the dashed line extensions of the refracted rays. For rays making a small angle with the normal to the surface, however, a reasonably good image can be located. In this approximation, where we allow only such *paraxial rays*² to form the image, the angles of incidence and refraction are both small, and the approximation

 $\sin \theta \cong \tan \theta \cong \theta$ (in radians)

is valid. From Eq. (2), Snell's law can be approximated by

$$n_1 \tan \theta_1 \cong n_2 \tan \theta_2 \tag{3}$$

and taking the appropriate tangents from Figure 9b, we have





Figure 9 Geometry of rays refracted by a plane interface.

²In general, a paraxial ray is one that remains near the central axis of the image-forming optical system, thus making small angles with the optical axis.

The image point occurs at the vertical distance s' below the surface given by

$$s' = \left(\frac{n_2}{n_1}\right)s\tag{4}$$

where s is the corresponding depth of the object. Thus, objects underwater, viewed from directly overhead, appear to be nearer the surface than they actually are, since in this case s' = (1/1.33) s = (3/4) s. Even when the viewing angle θ_2 is not small, a reasonably good retinal image of an underwater object is formed because the aperture or pupil of the eye admits only a small bundle of rays while forming the image. Since these rays differ very little in direction, they will appear to originate from approximately the same image point. However, the depth of this image will not be 3/4 the object depth, as for paraxial rays, and in general will vary with the angle of viewing.

Rays from the object that make increasingly larger angles of incidence with the interface must, by Snell's law, refract at increasingly larger angles, as shown in Figure 9c. A critical angle of incidence θ_c is reached when the angle of refraction reaches 90°. Thus, from Snell's law,

$$\sin \theta_c = \left(\frac{n_2}{n_1}\right) \sin 90 = \frac{n_2}{n_1}$$

or

$$\theta_c = \sin^{-1} \left(\frac{n_2}{n_1} \right) \tag{5}$$

For angles of incidence $\theta_1 > \theta_c$, the incident ray experiences *total internal re-flection*, as shown. For angle of incidence $\theta_1 < \theta_c$ both refraction and reflection occur. The reflected rays for this case are not shown in Figure 9c. This phenomenon is essential in the transmission of light along glass fibers by a series of total internal reflections. Note that the phenomenon does not occur unless $n_1 > n_2$, so that θ_c can be determined from Eq. (5).

We return to the nature of images formed by refraction at a plane surface when we deal with such refraction as a special case of refraction from a spherical surface.

6 IMAGING BY AN OPTICAL SYSTEM

We discuss now what is meant by an image in general and indicate the practical and theoretical factors that render an image less than perfect. In Figure 10, let the region labeled "optical system" include any number of reflecting and/or refracting surfaces, of any curvature, that may alter the direction of rays leaving an *object point O*. This region may include any number of intervening media, but we shall assume that each individual medium is homogeneous and isotropic, and so characterized by its own refractive index. Thus rays spread out radially in all directions from object point *O*, as shown, in real *object space*, which precedes the first reflecting or refracting surface of the optical system. The family of spherical surfaces normal to the rays are the *wavefronts*, the locus of



points such that each ray contacting a wavefront represents the same transit time of light from the source. In real object space the rays are diverging and the spherical wavefronts are expanding. Suppose now that the optical system redirects these rays in such a way that on leaving the optical system and entering real *image space*, the wavefronts are contracting and the rays are converging to a common point that we define to be the *image point*, *I*. In the spirit of Fermat's principle, we can say that since every such ray starts at *O* and ends at *I*, every such ray requires the same transit time. These rays are said to be *isochronous*. Further, by the *principle of reversibility*, if *I* is the object point, each ray will reverse its direction but maintain its path through the optical system, and *O* will be the corresponding image point. The points *O* and *I* are said to be *conjugate* points for the optical system. In an ideal optical system, every ray from *O* intercepted by the system—and only these rays also passes through *I*. To image an actual object, this requirement must hold for every object point and its conjugate image point.

Nonideal images are formed in practice because of (1) light scattering, (2) aberrations, and (3) diffraction. Some rays leaving O do not reach I due to reflection losses at refracting surfaces, diffuse reflections from reflecting surfaces, and scattering by inhomogeneities in transparent media. Loss of rays by such means merely diminishes the brightness of the image; however, some of these rays are scattered through I from nonconjugate object points, degrading the image. When the optical system itself cannot produce the one-to-one relationship between object and image rays required for perfect imaging of all object points, we speak of system aberrations. Finally, since every optical system intercepts only a portion of the wavefront emerging from the object, the image cannot be perfectly sharp. Even if the image were otherwise perfect, the effect of using a limited portion of the wavefront leads to diffraction and a blurred image, which is said to be *diffraction limited*. This source of imperfect image formation, discussed further in the sections under diffraction, represents a fundamental limit to the sharpness of an image that cannot be entirely overcome. This difficulty rises from the wave nature of light. Only in the unattainable limit of geometrical optics, where $\lambda \rightarrow 0$, would diffraction effects disappear entirely.

Reflecting or refracting surfaces that form perfect images are called *Cartesian surfaces*. In the case of reflection, such surfaces are the conic sections, as shown in Figure 11. In each of these figures, the roles of object and image points may be reversed by the principle of reversibility. Notice that in Figure 11b, the image is virtual. In Figure 11c, the parallel reflected rays are said to form an image "at infinity." In each case, one can show that Fermat's principle, requiring isochronous rays between object and image points, leads to a condition that is equivalent to the geometric definition of the corresponding conic section.

Cartesian surfaces that produce perfect imaging by refraction may be more complicated. Let us ask for the equation of the appropriate refracting surface that images object point O at image point I, as illustrated in Figure 12. There an arbitrary point P with coordinates (x, y) is on the required surface Σ . The requirement is that every ray from O, like OPI, refracts and passes through the image I. Another such ray is evidently OVI, normal to the surface at its vertex point V. By Fermat's principle, these are isochronous rays. Since the media on either side of the refracting surface are characterized by different refractive indices, however, the isochronous rays are not equal in length. The transit time of a ray through a medium of thickness x with refractive index n is



Innininininininininini

(a) Ellipsoid

Figure 11 Cartesian reflecting surfaces showing conjugate object and image points.

$$=\frac{x}{v}=\frac{nx}{c}$$

t



Therefore, equal times imply equal values of the product *nx*, called the *optical path length*. In the problem at hand, then, Fermat's principle requires that

$$n_o d_o + n_i d_i = n_o s_o + n_i s_i = \text{constant}$$
(6)

where the distances are defined in Figure 12. In terms of the (x, y)-coordinates of *P*, the first sum of Eq. (6) becomes

$$n_0(x^2 + y^2)^{1/2} + n_i[y^2 + (s_o + s_i - x)^2]^{1/2} = \text{constant}$$
(7)

The constant in the equation is determined by the middle member of Eq. (6), $n_o s_o + n_i s_i$, which can be calculated once the specific problem is defined. Equation (7) describes the *Cartesian ovoid* of revolution shown in Figure 13a.

In most cases, however, the image is desired in the same optical medium as the object. This goal is achieved by a lens that refracts light rays twice, once at each surface, producing a real image outside the lens. Thus it is of particular interest to determine the Cartesian surfaces that render every object ray parallel after the first refraction. Such rays incident on the second surface can then be refracted again to form an image. The solutions to this problem are illustrated in Figure 13b and c. Depending on the relative magnitudes of the refractive indices, the appropriate refracting surface is either a hyperboloid $(n_i > n_o)$ or an ellipsoid $(n_o > n_i)$, as shown.

