OPTICS

Optics, one of the oldest branches of physics, is the study of the properties of visible and near-visible light covering the spectrum of wavelengths from the far ultraviolet to the far infrared, a range of wavelengths from very roughly 1 nm (1 nm $=$ 1 nanometer $= 10^{-9}$ meter) to $1{,}000{,}000$ nm, as illustrated in Figure 0.1. Visible light is usually considered to have its spectrum from the violet of about 400 nm to the red of about 700 nm with the other colors between these boundaries (again, see Figure 0.1). Light is a part of the electromagnetic spectrum which is composed of radiation formed by combined electric and magnetic fields, a transverse wave motion because these fields are perpendicular to the direction of travel of the wave.

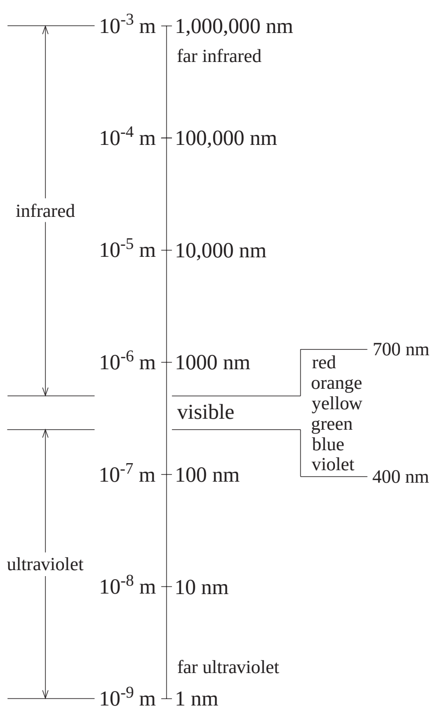


Figure 0.1

When Maxwell was formulating his theory of electricity and magnetism in the 1860s, he was able to show that his electric and magnetic fields obeyed a wave equation; this equation had the same mathematical form as the one for waves traveling on a string or for sound waves. The wave equation has an extremely useful coefficient that equals the square of the speed of travel of the waves. When Maxwell evaluated this coefficient in his wave equation, he obtained a value equal to the square of the speed of light, a speed that was already reasonably well known at that time. He made the inference that light was constructed of traveling electric and magnetic fields, or collectively, electromagnetic fields. Now it is known that light is a small part of an extremely large spectrum, the electromagnetic spectrum, of electromagnetic fields that travel with the speed of light and range from gamma to radio waves—from the very short wavelengths (10^{-13} m) to the very long (10^{10} m), a range of 23 factors of ten (and it is possible that the range is larger, since there is no theoretical limit at this time on how small or large the waves might be). The electromagnetic spectrum includes, of course, the spectrum of wavelengths shown in Figure 0.1.

The nature of light has been a matter of discussion since the time of the Greek philosophers starting in the 500s BC, and perhaps even earlier. However, although the properties of light have been studied for centuries, its exact nature still remains elusive. We mentioned before that light can be described as electromagnetic radiation that satisfies a wave equation; therefore, it's tempting to conclude that we always can model light as a wave. However, there are many experiments that show that light has particle-like properties, as well. To make a model of something that is both a wave and a particle is a difficult problem. One way out of this quandary is to say just because light satisfies a wave equation, that does not necessarily mean that light is a wave, but only a convenient mathematical way of describing it. As an example of a similar situation, consider the behavior of charge in an electronics circuit composed of a resistor, capacitor, and inductor in series; the behavior of charge in this circuit is described by a differential equation that has exactly the same form as one in mechanics that describes the motion of a mass in a viscous medium at the end of an elastic spring. Nevertheless, we do not say that therefore electricity and mechanics are the same.

But there is another problem: other waves, such as sound waves, require a medium in which to travel. In like manner, it was said that light waves traveled in a special medium called the ether. Also, the more rigid the medium in which sound waves travel, the greater the speed of travel. But no pulse of any kind travels faster than a light pulse in a vacuum, a speed that is measured to be very close to $c = 3 \times 10^8$ m/s. This value is many times greater than the speed of sound, which means that the ether is an extremely rigid medium. During the latter part of the 19th century and the early part of the 20th, great effort was expended in searching for the ether; however, no evidence of its existence was found. Thus, the concept of the ether was finally discarded, and light is now thought not to need anything in which to travel—in fact, it travels best through a region filled with nothing, a true vacuum.

And there is more: It is a fact that the speed of a light pulse in a vacuum is always measured as the constant c by an observer, independent of the motion of the source or the observer. In other words, no matter how fast the source moves when a light pulse is emitted, an observer measures the speed of the light pulse as c; also, no matter how fast the observer moves, the observer still measures c. For example, if the observer travels at approximately the speed of light in one direction, and a pulse of light travels in the opposite direction, the observer still measures the speed of the pulse as it passes by as c. This concept violates our common-sense ideas of relative velocities; nevertheless, it is in agreement with experiment. This property of the speed of light is one of the basic tenets in Einstein's special theory of relativity.

These properties of light are very difficult to model into something that agrees with our normal experience—a bit of energy that has a both a wave and a particle nature, and always has the speed c in a vacuum. But, let's try to make a simple model to help our thinking. We start with a picture of an ideal wave; namely, one that goes on forever from minus infinity to plus infinity. Now for light let's imagine that it is just a small part of such a wave so that it is finite and small enough in extent to approximate a particle. Such a piece of a wave we call a photon. Whenever we measure the speed of a photon in a vacuum, we obtain the value c. Now we have a simple model of light that reminds us that light has both a wave and a particle nature, and although this model is oversimplified, it can act as a starting point.

PART I
GEOMETRICAL OPTICS

It is convenient to divide the study of optics into parts, each part depending on the complexity of the model used to represent the properties of light. Because the properties of geometrical optics are described by the simplest model, we begin our study with that topic. In this model, light travels in straight lines through uniform media (or more precisely, homogeneous and isotropic media), and obeys the laws of reflection and refraction at surfaces between different media. With these simple concepts, we can work out the basic properties of many optical systems.

The earliest optical instruments were mirrors made of polished metal. Several mirrors, some in excellent condition, have been recovered that were used in ancient Egypt (ca. 1900 BC). Mirrors are mentioned in the Bible (ca. 1200 BC) in the building of the tabernacle; Exodus 38:8 in the New International Version reads: "They made the bronze basin and its bronze stand from the mirrors of the women who served at the entrance to the Tent of Meeting." Lenses, called burning glasses, were known to the Greek philosophers (ca. 400 BC), and the refraction effect of the apparent bending of objects when they were partly immersed in water was mentioned by Plato at about the same time. By 300 BC it was noted that a ring at the bottom of an empty vessel, when hidden by the lip of the vessel, became visible after the vessel was filled with water. The Romans also knew about burning glasses by the first century BC, and had observed that a glass globe filled with water would produce magnification.

Although other cultures in early times had similar social orders, it was in Greece that speculative thinking turned into rational thought, philosophy, and science by the 500s BC. This transformation did not happen quickly, but took place over a period of several centuries as the Greek city-states interacted in trade with the civilizations of Mesopotamia and Egypt that brought an acquaintance with their astronomy, medicine, and technology. The idea that light travels in straight lines (rectilinear propagation) was taught by Plato about 400 BC, and the equality of the angles of incidence and reflection in what came to be called the law of reflection appears in the works of Euclid as early as 300 BC.

Although Greek science largely ceased to develop after approximately 100 BC, Hero of Alexandria and Ptolemy (also of Alexandria) were exceptions. Just when Hero lived is open to question; sources give dates that range between 150 BC and 250 AD. However, because Hero mentions an eclipse in his writings that is known to have occurred in 62 AD, current thinking is that he lived in the first century AD. Hero is important because he attempted to derive the law of reflection from a principle that stated that reflected light traveled the shortest distance between two points. Although, the law does follow from this principle for a number of reflective surfaces, later workers showed that it was not true in general.

Ptolemy (100–165 AD) is significant because of his work on refraction. The discovery of the law of refraction took many centuries. But Ptolemy made an important beginning; he not only measured angles of incidence and refraction for rays of light, but also arranged them into tables. Examining his tables, he concluded that the angle of refraction was proportional to the angle of incidence, which we now know to be true only when the angles are small.

During this time period, the Roman empire ruled the western world. However, the Romans did little to advance science; they excelled at law, government, and warfare—in science they primarily collected what the Greeks had learned. The Roman empire declined and basically ceased to exist by 475 AD; then little happened in science until the Arabian empire became important by approximately 700 AD. The Arabian scholars translated and absorbed the then-known scientific knowledge, but did little original work, except in optics where Alhazen (Al hah ZAHN) in about 1000 AD made important contributions. He studied reflection and, like those before him, concluded that the angles of incidence and reflection were indeed equal, but also added the concept that these angles lay in the same plane normal to the interface—a plane that is now called the plane of incidence. He studied the law of refraction, and repeated the measurements made by Ptolemy. However, he correctly concluded that Ptolemy was in error when the angles were larger, and that the angle of refraction was not in general proportional to the angle of incidence. But, he failed to discover the true law of refraction.

Not until the 1200s did workers in Europe show interest in science and begin to assimilate the scientific concepts recorded by Arabian scholars; in particular, the writings of Alhazen on optics had great influence. By the 1300s, paintings showed monks wearing eyeglasses, and in the 1500s combinations of converging and diverging lenses were described. Then in the early 1600s the telescope and the microscope were invented. Although contested, the evidence indicates that the first telescope was made by a spectacle-maker, Hans Lippershey, of Holland in 1608. The first microscope was constructed at approximately the same time, most likely by another spectacle-maker, Zacharias Janssen, also in Holland.

With the invention of these instruments, the search for the law of refraction was renewed. By 1611, the German astronomer Kepler, had discovered total internal reflection, and had obtained the law of refraction for small angles, which was basically what Ptolemy and Alhazen had determined centuries before. The true form of the law of refraction is usually credited to Willebrord Snell some years later (some sources say 1621), who was a professor of mechanics at the University of Leyden in Holland, although he never published his discovery. It appears that he found the law experimentally, and expressed it in terms of cosecants of the angles rather than in the modern form using sines. In 1637, René Descartes of France, published the law using sines of the angles. He obtained the law theoretically, although some of his hypotheses were tenuous. In fact, the French mathematician Pierre de Fermat and others were quick to find fault with the hypotheses of Descartes, and in 1657 Fermat deduced the law of refraction using the assumption that light travels from a point in one medium to a point in another medium in the least time, a principle that came to be called Fermat's principle of least time. Years later, it was found that this principle could be derived from Maxwell's equations of electromagnetism. In current literature, the name assigned to the law of refraction is under some contention: In many countries it is called either Snell's law or simply the law of refraction; however, in France it is known as Descartes' law. We will opt on the side of clarity in this textbook, and call it the law of refraction.

In the main we will study the consequences of the law of refraction; we will also touch on the law of reflection, but most of our time will be spent on refraction. The law of refraction works for any surface, but because of the relative ease with which spherical surfaces are made, it is these surfaces that will be our primary concern (plane surfaces are also included, since a plane surface may be regarded as a spherical surface of infinite radius). Even though the law of refraction is simple to state, when rays are traced through spherical surfaces, many of the equations become quite complicated. Therefore, we will follow common practice and obtain approximate results using paraxial rays, rays that travel close to the symmetry axis and make small angles with normals to the surfaces, the same rays that Ptolemy, Alhazen, and Kepler had observed to have simple properties. This procedure will allow us to obtain the characteristics of the lens system under study reasonably quickly as a first approximation; then the details will be obtained by tracing the rays through the surfaces exactly without approximation.

Ray tracing, even with paraxial rays, can be very calculation intensive. In the days before electronic computers, the calculations were done by hand with pencil and paper using large tables of logarithms. To minimize the time spent in calculation, great effort was given to obtaining equations that would reduce the number of arithmetic operations, arranging the work in tabular form so that numerical values could be used efficiently, and in developing techniques for error checking. A skillful worker could use these techniques to trace a ray through a typical lens system in about five minutes. The same kind of ray tracing is now done with the modern computer in less then a second; even with a programmable scientific calculator it is done quickly. In many optics textbooks, because of these computation difficulties, results are frequently given without derivation. With the wide availability of personal computers, the tools are now in place such that most results can be obtained by everyone, even those that are calculation intensive. We shall usually avoid giving any programs in this textbook, because of the wide variety of programming languages now available. Instead, we shall try to provide guidelines in the form of flowcharts, or algorithms, to aid program writing in whatever programming language is chosen. These flowcharts will have the added advantage of bringing together and summarizing equations derived over many pages. Most of the programs are not long, and to write them is extremely informative, instructive, and helpful in understanding concepts.

Because many of the results are best understood by drawing a graph, a software language that has graphical capability is advantageous. Symbolic algebra software, such as *Mathematica*, is made to order for much of the work we shall need to perform. Such programs can perform calculations to high precision, draw graphs in two and three dimensions, and furthermore, manipulate expressions symbolically that arise when the rules of algebra or calculus are applied. The examples and graphs in this textbook are performed with *Mathematica* on a Macintosh computer. In the numerical calculations, the printed results will usually be rounded to an appropriate number of significant figures, but the calculations will be carried out to the default precision of *Mathematica*, which is normally 16 decimal digits.

In earlier works on ray tracing in lens systems, length units of inches were customarily used in English speaking countries. Now, metric units or no units at all are used. Expressing lengths with no units is popular because, as we shall show in the work to follow, the dimensions of length have the property of scaling linearly in lens systems. However, when no units are expressed, it is difficult for the beginner to tell if changes in a length quantity are big or small. Therefore, in this textbook, as an aid to understanding and comprehension, we shall usually state the length units in terms of millimeters (abbreviated as mm).

Chapter 1—Outline

Paraxial Optics—Refraction & Translation

Chapter 1

Paraxial Optics—Refraction & Translation

1.1 Introduction

1.1.1 Basic laws

Our model for light is composed of tiny bits of energy called photons which have both a wave and a particle nature; a diagram that might represent a photon is drawn in Figure 1.1. The crest-to-crest distance λ is called the wavelength, which our eyes interpret as color in the visible region. A photon travels with a very high speed: in a vacuum its speed is very close to $c = 3 \times 10^8$ m/s; in other media, the speed is smaller and also varies with wavelength. In the visible region, the wavelength ranges from the violet ($\lambda \approx 400$ nm) to the red ($\lambda \approx 700$ nm), where 1 nm = 1 nanometer = 10^{-9} meter. Another unit which was frequently used in the past, but is now no longer recommended, is the Angstrom (10 A = 1 nm).

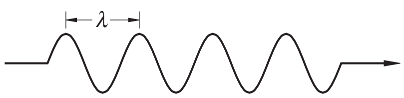


Figure 1.1

Figure 1.2

In geometrical optics, $\lambda <<$ (the dimensions of objects and openings), and we model photons traveling in a given rectilinear direction by simply drawing a ray, as shown in Figure 1.2. In this chapter, and several of the chapters to follow, we want to investigate how these rays interact with surfaces that represent boundaries between two different media where each medium is both homogeneous and isotropic. Such a medium's most important optical property is its index of refraction n which is defined to be

$$n = \frac{c}{v} \tag{1.1}$$

where c is the speed of photons in a vacuum, and v is the speed in the medium. For a medium that is a vacuum, this definition gives $n = 1$ exactly. A medium is homogeneous when it is the same throughout. So a crystal with its atoms (or molecules) in their lattice sites is homogeneous. But a crystal may have certain directions where photons travel at different speeds and therefore have different indices of refraction. A medium where all the atoms (or molecules) are mixed up—like in a liquid or glass—so there are no preferred directions is called isotropic. Thus, in a medium that is both homogeneous and isotropic, the speed of a photon is constant in all directions and the index of refraction is constant (for a given wavelength) everywhere in the medium. Then, we have

The law of rectilinear propagation. In a homogeneous and isotropic medium, light travels in a straight line at a constant speed.

