2. Lecture, 7 September 1999

2.1 Rays and Snell's Law

The first part of the job of an observational astronomer is to collect the light from a celestial source – incident on the earth in the form of plane waves – and concentrate it on a detector, making the wavefronts converge there. One can focus light in this fashion, of course, by the use of curved dielectric or conducting surfaces. Here we discuss the simplest way to prescribe the surfaces that would be required for a given focusing task.

In your E&M classes, you have learned (or will learn) how to treat the propagation of a plane electromagnetic wave incident upon a plane surface dividing two dielectric or conducting media, and solve for the resulting wavevectors and electric field amplitudes of transmitted and reflected light. The directions of the wavevectors are given by Snell's Law and the mirror-reflection rule, illustrated in Figure 2.1:

$$\begin{aligned} n_i \sin\theta_i &= n_t \sin\theta_t \\ \theta_i &= -\theta_r \end{aligned} , \qquad (2.1)$$

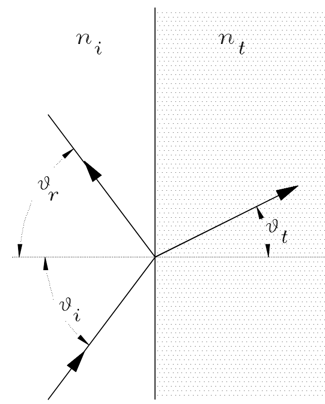

Figure 2.1: reflection and refraction.

where the angles θ between the wavevectors and the normal to the surface are taken to be positive if measured counterclockwise from the normal, and the refractive indices n_i and n_t are, in general, complex. We can use these expressions to deal with *curved* wavefronts and *non-planar* surfaces by considering separate, infinitesimal sections of wavefront and surface, on which scale everything is flat. The infinitesimal section of wavefront is generally represented graphically by the wavevector appropriate to that point on the wavefront, or even more conveniently, simply by the direction of that wavevector. We call this representation a *ray*.

A given ray is incident on a specific point on an interface; we can deal with the transmission or reflection of this ray by considering the properties of a plane wave incident on the plane surface tangent to that point. Ignore for the moment the fact that only a fraction of the power is transmitted, in general, at each interface (the rest getting reflected or absorbed). The propagation of a wave through a surface can then be described by the propagation of an appropriate bundle of rays, acting in accordance with Snell's Law and the mirror-reflection rule. The relationship between wavefronts and rays is shown in Figure 2.2 for a spherical wave. If we are interested in the nature of the wavefront, we can reconstruct it from the rays; it is a locus of points perpendicular to the rays, lying a given number of wavelengths along the rays from some reference surface at which the wave's phase was constant. (The latter requirement guarantees that, for example, a wavefront corresponding to a peak in the wave amplitude remains a peak.)

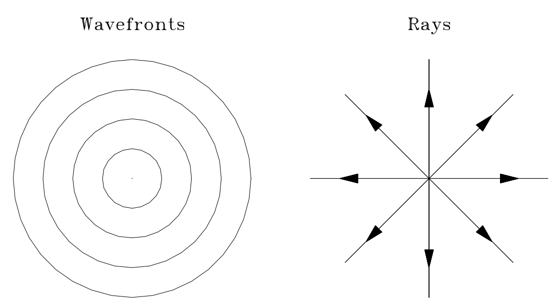

Figure 2.2: Expanding spherical wavefronts (the circles denoting peaks in field amplitude) and their ray representation.

2.2 Paraxial rays

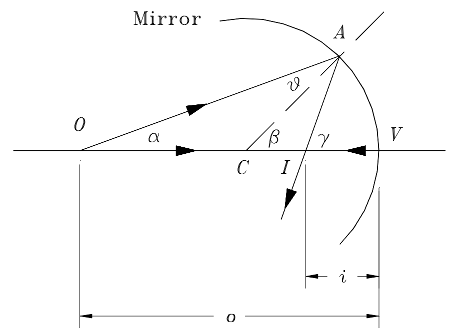

Figure 2.3: Concave spherical mirror.