The first of these corresponds to the usual case of an object in air. A double hyperbolic lens then functions as shown in Figure 14. Note, however, that the aberration-free imaging so achieved applies only to object point O at the correct distance from the lens and on axis. For nearby points, imaging is not perfect. The larger the actual object, the less precise is its image. Because images of actual objects are not free from aberrations and because hyperboloid surfaces are *difficult to grind* exactly, most optical surfaces are spherical.³ The *spherical aberrations* so introduced are accepted as a compromise when weighed against the relative ease of fabricating spherical surfaces. In the remainder of this chapter, we examine, in detail, spherical reflecting and refracting surfaces and, more briefly, cylindrical reflecting and refracting surfaces. Note that a plane surface can be treated as a special case of a cylindrical or a spherical surface in the limit that the *radius of curvature R* of either type of surface tends to infinity.

7 REFLECTION AT A SPHERICAL SURFACE

Spherical mirrors may be either concave or convex relative to an object point O, depending on whether the center of curvature C is on the same or opposite side of the reflecting surface. In Figure 15 the mirror shown is convex, and two

Figure 12 Cartesian refracting surface which images object point *O* at image point *I*.



Figure 13 Cartesian refracting surfaces. (a) Cartesian ovoid images *O* at *I* by refraction. (b) Hyperbolic surface images object point *O* at infinity when *O* is at one focus and $n_i > n_o$. (c) Ellipsoid surface images object point *O* at infinity when *O* is at one focus and $n_o > n_i$.



Figure 14 Aberration-free imaging of point object *O* by a double hyperbolic lens.

³The refinement of lens construction using injection molding technology has eased the production of lenses with aspherical surfaces.



Figure 15 Reflection at a spherical surface.

rays of light originating at O are drawn, one normal to the spherical surface at its vertex V and the other an arbitrary ray incident at P. The first ray reflects back along itself; the second reflects at P as if from a plane tangent at P, satisfying the law of reflection. The two reflected rays diverge as they leave the mirror. The intersection of the two rays (extended backward) determines the image point I conjugate to O. The image is *virtual*, located behind the mirror surface.

Object and image distances from the vertex are shown as s and s', respectively. A perpendicular of height h is drawn from P to the axis at Q. We seek a relationship between s and s' that depends only on the radius of curvature R of the mirror. As we shall see, such a relation is possible only to first-order approximation of the sines and cosines of the angles made by the object and image rays to the spherical surface. This means that in place of the expansions

 $\sin \varphi = \varphi - \frac{\varphi^3}{3!} + \frac{\varphi^5}{5!} - \cdots$

and

$$\cos \varphi = 1 - \frac{\varphi^2}{2!} + \frac{\varphi^4}{4!} + \cdots$$
 (8)

we consider the first terms only and write

$$\sin \varphi \cong \varphi \quad \text{and} \quad \cos \varphi \cong 1 \tag{9}$$

relations that can be accurate enough if the angle φ is small enough.⁴ This approximation leads to *first-order*, or *Gaussian*, optics, after Karl Friedrich Gauss, who in 1841 developed the foundations of the subject. Returning now to the problem at hand, notice that two angular relationships may be obtained from Figure 15, because the exterior angle of a triangle equals the sum of its interior angles. These are

$$\theta = \alpha + \varphi$$
 and $2\theta = \alpha + \alpha$

which combine to give

$$\alpha - \alpha' = -2\varphi \tag{10}$$

Using the small-angle approximation, the angles of Eq. (10) can be replaced by their tangents, yielding

$$\frac{h}{s} - \frac{h}{s'} = -2\frac{h}{R}$$

⁴For example, for angles φ around 10°, the approximation leads to errors around 1.5%.

where we have also neglected the axial distance VQ, small when angle φ is small. Cancellation of h produces the desired relationship,

$$\frac{1}{s} - \frac{1}{s'} = -\frac{2}{R} \tag{11}$$

If the spherical surface is chosen to be concave instead, the center of curvature would be to the left. For certain positions of the object point O, it is then possible to find a real image point also to the left of the mirror. In these cases, the resulting geometric relationship analogous to Eq. (11) consists of terms that are all positive. It is possible, by employing an appropriate sign convention, to represent all cases by the single equation

$$\frac{1}{s} + \frac{1}{s'} = -\frac{2}{R}$$
(12)

The sign convention to be used in conjunction with Eq. (12) is as follows. Assume the light propagates from left to right:

- 1. The *object distance s* is positive when *O* is to the left of *V*, corresponding to a real object. When *O* is to the right, corresponding to a virtual object, *s* is negative.
- 2. The *image distance s'* is positive when *I* is to the left of *V*, corresponding to a real image, and negative when *I* is to the right of *V*, corresponding to a virtual image.
- 3. The *radius of curvature R* is positive when *C* is to the right of *V*, corresponding to a convex mirror, and negative when *C* is to the left of *V*, corresponding to a concave mirror.

These rules⁵ can be quickly summarized by noticing that positive object and image distances correspond to real objects and real images and that convex mirrors have positive radii of curvature. Applying Rule 2 to Figure 15, we see that the general Eq. (12) becomes identical with Eq. (11), a special case derived in conjunction with Figure 15. Virtual objects occur only with a sequence of two or more reflecting or refracting elements and are considered later.

The spherical mirror described by Eq. (12) yields, for a plane mirror with $R \rightarrow \infty$, s' = -s, as determined previously. The negative sign implies a virtual image for a real object. Notice also in Eq. (12) that object distance and image distance appear symmetrically, implying their interchangeability as conjugate points. For an object at infinity, incident rays are parallel and s' = -R/2, as illustrated in Figure 16a and b for both concave (R < 0) and convex (R > 0) mirrors. The image distance in each case is defined as the *focal length f* of the mirrors. Thus,

$$f = -\frac{R}{2} \begin{cases} >0, & \text{concave mirror} \\ <0, & \text{convex mirror} \end{cases}$$
(13)

and the mirror equation can be written, more compactly, as

$$\frac{1}{s} + \frac{1}{s'} = \frac{1}{f}$$
(14)

⁵Although this set of sign conventions is widely used, the student is cautioned that other schemes exist. No one with a continuing involvement in optics can hope to escape confronting other conventions, nor should the matter be beyond the mental flexibility of the serious student to accommodate.



Figure 16 Location of focal points (a) and (b) and construction to determine magnification (c) of a spherical mirror.

The focal point F, located a focal length f from the vertex of the mirror, and shown in Figure 16a and b, serves as an important construction point in graphical ray-tracing techniques, which we discuss following Example 1.