In Figure 1.3 we illustrate what happens when a ray—the incident ray—falls on a surface that separates two homogeneous, isotropic media. In general, the incident ray separates into two rays: one is reflected and obeys the law of reflection, the other is refracted and obeys the law of refraction. Even when the second medium (the medium containing the refracted ray) is opaque, there is still a refracted ray; it is simply absorbed after it travels a short distance. However, in our study we are primarily interested in transparent media where most of the incident ray goes into the refracted ray, very little is reflected or absorbed.

The law of reflection. This law has two parts (see Figure 1.3):

1) $\theta'' = \theta$
2) the reflected ray lies in the plane of incidence

where θ is the angle of incidence, θ'' is the angle of reflection, and the plane of incidence is the plane that contains the incident ray and the normal to the surface.

The law of refraction (Snell's law). This law also has two parts (see Figure 1.3):

1) $n \sin\theta = n' \sin\theta'$ (1.2)
2) the refracted ray lies in the plane of incidence

where θ' is the angle of refraction, n is the index of refraction of the medium in which the incident ray is located, and n' is the index of refraction of the medium in which the refracted ray is located. Equation 1.2 is an extremely important equation, and many people searched for it over a period of many centuries; only after its discovery did lens design begin.

It is important to note that in both laws the angles are measured relative to the normal, not the surface.

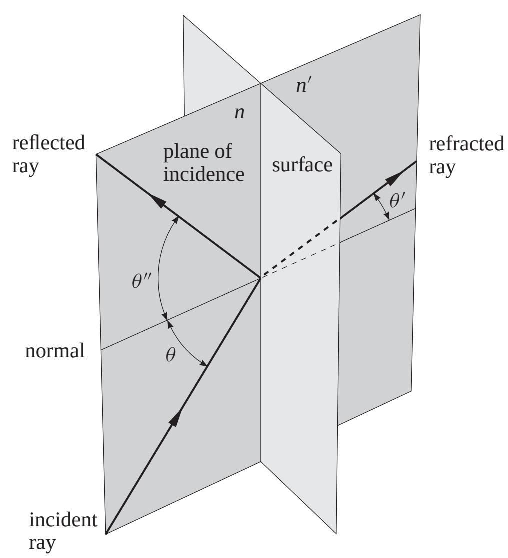


Figure 1.3

For a medium of air at 1 atm pressure and 0°C temperature, the value of n is 1.00029 measured to six significant figures, a value so close to one that we customarily say the n for air is unity, like that for a vacuum. For some other media, $n = 1.50$ for common glass, $n = 2.42$ for diamond, and $n = 1.33$ for water. Because the speed of a photon in a medium that is not a vacuum can vary with wavelength, the index of refraction varies with wavelength; it is usually a rather small variation. For solids and liquids, the variation occurs in the third or fourth significant figure; for a gas at normal pressures, such as air, the variation does not occur until the seventh significant figure.

1.1.2 Angles

The simplest and most effective definition of an angle θ is

$$\theta = \frac{s}{r} \tag{1.3}$$

where s is the arc length, and r is the radius, as shown in Figure 1.4. This definition is consistent with the relationship for small angles, $\sin\theta \approx \theta$, when $\sin\theta$ is given by the triangle definition. Because s and r have common length units, the angle θ is dimensionless. Thus, we might think that θ should have no units; however, an angle may be defined in other ways. To prevent confusion, the radian (abbreviated rad) unit is used when angles are defined by Equation 1.3. But the radian unit is not a physical unit like the meter, second, or kilogram; instead, it is called a supplementary unit—it helps to distinguish angles in radians from angles determined in some other way, say degrees. Because the radian is dimensionless, it may be dropped or erased in an equation when it is not needed.

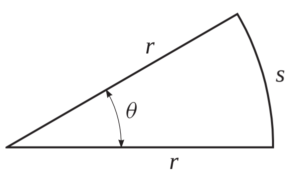


Figure 1.4

The degree is the more convenient unit for easier visualization. To obtain the conversion between the degree and the radian, consider the semicircle shown in Figure 1.5. Because the circumference of a circle of radius r is $2\pi r$, the arc length of a semicircle is πr, and Equation 1.3 gives

$$\theta = \frac{s}{r} = \frac{\pi r}{r} = \pi \text{ rad}$$

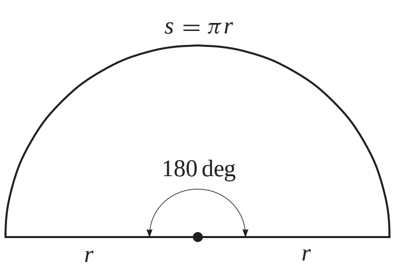


Figure 1.5

Because the angle of a semicircle is 180 deg, we have the desired conversion

$$\pi \text{ rad} = 180 \text{ deg} \tag{1.4}$$

where we have written the degree unit as the deg, rather than °, to make it easier to use when units are treated as fractions in converting units from one form to another.

1.2 Matrices and Determinants

Matrices and determinants are a shorthand representation of the numbers in a problem written as a rectangular array in rows and columns. Some evidence exists which shows that some aspects of matrices were used in the second century BC and perhaps even as far back as the fourth century BC. But the mathematician Arthur Cayley in 1858 gave the first abstract definition of a matrix, and showed how arrays applied to problems at that time were special cases of his more general concept. He developed an algebra of matrices, and in subsequent years many applications were found not only in mathematics, but also in physics and engineering. Matrices were first used in optics by Sampson in 1913, and their use was also suggested by the work of T. Smith in the 1930s, but not until an article by Halback in the *American Journal of Physics* on matrix representation in optics in the early 1960s, did their use become common in textbooks on geometrical optics.

A determinant is a less general concept than a matrix: it takes the numbers of a square matrix and calculates a single number by following a specific procedure. In our matrix development of paraxial optics, determinants will have the useful property that their value is always unity; this feature is important both theoretically and numerically.

In general, a matrix A is defined to be an array of numbers, called matrix elements, arranged in rows and columns; as an example

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \tag{1.5}$$

is called a 2×3 matrix because it has 2 rows and 3 columns. The matrix elements have numerical subscripts for easy reference; the first subscript refers to the row, and the second to the column—for example, a_{12} means that this element is in the first row and second column. A square matrix has as many rows as columns; the following example illustrates a 2×2 matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{1.6}$$

When a matrix A is a square 2×2 matrix, then its determinant $|A|$ is defined quite simply as

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \tag{1.7}$$

We say the elements a_{11} and a_{22} lie along the principal diagonal; the other two elements, a_{12} and a_{21}, we call the off-diagonal elements.

As we mentioned, an entire algebra for matrices exists, but because of the space requirement to present it, we will only introduce those parts of the algebra that we need. We will also simply illustrate some of the laws by example rather than to prove them rigorously. In paraxial optics, the multiplication of 2×2 square matrices occurs frequently, so we give the definition for multiplying two such matrices:

$$\begin{aligned} AB &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix} \end{aligned} \tag{1.8}$$

To remember the steps of matrix multiplication, observe that the rows of the first matrix are multiplied and added into the columns of the second. We next illustrate matrix multiplication with examples, and show at the same time some of the laws that matrix multiplication obeys. We sometimes will include more steps than usual to aid understanding.

Example 1.2.1 Performing matrix multiplication and illustrating the commutative law.

$$\begin{aligned} AB &= \begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 5 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 3 \cdot 1 + 1 \cdot 5 & 3 \cdot 4 + 1 \cdot 2 \\ 2 \cdot 1 + 5 \cdot 5 & 2 \cdot 4 + 5 \cdot 2 \end{pmatrix} \\ &= \begin{pmatrix} 8 & 14 \\ 27 & 18 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} BA &= \begin{pmatrix} 1 & 4 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix} \\ &= \begin{pmatrix} 1 \cdot 3 + 4 \cdot 2 & 1 \cdot 1 + 4 \cdot 5 \\ 5 \cdot 3 + 2 \cdot 2 & 5 \cdot 1 + 2 \cdot 5 \end{pmatrix} \\ &= \begin{pmatrix} 11 & 21 \\ 19 & 15 \end{pmatrix} \end{aligned}$$

This example illustrates the very important result, $AB \neq BA$; that is, matrix multiplication does not, in general, obey the commutative law.

Example 1.2.2 Illustrating the associative law.

$$\begin{aligned} A(BC) &= \begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix} \left(\begin{pmatrix} 1 & 4 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} 6 & 7 \\ 8 & 9 \end{pmatrix} \right) \\ &= \begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} 38 & 43 \\ 46 & 53 \end{pmatrix} = \begin{pmatrix} 160 & 182 \\ 306 & 351 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} (AB)C &= \left(\begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 5 & 2 \end{pmatrix} \right) \begin{pmatrix} 6 & 7 \\ 8 & 9 \end{pmatrix} \\ &= \begin{pmatrix} 8 & 14 \\ 27 & 18 \end{pmatrix} \begin{pmatrix} 6 & 7 \\ 8 & 9 \end{pmatrix} = \begin{pmatrix} 160 & 182 \\ 306 & 351 \end{pmatrix} \end{aligned}$$

We observe in Example 1.2.2 that $A(BC) = (AB)C$, a relationship called the associative law; in fact, matrix multiplication always obeys this law. Because the associative law says that it does not matter how the matrices are grouped together for multiplication, it is acceptable to write the multiplication of three matrices simply as ABC. This law can be extended to the multiplication of any number of matrices, not just three. However, as Example 1.2.1 shows, the order in which the matrices are written must in general not be changed.

Another matrix of importance is a 2×1 rectangular matrix, denoted by

$$\begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix}$$

which is also called a column matrix or a column vector. Matrix multiplication of a 2×2 square matrix and a 2×1 column vector is defined as

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} \\ a_{21}b_{11} + a_{22}b_{21} \end{pmatrix} \tag{1.9}$$

a result that is another column vector. However, a column vector times a square matrix—the reverse order of the multiplication in Equation 1.9—is not defined.

Example 1.2.3 Multiplication of a square matrix and a column vector.

$$\begin{pmatrix} 5 & 8 \\ 6 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 7 \end{pmatrix} = \begin{pmatrix} 71 \\ 32 \end{pmatrix}$$

A matrix with special properties is the identity or unit matrix I; for 2×2 matrices, it is defined as

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{1.10}$$

One of its special properties is that it commutes with all 2×2 matrices:

$$AI = IA = A \tag{1.11}$$

because

$$AI = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = A$$

$$IA = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = A$$

An identity matrix can be defined for other square matrices as well, not just 2×2 matrices, and Equation 1.11 is also true for these matrices.

The identity matrix I is used to define an inverse or reciprocal A^{-1} of the matrix A by writing

$$AA^{-1} = A^{-1}A = I \tag{1.12}$$

that is, A^{-1} is that matrix which when multiplied by A gives I. The properties of a matrix are such that if $AA^{-1} = I$, then it can be shown that $A^{-1}A = I$ must also hold.

The inverse of a 2×2 matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

has the compact and useful form

$$A^{-1} = \begin{pmatrix} \dfrac{a_{22}}{|A|} & -\dfrac{a_{12}}{|A|} \\ -\dfrac{a_{21}}{|A|} & \dfrac{a_{11}}{|A|} \end{pmatrix} \tag{1.13}$$

where $|A|$ is the determinant of A, namely,

$$|A| = a_{11}a_{22} - a_{12}a_{21}$$

As we mentioned before, the determinants of the 2×2 matrices that we use in paraxial optics have the value of one; thus, $|A| = 1$ in Equation 1.13, and we have the simple inverse

$$A^{-1} = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \tag{1.14}$$

In words, the diagonal elements are interchanged, and the off-diagonal elements have their signs changed.

Methods to calculate inverses are found in the literature. Since the use of matrices has increased, routines for the numerical calculation of inverses are included on many scientific pocket calculators. Symbolic algebra systems, such as *Mathematica*, can determine inverses either symbolically or numerically.

Example 1.2.4 The inverse of a matrix.
Suppose we start with the simple matrix

$$A = \begin{pmatrix} 2 & 5 \\ -1 & 1 \end{pmatrix}$$

Then the determinant of this matrix

$$|A| = 7$$

and Equation 1.13 give the inverse of A as

$$A^{-1} = \begin{pmatrix} \frac{1}{7} & -\frac{5}{7} \\ \frac{1}{7} & \frac{2}{7} \end{pmatrix}$$

We can verify Equation 1.12 with matrix multiplication:

$$AA^{-1} = \begin{pmatrix} 2 & 5 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{7} & -\frac{5}{7} \\ \frac{1}{7} & \frac{2}{7} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$A^{-1}A = \begin{pmatrix} \frac{1}{7} & -\frac{5}{7} \\ \frac{1}{7} & \frac{2}{7} \end{pmatrix} \begin{pmatrix} 2 & 5 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Because a determinant represents (or equals) a number, determinants obey laws similar to the laws numbers obey. We evaluate the determinants with Equation 1.7.

Example 1.2.5 Multiplication of determinants.
Unlike matrices, determinants obey the commutative law; we use the matrices in Example 1.2.1:

$$|A|\,|B| = \begin{vmatrix} 3 & 1 \\ 2 & 5 \end{vmatrix} \begin{vmatrix} 1 & 4 \\ 5 & 2 \end{vmatrix} = (13)(-18) = -234$$

$$|B|\,|A| = \begin{vmatrix} 1 & 4 \\ 5 & 2 \end{vmatrix} \begin{vmatrix} 3 & 1 \\ 2 & 5 \end{vmatrix} = (-18)(13) = -234$$

Even when we evaluate the matrix product first before we calculate the determinant, we still get the same result:

$$|AB| = \begin{vmatrix} 8 & 14 \\ 27 & 18 \end{vmatrix} = -234$$

$$|BA| = \begin{vmatrix} 11 & 21 \\ 19 & 15 \end{vmatrix} = -234$$

Determinants obey the associative law; we use the matrices in Example 1.2.2:

$$|A|\,(\,|B|\,|C|\,) = \begin{vmatrix} 3 & 1 \\ 2 & 5 \end{vmatrix} \left(\begin{vmatrix} 1 & 4 \\ 5 & 2 \end{vmatrix} \begin{vmatrix} 6 & 7 \\ 8 & 9 \end{vmatrix} \right)$$

$$= (13)\,((-18)(-2)) = 468$$

$$(\,|A|\,|B|\,)\,|C| = \left(\begin{vmatrix} 3 & 1 \\ 2 & 5 \end{vmatrix} \begin{vmatrix} 1 & 4 \\ 5 & 2 \end{vmatrix} \right) \begin{vmatrix} 6 & 7 \\ 8 & 9 \end{vmatrix}$$

$$= ((13)(-18))\,(-2) = 468$$

$$|A|\,|B|\,|C| = \begin{vmatrix} 3 & 1 \\ 2 & 5 \end{vmatrix} \begin{vmatrix} 1 & 4 \\ 5 & 2 \end{vmatrix} \begin{vmatrix} 6 & 7 \\ 8 & 9 \end{vmatrix}$$

$$= (13)(-18)(-2) = 468$$

No matter how we combine the matrices before we calculate the determinants, we get the same result:

$$|A||BC| = \begin{vmatrix} 3 & 1 \\ 2 & 5 \end{vmatrix} \begin{vmatrix} 38 & 43 \\ 46 & 53 \end{vmatrix}$$

$$= (13)(36) = 468$$

$$|AB||C| = \begin{vmatrix} 8 & 14 \\ 27 & 18 \end{vmatrix} \begin{vmatrix} 6 & 7 \\ 8 & 9 \end{vmatrix}$$

$$= (-234)(-2) = 468$$

$$|ABC| = \begin{vmatrix} 160 & 182 \\ 306 & 351 \end{vmatrix} = 468$$

$$|CBA| = \begin{vmatrix} 199 & 231 \\ 259 & 303 \end{vmatrix} = 468$$

1.3 The Paraxial Approximation and Its Matrix Representation—Paraxial Mathematics

1.3.1 Introduction

In this chapter, an optical system means a system of one or more refracting surfaces composed of sections of spherical surfaces with all their centers of curvature on the same axis (the z axis), and with the sections centered on this axis—a centered optical system. The outline of a spherical section is shown in Figure 1.6. Not only do spherical surfaces have important physical properties when they interact with light, but mathematically they are the simplest curved surface to describe, requiring only a center and a radius (the radius of curvature). To describe the nature of the spherical section more completely, we imagine the arc shown in Figure 1.6 to rotate about the z axis to obtain a surface of revolution. The z axis has its positive direction to the right, and because it is such an important reference axis, it is given special names such as the symmetry axis, the optical axis, or the principal axis. Since the refracting surfaces are symmetrically placed with reference to the z axis, this property suggests the symmetry axis as the most descriptive name. The plane determined by the y and z axes is called the meridional or tangential plane. The position of the axis names, here y and z, marks the positive end of the axis.