The use of the ray formalism is illustrated well in the following simple example. Consider a point source of light (emitting spherical waves), located at point *O* in Figure 2.3, and a concave, spherical reflecting surface whose apex lies a distance *o* away from the *O*. What does this mirror do to the light? It focuses it, of course, and we can follow the rays to find out where the focused *image* lies.

Let's follow two different rays that start at *O*, one passing through the center of the sphere, *C*, and striking the sphere's apex, *V*, and the other striking the mirror at point *A*. The first of these follows a radius, since it passes through the sphere's center. The plane tangent to the sphere at its intersection with a radius is perpendicular to that radius, so this ray must reflect right back on itself ($\theta_i = \theta_r = 0$). By the same token, the dashed line bisects the angle *OAI* (so that $\theta_i = -\theta_r = \theta$). This, and the fact that the sum of the angles of a triangle is π, gives us the following two relations between the angles in the figure:

$$\begin{aligned} \alpha + \theta &= \beta \\ \alpha + 2\theta &= \gamma \end{aligned} \quad . \tag{2.2}$$

Eliminating θ from these two equations, we find an expression that holds for all of the rays leaving *O* that encounter the mirror:

$$\alpha = 2\beta - \gamma \Rightarrow \alpha + \gamma = 2\beta \quad . \tag{2.3}$$

We now make an approximation. Suppose that all of the angles in Figure 2.3 are small enough that the arc-length AV can be used for their radian measure. This means that we only consider rays that are close to the line through the axis *OCV* of the system; these are therefore called *paraxial rays*, and this invocation of the small-angle approximation is known as the *paraxial limit*. If the sphere has radius *r*,

$$\alpha \approx \frac{AV}{o} \qquad \beta \equiv \frac{AV}{r} \quad \text{and} \quad \gamma \approx \frac{AV}{i} \quad . \tag{2.4}$$

Putting these into the previous equation, and canceling the common factor of AV, we have

$$\frac{1}{o} + \frac{1}{i} = \frac{2}{r} \quad . \tag{2.5}$$

What we have found here is the location of the *intersection* of our two rays. Note, however, that the result for the distance *i* is independent of the location of point *A*, as long as the small-angle approximation is valid. Thus *all* such rays leaving *O* and hitting the mirror pass through *I*: the light is *focused* there. This form of focusing is called a *real image*; real, because a detector placed at that point would detect light. In addition, all of the rays passing through point *I* diverge from it just as they would if they were coming from a real point source, so to an observer looking at the mirror, there would appear to be a light source there.

If the light source were very far away, so that the rays encountering the mirror would be essentially parallel to one another, we would expect them to reflect through the mirror's focal point. Thus we can get the focal length of the mirror by deriving the distance from the mirror's apex at which the paraxial rays intersect, for a very large object distance:

$$\frac{1}{\infty}+\frac{1}{i}=\frac{1}{i}=\frac{2}{r} \equiv \frac{1}{f} \quad . \tag{2.6}$$

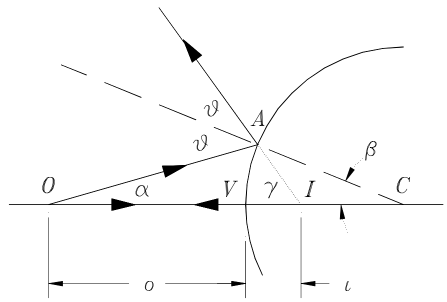

Figure 2.4: a convex, spherical mirror.

The fact that paraxial rays from a point object all focus to a point image is in itself very useful, and it forms the basis of the simplest approach to geometrical optics, which most workers use as a first approximation to any optical system they design: the paraxial theory of *thin**, spherical-surface, lenses and mirrors. Consider a convex spherical mirror, in the same situation as the previous example (Figure 2.4). Once again, we consider two rays originating in a point source of light, one aimed at the apex and center of the sphere, and one at an off-center point. The axial ray reflects back upon itself again, since it is normally incident on the spherical surface. The other ray, incident at angle θ, leaves the picture deviated by an angle γ from the axis. This time they do not intersect, but the reflected rays appear to diverge from a point a distance *i* from the sphere's apex *V*. The relations between the angles in Figure 2.4 are now

$$\alpha+\beta=\theta \quad \text{and} \quad \alpha+\gamma=2\theta \quad , \tag{2.7}$$

whence

$$\alpha-\gamma=-2\beta \quad . \tag{2.8}$$

Restricting ourselves to the paraxial rays, we can once again use the arclength AV in the radian measure of the angles ($\alpha \approx AV/o$, $\beta \equiv AV/r$, $\gamma \approx AV/i$) and obtain

$$\frac{1}{o}-\frac{1}{i}=-\frac{2}{r} \quad . \tag{2.9}$$

An observer looking at the reflected paraxial rays would thus see an image, located at point *I*, but since light doesn't actually converge there, a detector placed there would not detect anything. This sort of image is called a *virtual image*.