In Figure 16c, a construction is shown that allows the determination of the transverse magnification. The object is an extended object of transverse dimension h_o . The image of the top of the object arrow is located by two rays whose behavior on reflection is known. The ray incident at the vertex must reflect to make equal angles with the axis. The other ray is directed toward the center of curvature along a normal and so must reflect back along itself. The intersection of the two reflected rays occurs behind the mirror and locates a virtual image of dimension h_i there. Because of the equality of the three angles shown, it follows that

$$\frac{h_o}{s} = \frac{h_i}{s'}$$

The lateral magnification m is defined by the ratio of lateral image size to corresponding lateral object size, so that

$$|m| = \frac{h_i}{h_o} = \frac{s'}{s} \tag{15}$$

Extending the sign convention to include magnification, we assign a (+) magnification to the case where the image has the same orientation as the object and a (-) magnification where the image is inverted relative to the object. To produce a (+) magnification in the construction of Figure 16c, where s' must itself be negative, we modify Eq. (15) to give the general form

$$m = -\frac{s'}{s} \tag{16}$$

The following example illustrates the correct use of the sign convention.

Example 1

An object 3 cm high is placed 20 cm from (a) a convex and (b) a concave spherical mirror, each of 10-cm focal length. Determine the position and nature of the image in each case.

Solution

a. Convex mirror: f = -10 cm and s = +20 cm.

$$\frac{1}{s} + \frac{1}{s'} = \frac{1}{f} \quad \text{or} \quad s' = \frac{fs}{s-f} = \frac{(-10)(20)}{(20) - (-10)} = -6.67 \text{ cm}$$
$$m = -\frac{s'}{s} = -\frac{-6.67}{20} = +0.333 = \frac{1}{3}$$

The image is virtual (because s' is negative), 6.67 cm to the right of the mirror vertex, and is erect (because m is positive) and $\frac{1}{3}$ the size of the object, or 1 cm high.

b. Concave mirror: f = +10 cm and s = +20 cm.

$$s' = \frac{fs}{s - f} = \frac{(10)(20)}{20 - 10} = +20 \text{ cm}$$
$$m = -\frac{s'}{s} = -\frac{20}{20} = -1$$

The image is real (because s' is positive), 20 cm to the left of the mirror vertex, and is inverted (because m is negative) and the same size as the object, or 3 cm high. Image and object happen to be at 2f = 20 cm, the center of curvature of the mirror.

The location and nature of the image formed by a mirror can be determined by graphical ray-trace techniques. Figure 17 illustrates how three key rays—labeled 1, 2, and 3—each leaving a point P at the tip of an object, can be drawn to locate the conjugate image point P'. In fact, under the conditions for which Eqs. (12) through (16) are valid, the paths of *any* two rays leaving P are sufficient to locate the conjugate image point P'. A third ray serves as a convenient check on the accuracy of the first two chosen rays. The three key rays discussed in connection with Figure 17 are chosen as the basis of the graphical ray-trace technique because, once the mirror center of curvature C, the focal point F, and vertex V are located along the optical axis of a spherical mirror, these three rays can be drawn using only a straightedge device. The conjugate image point P' marks the tip of the image—the entire image then lies between P' and the point on the optical axis directly above or below P'.

Refer to Figure 17a, b, and c in connection with the following description of how the three key rays can be drawn. Note the difference in each ray



Figure 17 Ray diagrams for spherical mirrors. (a) Real image, concave mirror. The object distance is greater than the focal length. (b) Virtual image, concave mirror. The object distance is less than the focal length. (c) Virtual image, convex mirror.

trace, depending on the object location before or after points C and F, and on the geometry of the mirror surface, concave or convex.

- *Ray 1*. This ray leaves point *P* as a ray parallel to the optical axis, strikes the mirror, reflects and passes through the focal point *F* of a *concave* mirror—as in Figure 17a and b. Or, as in Figure 17c, it strikes a *convex* mirror and reflects as if it came from the focal point *F* behind the mirror. In each case, after reflection this ray is labeled 1'.
- *Ray 2.* This ray leaves point *P*, passes through *F*, strikes a *concave* mirror, and is reflected as a ray parallel to the optical axis, as in Figure 17a. Or, as in Figure 17b, it leaves point *P* as if it is coming from the point *F* to its left (dotted line), strikes the *concave* mirror, and reflects as a parallel ray. Or, as in Figure 17c, for a *convex* mirror, the ray leaves point *P* heading toward focal point *F* behind the mirror, strikes the mirror, and reflects as a parallel ray. In each case, after reflection, this ray is labeled 2'.
- *Ray 3*. This ray leaves point *P* in Figure 17a, passes through point *C* for the *concave* mirror, strikes the mirror, and reflects back along itself. Or, as in Figure 17b—still for a *concave* mirror—ray 3 appears to come from the point *C* to its left, strikes the mirror, and reflects back along itself. Or, as in Figure 17c, for a *convex* mirror, it heads toward point *C* behind the mirror, strikes the mirror, and reflects back along itself. In each case, after reflection, this ray is labeled 3'.

To understand how these rays locate the conjugate image point P' that marks the tip of the image, it is useful to imagine that these three rays arrive at the eye of one viewing the image. For the case shown in Figure 17a, the three rays 1', 2', and 3' intersect at a real image point as they progress away from the mirror and toward the viewer. For the arrangements shown in Figure 17b and 17c, the rays 1', 2', and 3' appear to originate from a point of intersection (a virtual image point) located behind the mirror. The real or apparent point of intersection is interpreted as the emanation point of these rays. That is, the viewer "sees" the tip of an image at point P'.

8 REFRACTION AT A SPHERICAL SURFACE

We turn now to a similar treatment of refraction at a spherical surface, choosing in this case the concave surface of Figure 18. Two rays are shown emanating from object point O. One is an axial ray, normal to the surface at its vertex and so refracted without change in direction. The other ray is an arbitrary ray incident at P and refracting there according to Snell's law,

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \tag{17}$$

The two refracted rays appear to emerge from their common intersection, the image point *I*. In triangle *CPO*, the exterior angle $\alpha = \theta_1 + \varphi$. In triangle *CPI*, the exterior angle $\alpha' = \theta_2 + \varphi$. Approximating for paraxial rays and substituting for θ_1 and θ_2 in Eq. (17), we have

$$n_1(\alpha - \varphi) = n_2(\alpha' - \varphi) \tag{18}$$

Next, writing the tangents for the angles by inspection of Figure 18, where again we may neglect the distance QV in the small angle approximation,

$$n_1\left(\frac{h}{s}-\frac{h}{R}\right) = n_2\left(\frac{h}{s'}-\frac{h}{R}\right)$$





or

$$\frac{n_1}{s} - \frac{n_2}{s'} = \frac{n_1 - n_2}{R} \tag{19}$$

Employing the *same sign convention* as introduced for mirrors (i.e., positive distances for real objects and images and negative distances for virtual objects and images), the virtual image distance s' < 0 and the radius of curvature R < 0. If these negative signs are understood to apply to these quantities for the case of Figure 18, a general form of the refraction equation may be written as

$$\frac{n_1}{s} + \frac{n_2}{s'} = \frac{n_2 - n_1}{R} \tag{20}$$

which holds equally well for convex surfaces. When $R \rightarrow \infty$, the spherical surface becomes a plane refracting surface, and

$$s' = -\left(\frac{n_2}{n_1}\right)s\tag{21}$$

where s' is the apparent depth determined previously. For a real object (s > 0), the negative sign in Eq. (21) indicates that the image is virtual. The lateral magnification of an extended object is simply determined by inspection of Figure 19. Snell's law requires, for the ray incident at the vertex V and in the small-angle approximation, $n_1\theta_1 = n_2\theta_2$ or, using tangents for angles,

$$n_1\left(\frac{h_o}{s}\right) = n_2\left(\frac{h_i}{s'}\right)$$

The lateral magnification is, then,

$$m = \frac{h_i}{h_o} = -\frac{n_1 s'}{n_2 s} \tag{22}$$

where the negative sign is attached to give a negative value corresponding to an inverted image. For the case of a plane refracting surface, Eq. (21) may 34 Chapter 2



Figure 19 Construction to determine lateral magnification at a spherical refracting surface.

be incorporated into Eq. (22), giving m = +1. Thus, the images formed by plane refracting surfaces have the same lateral dimensions and orientation as the object.