We will generalize our definition of a centered optical system to include reflecting surfaces later. Other shapes, such as cylindrical, are permitted in an optical system, but we limit our discussion to spherical sections.

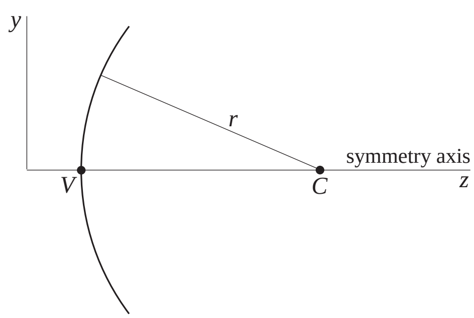
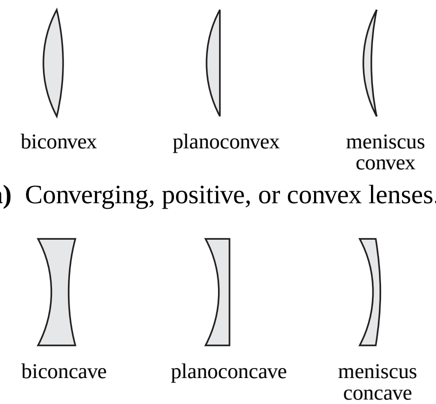


Figure 1.6 The radius of curvature r is positive.

As shown in Figure 1.6, the y axis is the vertical axis with positive direction up, and in later chapters when we need to have a three-dimensional reference system, the x axis will be drawn perpendicular to the page with positive direction pointing into the page. We choose this somewhat awkward reference system because in later chapters when we want to describe what is happening in the xy plane, it gives a very convenient reference frame. Positive and negative signs are extremely important in geometrical optics, and we will have to pay special attention to them; in Figure 1.6, the radius of curvature r is defined to be positive when the center of curvature C lies to the right of the refracting surface, negative when it lies to the left. The point V of the refracting surface is called the vertex.

An optical system of two or more refracting surfaces with at least one surface curved, we name a lens system; when the system has just two refracting surfaces, we call it a lens, or for emphasis, a simple lens. The class of simple lenses is divided into two subclasses: in the convex subclass, shown in Figure 1.7(a), incident rays parallel to the symmetry axis bend or converge toward the symmetry axis after passing through the lens; in the concave subclass, illustrated in Figure 1.7(b), similar incident rays diverge from the symmetry axis—assuming that these lenses are in air or another medium with an index of refraction less than that of the lens material. We also call these lenses positive or negative, as we explain later in this chapter when we talk about the focal length of a lens. We observe that a converging lens is thicker in the center; that is, it bulges outward, the basic meaning of convex. In a similar way, the diverging lens is thinner at the center, it is depressed, or concave.



(a) Converging, positive, or convex lenses.

(b) Diverging, negative, or concave lenses.

Figure 1.7 Cross sections of some simple lenses.

Some of the lenses in each subclass are given special names, as we indicate in Figure 1.7. A biconvex lens, or double-convex lens, has each of its surfaces curved outward; when a biconvex lens has surfaces of equal radii in magnitude, the lens is called equiconvex. A planoconvex lens, or planar-convex lens, has one surface that is plane. The meniscus-convex lens is a crescent-shaped object. We can give a similar description for the concave lenses. Other names are sometimes used for these lenses, but we have given the ones frequently used.

The first practical use for lenses was to assist in reading, and was described in writing after the 1250s. Although these early lenses were crude, they worked and enabled individuals to perceive detail that they could not see with the unaided eye. Because the lenses were of the convex shape, they reminded people at that time of lentils, the Latin name for which was lens, and thereby became the popular name for these reading aids. With time a skill was developed in making lenses of better quality. These lenses usually had a spherical shape because that is the natural surface that develops when two materials are rotated and rubbed against each other over all sections of both surfaces.

1.3.2 The refraction matrix

Consider a ray in the meridional plane that undergoes refraction at the point P_1 of a spherical refracting surface of radius r_1, as shown in Figure 1.8; for convenience, we draw the rays such that they travel from left to right (the positive z direction). The diagram in Figure 1.8 is a reference diagram: it is used to illustrate the definition of quantities and to show when they are positive, and by inference, when negative. By inspection of the diagram, we observe this rule means that angles drawn counterclockwise from the reference lines (the straight lines touched by the tails of the arcs with arrows) are positive; when drawn clockwise, they are negative. Because the radius r_1 (and its extension) form the normal to the surface, the angle θ_1 is the angle of incidence, and the angle θ'_1 is the angle of refraction. The angles δ_1 and δ'_1 are the angles of inclination (or slope angles), angles relative to the horizontal; it is these angles that will be the most important in describing the directions of a ray in paraxial optics. A quick way to obtain the sign of the angle of inclination is to note that when a ray travels from left to right, it is positive when the ray travels "uphill," negative when "downhill."

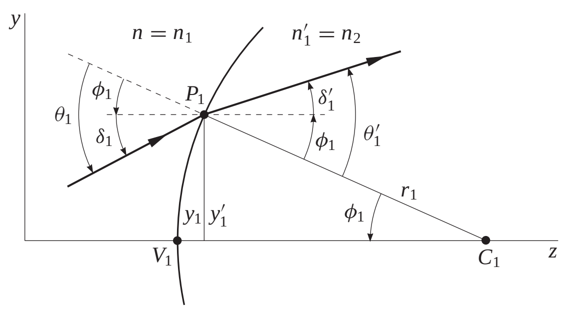


Figure 1.8 The reference diagram for refraction drawn with all the quantities positive.

We obtain several mathematical equations by inspection of the diagram in Figure 1.8:

$$\sin\phi_1 = \frac{y_1}{r_1} = \frac{y'_1}{r_1} \tag{1.15a}$$

$$\theta_1 = \phi_1 + \delta_1 \tag{1.15b}$$

$$\theta'_1 = \phi_1 + \delta'_1 \tag{1.15c}$$

$$n_1 \sin\theta_1 = n'_1 \sin\theta'_1 \tag{1.15d}$$

where the last equation is the law of refraction (Snell's law), our key equation to be simplified.

It is important to note that in the diagram the y coordinate of the point P_1 before refraction is called y_1, immediately after refraction it is called y'_1; these coordinates are positive when above the z axis, negative when below. Clearly, $y'_1 = y_1$. We also observe the rather strange equalities for the indices of refraction, $n = n_1$ and $n'_1 = n_2$; these apparently redundant relations will allow us to write practical iterative equations—that is, equations which can be performed repeatedly—in tracing a ray through an optical system of several refracting surfaces. It is also correct to write the relation $n = n_1$ as $n'_0 = n_1$, and there are occasions when we shall use the latter notation. However, here we use the $n = n_1$ notation because we can construct simpler equations by treating the first medium (and the last) in a distinctive manner.

We can greatly simplify the description of rays that pass through an optical system by the requirement that the rays are paraxial, where "par" means "beside" and "axial" means, of course, "axis." Thus, paraxial rays are "beside axis" rays; that is, they travel close to the symmetry axis along the entire length of the optical system. Then,

> the paraxial approximation means that
> rays travel close enough to the symmetry axis to make all angles used in the description of the rays small.

When we make the angles small, we can replace the sine functions in Equations 1.15a and 1.15d by simpler relationships. In calculus it is shown that when an angle A is measured in radians, then the sine function can be written in terms of the infinite series

$$\sin A = A - \frac{A^3}{3!} + \frac{A^5}{5!} - \cdots \tag{1.16a}$$

and we see that when A is small, we have

$$\sin A \approx A \tag{1.16b}$$

For example, sin (0.1 rad) gives 0.09983 for an error of 0.17% (error $= 100\,\%\,|\text{approximate} - \text{exact}|/\text{exact}$). Using Equation 1.4, we find that 0.1 rad corresponds to 5.73 deg. Thus, for angles that are approximately 6 deg or less, the error when Equation 1.16b is used to determine the sine is about 0.2 % or less. Therefore, when ϕ_1, θ_1, and θ'_1 are small, Equations 1.15a and 1.15d become

$$\phi_1 \approx \frac{y_1}{r_1} = \frac{y'_1}{r_1} \tag{1.17a}$$

$$n_1\,\theta_1 \approx n'_1\,\theta'_1 \tag{1.17b}$$

where the angles are measured in radians—very important. Substituting Equations 1.15b and 1.15c into Equation 1.17b, and then substituting Equation 1.17a for ϕ_1, we obtain

$$n_1(\phi_1 + \delta_1) = n'_1(\phi_1 + \delta'_1)$$

$$n_1\left(\frac{y_1}{r_1} + \delta_1\right) = n'_1\left(\frac{y_1}{r_1} + \delta'_1\right) \tag{1.18}$$

Solving Equation 1.18 for $n'_1\delta'_1$ gives

$$\begin{aligned} n'_1\delta'_1 &= n_1\delta_1 - \frac{n'_1 - n_1}{r_1}y_1 \\ &= n_1\delta_1 - p_1 y_1 \end{aligned} \tag{1.19}$$

where

$$p_1 = \frac{n'_1 - n_1}{r_1} = (n'_1 - n_1)c_1 \tag{1.20}$$

is called the power of the refracting surface, and $c_1 = 1/r_1$ is the curvature of the surface. Using the curvature c instead of the radius r is especially useful when describing a plane surface, since in place of writing $r = \infty$, we can write $c = 0$; working with zeros is usually easier than working with infinities, especially in a computer program.

The power p_1 is the only term in Equation 1.19 that contains the radius r_1—we use p_1 instead of r_1 because we obtain a linear relationship between δ_1' and p_1. This relationship gives an easier intuitive feel to the change in δ_1' that occurs when we change p_1 rather than thinking in terms of a change in r_1. Since a similar equation to Equation 1.19 can be written for any surface of the optical system, what we have just said is true for any surface.

Finally, we write the obvious relation, obtained by looking at Figure 1.8,

$$y_1' = y_1 \tag{1.21}$$

We now observe by inspection that Equations 1.19 and 1.21 are summarized by the matrix equation

$$\begin{pmatrix} n_1'\delta_1' \\ y_1' \end{pmatrix} = \begin{pmatrix} 1 & -p_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} n_1\delta_1 \\ y_1 \end{pmatrix} \tag{1.22}$$

or more simply as

$$V_1' = R_1 V_1 \tag{1.23}$$

where

$$R_1 = \begin{pmatrix} 1 & -p_1 \\ 0 & 1 \end{pmatrix} \tag{1.24}$$

is called the refraction matrix of surface 1, and

$$V_1 = \begin{pmatrix} n_1\delta_1 \\ y_1 \end{pmatrix}$$

$$V_1' = \begin{pmatrix} n_1'\delta_1' \\ y_1' \end{pmatrix}$$

are column vectors. Because these column vectors V contain information on the ray as it traverses the optical system, we call them ray vectors. The values of δ and y contained in the ray vectors are called the ray data (or ray parameters). Although we have already used the symbol V for a vertex of a refracting surface in our diagrams, the context tells us when V represents a vertex or a ray vector.

Referring to the ray vectors in Equation 1.22, we see that the $n\delta$ quantity is in the first row and the y quantity is in the second. This order could be reversed, as long as the off diagonal elements in the R matrix are also interchanged; it is primarily a choice of style.

When we calculate the determinant of the refraction matrix R_1, we obtain the simple but important result

$$|R_1| = \begin{vmatrix} 1 & -p_1 \\ 0 & 1 \end{vmatrix} = 1$$

To provide this result of one is the reason that the indices of refraction n_1 and n_1' are kept with the angles of inclination δ_1 and δ_1' in Equations 1.19 and 1.22.

Every refraction matrix in paraxial optics has the form of the matrix in Equation 1.24 and has the property that its determinant equals one.

1.3.3 The translation matrix

Now we look at the translation (or transfer) of a paraxial ray from one point to another in a given medium, say between the points P_1 and P_2 on the two surfaces shown in Figure 1.9 (the diagram in this figure is a continuation of the one in Figure 1.8). By inspection of this diagram we see that

$$\delta_2 = \delta_1'$$

$$n_2\delta_2 = n_1'\delta_1' \tag{1.25}$$

where in the last step we have used the relation $n_1' = n_2$ to make an equation that lends itself more easily to iteration. From Figure 1.9, we have

$$y_2 = y_1' + \left(t_1' - \Delta t_1' + \Delta t_2'\right) \tan \delta_1' \tag{1.26}$$

Because the paraxial approximation says that paraxial rays travel close to the symmetry axis (the z axis), we see that $\Delta t_1'$ and $\Delta t_2'$ are small compared to t_1' so Equation 1.26 becomes

$$y_2 = y_1' + t_1' \tan \delta_1' \tag{1.27}$$

Thus, we can simply use the thickness of a lens to describe the translation of a ray between its surfaces. In parallel with $n_1' = n_2$, we write $t_1' = t_2$ in Figure 1.9 to provide flexibility in creating equations and computer programs.

In Figure 1.9, we draw the dimension lines for the translations $t_1' = t_2$, $\Delta t_1'$, and $\Delta t_2'$ as arrows with a tail and a head to make them look like vectors. The tail indicates the reference from which we measure the quantity, and the head gives the direction so that we can infer the positive or negative status of the quantity. In Figure 1.9, the dimension lines (or dimension arrows) point in the positive z-axis direction; therefore, all three of these quantities are positive. This practice, which we shall follow in our text, helps us to write relationships with the correct plus and minus signs by inspection of the diagram—a very useful and time-saving method.