The equations relating *o*, *i* and *r* for concave and convex spherical mirrors can be combined into a single expression if we introduce a *sign convention*. We can work out image and object positions for either situation if we agree to use the following rules:

$$\frac{1}{o}+\frac{1}{i}=\frac{1}{f} \quad \text{(the } \textit{thin mirror} \text{ or } \textit{thin surface} \text{ equation);} \tag{2.10}$$

$$f=r/2 \quad \text{for spherical mirrors;} \tag{2.11}$$

* By *thin* we mean that the surface normals always make small angles with the optical axis.

$$\begin{aligned} &o, i < 0 && \text{for virtual objects (more on these later) and images;} \\ &o, i > 0 && \text{for real objects;} \\ &r < 0 && \text{for diverging (convex) mirrors;} \\ &r > 0 && \text{for converging (concave) mirrors.} \end{aligned} \qquad (2.12)$$

Similar steps can be taken to obtain the relationship among o, i and r for a spherical surface bounding two dielectric media. Consider the situations, shown in Figure 2.5, of dielectric media with refractive indices n_1 and $n_2 > n_1$. For the convex surface (Figure 2.5 a) we would find that

$$\frac{n_1}{o} + \frac{n_2}{i} = \frac{n_2 - n_1}{r} \quad , \qquad (2.13)$$

and for the concave surface,

$$\frac{n_1}{o} + \frac{n_2}{i} = -\frac{n_2 - n_1}{r} \quad , \qquad (2.14)$$

from which in either case we can define a focal length (the image distance corresponding to $o \to \infty$) given by

$$f = \frac{n_2}{n_2 - n_1} r \quad . \qquad (2.15)$$

subject to the sign convention $r > 0$ for the convex shape and $r < 0$ for the concave one. This is opposite the convention just established for mirrors, but makes sense in terms of the usual applications of converging and diverging lenses, as we shall see.

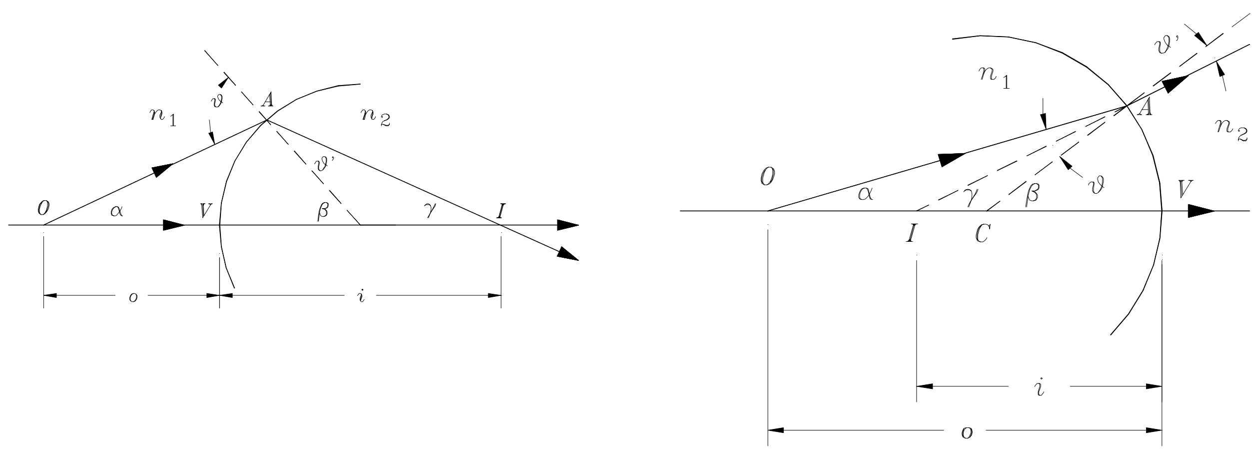

Figure 2.5: a (left) a convex spherical dielectric; b (right) a concave spherical dielectric.