Example 2

As an extended example of refraction by spherical surfaces, refer to Figure 20. In (a), a real object is positioned in air, 30 cm from a convex spherical surface of radius 5 cm. To the right of the interface, the refractive index is that of water. Before constructing representative rays, we first find the image distance and lateral magnification of the image, using Eqs. (20) and (22). Equation (20) becomes

$$\frac{1}{30} + \frac{1.33}{s_1'} = \frac{1.33 - 1}{5}$$

giving $s'_1 = +40$ cm. The positive sign indicates that the image is real and so is located to the right of the surface, where real rays of light are refracted. Equation (22) becomes

$$m = -\frac{(1)(+40)}{(1.33)(+30)} = -1$$



Figure 20 Example of refraction by spherical surfaces. (All distances are in cm.) (a) Refraction by a single spherical surface. (b) Refraction by a thick lens. Subscripts 1 and 2 refer to refractions at the first and second surfaces, respectively.

indicating an inverted image, equal in size to that of the object. Figure 20a shows the image, as well as several rays, which are now determined. In this example we have assumed that the medium to the right of the spherical surface extends far enough so that the image is formed inside it, without further refraction. Let us suppose now (Figure 20b) that the second medium is only 10 cm thick, forming a *thick lens*, with a second, concave spherical surface, also of radius 5 cm. The refraction by the first surface is, of course, unaffected by this change. Inside the lens, therefore, rays are directed as before to form an image 40 cm to the right of the first surface. However, these rays are intercepted and refracted by the second surface to produce a different image, as shown. Since the convergence of the rays striking the second surface is determined by the position of the first image, its location now specifies the appropriate object distance to be used for the second refraction. We call the real image formed by surface (1) a virtual object for surface (2). Then, by the sign convention established previously, we make the virtual object distance, relative to the second surface, a negative quantity when using Eqs. (20) and (22). For the second refraction, then, Eq. (20) becomes

$$\frac{1.33}{-30} + \frac{1}{s'_2} = \frac{1 - 1.33}{-5}$$

or s' = +9 cm. The magnification, according to Eq. (22), is

$$m = \frac{(-1.33)(+9)}{(1)(-30)} = +\frac{2}{5}$$

The final image is, then, 2/5 the lateral size of its (virtual) object and apears with the same orientation. Relative to the original object, the final image is 2/5 as large and inverted.

In general, whenever a train of reflecting or refracting surfaces is involved in the processing of a final image, the individual reflections and/or refractions are considered in the order in which light is actually incident upon them. The object distance of the *n*th step is determined by the image distance for the (n-1)th step. If the image of the (n-1) step is not actually formed, it serves as a *virtual object* for the *n*th step.

9 THIN LENSES

We now apply the preceding method to discover the thin-lens equation. As in the example of Figure 20, two refractions at spherical surfaces are involved. The simplification we make is to neglect the thickness of the lens in comparison with the object and image distances, an approximation that is justified in most practical situations. At the first refracting surface, of radius R_1 ,

$$\frac{n_1}{s_1} + \frac{n_2}{s_1'} = \frac{n_2 - n_1}{R_1}$$
(23)

and at the second surface, of radius R_2 ,

$$\frac{n_2}{s_2} + \frac{n_1}{s_2'} = \frac{n_1 - n_2}{R_2} \tag{24}$$

We have assumed that the lens faces the same medium of refractive index n_1 on both sides. Now the second object distance, in general, is given by

$$s_2 = t - s_1' \tag{25}$$

where t is the thickness of the lens. Notice that this relationship produces the correct sign of s_2 , as in Figure 20, and also when the intermediate image falls inside or to the left of the lens. In the thin-lens approximation, neglecting t,

$$s_2 = -s_1'$$
 (26)

When this value of s_2 is substituted into Eq. (24) and Eqs. (23) and (24) are added, the terms n_2/s'_1 cancel and there results

$$\frac{n_1}{s_1} + \frac{n_1}{s_2'} = (n_2 - n_1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right)$$

Now s_1 is the original object distance and s'_2 is the final image distance, so we may drop their subscripts and write simply

$$\frac{1}{s} + \frac{1}{s'} = \frac{n_2 - n_1}{n_1} \left(\frac{1}{R_1} - \frac{1}{R_2} \right)$$
(27)

The *focal length* of the thin lens is defined as the image distance for an object at infinity, or the object distance for an image at infinity, giving

$$\frac{1}{f} = \frac{n_2 - n_1}{n_1} \left(\frac{1}{R_1} - \frac{1}{R_2} \right)$$
(28)

Equation (28) is called the *lensmaker's equation* because it predicts the focal length of a lens fabricated with a given refractive index and radii of curvature and used in a medium of refractive index n_1 . In most cases, the ambient medium is air, and $n_1 = 1$. The thin-lens equation, in terms of the focal length, is then

$$\frac{1}{s} + \frac{1}{s'} = \frac{1}{f}$$
(29)

Wavefront analysis for plane wavefronts, as shown in Figure 21, indicates that a lens thicker in the middle causes convergence, and one thinner in the middle causes divergence of the incident parallel rays. The portion of the wavefront that must pass through the thicker region is delayed relative to the other portions. Converging lenses are characterized by positive focal lengths and diverging lenses by negative focal lengths, as is evident from the figure, where the images are real and virtual, respectively.

Sample ray diagrams for converging (or *convex*) and diverging (or *concave*) lenses are shown in Figure 22. The thin lenses are best represented, for purposes of ray construction, by a vertical line with edges suggesting the general shape of the lens—ordinary arrowheads for converging lenses, inverted arrowheads for diverging lenses. Graphical methods of locating images, as with spherical mirrors in Figure 17, make use of *three key rays*. This procedure is outlined next and illustrated in Figures 22 and 23. The three rays leaving the tip of the object change direction due to refraction at the thin-lens interfaces. The redirected rays can be used to locate the image.

- *Ray 1*. A ray leaving the tip of the object, parallel to the optical axis, undergoing refraction at the lens surfaces and passing through the *right* focal point *F* of a *converging* lens, as in Figure 22a. Or, as in Figure 22b, a parallel ray which refracts at the lens surfaces as if coming directly from the *left* focal point *F* of a *diverging* lens.
- *Ray 2*. A ray leaving the tip of the object and passing through the *left* focal point *F* of a *converging* lens, undergoing refraction at the lens surfaces,



Figure 21 Lens action on plane wavefronts of light. (a) Converging lens (positive focal length). (b) Diverging lens (negative focal length).



Figure 22 Ray diagrams for image formation by a convex lens (a) and a concave lens (b).

and emerging parallel to the axis as in Figure 22a. Or, as in Figure 22b, a ray leaving the tip of the object, directed toward the *right* focal point F of a *diverging* lens, undergoing refraction at the lens and emerging parallel to the axis.