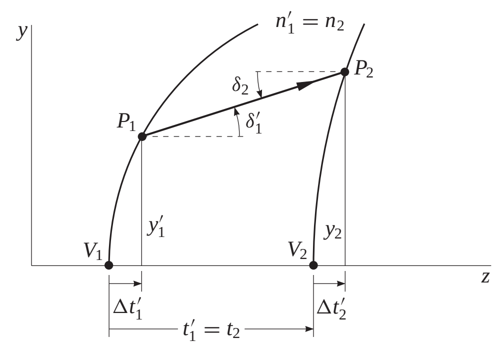


Figure 1.9 The reference diagram for translation drawn with all the quantities positive.

Just as there is an infinite series representation for the sine function (see Equation 1.16a), there is one for the tangent function:

$$\tan A = A + \frac{A^3}{3} + \frac{2A^5}{15} + \cdots \tag{1.28a}$$

where A is in radians. When A is small we have

$$\tan A \approx A \tag{1.28b}$$

By the paraxial approximation, the angle of inclination δ_1' in Equation 1.27 is small; thus, Equation 1.28b allows us to write Equation 1.27 as

$$y_2 = y_1' + t_1'\delta_1' \tag{1.29}$$

Since our practice is to group the index of refraction with the angle of inclination to make the determinant of a refraction matrix R unity, we rewrite Equation 1.29 as

$$y_2 = y_1' + \left(\frac{t_1'}{n_1'}\right)\left(n_1'\delta_1'\right) \tag{1.30}$$

Then we observe that we can write both Equations 1.25 and 1.30 in matrix form as

$$\begin{pmatrix} n_2\delta_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{t_1'}{n_1'} & 1 \end{pmatrix} \begin{pmatrix} n_1'\delta_1' \\ y_1' \end{pmatrix} \tag{1.31}$$

or more simply as

$$V_2 = T_{21}V_1' \tag{1.32}$$

where

$$T_{21} = \begin{pmatrix} 1 & 0 \\ \frac{t_1'}{n_1'} & 1 \end{pmatrix} \tag{1.33}$$

is the translation matrix (or transfer matrix) from surface 1 to 2. We write the subscripts of the translation matrix in reverse order because of the subscript pattern that emerges in Equation 1.32: when we read the subscripts from right to left on the right side of Equation 1.32, we have a description of the progress of the ray as it traverses the system. The subscript 1 in the column matrix V_1' indicates that the ray starts at surface 1 (the prime indicates that the ray has already been refracted at that surface), and the 21 subscripts of the translation matrix T says that the ray translates from surface 1 to 2. On the left side of Equation 1.32, the subscript 2 of the column matrix simply says that the ray is now at surface 2; since there is no prime attached to V_2, we understand that the ray has not yet undergone refraction at that surface.

We easily see that the translation matrix in Equation 1.33 has a determinant equal to one, just like a refraction matrix. Every translation matrix in paraxial optics has the form of the matrix in Equation 1.33, and has the property that its determinant is unity.

1.3.4 The plane-to-plane matrix

We can make a complete description of the ray travel so far by substituting Equation 1.23 into Equation 1.32 to obtain the extended matrix equation

$$V_2 = T_{21}R_1V_1 \tag{1.34}$$

which now includes the refraction that takes place at surface 1. Again we note that when we read the right side of Equation 1.34 from right to left, we can follow the advance of the ray through the optical system. Equation 1.34 keeps the same form when it is generalized to describe the ray travel through a system of one, two, or many refracting surfaces.

To bring together what we have already done, to set the pattern for tracing a ray through any number of refracting surfaces from beginning to end, and to finally summarize the entire ray travel in a single matrix called the plane-to-plane matrix, we draw the reference diagram of Figure 1.10. This diagram illustrates how any system of refracting surfaces—from one to many—is represented. Following our usual practice, we draw the diagram such that all the quantities are positive. The radii r_1 and r_2 are positive, but are not shown to

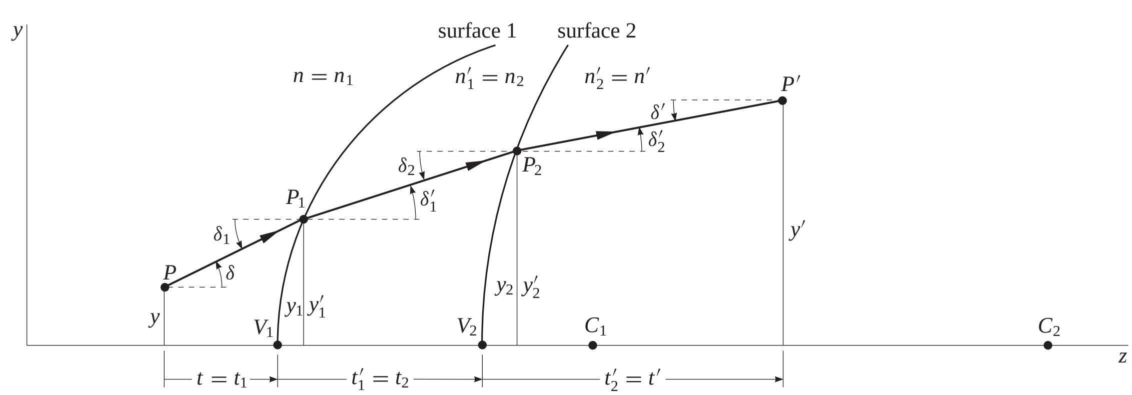


Figure 1.10 The reference diagram for an optical system of two refracting surfaces—all quantities are positive.

keep the diagram simple; however, we do show the centers of curvature C_1 and C_2. The diagram reiterates what we have said before: whenever a center of curvature is to the right of its refracting surface the corresponding radius is positive; if the center of curvature is to the left, then the radius is negative. We shall use examples to illustrate these features for the radii, as well as to illustrate the properties of the other quantities. The middle part of this diagram, the part between the surfaces 1 and 2, shows the pattern by which quantities would be renamed if there were another refracting surface, say surface 3. For example, the index of refraction relation $n'_1 = n_2$ would become $n'_2 = n_3$ between surfaces 2 and 3; in a similar way, the translation relation would become $t'_2 = t_3$. We also note that quantities that have not yet been refracted at a surface are unprimed, after refraction they are primed. As we have already mentioned, this somewhat messy notation is used because it lends itself to convenient iteration. Finally, it proves useful to treat the initial and final media in a special way using quantities without subscripts: for the first medium we use the unprimed quantities (n, δ, y, t, P), for the final medium we use the primed quantities $(n', \delta', y', t', P')$.

Lastly, we write the matrix equation that represents the ray travel shown in Figure 1.10 using symbols to represent the matrices:

$$V_{P'} = T_{P'2} R_2 T_{21} R_1 T_{1P} V_P \tag{1.35a}$$

$$= M_{P'P} V_P \tag{1.35b}$$

where

$$M_{P'P} = T_{P'2} R_2 T_{21} R_1 T_{1P} \tag{1.36}$$

is called the plane-to-plane matrix, a square matrix that allows us to calculate where a ray on the plane containing the point P will terminate on a plane containing the point P'.

Again we note that by reading Equation 1.35a from right to left we obtain the left-to-right travel of the ray through the optical system of centered refracting surfaces shown in Figure 1.10. Filling out the matrix symbols in Equation 1.35a with what each one represents, we obtain the matrix equation

$$\begin{pmatrix} n'\delta' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \dfrac{t'}{n'} & 1 \end{pmatrix} \begin{pmatrix} 1 & -p_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \dfrac{t'_1}{n'_1} & 1 \end{pmatrix} \begin{pmatrix} 1 & -p_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \dfrac{t}{n} & 1 \end{pmatrix} \begin{pmatrix} n\delta \\ y \end{pmatrix} \tag{1.37}$$

where the powers of the surfaces are

$$p_1 = \frac{n'_1 - n_1}{r_1} \quad \text{and} \quad p_2 = \frac{n'_2 - n_2}{r_2} \tag{1.38}$$

No matter how complicated the system of refracting surfaces is, we can represent the system with a diagram which has the form of Figure 1.10, and represent the paraxial ray travel with equations that follow the pattern shown in Equations 1.35 through 1.38.

Using examples, we now trace a ray through several lens systems to illustrate the concepts introduced so far.

Example 1.3.1 Tracing a ray through a converging lens. In Figure 1.11, we show an equiconvex lens; that is, a converging lens with radii that are equal in magnitude. We wish to trace a ray from a point P through the lens to another point P' (not shown) located on a plane normal to the symmetry axis a distance of 40.00 mm to the right of V_2.

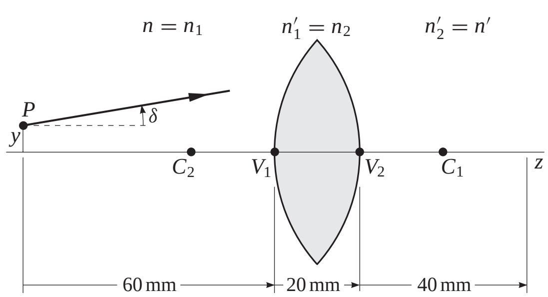


Figure 1.11

Initial ray-data values (the δ and y) and the values of the other quantities that refer to the lens system are given in tabular fashion in Figure 1.12. The headers of the columns (the r, n, and t) are the generic symbols for radii, indices of refraction, and translations, respectively.

	r (mm)	n	t (mm)
$\delta = 0.100$ rad		1.000	60.00
$= 5.73$ deg	40.00		
		1.500	20.00
$y = 5.00$ mm	−40.00		
		1.000	40.00

Figure 1.12

We understand the table is shorthand for the following (refer to Figures 1.10 and 1.11 for the meaning of the symbols):

$$r_1 = 40.00\,\text{mm}$$
$$r_2 = -40.00\,\text{mm}$$
$$n = n_1 = 1.000$$
$$n'_1 = n_2 = 1.500$$
$$n'_2 = n' = 1.000$$
$$t = t_1 = 60.00\,\text{mm}$$
$$t'_1 = t_2 = 20.00\,\text{mm}$$
$$t'_2 = t' = 40.00\,\text{mm}$$

Using Equation 1.38, we calculate the powers of the surfaces:

$$\left.\begin{aligned} p_1 &= \frac{n_1' - n_1}{r_1} = \frac{1.5 - 1}{40} = 0.0125\,\text{mm}^{-1} \\ p_2 &= \frac{n_2' - n_2}{r_2} = \frac{1 - 1.5}{-40} = 0.0125\,\text{mm}^{-1} \end{aligned}\right\} \tag{1.39}$$

and then, starting with the initial ray vector V_P, we determine the other ray vectors (the column vectors) after each translation or refraction to trace the ray through the lens to the point P'. First in symbol form, we have

$$V_P = \begin{pmatrix} n\delta \\ y \end{pmatrix}$$

$$V_1 = T_{1P} V_P \qquad V_1' = R_1 V_1$$

$$V_2 = T_{21} V_1' \qquad V_2' = R_2 V_2$$

$$V_{P'} = T_{P'2} V_2'$$

and then substituting

$$V_P = \begin{pmatrix} n\delta \\ y \end{pmatrix} = \begin{pmatrix} 0.100 \\ 5.00 \end{pmatrix}$$

$$V_1 = \begin{pmatrix} n_1\delta_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \dfrac{t}{n} & 1 \end{pmatrix} \begin{pmatrix} n\delta \\ y \end{pmatrix} = \begin{pmatrix} 0.100 \\ 11.00 \end{pmatrix}$$

$$V_1' = \begin{pmatrix} n_1'\delta_1' \\ y_1' \end{pmatrix} = \begin{pmatrix} 1 & -p_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} n_1\delta_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} -0.0375 \\ 11.00 \end{pmatrix}$$

$$V_2 = \begin{pmatrix} n_2\delta_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \dfrac{t_1'}{n_1'} & 1 \end{pmatrix} \begin{pmatrix} n_1'\delta_1' \\ y_1' \end{pmatrix} = \begin{pmatrix} -0.0375 \\ 10.50 \end{pmatrix}$$

$$V_2' = \begin{pmatrix} n_2'\delta_2' \\ y_2' \end{pmatrix} = \begin{pmatrix} 1 & -p_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} n_2\delta_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} -0.169 \\ 10.50 \end{pmatrix}$$

$$V_{P'} = \begin{pmatrix} n'\delta' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \dfrac{t'}{n'} & 1 \end{pmatrix} \begin{pmatrix} n_2'\delta_2' \\ y_2' \end{pmatrix} = \begin{pmatrix} -0.169 \\ 3.75 \end{pmatrix}$$

It is easy to read off the y values from the above ray vectors, but we need to apply another step to get the δ values (refer to Figure 1.10 for the meaning of the δs):

$$\delta = \delta_1 = \frac{0.100}{n} = \frac{0.100}{1.000} = 0.100\,\text{rad}$$

$$\delta_1' = \delta_2 = \frac{-0.0375}{n_1'} = \frac{-0.0375}{1.500} = -0.0250\,\text{rad}$$

$$\delta_2' = \delta' = \frac{-0.169}{n'} = \frac{-0.169}{1.000} = -0.169\,\text{rad}$$

We summarize the values of the angles of inclination δ and coordinates y for the ray—the ray data—as it passes through the lens in the table of Figure 1.13 (the δ values are also expressed in degrees for convenience). We refer to tables as figures to make it easier to find them, because tables are usually few in number making them difficult to locate by a table number. The numbers 1 and 2 refer to the first and second surfaces, respectively, of the lens in Figure 1.14; this diagram displays a scale drawing of the path the ray follows in traveling from P to P' (the horizontal and vertical scales are not the same). Because there is not enough space to show the δ angles in Figure 1.14, we must refer to Figure 1.10 for that information.

	δ (rad)	δ (deg)	y (mm)
P	0.100	5.73	5.00
1	0.100 −0.0250	5.73 −1.43	11.00 11.00
2	−0.0250 −0.169	−1.43 −9.67	10.50 10.50
P'	−0.169	−9.67	3.75

Figure 1.13

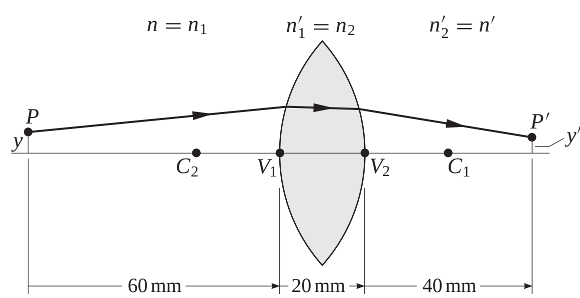


Figure 1.14

In the above calculations, we have rounded the results to an appropriate number of significant figures; however, it is important to remember that 16 decimal digits (actually, the binary equivalent) are retained for each of the values in *Mathematica*'s computer memory, and all calculations are performed with these 16 decimal digits. This practice occasionally produces small discrepancies; for example, if we use a calculator to convert the printed values in Figure 1.13 for δ_2' (or δ') of -0.169 rad to degrees, we obtain a rounded value of -9.68 deg rather than the printed -9.67 deg value obtained with *Mathematica* after rounding (using $\pi = 3.1416$).

In ray tracing it is important to do all calculations with fairly high precision because the many subtractions that take place can cause loss of significant figures, and produce inaccurate results—normally, the 16 decimal digits that *Mathematica* maintains is more than enough.