Homework Problem 1.1. Show that Equations 2.13-15 are obtained for the convex and concave dielectric surfaces shown in Figure 2.5.

2.3 Thin lenses

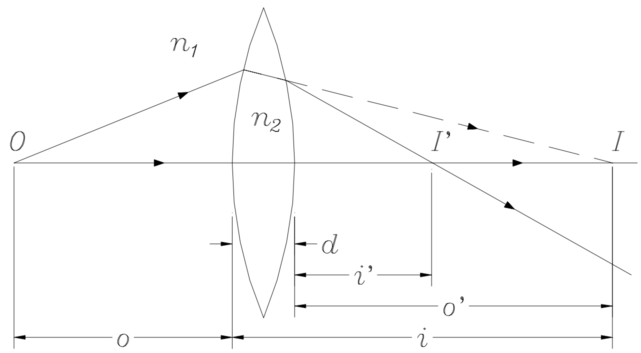

Figure 2.6: Biconvex lens: front surface with radius r_1 (> 0), second surface with radius r_2 (< 0).

Now suppose that the convex spherical dielectric surface we just considered is followed immediately by another surface, also spherical and with center and apex (a distance *d* from the first surface's apex) lying along the original optical axis. If the index returns to the value n_1 following this surface, and if $n_2 > n_1$, then this is just the familiar situation of the *lens*, as shown for a biconvex shape in Figure 2.6. We would like an account of the effect of the second surface on the position of the focus, independent of the algebraic signs of the radii of curvature and n_2 - n_1. We could find one out by tracing rays geometrically, as we have been, but within the paraxial approximation we could also use Equation 2.13, and employ the image formed by the first surface as the object for the second.

According to our previous results, equations 2.13 and 2.15, we can write, for the position of the image (*I*) formed by the first surface in the absence of the second,

$$\frac{n_1}{o} + \frac{n_2}{i} = \frac{n_2 - n_1}{r_1} = \frac{n_2}{f_1} \quad , \tag{2.16}$$

and, for the second image (*I*),

$$\frac{n_2}{o'} + \frac{n_1}{i'} = \frac{n_1 - n_2}{r_2} = \frac{n_1}{f_2} \quad . \tag{2.17}$$

The position of the first image is

$$i = \frac{n_2 o f_1}{n_2 o - n_1 f_1} \quad . \tag{2.18}$$

Now let's introduce the *thin lens approximation:* assume that the lens thickness *d* is much smaller than any of the object or image distances. Then the magnitude of the second object distance o' is equal to that of the first image distance *i*. Depending upon the sign of f_1, the first image will be virtual or real. If virtual ($i < 0$), the image lies to the left of the second surface, the normal situation for an object, and to take account of the sign of *i* we must write $o' = -i$. If on the other hand the image is real ($i > 0$), it lies to the right of the second surface; light has not reached a focus before the second surface is encountered. In this case the first image would be a *virtual object* to the second surface, for which its distance to the apex, o', needs to be entered in Equation 2.17 as a negative number; since *i* is positive, $o' = -i$ once again. Equation 2.17 becomes

$$-\frac{n_2}{i} + \frac{n_1}{i'} = \frac{n_1 - n_2}{r_2} \quad , \tag{2.19}$$

which according to Equation 2.18 can be written as

$$\frac{n_1}{o} - \frac{n_2 - n_1}{r_1} + \frac{n_1}{i'} = \frac{n_1 - n_2}{r_2} \quad , \tag{2.20}$$

or

$$\frac{n_1}{o} + \frac{n_1}{i'} = (n_2 - n_1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) \quad . \tag{2.21}$$

We can find the focal length of the system, by taking $o \to \infty$. The reciprocal, however, has a simpler form:

$$\frac{1}{f} = \lim_{o\to\infty} \frac{1}{i'} = \frac{(n_2 - n_1)}{n_1}\left(\frac{1}{r_1} - \frac{1}{r_2}\right) \quad . \tag{2.22}$$