• *Ray 3*. A ray leaving the tip of the object and passing directly through the center of a converging or diverging lens, emerging unaltered, as in Figure 22a or 22b.

The viewer, located at the far right in Figure 22a and 22b, receives these rays as if they have come directly from an object and so "sees" the tip of the image at the point where the backwards extensions of these rays either intersect or appear to intersect. Any two rays are sufficient to locate the image; the third ray may be drawn as a check on the accuracy of the graphical trace.

In constructing ray diagrams, as in Figure 22, observe that, except for the central ray (ray 3), each ray refracted by a convex lens bends toward the axis and each ray refracted by a concave lens bends away from the axis. From either diagram, the angles subtended by object and image at the center of the lens are seen to be equal. For either the real image RI in (a) or the virtual image VI in (b), it follows that

$$\frac{h_o}{s} = \frac{h_i}{s'}$$

and lateral magnification

$$|m| = \left|\frac{h_i}{h_o}\right| = \left|\frac{s'}{s}\right|$$

In accordance with the sign convention adopted here, the magnification should be the negative of the ratio of the image and object distances since, in case (a), s > 0, s' > 0, and m < 0 because the image is inverted; in case (b),



Figure 23 (a) Formation of a virtual image VI_2 by a two-element train of a convex lens (1) and concave lens (2). (b) Formation of a real image RI_2 by a train of two convex lenses. The intermediate image RI_1 serves as a virtual object VO_2 for the second lens.

s > 0, s' < 0, and m > 0. In either case, then,

$$m = -\frac{s'}{s} \tag{30}$$

Further ray-diagram examples for *a train of two lenses* are illustrated in Figure 23 and a calculation involving image formation in two lenses is given in Example 3.

Example 3

Find and describe the intermediate and final images produced by a two-lens system such as the one sketched in Figure 23a. Let $f_1 = 15$ cm, $f_2 = 15$ cm, and their separation be 60 cm. Let the object be 25 cm from the first lens, as shown.

Solution

The first lens is convex: $f_1 = +15$ cm, $s_1 = 25$ cm.

$$\frac{1}{s_1} + \frac{1}{s_1'} = \frac{1}{f} \quad \text{or} \quad s_1' = \frac{s_1 f}{s_1 - f} = \frac{(25)(15)}{25 - 15} = +37.5 \text{ cm}$$
$$m_1 = -\frac{s_1'}{s_1} = -\frac{37.5}{25} = -1.5$$

Thus, the first image is real (because s'_1 is positive), 37.5 cm to the right of the first lens, inverted (because *m* is negative), and 1.5 times the size of the object.

The second lens is concave: $f_2 = -15$ cm. Since real rays of light diverge from the first real image, it serves as a real object for the second lens, with $s_2 = 60 - 37.5 = +22.5$ cm to the left of the lens. Then,

$$s'_{2} = \frac{s_{2}f}{s_{2} - f} = \frac{(22.5)(-15)}{(22.5) - (-15)} = -9 \text{ cm}$$
$$m_{2} = -\frac{s'_{2}}{s_{2}} = -\frac{-9}{22.5} = +0.4$$

Thus, the final image is virtual (because s'_2 is negative), 9 cm to the *left* of the seconds lens, erect *with respect to its own object* (because *m* is positive), and 0.4 times its size. The *overall* magnification is given by $m = m_1m_2 = (-1.5)(0.4) = -0.6$. Thus, the final image is inverted relative to the *original* object and 6/10 its lateral size. All these features are exhibited qualitatively in the ray diagram of Figure 23a.

Table 1 and Figure 24 provide a convenient summary of image formation in lenses and mirrors.

10 VERGENCE AND REFRACTIVE POWER

Another way of interpreting the thin-lens equation is useful in certain applications, including optometry. The interpretation is based on two considerations. In the thin-lens equation,

$$\frac{1}{s} + \frac{1}{s'} = \frac{1}{f}$$
(31)

	Spherical surface	Plane surface
Reflection	$\frac{1}{s} + \frac{1}{s'} = \frac{1}{f}, f = -\frac{R}{2}$	s' = -s
	$m = -\frac{s'}{s}$	m = +1
	Concave: $f > 0, R < 0$	
	Convex : $f < 0, R > 0$	
Refraction Single surface	$\frac{n_1}{s} + \frac{n_2}{s'} = \frac{n_2 - n_1}{R}$	$s' = -\frac{n_2}{n_1}s$
	$m = -\frac{n_1 s'}{n_2 s}$	m = +1
	Concave: $R < 0$	
	Convex: R > 0	
Refraction Thin lens	$\frac{1}{s} + \frac{1}{s'} = \frac{1}{f}$	
	$\frac{1}{f} = \frac{n_2 - n_1}{n_1} \left(\frac{1}{R_1} - \frac{1}{R_2} \right)$	
	$m = -\frac{s'}{s}$	
	Concave: $f < 0$	
	Convex : f > 0	

TABLE 1 SUMMARY OF GAUSSIAN MIRROR AND LENS FORMULAS



Figure 24 Summary of image formation by (a) spherical mirrors and (b) thin lenses. The location, nature, magnification, and orientation of the image are indicated or suggested. The letters R and V refer to *real* and *virtual*, O and I to *object* and *image*. Changes in elevation of the horizontal lines suggest the magnification in the various regions.

notice that (1) the reciprocals of distances in the left member add to give the reciprocal of the focal length and (2) the reciprocals of the object and image distances describe the curvature of the wavefronts incident at the lens and centered at the object and image positions O and I, respectively. A plane wavefront, for example, has a curvature of zero. In Figure 25 spherical waves expand from the object point O and attain a curvature, or *vergence*, V, given by 1/s, when they intercept the thin lens. On the other hand, once refracted by the lens, the wavefronts contract, in Figure 25a, and expand further, in Figure 25b, to locate the real and virtual image points shown. The curvature, or vergence, V', of the wavefronts as they emerge from the lens is 1/s'. The change in curvature from object space to image space is due to the *refracting power* P of the lens, given by 1/f. With these definitions, Eq. (31) may be written

$$V + V' = P \tag{32}$$



Figure 25 Change in curvature of wavefronts on refraction by a thin lens. (a) Convex lens. (b) Concave lens.

The units of the terms in Eq. (32) are reciprocal lengths. When the lengths are measured in *meters*, their *reciprocals* are said to have units of *diopters* (D). Thus, the refracting power of a lens of focal length 20 cm is said to be

 $\frac{1}{0.2 \text{ m}}$ = 5 diopters. This alternative point of view emphasizes the degree of

wave curvature or ray convergence rather than object and image distances. Accordingly, the degree of convergence V' of the image rays is determined by the original degree of convergence V of the object rays and the refracting power P of the lens, that is, the power to change incident wave curvature. Eq. (32) can also be applied to the case of refraction at a single surface, Eq. (20), in which case the refractive indices in object and image space need not be 1. In this event, the *power of the refracting surface* is $(n_2 - n_1)/R$.