If we don't need the intermediate δ and y ray-data values, but just want to obtain the δ' and y' values at the final point P', then one of the very useful features of matrices in paraxial optics is that we can calculate just one matrix that connects the final ray-vector values with the initial ones. Thus, we first calculate the plane-to-plane matrix $M_{P'P}$ given in Equation 1.36 by substituting the table values in Figure 1.12 and the power values of Equation 1.39 into the square matrices:

$$\begin{aligned} M_{P'P} &= T_{P'2}R_2T_{21}R_1T_{1P} \\ &= \begin{pmatrix} -0.541667 & -0.0229167 \\ 41.6667 & -0.0833333 \end{pmatrix} \end{aligned} \tag{1.40}$$

where we have used *Mathematica* to calculate the product of these five matrices, and showed the results to *Mathematica*'s default decimal display of rounded six decimal digits. We can now use Equation 1.35b to calculate

$$\begin{aligned} V_{P'} &= M_{P'P}V_P \\ &= \begin{pmatrix} -0.541667 & -0.0229167 \\ 41.6667 & -0.0833333 \end{pmatrix} \begin{pmatrix} 0.100 \\ 5.00 \end{pmatrix} \\ &= \begin{pmatrix} -0.169 \\ 3.75 \end{pmatrix} \end{aligned} \tag{1.41}$$

or more explicitly,

$$V_{P'} = \begin{pmatrix} n'\delta' \\ y' \end{pmatrix} = \begin{pmatrix} -0.169 \\ 3.75 \end{pmatrix} \tag{1.42}$$

and quickly determine the final ray data

$$\begin{aligned} \delta' &= \frac{-0.169}{n'} = \frac{-0.169}{1.000} = -0.169\,\text{rad} \\ y' &= 3.75\,\text{mm} \end{aligned}$$

in agreement with the results we obtained before for the ray data at the point P' (see the table in Figure 1.13).

Example 1.3.2 Tracing a ray backward.

One of the remarkable properties of the matrix approach is that rays can be traced backward easily using the concept of the matrix inverse. For example, if we know the final ray vector $V_{P'}$ and the plane-to-plane matrix inverse $M_{P'P}^{-1}$, we can determine the initial ray vector V_P. Using the matrix algebra presented in Section 1.2, we start with

$$V_{P'} = M_{P'P}V_P$$

multiply both sides by the inverse $M_{P'P}^{-1}$ of the matrix $M_{P'P}$ to obtain

$$M_{P'P}^{-1}V_{P'} = (M_{P'P}^{-1}M_{P'P})V_P = IV_P = V_P$$

and then interchange sides to get the desired result

$$V_P = M_{P'P}^{-1}V_{P'} \tag{1.43}$$

All the determinants of refraction matrices, translation matrices, and their products, are equal to one; therefore, the determinant of $M_{P'P}$ is unity. Then, according to Equation 1.14, the inverse of $M_{P'P}$ is easy to calculate: simply interchange the diagonal elements, and change the signs of the off-diagonal elements. We obtain

$$\begin{aligned} V_P &= M_{P'P}^{-1}V_{P'} \\ &= \begin{pmatrix} -0.0833333 & 0.0229167 \\ -41.6667 & -0.541667 \end{pmatrix} \begin{pmatrix} -0.169 \\ 3.75 \end{pmatrix} \\ &= \begin{pmatrix} 0.100 \\ 5.00 \end{pmatrix} \end{aligned} \tag{1.44}$$

which agrees with the ray vector V_P we used initially in the previous example. Thus, in one short matrix multiplication, we have traced the ray back to where it began. Again we should remember that all the numbers displayed in Equation 1.44 are rounded, but that in *Mathematica*'s internal memory, all the numbers are recorded to 16 decimal digit precision—it is these high precision numbers that are used in the matrix multiplication of Equation 1.44. This information is important for it explains why the use of the displayed values for $M_{P'P}^{-1}$ and V_P' in the matrix multiplication, say with a calculator, do not quite give the same result for V_P.

We can also trace the ray backward in steps, very much like the way we traced it forward in the previous example. To trace the ray backward from the final point P' to the initial point P, we have

$$\begin{aligned} V_2' &= T_{P'2}^{-1}V_{P'} \\ &= \begin{pmatrix} 1 & 0 \\ -40 & 1 \end{pmatrix} \begin{pmatrix} -0.169 \\ 3.75 \end{pmatrix} = \begin{pmatrix} -0.169 \\ 10.50 \end{pmatrix} \\ V_2 &= R_2^{-1}V_2' \\ &= \begin{pmatrix} 1 & 0.0125 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -0.169 \\ 10.50 \end{pmatrix} = \begin{pmatrix} -0.0375 \\ 10.50 \end{pmatrix} \\ V_1' &= T_{21}^{-1}V_2 \\ &= \begin{pmatrix} 1 & 0 \\ -13.3333 & 1 \end{pmatrix} \begin{pmatrix} -0.0375 \\ 10.50 \end{pmatrix} = \begin{pmatrix} -0.0375 \\ 11.00 \end{pmatrix} \\ V_1 &= R_1^{-1}V_1' \\ &= \begin{pmatrix} 1 & 0.0125 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -0.0375 \\ 11.00 \end{pmatrix} = \begin{pmatrix} 0.100 \\ 11.00 \end{pmatrix} \\ V_P &= T_{1P}^{-1}V_1 \\ &= \begin{pmatrix} 1 & 0 \\ -60.00 & 1 \end{pmatrix} \begin{pmatrix} 0.100 \\ 11.00 \end{pmatrix} = \begin{pmatrix} 0.100 \\ 5.00 \end{pmatrix} \end{aligned}$$

When we compare these results with the ones we got in the previous example, we see that we have indeed traced the ray backward to where it began.

Example 1.3.3 Tracing a ray through a diverging lens. We worked with a converging lens in the previous two examples. Here we discuss the opposite type, a diverging or concave lens, a lens thinner in the middle than at the edges. We show an equiconcave lens in Figure 1.15; that is, a diverging lens with radii that are equal in magnitude.

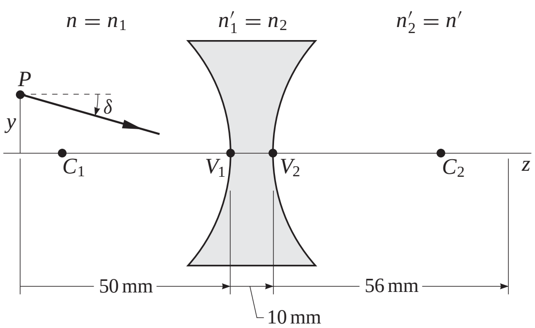


Figure 1.15

We follow the same format used in Example 1.3.1 to list in Figure 1.16 the data pertinent to this example.

	r (mm)	n	t (mm)
$\delta = -0.127$ rad		1.000	50.00
$= -7.28$ deg	−40.00		
		1.500	10.00
$y = 14.00$ mm	40.00		
		1.000	56.00

Figure 1.16

As Figure 1.15 indicates, we wish to trace the ray from the point P through the lens to a point located on a plane normal to the symmetry axis and 56.00 mm to the right of V_2. We first calculate the powers of the two surfaces as

$$p_1 = \frac{n_1' - n_1}{r_1} = \frac{1.5 - 1}{-40} = -0.0125\,\text{mm}^{-1}$$

$$p_2 = \frac{n_2' - n_2}{r_2} = \frac{1 - 1.5}{40} = -0.0125\,\text{mm}^{-1}$$

and then determine the ray vectors after each translation or refraction to trace the ray through the lens (if you've forgotten the meaning of the symbols, refer to the reference diagram in Figure 1.10):

$$V_P = \begin{pmatrix} n\delta \\ y \end{pmatrix} = \begin{pmatrix} -0.127 \\ 14.00 \end{pmatrix}$$

$$V_1 = \begin{pmatrix} n_1\delta_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{t}{n} & 1 \end{pmatrix} \begin{pmatrix} n\delta \\ y \end{pmatrix} = \begin{pmatrix} -0.127 \\ 7.65 \end{pmatrix}$$

$$V_1' = \begin{pmatrix} n_1'\delta_1' \\ y_1' \end{pmatrix} = \begin{pmatrix} 1 & -p_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} n_1\delta_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} -0.0314 \\ 7.65 \end{pmatrix}$$

$$V_2 = \begin{pmatrix} n_2\delta_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{t_1'}{n_1'} & 1 \end{pmatrix} \begin{pmatrix} n_1'\delta_1' \\ y_1' \end{pmatrix} = \begin{pmatrix} -0.0314 \\ 7.44 \end{pmatrix}$$

$$V_2' = \begin{pmatrix} n_2'\delta_2' \\ y_2' \end{pmatrix} = \begin{pmatrix} 1 & -p_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} n_2\delta_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0.0616 \\ 7.44 \end{pmatrix}$$

$$V_{P'} = \begin{pmatrix} n'\delta' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{t'}{n'} & 1 \end{pmatrix} \begin{pmatrix} n_2'\delta_2' \\ y_2' \end{pmatrix} = \begin{pmatrix} 0.0616 \\ 10.89 \end{pmatrix}$$

Just as it was in Example 1.3.1, it's easy to read off the y values, but we must perform an extra step to get the δ values. We calculate

$$\delta = \delta_1 = \frac{-0.127}{n} = \frac{-0.127}{1.000} = -0.127\,\text{rad}$$

$$\delta_1' = \delta_2 = \frac{-0.0314}{n_1'} = \frac{-0.0314}{1.500} = -0.0209\,\text{rad}$$

$$\delta_2' = \delta' = \frac{0.0616}{n'} = \frac{0.0616}{1.000} = 0.0616\,\text{rad}$$

Finally, we extract the values of δ and y from the calculations we just made, and summarize these ray data in the table in Figure 1.17. Then we make a scale drawing of the path followed by the ray to the point P' in Figure 1.18.

	δ (rad)	δ (deg)	y (mm)
P	−0.127	−7.28	14.00
1	−0.127 −0.0209	−7.28 −1.20	7.65 7.65
2	−0.0209 0.0616	−1.20 3.53	7.44 7.44
P'	0.0616	3.53	10.89

Figure 1.17

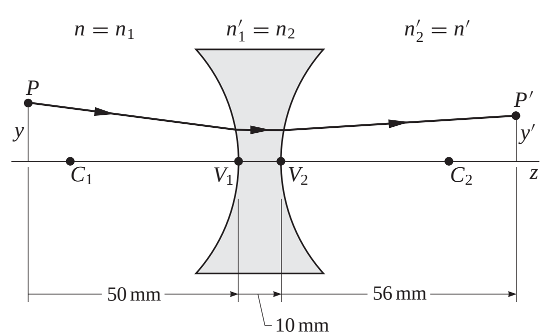


Figure 1.18

For the plane-to-plane matrix, we obtain

$$M_{P'P} = T_{P'2}R_2T_{21}R_1T_{1P} = \begin{pmatrix} 2.38542 & 0.0260417 \\ 194.417 & 2.54167 \end{pmatrix} \quad (1.45)$$

so we can quickly calculate the ray vector $V_{P'}$ at the point P': first, symbolically,

$$V_{P'} = M_{P'P}V_P$$

and, second, in matrix form

$$\begin{pmatrix} n'\delta' \\ y' \end{pmatrix} = \begin{pmatrix} 2.38542 & 0.0260417 \\ 194.417 & 2.54167 \end{pmatrix}\begin{pmatrix} -0.127 \\ 14.00 \end{pmatrix} = \begin{pmatrix} 0.0616 \\ 10.89 \end{pmatrix} \quad (1.46)$$

from which we determine the ray data

$$\delta' = \frac{0.0616}{n'} = \frac{0.0616}{1.000} = 0.0616\,\text{rad}$$
$$y' = 10.89\,\text{mm}$$

in agreement with the results we obtained before for these quantities at the point P' (see the table in Figure 1.17).

Example 1.3.4 Ray tracing that ends with a backward extension of the ray.
In our preceding examples, the final point P' was located to the right of the final lens surface. However, our matrix approach easily describes the situation when the point P' is to the left. To illustrate this case, we work with the converging lens shown in Figure 1.19. This lens system is the same as the one in Example 1.3.1, except for the last surface which is placed a distance of −30 mm to the left of V_2—we place the final point P' in this imaginary surface. We display the numeric data for this system in Figure 1.20. To shorten our calculations, we refer to Example 1.3.1, and use the result obtained after refraction at the second lens surface:

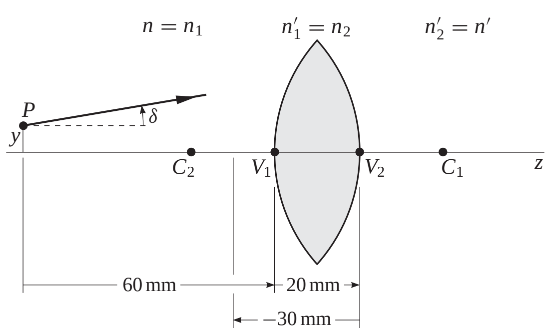


Figure 1.19

	r (mm)	n	t (mm)
$\delta = 0.100$ rad		1.000	60.00
$= 5.73$ deg	40.00		
		1.500	20.00
$y = 5.00$ mm	−40.00		
		1.000	−30.00

Figure 1.20

$$V_2' = \begin{pmatrix} n_2'\delta_2' \\ y_2' \end{pmatrix} = \begin{pmatrix} -0.169 \\ 10.50 \end{pmatrix}$$

Then, for the final translation, we have

$$V_{P'} = \begin{pmatrix} n'\delta' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{t'}{n'} & 1 \end{pmatrix}\begin{pmatrix} n_2'\delta_2' \\ y_2' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{-30}{1} & 1 \end{pmatrix}\begin{pmatrix} -0.169 \\ 10.50 \end{pmatrix} = \begin{pmatrix} -0.169 \\ 15.56 \end{pmatrix}$$

We summarize the complete travel of the ray from P to P' in the table shown in Figure 1.21 and in the ray diagram of Figure 1.22.

	δ (rad)	δ (deg)	y (mm)
P	0.100	5.73	5.00
1	0.100 −0.0250	5.73 −1.43	11.00 11.00
2	−0.0250 −0.169	−1.43 −9.67	10.50 10.50
P'	−0.169	−9.67	15.56

Figure 1.21

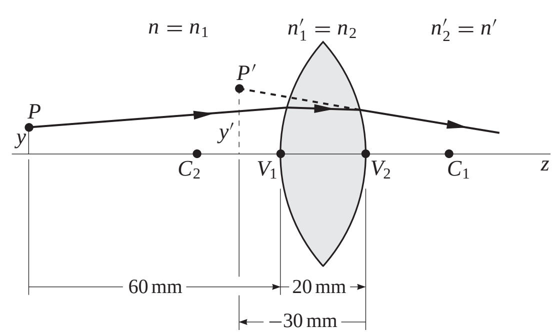


Figure 1.22

The diagram in Figure 1.22 makes clear that when a surface is to the left of the previous surface, the ray doesn't go there; instead, it's the backward extension of the ray that does—the ray just appears to come from that point.

Following our usual practice, we calculate the plane-to-plane matrix

$$M_{P'P} = T_{P'2}R_2T_{21}R_1T_{1P} = \begin{pmatrix} -0.541667 & -0.0229167 \\ 79.5833 & 1.52083 \end{pmatrix} \quad (1.47)$$

and then obtain the ray vector $V_{P'}$ at the point P' as

$$V_{P'} = M_{P'P}V_P = \begin{pmatrix} -0.541667 & -0.0229167 \\ 79.5833 & 1.52083 \end{pmatrix}\begin{pmatrix} 0.100 \\ 5.00 \end{pmatrix} = \begin{pmatrix} -0.169 \\ 15.56 \end{pmatrix} \quad (1.48)$$

We finally determine the ray data

$$\delta' = \frac{-0.169}{n'} = \frac{-0.169}{1.000} = -0.169\,\text{rad}$$
$$y' = 15.56\,\text{mm}$$

in agreement with the ray data for P' in Figure 1.21.