With $n_1 = 1$ (the refractive index of vacuum), this equation is called the *lensmakers' formula*. With this definition of focal length, the thin lens with two spherical surfaces focuses or diverges light according to

$$\frac{1}{o} + \frac{1}{i} = \frac{1}{f} \quad , \tag{2.23}$$

the same expression as we obtained for spherical mirrors in the paraxial approximation. In its present incarnation this is called the *thin-lens equation*. It is, of course, the same as the thin mirror equation, and since we have explicitly used the sign conventions we had before (Equations 2.12), those still apply as well.

2.4 Use of the thin lens equation

The process of using the image formed by one optical element as the object for the next is the fundamental problem-solving method offered in the paraxial approximation, and can be used for any sequence of mirrors and lenses, as follows.

Example 2.1
An 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 concave mirror of focal length f_2 separated from the lens by a distance $2(f_1 + f_2)$. Find the location of the final image.

The first object distance is $o_1 = 2f_1$, so by solving Equation 3.8 we find the first image at

$$i_1 = \frac{o_1 f_1}{o_1 - f_1} = 2 f_1 \quad .$$

This image is now $o_2 = 2f_2$ in front of the mirror, so

$$i_2 = \frac{o_2 f_2}{o_2 - f_2} = 2 f_2 \quad .$$

This image is on the same side of the mirror as the first image; in fact, it's in the same place. Now this second image is $o_3 = 2f_1$ from the lens and the rays are heading toward the lens again:

$$i_3 = \frac{o_3 f_1}{o_3 - f_1} = 2 f_1 \quad .$$

This is the final image, and it winds up at the same location along the axis as the original object.

Example 2.2
Consider two thin lenses in vacuum, with focal lengths f_1 *and* f_2*, separated by a distance* d. *Obtain expressions for the position of the image for arbitrary object distance, and for the effective focal length of the combination.*

For an object at distance *o*, the first lens would produce an image at distance *i*, given by solution to thin-lens equation:

$$i = \frac{o f_1}{o - f_1} \quad .$$

The situation and status of the first image/second object is the same as we encountered above with the thick lens, and the result is the same: no matter which side of the second lens the first image falls, the second object distance amounts to $d - i$, so the final image distance is given by

$$i' = \frac{o' f_2}{o' - f_2} = \frac{\left(d - \frac{o f_1}{o - f_1}\right) f_2}{d - \frac{o f_1}{o - f_1} - f_2} = \frac{\left(o\left(f_1 - d\right) + d f_1\right) f_2}{o\left(f_1 + f_2 - d\right) + d f_1 - f_1 f_2} \quad .$$

From this one may obtain the effective focal length by moving the object to large distances:

$$f = \lim_{o \to \infty} i' = \frac{\left(f_1 - d\right) f_2}{f_1 + f_2 - d} \quad . \tag{2.24}$$

Since it is a limiting value of the final object distance, this effective focal length should be measured from the second lens. Following directly from this result is the useful case of two thin lenses in contact ($d = 0$):

$$f = \frac{f_1 f_2}{f_1 + f_2} \quad , \tag{2.25}$$

which can be used to expand considerably the choices of focal lengths available from a limited collection of thin lenses.

2.5 Off-axis objects and images

So far we have considered only point objects lying on the optical axis. Suppose we have instead an extended object; what does its image look like? The geometry of this situation is easy to sketch if you apply two simple facts about lenses, mirrors and their foci, that are implicit in our discussion above.

1. *Rays incident parallel to the optical axis are refracted or reflected through a focus; rays incident through a focus emerge parallel to the optical axis.* This is another way to phrase the definition of the focal point; rays from a very distant object lying along the optical axis all intersect at the focus, and the process works in reverse as well.