This approach is useful for another reason. When thin lenses are placed together, *in contact*, the focal length f of the combination, treated as a single thin lens, can be found in terms of the focal lengths f_1, f_2, \ldots of the individual lenses. For example, with two such lenses back-to-back, we write the lens equations

$$\frac{1}{s_1} + \frac{1}{s_1'} = \frac{1}{f_1}$$
 and $\frac{1}{s_2} + \frac{1}{s_2'} = \frac{1}{f_2}$

Since the image distance for the first lens plays the role of the object distance for the second lens, we may write

$$s_2 = -s'_1$$

and, adding the two equations,

$$\frac{1}{s_1} + \frac{1}{s_2'} = \frac{1}{f_1} + \frac{1}{f_2} = \frac{1}{f}$$

The reciprocals of the individual focal lengths, therefore, add to give the reciprocal of the overall focal length *f* of the pair. In general, for several thin lenses, *in direct contact*,

$$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2} + \frac{1}{f_3} + \dots$$
(33)

Expressed in diopters, the refractive powers simply add:

$$P = P_1 + P_2 + P_3 + \cdots$$
 (34)

In a nearsighted eye, the refracted (converging) power of the eye is too great, so that a real image is formed in front of the retina. By reducing the convergence with a number of diverging lenses placed in front of the eye, until an object is clearly focused, an optometrist can determine the net diopter specification of the single corrective lens needed by simply adding the diopters of these test lenses. In a farsighted eye, the natural converging power of the eye is not strong enough, and additional converging power must be added in the form of spectacles with a converging lens.

11 NEWTONIAN EQUATION FOR THE THIN LENS

When object and image distances are measured relative to the focal points F of a lens, as by the distances x and x' in Figure 26, an alternative form of the thin-lens equation results, called the *Newtonian form*. In the figure, the two rays shown determine two right triangles, joined by the focal point, on each side of the lens. Since each pair constitutes similar triangles, we may set up proportions between sides that represent the lateral magnification:

$$\frac{h_i}{h_o} = \frac{f}{x}$$
 and $\frac{h_i}{h_o} = \frac{x'}{f}$

Introducing a negative sign for the magnification, due to the inverted image,

$$m = -\frac{f}{x} = -\frac{x'}{f} \tag{35}$$

The two parts of Eq. (35) also constitute the Newtonian form of the thin-lens equation,

$$xx' = f^2 \tag{36}$$

This equation is somewhat simpler than Eq. (29) and is found to be more convenient in certain applications.

12 CYLINDRICAL LENSES

Spherical lenses and mirrors with circular cross sections are far more common in optical systems than are *cylindrical lenses*. Nevertheless, cylindrical lenses are important, for example, in the field of optometry for correcting the visual defect known as astigmatism, as well as in novel visual displays where it is useful to image points as lines. We close this chapter on geometrical optics with a brief look at this special type of lens.

The *optical axis* for a spherical lens is an *axis of symmetry* since rotation of the lens through an arbitrary angle about the optical axis leaves the lens looking just as it did before the rotation. Because the orientation of the surface curvature does not change in such a rotation, its optical behavior remains unchanged. This rotational symmetry simplifies the analysis of the imaging properties of such a spherical lens. On the other hand, a cylindrical lens—shaped like a section of a soft drink can, sliced down the side from top to bottom—lacks rotational symmetry about its optical axis. As a consequence, a cylindrical lens has asymmetric focusing properties, as will be seen later in greater detail. Whereas a spherical lens produces a point image of a point



Figure 26 Construction used to derive Newton's equations for the thin lens.

object, a cylindrical lens produces a line image of a point object. Because of these properties, a spherical lens is said to be *stigmatic*, and the cylindrical lens *astigmatic*.

Consider first a spherical lens, as shown in Figure 27a and b. Orthogonal vertical and horizontal axes are shown as solid diametrical lines through the geometric lens center. Parallel rays of light passing through the vertical axis (see Figure 27a) and through the horizontal axis (see Figure 27b) are handled identically by the lens, converging them to a common focus at F.

The lens can be rotated through an arbitrary angle about its optical axis with the same result. Thus, the focusing properties of a spherical lens are invariant to rotation about its central (optical) axis.

Next, consider the convex and concave cylindrical lenses shown in Figure 28. One surface of the lens is cylindrical while the opposite is plane.⁶ Thus, the curved surface has a definite, finite radius of curvature, whereas the plane surface has an infinite radius of curvature. In Figure 29, two vertical slices or sections are shown perpendicular to the axis of a convex cylindrical lens. Through each section, three representative rays are drawn. The operation of this lens is clearly asymmetric. Focusing occurs for rays along a vertical section but not for rays along a horizontal section, where the lens presents no curvature. Thus, rays 1, 2 and 3 focus to point A, and rays 4, 5 and 6 focus to point B. However, there is no focusing of rays in a horizontal section, such as the pairs of rays 1 and 4, 2 and 5, or 3 and 6. Other vertical sections would produce other points along the focused line image AB in the same way. Notice that the line image AB so formed is always parallel to the cylinder axis. This important feature is also shown in Figure 30, where the line image is real for a cylindrical convex lens and virtual for a cylindrical concave lens. From these figures, it is evident that the length of the line image is equal to the axial length of the lens, assuming that rays of light parallel to the optical axis enter along the entire extent of the lens. If an aperture is placed in front of the lens to limit the bundle of light rays through the lens, the height of the line image is just the aperture dimension along the cylinder axis, or the effective height of the lens.

In Figure 29, the line image formed is the result of an object point "at infinity," which produces parallel rays at the lens. In Figure 31, the object point O is near the lens, producing diverging rays of light incident on the lens. Still, if the lens is thin, focusing occurs along the vertical sections, as shown. Rays 1 and 3, in the left vertical section of Figure 31, focus at A; rays 2 and 4 in the right vertical section focus at B. However, no focusing occurs for rays 1, 5, and 2 along the horizontal section. Because of the divergence of the rays entering the lens, however, the length of the focused line image AB is no longer equal to the effective length CL of the lens. The divergence of the extreme rays at each end of the lens now determines an image that is *longer* than the length of the lens. The image length AB can be found from a simple, geometrical argument that is apparent in Figure 32a, a view of the central horizontal section in Figure 31 as seen from above. If the effective length of the cylindrical lens is CL, then by similar triangles it follows that

$$\frac{AB}{CL} = \frac{s+s'}{s}$$





Figure 27 Parallel rays of light focused by a spherical lens. Because of its axis of symmetry relative to rotation about an axis through its center, the lens treats (a) vertical and (b) horizontal fans of rays similarly, producing in each case a point image at the same location. Each ray refracts twice through the lens, once at each surface. For simplicity, only one refraction is shown.

Convex (a)



Figure 28 Cylindrical lenses shown as sections of a solid and hollow cylindrical rod.

Geometrical Optics

or

$$AB = \left(\frac{s+s'}{s}\right)CL \tag{37}$$

Subject to our sign convention, this relation is a general form for the planocylindrical lens that handles all cases, with s and s' object and image distances, respectively, and AB always positive.

Example 4

A thin plano-cylindrical lens in air has a radius of curvature of 10 cm, a refractive index of 1.50, and an axial length of 5 cm. Light from a point object is incident on the convex cylindrical surface from a distance of 25 cm to the left of the lens. Find the position and length of the line image formed by the lens.