Example 1.3.5 Tracing a ray through a doublet. Two lenses with a pair of surfaces in contact form a doublet, a system of three refracting surfaces, as shown in Figure 1.23; we note that this doublet consists of a converging and a diverging lens. Because the second surface of the diverging lens is almost plane, the magnitude of the radius for this surface is very large; thus, we cannot show the center of curvature C_3 on the diagram—we should imagine it 1524 mm to the left of V_3. Following our usual practice, the initial ray data and the properties of the lens system are shown in Figure 1.24.

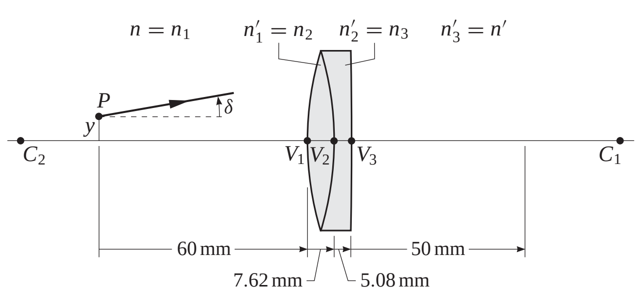


Figure 1.23

	r (mm)	n	t (mm)
$\delta = 0.110$ rad		1.0000	60.00
$= 6.30$ deg	90.17		
		1.5166	7.62
$y = 7.00$ mm	−90.17		
		1.6256	5.08
	−1524.00		
		1.0000	50.00

Figure 1.24

To trace the ray from P to the plane 50 mm to the right of V_3, we first calculate the powers of the three surfaces:

$$p_1 = \frac{n_1' - n_1}{r_1} = \frac{1.5166 - 1.0000}{90.17} = 0.005729\,\text{mm}^{-1}$$
$$p_2 = \frac{n_2' - n_2}{r_2} = \frac{1.6256 - 1.5166}{-90.17} = -0.001209\,\text{mm}^{-1}$$
$$p_3 = \frac{n_3' - n_3}{r_3} = \frac{1.0000 - 1.6256}{-1524.00} = 0.0004105\,\text{mm}^{-1}$$

We then calculate the ray vectors as we did in the previous examples, but do not show the work; instead, we simply list the results in the table in Figure 1.25. We show the ray diagram that corresponds to these results in Figure 1.26.

	δ (rad)	δ (deg)	y (mm)
P	0.110	6.30	7.00
1	0.110 0.0212	6.30 1.21	13.60 13.60
2	0.0212 0.0300	1.21 1.72	13.76 13.76
3	0.0300 0.0430	1.72 2.46	13.91 13.91
P'	0.0430	2.46	16.06

Figure 1.25

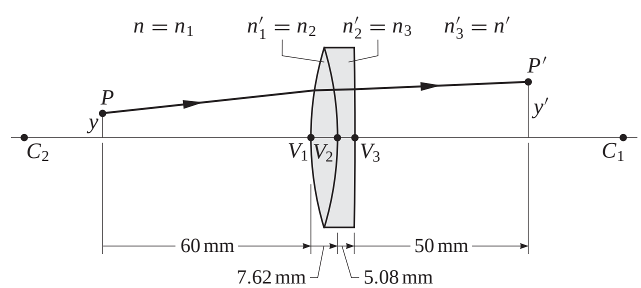


Figure 1.26

In all the ray diagrams of these examples, we have drawn the rays to intersect at the lens surfaces. Actually, we should draw the rays to intersect at vertical planes that pass through the V vertices of the lens surfaces, as shown in Figure 1.27—all because of the approximation we discussed in Section 1.3.3 and illustrated in Figure 1.9. Remember, in the

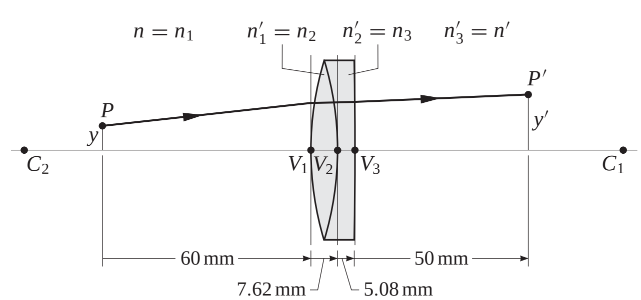


Figure 1.27

paraxial approximation, we say that the ray travels so close to the symmetry axis that we can ignore quantities like the $\Delta t_1'$ and $\Delta t_2'$ shown in Figure 1.9, and treat the lens thickness as equal to the distance between its vertices; that is, we think of the lens surfaces as approximating planes. However, in our examples, so that we can see the ray more easily, we have chosen initial values of the ray data to cause the ray to travel some distance from the symmetry axis; thus, the ray may become less paraxial than we might like. Therefore, from the rigorous point of view, a ray of this type should be drawn to vertical planes, as we have done in Figure 1.27. However, when we compare the ray diagrams of Figures 1.26 and 1.27, we see that the difference between the two is not great; therefore, we opt on the side of simplicity and draw our ray diagrams like the one shown in Figure 1.26. However, to accurately apply the data listed in our tables, such as the table in Figure 1.25, we must remember that the data applies to vertical planes like those shown in Figure 1.27.

It is important to understand that paraxial mathematics is a consistent mathematical formalism; it gives valid mathematical results even when the ray is physically no longer paraxial. We simply have to remember that the ray data in such cases may not describe the physical rays very well; the results are still useful, but we must apply them with care.

Finally, we want to finish this example in the same manner as the previous examples in this section by calculating the plane-to-plane matrix $M_{P'P}$. The product of the matrices forming $M_{P'P}$ is longer than before because there is an additional lens surface; thus, there is another translation-refraction set of matrices that refer to surface 3 in the system. We calculate the plane-to-plane matrix

$$\begin{aligned} M_{P'P} &= T_{P'3}R_3T_{32}R_2T_{21}R_1T_{1P} \\ &= \begin{pmatrix} 0.705841 & -0.00494799 \\ 100.879 & 0.709580 \end{pmatrix} \end{aligned} \tag{1.49}$$

and the ray vector $V_{P'}$

$$\begin{aligned} V_{P'} &= M_{P'P}V_P \\ &= \begin{pmatrix} 0.705841 & -0.00494799 \\ 100.879 & 0.709580 \end{pmatrix} \begin{pmatrix} 0.110 \\ 7.00 \end{pmatrix} \\ &= \begin{pmatrix} 0.0430 \\ 16.06 \end{pmatrix} \end{aligned} \tag{1.50}$$

from which we get the ray data

$$\delta' = \frac{0.0430}{n'} = \frac{0.0430}{1.000} = 0.0430\,\text{rad}$$
$$y' = 16.06\,\text{mm}$$

in agreement with the ray data shown for P' in Figure 1.25.

1.3.5 A recursion relation for $M_{P'P}$

Computers have made it practical to use the matrix approach in paraxial optics. To write elegant computer programs is a satisfying activity in any field including optics. In particular, the calculation of the plane-to-plane matrix $M_{P'P}$ can involve the multiplication of many matrices, and we would like a short way to specify this multiplication. To start, we use the symbolic part of Equation 1.49 as a guide, which we rewrite for convenience,

$$M_{P'P} = T_{P'3}R_3T_{32}R_2T_{21}R_1T_{1P} \tag{1.51a}$$

Reading from left to right, we note a pattern in the multiplication: a translation-refraction pair followed by another such pair and yet another, finishing with the special case of a single translation; that is, using parentheses,

$$M_{P'P} = (T_{P'3}R_3)(T_{32}R_2)(T_{21}R_1)(T_{1P}) \tag{1.51b}$$

All calculations of $M_{P'P}$ follow this pattern, which suggests the use of a special formula, a recursion formula. But before we write this formula, we must rewrite Equation 1.51b in a more convenient form for use in a computer program. To this end, we write

$$M(4) = [T(4)R(3)]\,[T(3)R(2)]\,[T(2)R(1)]\,[T(1)] \tag{1.51c}$$

where we think of P' as on surface 4, $T(4)$ as the translation to surface 4, $R(3)$ as the refraction on surface 3, and so on, until we reach the first translation $T(1)$ to surface 1 from P.

To help us write Equation 1.51c as a recursion formula, we look at the one for the calculation of the factorial function for positive integers:

$$f(i) = i\ f(i-1) \qquad f(1) = 1 \tag{1.52}$$

The formula involving i is repeated until $f(1)$ is reached, then $f(1)$ is replaced by 1 and the calculation stops. Using Equation 1.52 as a guide, we see that we can generalize Equation 1.51c to the recursion formula

$$M(i) = T(i)R(i-1)M(i-1) \quad M(1) = T(1) \tag{1.53}$$

We have now obtained a simple and effective way to describe the calculation of the plane-to-plane matrix for a system which is small or large.

1.3.6 Optical systems scale linearly

As the examples in Section 1.3.4 show, we specify the characteristics of an optical system by listing the values of r, n, t, and start the ray tracing by giving the initial values of the ray data δ, y; for example, see the table in Figure 1.12. Of these quantities only the r, t, y have a physical dimension; namely, the length dimension. The other two quantities n, δ don't have a physical dimension—n has no units and δ has only a supplementary unit (the radian) that can be erased (see Section 1.1.2).

When we want to know how a system scales, we examine how quantities that have a common physical dimension—in our case, length—change when they are multiplied by some factor, say e (where e represents some number such as $1/2$, $1/3$, 2, 5, etc.). We look first at the translation from the initial point P to the first surface of an optical system—we multiply the two length quantities involved, t and y, by e:

$$V_1 = T_{1P} V_P$$

$$\begin{pmatrix} n_1\delta_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \dfrac{et}{n} & 1 \end{pmatrix} \begin{pmatrix} n\delta \\ ey \end{pmatrix} = \begin{pmatrix} n\delta \\ e(t\delta + y) \end{pmatrix}$$

from which we read off (remember $n_1 = n$)

$$n_1\delta_1 = n\delta \qquad \text{or} \qquad \delta_1 = \delta$$

and

$$y_1 = e(t\delta + y)$$

We see that the angle of inclination δ_1 doesn't change, and the y_1 coordinate is e times greater than it is when there is no scale factor (that is, when $e = 1$). Therefore, the y coordinate scales linearly under a translation.

Next, we consider refraction at the first surface; again we multiply all the length quantities by e. For the power p, we have

$$\frac{p_1}{e} = \frac{n_1' - n_1}{er_1}$$

and then

$$V_1' = R_1 V_1$$

$$\begin{pmatrix} n_1'\delta_1' \\ y_1' \end{pmatrix} = \begin{pmatrix} 1 & -\dfrac{p_1}{e} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} n\delta \\ e(t\delta + y) \end{pmatrix}$$

$$= \begin{pmatrix} n\delta - p_1(t\delta + y) \\ e(t\delta + y) \end{pmatrix}$$

We see that the e factors cancel when we multiply $-p_1/e$ and $e(t\delta + y)$ to give

$$n_1'\delta_1' = n\delta + p_1(y + t\delta)$$

an equation for the new angle of inclination with no e factor in it. Of course, as always happens with refraction, the y coordinate remains the same, namely,

$$y_1' = y_1 = e(t\delta + y)$$

This behavior continues throughout the remainder of the optical system: 1) there is no e factor in any of the equations that relate to the angle of inclination δ, and 2) the y coordinate is always multiplied by a single e factor. It is this behavior that allows us to say that an optical system scales linearly.

Because an optical system scales linearly in the manner shown, we can multiply by e to convert easily from one length unit to another, say from millimeters to inches. Or, we can keep a given length unit, and multiply by e to make the optical system larger or smaller. Thus, there is no reason to give a length unit when we specify the quantities that have a length dimension, and this practice is frequently followed in optics. However, we shall continue to use the millimeter (mm) as the length unit in this text, because we can then gain more easily a feeling for the size and properties of the optical system.

1.4 The Gaussian Constants

1.4.1 The system matrix

As we discussed in Section 1.3.4, the plane-to-plane matrix $M_{P'P}$ of any two surface optical system (see Figure 1.28) is given by Equation 1.36, which we rewrite for convenience,

$$M_{P'P} = T_{P'2} R_2 T_{21} R_1 T_{1P} \tag{1.54}$$

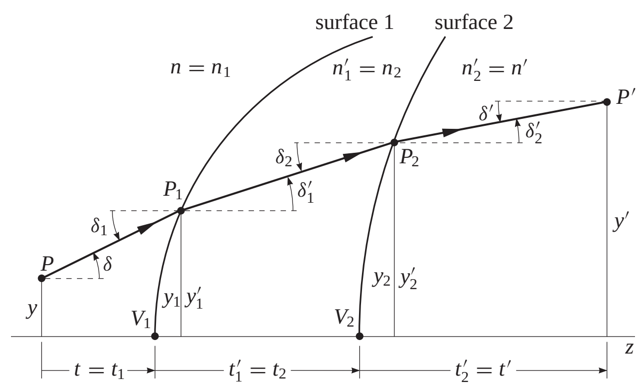


Figure 1.28

The inner refraction R and translation T matrices in Equation 1.54 can be grouped together to give

$$M_{P'P} = T_{P'2}(R_2 T_{21} R_1) T_{1P}$$

or

$$M_{P'P} = T_{P'2} S_{21} T_{1P} \tag{1.55}$$

where

$$S_{21} = R_2 T_{21} R_1 \tag{1.56}$$

is called the system matrix for the two surfaces making up the optical system. We see that the system matrix S_{21} is composed only of the matrices that refer to the system of refracting surfaces and their separation; the first and last translation matrices in the expression for $M_{P'P}$ are excluded. The system matrix is a useful way to describe the properties of an optical system for paraxial work.

We now generalize this idea to a system of N refracting surfaces, where $N \geq 1$, and illustrate such a system in Figure 1.29; the wavy line in the diagram indicates the possible presence of other surfaces in the system. We use our

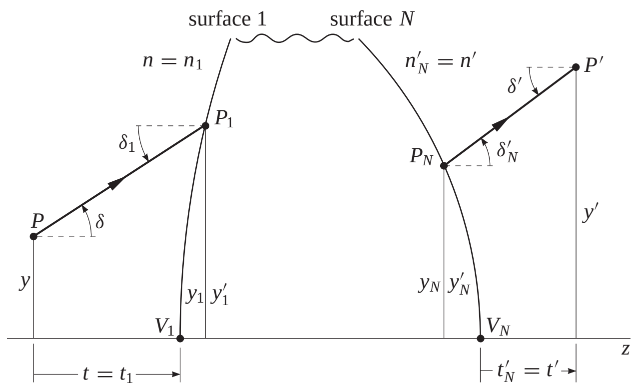


Figure 1.29 The generalized optical system of N surfaces, where $N \geq 1$.

work for two surfaces as a guide to write the plane-to-plane matrix $M_{P'P}$ for N surfaces as

$$M_{P'P} = T_{P'N}(R_N T_{N\,N-1} \cdots T_{21} R_1) T_{1P}$$
$$= T_{P'N} S_{N1} T_{1P} \tag{1.57}$$

where

$$S_{N1} = R_N T_{N\,N-1} \cdots T_{21} R_1 \tag{1.58}$$

is the system matrix for this generalized system. Because the R and T matrices are square matrices of four elements, S_{N1} is a square matrix of four elements; we can write it as

$$S_{N1} = \begin{pmatrix} b & -a \\ -d & c \end{pmatrix} \tag{1.59}$$

where minus signs are inserted simply to make some later equations have a better form. We call the elements a, b, c, d the Gaussian constants, and they represent the translation and refraction properties of the entire optical system of N refracting surfaces. These elements a, b, c, d are not all independent. Because S_{N1} is composed of the product of refraction and translation matrices, each of which has a determinant of unity, the determinant of S_{N1} must be unity as well. Therefore, we have the important result

$$|S_{N1}| = bc - ad = 1 \tag{1.60}$$

This result is very useful because we can use it as a check on the correctness of the a, b, c, d values in problems, and also to simplify formulas.