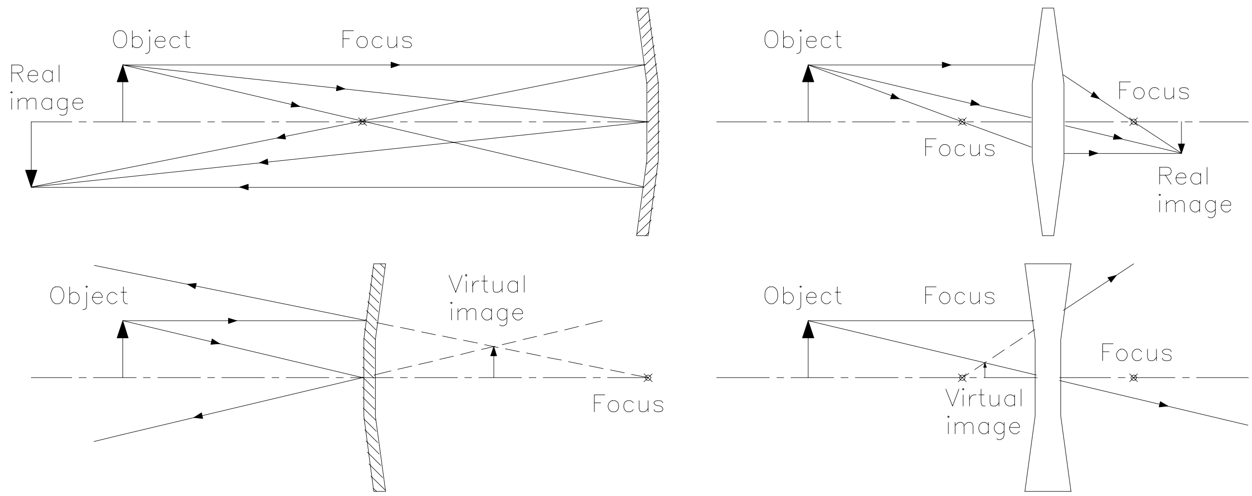

Figure 2.7: off-axis images. Mirrors are on the left and lenses are on the right; positive focal lengths are illustrated at the top and negative ones underneath.

2. *Rays incident on the center of the mirror reflect with the same angle with respect to the optical axis that they had before; rays incident on the center of a lens pass through without deviation*. The first part of this statement owes its truth to the fact that the tangent to the mirror's surface is perpendicular to the optical axis in the center. In the case of the lens, the front and back surfaces are parallel at the center, so since the lens is thin the ray is neither deviated nor displaced.

These considerations provide two or three rays to draw that can determine the position and orientation of an image; two are sufficient. The way it works is to start rays from one off-axis point on the object, and send them through the lens in the ways prescribed; they will intersect at the corresponding point on the image, for example, as shown in Figure 2.7.

From the second of these simple considerations above also follow two trivial results of importance for optical elements: the *lateral magnification* of images and objects at finite distances, and the *plate scale* for objects at very large distances. Consider first the object, image and lens in Figure 2.7. Since rays incident on the center of the lens pass through undeviated (i.e. its angle θ with respect to the axis is the same on both sides of the lens), the transverse extent of object and image are related to the object and image distance by:

$$\frac{h_o}{o} = \tan\theta = \frac{h_i}{i} \quad , \qquad \text{or} \quad \frac{h_i}{h_o} = \frac{i}{o} \quad . \tag{2.26}$$

Real objects and their real images have positive object and image distances (as in Figure 2.8) and are inverted with respect to each other, while virtual images of real objects have the same orientation. The lateral magnification m is thus defined with a minus sign:

$$m \equiv -\frac{h_i}{h_o} = -\frac{i}{o} \quad ; \tag{2.27}$$

$m < 0$ if the image is inverted, and $m > 0$ if it is erect.

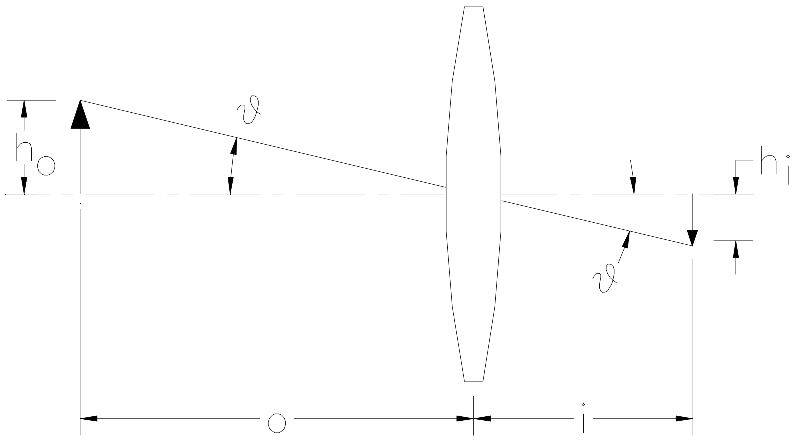

Figure 2.8*:* lateral magnification.