Solution

As given, s = 25 cm, R = 10 cm, n(lens) = 1.50 and CL = 5 cm. Using the spherical surface refraction equation (see Table 1),

$$\frac{n_1}{s} + \frac{n_2}{s'} = \frac{n_2 - n_1}{R}$$

and

$$AB = \left(\frac{s+s'}{s}\right)CL$$

together with the sign convention-positive for real objects and images, negative for virtual objects and images, positive R for convex surface.

Entering values, we have for the first convex surface at entry, $\frac{1}{25}$ + $\frac{1.50}{s'} = \frac{1.50 - 1.00}{10}$, which gives s' = 150 cm, real. And for the second



Figure 29 Focusing property of a convex cylindrical lens. Rays through a vertical section, such as rays 1, 2, and 3, come to a common focus, but rays through a horizontal section, such as rays 1 and 4, do not. Parallel rays form a line image that is parallel to the cylinder axis. For simplicity, refraction is shown only at the front surface and spherical aberration for non-paraxial rays is ignored.



Figure 30 Formation of line images by cylindrical lenses for light incident from a distant object. In (a) the convex lens forms a real image. In (b), the concave lens forms a virtual image. In either case, the line image is parallel to the cylinder axis.



Figure 31 Formation of a line image AB by a convex cylindrical lens when the object is a point O at a finite distance from the lens. In this case, the line image AB is longer than the axial length of the lens, CL.

(plane) surface at exit, we obtain $\frac{1.50}{-150} + \frac{1.0}{s'} = 0$, which gives s' = 100 cm. Then with Eq. (37), $AB = \left(\frac{25 + 100}{25}\right)5$ cm = 25 cm.

Thus, the line image is parallel to the cylindrical axis, enlarged to 25 cm, and located 1 m from the lens. If the lens is rotated about its optical axis, the line image also rotates, remaining always parallel to the cylindrical axis.

Looking again at Figure 31, imagine a screen placed on the exit side of the lens so as to capture the light from the lens. We have argued that when the screen is at the distance s' from the lens, one sees a focused line image AB on the screen, in this case with a horizontal orientation. As the screen is moved





Figure 32 (a) Light rays in a top view of the horizontal (nonfocusing) section of the lens in Figure 31. (b) Light rays in a side view of the vertical (focusing) section of the lens in Figure 31.

further than s' from the lens, one sees an unfocused blur that has the general shape of the aperture—either of the rectangular cross section of the lens alone or of the lens with an aperture placed against it. Further, it should be evident from Figure 31 that, as a screen, initially positioned just behind the lens, moves toward the line image *AB*, the horizontal dimension (width) of the blur increases and its vertical dimension (height) decreases. As the screen moves beyond the line image, its width continues to increase but its height now also increases due to the divergence of the rays after focusing. If an aperture placed in front of the lens is circular, these blur images are elliptical in shape, with changing major and minor axes formed by the width and height of the blur. If the aperture is square, the blurs are rectangular in shape. Widths and heights of the blur pattern can be found at any position of the screen using the geometry apparent in Figure 32a and b, respectively. This behavior can be observed easily in the laboratory.

Up to this point we have been dealing with a cylindrical lens whose axis is either horizontal or vertical. Of course, the cylinder axis can be oriented at any angle. An astigmatic eye, for example, while it possesses predominantly spherical optics, might have a cylindrical axis component whose axis could be horizontal, vertical, or some angle in between. To deal with cylindrical lenses and astigmatism in a general way, then, we must be able to determine the effect of combining cylindrical lenses having arbitrary orientations with each other and with spherical lenses. It turns out that two cylindrical lenses can produce the same effect as a sphero-cylindrical lens. Lens prescriptions for vision correction are, in fact, expressed in terms of combinations of spherical and cylindrical lenses. This subject is treated further elsewhere.⁷

PROBLEMS

- 1 Derive an expression for the transit time of a ray of light that travels a distance x_1 through a medium of index n_1 , a distance x_2 through a medium of index n_2, \ldots , and a distance x_m through a medium of index n_m . Use a summation to express your result.
- 2 Deduce the Cartesian oval for perfect imaging by a refracting surface when the object point is on the optical *x*-axis 20 cm from the surface vertex and its conjugate image point lies 10 cm inside the second medium. Assume the refracting medium to have an index of 1.50 and the outer medium to be air. Find the equation of the intersection of the oval with the *xy*-plane, where the origin of the coordinates is at the object point. Generate a table of (x, y)-coordinates for the surface and plot, together with sample rays.
- 3 A double convex lens has a diameter of 5 cm and zero thickness at its edges. A point object on an axis through the center of the lens produces a real image on the opposite side. Both object and image distances are 30 cm, measured from a plane bisecting the lens. The lens has a refractive index of 1.52. Using the equivalence of optical paths through the center and edge of the lens, determine the thickness of the lens at its center.

- 4 Determine the minimum height of a wall mirror that will permit a 6-ft person to view his or her entire height. Sketch rays from the top and bottom of the person, and determine the proper placement of the mirror such that the full image is seen, regardless of the person's distance from the mirror.
- 5 A ray of light makes an angle of incidence of 45° at the center of the top surface of a transparent cube of index 1.414. Trace the ray through the cube.
- **6** To determine the refractive index of a transparent plate of glass, a microscope is first focused on a tiny scratch in the upper surface, and the barrel position is recorded. Upon further lowering the microscope barrel by 1.87 mm, a focused image of the scratch is seen again. The plate thickness is 1.50 mm. What is the reason for the second image, and what is the refractive index of the glass?
- 7 A small source of light at the bottom face of a rectangular glass slab 2.25 cm thick is viewed from above. Rays of light totally internally reflected at the top surface outline a circle of 7.60 cm in diameter on the bottom surface. Determine the refractive index of the glass.

⁷See F. L. Pedrotti and L. S. Pedrotti, *Optics and Vision* (Upper Saddle River, N. J.: Prentice Hall, Inc., 1998).





8 Show that the lateral displacement *s* of a ray of light penetrating a rectangular plate of thickness *t* is given by

$$s = \frac{t\sin(\theta_1 - \theta_2)}{\cos\theta_2}$$

where θ_1 and θ_2 are the angles of incidence and refraction, respectively. Find the displacement when t = 3 cm, n = 1.50, and $\theta_1 = 50^\circ$.

- **9** A meter stick lies along the optical axis of a convex mirror of focal length 40 cm, with its nearer end 60 cm from the mirror surface. How long is the image of the meter stick?
- **10** A glass hemisphere is silvered over its curved surface. A small air bubble in the glass is located on the central axis through the hemisphere 5 cm from the plane surface. The radius of curvature of the spherical surface is 7.5 cm, and the glass has an index of 1.50. Looking along the axis into the plane surface, one sees two images of the bubble. How do they arise and where do they appear?



Figure 34 Problem 10

- **11** A concave mirror forms an image on a screen twice as large as the object. Both object and screen are then moved to produce an image on the screen that is three times the size of the object. If the screen is moved 75 cm in the process, how far is the object moved? What is the focal length of the mirror?
- 12 A sphere 5 cm in diameter has a small scratch on its surface. When the scratch is viewed through the glass from a position directly opposite, where does the scratch appear and what is its magnification? Assume n = 1.50 for the glass.