Just as we wrote a recursion formula for the calculation of the plane-to-plane matrix $M_{P'P}$ in Equation 1.53, we can write one for the calculation of the system matrix S_{N1}:

$$S(i) = R(i)T(i)S(i-1) \quad S(1) = R(1) \tag{1.61}$$

where i equals the number of surfaces N in the optical system, $N \geq 1$. It is important to note that the value of N can be as small as one or a very large number. But, whether N is large or small, just four numbers in the S matrix encompass the properties of the entire optical system. We note that the pattern of recursion in Equation 1.61 is not exactly the same as that for the plane-to-plane matrix $M_{P'P}$ of Equation 1.53, but it is similar.

Example 1.4.1 System matrix, numerical example. We wish to calculate the system matrix, and the Gaussian constants, of the converging lens described in Example 1.3.1. The properties of this lens are given in the table of Figure 1.30,

r (mm)	n	t (mm)
	1.000	
40.00		
	1.500	20.00
−40.00		
	1.000	

Figure 1.30

and with these values, we calculate the system matrix:

$$S_{21} = R_2 T_{21} R_1$$
$$= \begin{pmatrix} 1 & -p_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \dfrac{t'_1}{n'_1} & 1 \end{pmatrix} \begin{pmatrix} 1 & -p_1 \\ 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 0.833333 & -0.0229167 \\ 13.3333 & 0.833333 \end{pmatrix}$$

where the calculation of p_1, p_2 is given by Equation 1.39. Because this lens has two surfaces, $N = 2$; therefore, Equation 1.59 gives

$$S_{21} = \begin{pmatrix} b & -a \\ -d & c \end{pmatrix}$$

and we read off the Gaussian constants as

$$a = 0.0229167$$
$$b = 0.833333$$
$$c = 0.833333$$
$$d = -13.3333$$

We check these values using Equation 1.60, and obtain

$$bc - ad = 1$$

Again, we must remember that we are performing all our calculations with *Mathematica* to a precision of 16 decimal digits, and not simply with the six decimal digits displayed.

1.4.2 The four optical system cases

Although the numerical calculation of the Gaussian constants is useful, the concept is even more important in theoretical applications. We can obtain several extremely important results for an optical system of any number of surfaces by considering what happens to a set of rays that travels through the system: we shall give these results in terms of the Gaussian constants a, b, c, d of the system.

We begin the investigation with the diagram in Figure 1.31, where we show an arbitrary set of rays which begin at or pass through a plane P that is located at the distance ℓ from the first vertex V_1. After passing through the optical system, which can be any system of $N \geq 1$ surfaces, the set of rays emerge to reach or pass through a plane P' located a distance ℓ' from the last vertex V_N.

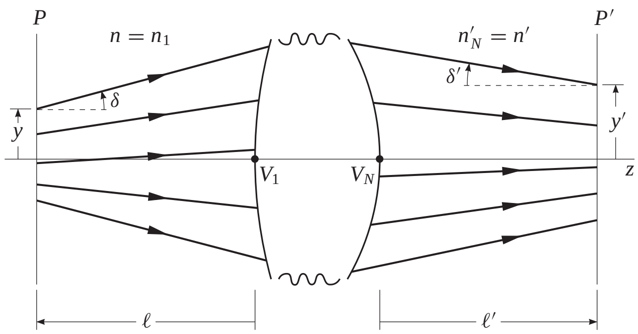


Figure 1.31 The reference diagram for ℓ and ℓ': the distance ℓ is negative and the distance ℓ' is positive by the rules of analytic geometry.

The distances ℓ and ℓ' can be positive, negative, or zero; the appropriate signs in this textbook are determined by the rules of analytic geometry. Thus, in Figure 1.31, ℓ is negative because its dimension arrow points to the left; on the other hand, ℓ' is positive for its dimension arrow points to the right—assuming the positive direction of the reference axis, here the z axis, is to the right, as is common. This sign rule for ℓ and ℓ' is the one followed by professional lens designers and is used in a variety of textbooks. The advantage is that it makes the sign rules easier to remember; the disadvantage is that some equations and definitions are less symmetric. To remove these latter difficulties, most introductory physics books that discuss optics define both ℓ and ℓ' as positive in a diagram like Figure 1.31; then authors of advanced texts in optics for college juniors and seniors feel forced to continue this practice for consistency. However, to do real work in optics, experience shows that it is best to follow the rules of analytic geometry, and practice readily removes the difficulties encountered with the lack of symmetry in some equations and definitions.

We next want to formulate the plane-to-plane matrix $M_{P'P}$ in terms of ℓ, ℓ' for rays that travel from points on the plane P to points on the plane P' rather than using t, t', as

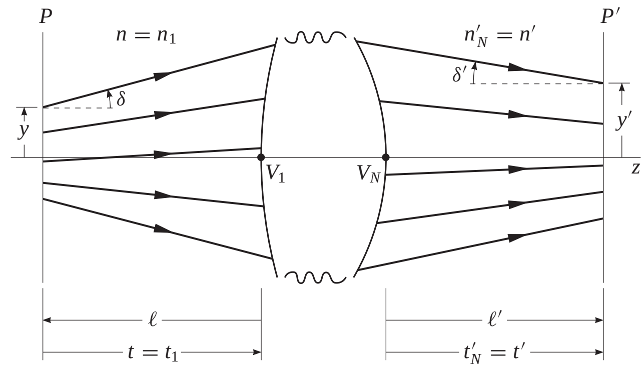


Figure 1.32 Diagram to determine the relationship between $t = t_1$ and ℓ; similarly, between $t'_N = t'$ and ℓ'.

we have done so far. To help us get the required relationships, we draw Figure 1.32: by inspection of the dimension arrows, we quickly obtain

$$\left.\begin{aligned} t &= -\ell \\ t' &= \ell' \end{aligned}\right\} \tag{1.62}$$

Therefore, using Equations 1.57 and 1.59, and substituting for t and t' with Equation 1.62, we obtain the $ABCD$ matrix:

$$\begin{aligned} M_{P'P} &= T_{P'N} S_{N1} T_{1P} \\ &= \begin{pmatrix} 1 & 0 \\ \dfrac{t'}{n'} & 1 \end{pmatrix} \begin{pmatrix} b & -a \\ -d & c \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \dfrac{t}{n} & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ \dfrac{\ell'}{n'} & 1 \end{pmatrix} \begin{pmatrix} b & -a \\ -d & c \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \dfrac{-\ell}{n} & 1 \end{pmatrix} \\ &= \begin{pmatrix} b + \dfrac{a\ell}{n} & -a \\ \dfrac{\ell'}{n'}\left(b + \dfrac{a\ell}{n}\right) - \left(d + \dfrac{c\ell}{n}\right) & c - \dfrac{a\ell'}{n'} \end{pmatrix} \\ &= \begin{pmatrix} B & A \\ D & C \end{pmatrix} \end{aligned} \tag{1.63}$$

By inspection, we see that

$$A = -a \tag{1.64a}$$

$$B = b + \frac{a\ell}{n} \tag{1.64b}$$

$$C = c - \frac{a\ell'}{n'} \tag{1.64c}$$

$$D = \frac{\ell'}{n'}\left(b + \frac{a\ell}{n}\right) - \left(d + \frac{c\ell}{n}\right) \tag{1.64d}$$

The incoming or incident rays in Figure 1.32 can be viewed as starting on the plane P, or starting at points to the left of the plane and then passing through the plane; likewise the emergent rays can stop at the plane P' or pass through. Each of the incident rays is described at the plane P by an angle of inclination δ and a coordinate y; similarly, at the plane P' each of the emergent rays is described by a δ' and y'. The ray vector V_P is composed of the $n\delta$ and y values of the rays at the plane P; likewise, the ray vector $V_{P'}$ is composed of the $n'\delta'$ and y' values of the rays at the plane P'. As we indicated in Section 1.3.4, these ray vectors are related to each other by the plane-to-plane matrix $M_{P'P}$; thus (see Equation 1.35b), $V_{P'} = M_{P'P}V_P$, or filling in the ray vectors and using Equation 1.63 to replace $M_{P'P}$,

$$\begin{pmatrix} n'\delta' \\ y' \end{pmatrix} = \begin{pmatrix} B & A \\ D & C \end{pmatrix} \begin{pmatrix} n\delta \\ y \end{pmatrix} = \begin{pmatrix} Bn\delta + Ay \\ Dn\delta + Cy \end{pmatrix} \qquad (1.65)$$

Solving for δ' and y' in Equation 1.65, we get the ray data

$$\delta' = \frac{Bn}{n'}\delta + \frac{A}{n'}y \qquad (1.66a)$$

$$y' = Dn\delta + Cy \qquad (1.66b)$$

The optical system, the planes P and P', and the set of rays in Figure 1.32 are quite general with nothing special about them. However, certain sets of rays have very interesting properties when A, B, C, or D are separately set to zero. We now want to investigate these four cases.

Case 1 ($A = 0$): The afocal or telescopic system. When we set A equal to zero—which according to Equation 1.64a, is the same as setting the Gaussian constant a to zero—we see that Equations 1.66 become

$$\delta' = \frac{bn}{n'}\delta \qquad (1.67a)$$

$$y' = Dn\delta + cy \qquad (1.67b)$$

Also, $B = b$ and $C = c$ by Equations 1.64b and 1.64c when $a = 0$. We have left D alone, since it simplifies little when a is replaced by zero in Equation 1.64d. The important equation is Equation 1.67a: It says that if the incident rays form a parallel set having an angle of inclination δ, then the emergent rays also form a parallel set, but with an angle of inclination δ' which in general is different from δ (see Figure 1.33). An optical system which transforms a set of parallel rays into another set of parallel rays is called an afocal or telescopic system—setting the Gaussian constant a to zero is all that is needed for a system to have this property. In actual practice, the incident set of parallel rays may come from a distant star, and the emergent set of parallel rays enters a human eye or other detector.

Equation 1.67b does not express a special property, it simply specifies how the y coordinate of a ray with angle of inclination δ at the plane P is transformed into a y' coordinate on the plane P', as illustrated in Figure 1.33.

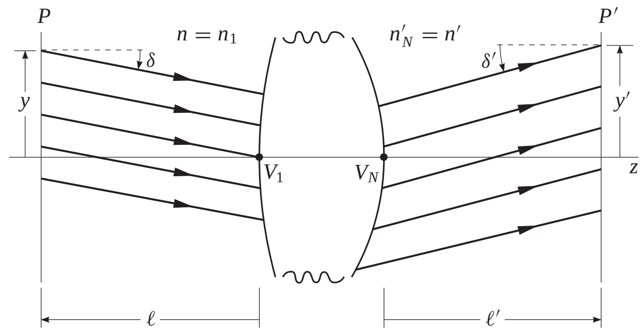


Figure 1.33

Case 2 ($B = 0$): The unprimed focal point F. When we set B equal to zero in Equation 1.64b, we obtain

$$0 = b + \frac{a\ell}{n}$$

which we quickly solve for ℓ to get

$$\ell_F = \ell = -n\frac{b}{a} \qquad (1.68)$$

where ℓ_F is the special name for this particular ℓ.

Substituting $B = 0$ into Equations 1.66, we get

$$\delta' = \frac{A}{n'}y = \frac{-a}{n'}y \qquad (1.69a)$$

$$y' = Dn\delta + Cy \qquad (1.69b)$$

where we have replaced A by $-a$, as Equation 1.64a allows.

The important equations are Equations 1.68 and 1.69a. First, Equation 1.68 says that when the plane P is located a distance ℓ_F from the vertex V_1, then $B = 0$: The plane P is called the unprimed focal plane for the optical system, as illustrated in Figure 1.34. Second, suppose every ray of a set comes from (or passes through) a point F_P on the unprimed focal plane (see Figure 1.34). Clearly, each of the rays has the same y coordinate, but different angles of inclination δ; Equation 1.69a then shows that this set of rays has the remarkable property that when it passes through the optical system, it emerges as a parallel set with the angle of inclination δ'.

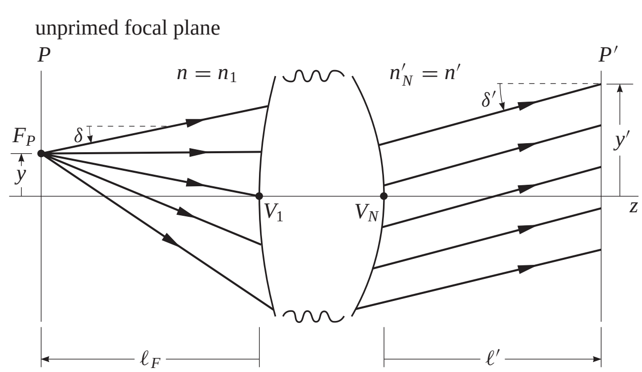


Figure 1.34

Equation 1.69b simply says that, given the δ and y values on the plane P of each ray in the set, then the value of y' can be calculated on the plane P'—of course, as Figure 1.34 shows, the value of y is the same for all the incident rays.

The diagram in Figure 1.35 illustrates what happens when y has the particular value of zero: the F_P point falls on the symmetry axis (the z axis)—we give this point the special name of the unprimed focal point F. Furthermore, Equation 1.69a says that because y is zero, the value of δ' must be zero, also; that is, the set of rays emerges from the system parallel to the symmetry axis. The unprimed focal point F is one of the important properties of an optical system.

Note that in both Figures 1.34 and 1.35 the dimension arrow for ℓ_F points to the left, which means that ℓ_F is negative, as is the case in many optical systems. However, ℓ_F is also positive in many systems (and it can be zero); we have chosen to draw the diagrams with $\ell_F < 0$ to obtain a neater looking diagram.

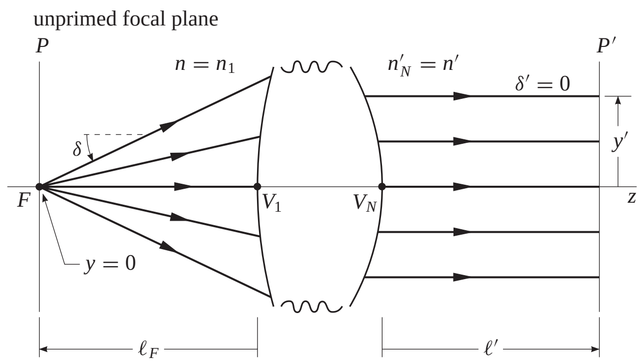


Figure 1.35

Case 3 ($C = 0$): The primed focal point F'. This case is the analog of the previous one. If we substitute $C = 0$ into Equation 1.64c, we obtain

$$0 = c - \frac{a\ell'}{n'}$$

which we solve for ℓ' to get

$$\ell'_F = \ell' = n'\frac{c}{a} \tag{1.70}$$

where we have given the special name of ℓ'_F to this particular value of ℓ'.