Example 2.3
A luminous object and a screen are a fixed distance L apart.

a. *Show that a converging lens of focal length* f *will form a real image on the screen for two positions of the lens, that are separated by*

$$\ell = \sqrt{L(L-4f)} \ .$$

(This comprises a nice way to determine the focal length of a converging lens; solve the expression for f*, and insert measurements of* ℓ *and* L.*)*

For the image and object to be a distance L apart, we need $i + o = L$. We also have

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

or

$$i = \frac{of}{o-f} = L - o \ .$$

Multiplying through and solving for o, we get

$$o^2 - Lo + Lf = 0$$
$$o = \frac{L}{2} \pm \frac{1}{2}\sqrt{L^2 - 4fL} \ .$$

The distance between the two lens positions is the difference of these two values of the object distance:

$$\begin{aligned}\ell &= \frac{L}{2}+\frac{1}{2}\sqrt{L^2-4fL} - \frac{L}{2}+\frac{1}{2}\sqrt{L^2-4fL} \\ &= \sqrt{L^2-4fL} = \sqrt{L(L-4f)}\end{aligned} \ .$$

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

$$\left(\frac{L-\ell}{L+\ell}\right)^2 \ .$$

For the first object position,

$$o_1 = \frac{L}{2}+\frac{1}{2}\sqrt{L^2-4fL} = \frac{L+\ell}{2} \ ,$$

the image distance is

$$i_1 = L - o_1 = \frac{L-\ell}{2} = o_2 \quad .$$

Similarly, $i_2 = o_1$. The ratio of magnifications is therefore

$$\frac{m_1}{m_2} = \frac{i_1}{o_1} \times \frac{o_2}{i_2} = \frac{o_2^2}{o_1^2} = \left(\frac{L-\ell}{L+\ell}\right)^2 \quad .$$

Example 2.4
A short linear object of length d lies along the axis of a spherical mirror of focal length f, a distance o from the mirror.

a. *Show that its image will have length d', where*

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

Since the object and image are presumed to be small compared to o, i and f, we can write

$$d' = \Delta i = -\frac{di}{do}\Delta o = -\frac{di}{do}d \quad ,$$

where the minus sign represents the fact that the near (far) end of the object corresponds to the far (near) end of the image. Since the thin-lens equation can be expressed as $i = of/(o-f)$,

$$\frac{di}{do} = \frac{d}{do}\left(\frac{of}{o-f}\right) = \frac{f}{o-f} - \frac{of}{(o-f)^2} = -\left(\frac{f}{o-f}\right)^2 \quad ,$$

or

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

b. *Show that the* longitudinal *magnification, $m' = d'/d$, is equal to m^2, where m is the* lateral *magnification (in Equation 2.27).*

The lateral magnification m is given by

$$m = -\frac{i}{o} = -\frac{1}{o}\left(\frac{of}{o-f}\right) = -\frac{f}{o-f} \quad ,$$

so

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

c. *Is there any condition such that, neglecting mirror defects, the image of a cube would also be a cube?*

For the image of a cube to look like a cube, the longitudinal and lateral magnification either have to be equal or opposite:

$$m = \pm m' = \pm m^2 \quad ,$$

or

$$m = \pm 1 \quad .$$

The lateral magnification, again, is

$$m = -\frac{f}{o - f} \quad ,$$

so the two possible values for lateral magnification correspond to two object distances:

$$o = f - \frac{f}{m} = 0 \quad \text{or} \quad 2f$$

for m = 1 and m = -1, respectively. With an object distance of $2f$, the image of a small cube looks like an inverted cube of the same size. The other object distance doesn't represent a practical situation. It corresponds to placement of an object at the center of a mirror, which violates the assumptions made in the derivation of the thin mirror/lens equation.