- **13 a.** At what position in front of a spherical refracting surface must an object be placed so that the refraction produces parallel rays of light? In other words, what is the focal length of a single refracting surface?
 - **b.** Since real object distances are positive, what does your result imply for the cases $n_2 > n_1$ and $n_2 < n_1$?
- 14 A small goldfish is viewed through a spherical glass fishbowl 30 cm in diameter. Determine the apparent position and magnification of the fish's eye when its actual position is (a) at the center of the bowl and (b) nearer to the oberver, halfway from center to glass, along the line of sight. Assume that the glass is thin enough so that its effect on the refraction may be neglected.
- **15** A small object faces the convex spherical glass window of a small water tank. The radius of curvature of the window is 5 cm. The inner back side of the tank is a plane mirror, 25 cm from the window. If the object is 30 cm outside the window, determine the nature of its final image, neglecting any refraction due to the thin glass window itself.



Figure 35 Problem 15

- **16** A plano-convex lens having a focal length of 25.0 cm is to be made with glass of refractive index 1.520. Calculate the radius of curvature of the grinding and polishing tools to be used in making this lens.
- 17 Calculate the focal length of a thin meniscus lens whose spherical surfaces have radii of curvature of magnitude 5 and 10 cm. The glass is of index 1.50. Sketch both positive and negative versions of the lens.
- 18 One side of a fish tank is built using a large-aperture thin lens made of glass (n = 1.50). The lens is equiconvex, with radii of curvature 30 cm. A small fish in the tank is 20 cm from the lens. Where does the fish appear when viewed through the lens? What is its magnification?

- 19 Two thin lenses have focal lengths of -5 and +20 cm. Determine their equivalent focal lengths when (a) cemented together and (b) separated by 10 cm.
- **20** Two identical, thin, plano-convex lenses with radii of curvature of 15 cm are situated with their curved surfaces in contact at their centers. The intervening space is filled with oil of refractive index 1.65. The index of the glass is 1.50. Determine the focal length of the combination. (*Hint*: Think of the oil layer as an intermediate thin lens.)



Figure 36 Problem 20

- **21** An eyepiece is made of two thin lenses each of +20-mm focal length, separated by a distance of 16 mm.
 - **a.** Where must a small object be positioned so that light from the object is rendered parallel by the combination?
 - **b.** Does the eye see an image erect relative to the object? Is it magnified? Use a ray diagram to answer these questions by inspection.
- **22** A diverging thin lens and a concave mirror have focal lengths of equal magnitude. An object is placed (3/2)*f* from



Figure 37 Problem 22

the diverging lens, and the mirror is placed a distance 3f on the other side of the lens. Using Gaussian optics, determine the *final* image of the system, after two refractions (a) by a three-ray diagram and (b) by calculation.

- 23 A small object is placed 20 cm from the first of a train of three lenses with focal lengths, in order, of 10, 15, and 20 cm. The first two lenses are separated by 30 cm and the last two by 20 cm. Calculate the final image position relative to the last lens and its linear magnification relative to the original object when (a) all three lenses are positive, (b) the middle lens is negative, (c) the first and last lenses are negative. Provide ray diagrams for each case.
- **24** A convex thin lens with refractive index of 1.50 has a focal length of 30 cm in air. When immersed in a certain transparent liquid, it becomes a negative lens with a focal length of 188 cm. Determine the refractive index of the liquid.
- 25 It is desired to project onto a screen an image that is four times the size of a brightly illuminated object. A planoconvex lens with n = 1.50 and R = 60 cm is to be used. Employing the Newtonian form of the lens equations, determine the appropriate distance of the object and screen from the lens. Is the image erect or inverted? Check your results using the ordinary lens equations.
- 26 Three thin lenses of focal lengths 10 cm, 20 cm, and -40 cm are placed in contact to form a single compound lens.
 - **a.** Determine the powers of the individual lenses and that of the unit, in diopters.
 - **b.** Determine the vergence of an object point 12 cm from the unit and that of the resulting image. Convert the result to an image distance in centimeters.
- 27 A lens is moved along the optical axis between a fixed object and a fixed image screen. The object and image positions are separated by a distance L that is more than four times the focal length of the lens. Two positions of the lens are found for which an image is in focus on the screen, magnified in one case and reduced in the other. If the two lens positions differ by distance D, show that the focal length of the lens is given by $f = (L^2 D^2)/4L$. This is *Bessel's method* for finding the focal length of a lens.



Figure 38 Problem 27

28 An image of an object is formed on a screen by a lens. Leaving the lens fixed, the object is moved to a new position and the image screen moved until it again receives a focused image. If the two object positions are S_1 and S_2 and if the transverse magnifications of the image are M_1 and M_2 , respectively, show that the focal length of the lens is given by

$$f = \frac{(S_2 - S_1)}{\left(\frac{1}{M_1} - \frac{1}{M_2}\right)}$$

This is Abbe's method for finding the focal length of a lens.

29 Derive the law of reflection from Fermat's principle by minimizing the distance of an arbitrary (hypothetical) ray from a given source point to a given receiving point.

- **30** Determine the ratio of focal lengths for two identical, thin, plano-convex lenses when one is silvered on its flat side and the other on its curved side. Light is incident on the unsilvered side.
- **31** Show that the minimum distance between an object and its image, formed by a thin lens, is 4*f*. When does this occur?
- **32** A ray of light traverses successively a series of plane interfaces, all parallel to one another and separating regions of differing thickness and refractive index.
 - **a.** Show that Snell's law holds between the first and last regions, as if the intervening regions did not exist.
 - **b.** Calculate the net lateral displacement of the ray from point of incidence to point of emergence.



Figure 39 Problem 32

- **33** A parallel beam of light is incident on a plano-convex lens that is 4 cm thick. The radius of curvature of the spherical side is also 4 cm. The lens has a refractive index of 1.50 and is used in air. Determine where the light is focused for light incident on each side.
- 34 A spherical interface, with radius of curvature 10 cm, separates media of refractive index 1 and $\frac{4}{3}$. The center of curvature is located on the side of the higher index. Find the focal lengths for light incident from each side. How do the results differ when the two refractive indices are interchanged?
- **35** An airplane is used in aerial surveying to make a map of ground detail. If the scale of the map is to be 1:50,000 and the camera used has a focal length of 6 in., determine the proper altitude for the photograph.
- **36** Light rays emanating in air from a point object on axis strike a plano-cylindrical lens with its convex surface facing the object. Describe the line image by length and location if the lens has a radius of curvature of 5 cm, a refractive index of 1.60, and an axial length of 7 cm. The point object is 15 cm from the lens.

- **37** A plano-cylindrical lens in air has a curvature of 15 cm and an axial length of 2.5 cm. The refractive index of the lens is 1.52. Find the position and length of the line image formed by the lens for a point object 20 cm from the lens. Light from the object is incident on the convex cylindrical surface of the lens.
- 38 A plano-cylindrical lens in air has a radius of curvature of 10 cm, a refractive index of 1.50, and an axial length of 5 cm. Light from a point object is incident on the concave, cylindrical surface from a distance of 25 cm to the left of the lens. Find the position and length of the image formed by the lens.
- **39** A plano-concave cylindrical lens is used to form an image of a point object 20 cm from the lens. The lens has a refractive index of 1.50, a radius of curvature of 20 cm, and an axial length of 2 cm. Describe as completely as possible the line image of the point.
- **40** Consider the plano-convex cylindrical lens in problem 36. If the point object is only 6 cm from the lens, describe the line image.