Next, we substitute $C = 0$ into Equations 1.66, and we obtain the equations

$$\delta' = \frac{Bn}{n'}\delta + \frac{A}{n'}y \tag{1.71a}$$

$$y' = Dn\delta = \frac{n}{a}\delta \tag{1.71b}$$

where we have replaced D in the last equation by $1/a$; this replacement is obtained when Equation 1.70 is used to substitute $n'c/a$ for ℓ' in Equation 1.64d, and then simplifying with the help of the $bc - ad = 1$ expression of Equation 1.60.

The expressions that play the important role in this case are Equations 1.70 and 1.71b. First, as we illustrate in Figure 1.36, Equation 1.70 says that the plane P' should be positioned a distance ℓ'_F from the vertex V_N; the plane P' is then given the special name of the primed focal plane. Second, consider a set of parallel rays, each of which has the same angle of inclination δ but different coordinates y on the plane P, as shown in Figure 1.36. This set of rays, according to Equation 1.71b, emerges from the Nth surface with angles of inclination δ' such that all the rays pass through (or stop at) a point $F_{P'}$ with coordinate y' on the primed focal plane (see Figure 1.36). The reason Equation 1.71b predicts this behavior is that there is no y in this equation, only δ, and δ has the same value for all the incident rays.

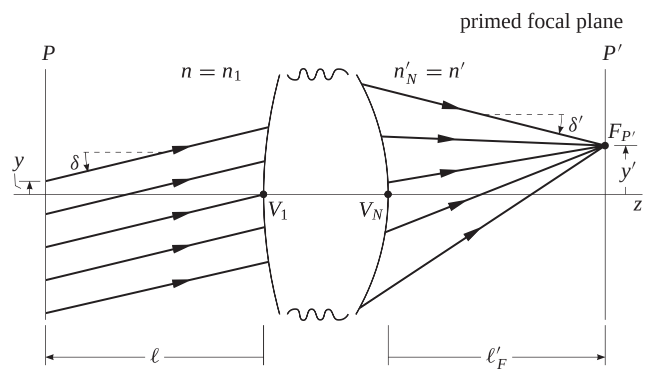


Figure 1.36

The equation we must discuss yet is Equation 1.71a. By inspection, we see that it gives the δ' that the rays have when they emerge from the Nth surface in terms of the incident ray values of δ and y The y values are the ones the rays have on the plane P, as indicated in Figure 1.36.

When δ is set to zero, we see that Equation 1.71b predicts that $y' = 0$ on the primed focal plane, and we obtain the diagram in Figure 1.37. This point is on the symmetry axis and is given the special name of the primed focal point F'. Just like the unprimed focal point F, it is another of the special points which helps summarize the properties of an optical system.

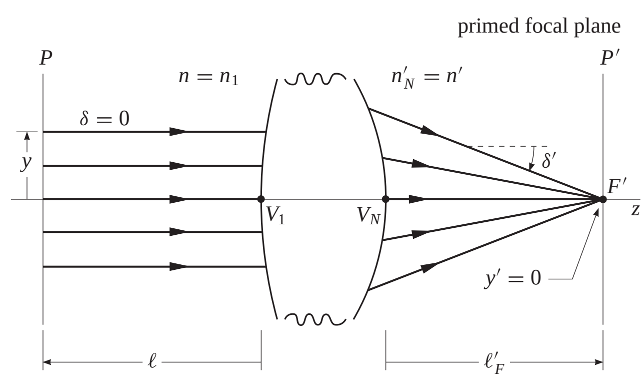


Figure 1.37

Case 4 ($D = 0$): The object-image condition. In this final case, we consider one of the most important properties of an optical system, namely, its ability to form an image of an object. Substituting $D = 0$ into Equation 1.64d, we obtain

$$0 = \frac{\ell'}{n'}\left(b + \frac{a\ell}{n}\right) - \left(d + \frac{c\ell}{n}\right) \tag{1.72}$$

which is called the object-image condition. Solving Equation 1.72 for ℓ', we get

$$\ell' = n'\left(\frac{d + \frac{c\ell}{n}}{b + \frac{a\ell}{n}}\right) \tag{1.73}$$

It is useful to give this equation a special name, also: we call it the ℓ, ℓ' relation. When ℓ and ℓ' are related by Equation 1.73, the distance ℓ is called the object distance, and the distance ℓ' is called the image distance.

To obtain additional information on the object-image properties, we substitute $D = 0$ into Equations 1.66 to get

$$\delta' = \frac{Bn}{n'}\delta + \frac{A}{n'}y \tag{1.74a}$$

$$y' = Cy \tag{1.74b}$$

The important equation is the last one, Equation 1.74b. First of all, because Equation 1.64c says that $C = c - a\ell'/n'$, we see that C is a constant once ℓ' has been determined from the ℓ, ℓ' relation of Equation 1.73. Next, we observe that there is no δ in Equation 1.74b, which means that all rays of a set emanating from a point of coordinate y on the P plane (no matter what the value of δ is) will pass through another point of coordinate y' on the P' plane, as shown in Figure 1.38. In other words, rays incident on an optical system from an object point are brought to a focus at an image point by the system. The point on the P plane is called the object point O, and the plane itself is called the object plane; similarly, the point on the P' plane is called the image point I, and the plane is called the image plane. Collectively, the object-image points are called conjugate points.

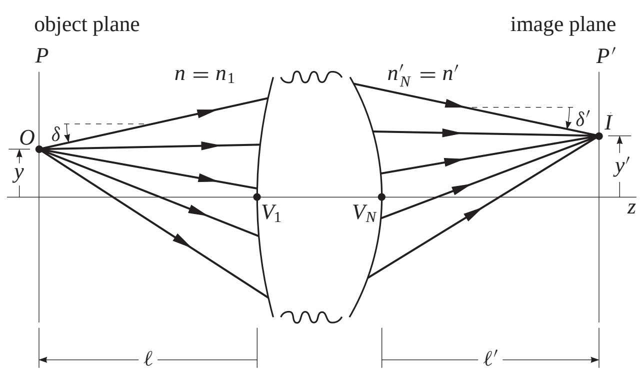


Figure 1.38

The other equation of our pair, namely Equation 1.74a, simply gives the δ' of an emergent ray in terms of the corresponding incident ray δ and y (see Figure 1.38).

Suppose there are several object points on the object plane, as shown in Figure 1.39, with coordinates y_1, y_2, and y_3, where $y_1 = 0$. First, because the ℓ, ℓ' relation in Equation 1.73 is independent of y, the image distance ℓ' is the same for all the points on the object plane. Therefore, all image points corresponding to the points on the object plane will be on the same image plane. Second, because C is a constant in Equation 1.74b, the y' of an image point is directly proportional to the y of an object point, and so the image points may be arranged on the image plane as shown in Figure 1.39. Also, since $y_1 = 0$, we must have $y_1' = 0$; that is, an on-axis object point must have a corresponding on-axis image point. In the diagram of Figure 1.39, we have assumed that C is positive, so that the y' coordinates are positive when the y coordinates are; when C is negative, this relationship changes to make negative y' values correspond to positive y values.

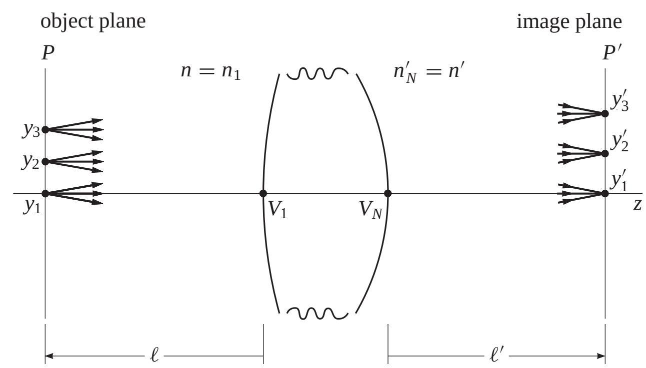


Figure 1.39

Now suppose that instead of just the three object points we used in Figure 1.39, we imagine there are a large number of points infinitesimally close together from the bottom to top point; then we get a continuous object O that produces a continuous image I, as shown in Figure 1.40. We have added arrowheads to make it easy to determine the object or image orientation, whether erect or inverted. It is convenient to

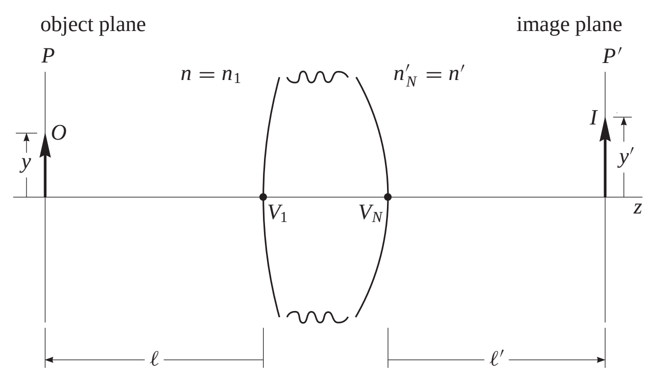


Figure 1.40

describe a continuous object and image in terms of the y, y' coordinates of their respective arrowhead points, as shown in the diagram of Figure 1.40. We can then use the value of y to represent the object height, and the value of y' to represent the image height.

Observe in Figure 1.38 that the rays actually emanate from the object point and fall on the image point; the same property is true for all the points making up the object and image in Figure 1.40—we call such objects and images real. We observe from these diagrams that this property is obtained when $\ell < 0$ and $\ell' > 0$.

However, objects and images are not always real. In Figure 1.41, we show a real object and a virtual image; that is, an image that is to the left of last surface and occurs when $\ell' < 0$. For such an image, the rays do not actually pass through it, but appear to come from it, as illustrated by the dashed lines which are backward extensions of the rays.

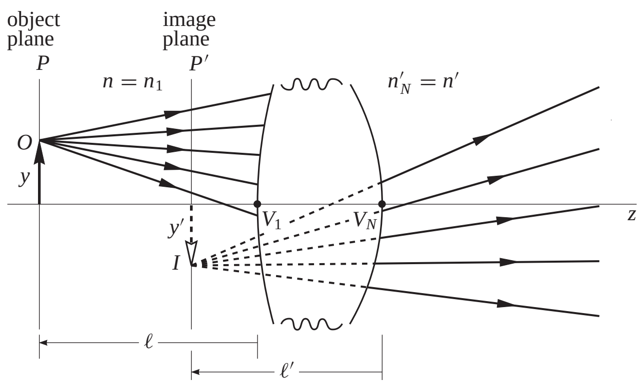


Figure 1.41

An object can also be virtual, as shown in Figure 1.42. As the diagram indicates, a virtual object has the property that $\ell > 0$. Such an object is produced by another optical system somewhere to the left of the diagram: As the dashed lines show, the rays are focused on the arrowhead of the object, but never reach it because the optical system gets in the way. Nevertheless, a virtual object produces an image—a real and erect one in the diagram; however, it could be a virtual image, it all depends on the properties of the optical system.

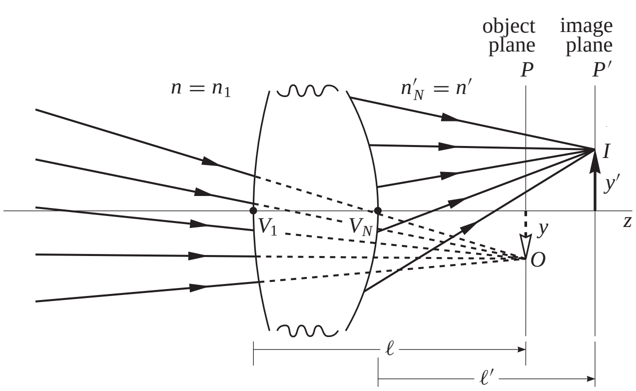


Figure 1.42

We can tell when an object or an image is real or virtual by looking at the sign of ℓ or ℓ', respectively. Similarly, the respective sign of y or y' indicates when an object or image is erect or inverted. The table in Figure 1.43 summarizes these properties.

	real	**virtual**	**erect**	**inverted**
object	$\ell < 0$	$\ell > 0$	$y > 0$	$y < 0$
image	$\ell' > 0$	$\ell' < 0$	$y' > 0$	$y' < 0$

Figure 1.43

Sometimes we find it useful to talk about the space that objects occupy as object space; similarly, the space occupied by images, as image space. Since objects and images can be located anywhere in the ranges of $-\infty < \ell < \infty$ and $-\infty < \ell' < \infty$, respectively, object and image space occupy the same physical space. This property can appear confusing, but it simply means that when we talk about object space, we are talking about objects; similarly, when we talk about image space, images.

1.4.3 The transverse magnification m_T

To compare the image height to that of the object, we define the transverse or lateral magnification as

$$m_T = \frac{y'}{y} \tag{1.75}$$

With this definition, we can determine by the sign of m_T whether the image points in the same direction as the object, or is inverted. For example, if the object is erect so that $y > 0$, then $m_T > 0$ means that the image is erect, also; if $m_T < 0$, then the image is inverted.

We can write the m_T definition of Equation 1.75 in terms of the Gaussian constants by first using Equation 1.74b and then Equation 1.64c to obtain

$$m_T = \frac{y'}{y} = C = c - \frac{a\ell'}{n'} \tag{1.76}$$

If we substitute into Equation 1.76 the expression for ℓ' from the ℓ, ℓ' relation of Equation 1.73, rearrange algebraically, and use the $bc - ad = 1$ relationship of Equation 1.60, we obtain another expression for m_T:

$$m_T = \frac{y'}{y} = C = \frac{1}{b + \dfrac{a\ell}{n}} = \frac{1}{B} \tag{1.77}$$

where in the last step we have used Equation 1.64b. By inspection of Equation 1.77, we see that

$$B = \frac{1}{m_T} \tag{1.78a}$$

and

$$C = m_T \tag{1.78b}$$

1.4.4 The object-image matrix

We can make a useful summary of the properties of the object-image relationship that we have just discussed in terms of a matrix. Recall that Equation 1.63 says that the plane-to-plane matrix from points on a plane P to points on a plane P' is given by the $ABCD$ matrix:

$$M_{P'P} = \begin{pmatrix} B & A \\ D & C \end{pmatrix} \tag{1.79}$$

This matrix takes on a special form when the plane P is an object plane, and the plane P' is the corresponding image plane. As we have shown in Case 4 starting with Equation 1.72, the object-image relationship arises when we require $D = 0$, and it then follows that B and C are related to the transverse magnification m_T by Equations 1.78a and 1.78b. Finally, the $A = -a$ relationship of Equation 1.64a is true in general. Substituting these expressions for A, B, C, and D into Equation 1.79, we get

$$M_{IO} = \begin{pmatrix} 1/m_T & -a \\ 0 & m_T \end{pmatrix} \tag{1.80}$$

where this matrix is called the object-image matrix, and is represented by the symbol M_{IO}. Therefore, instead of writing $V_{P'} = M_{P'P} V_P$, we can now write for the particular case of object-image ray vectors

$$V_I = M_{IO} V_O \tag{1.81}$$

to describe the rays traveling from the object plane to the image plane.

Example 1.4.2 The system matrix, the Gaussian constants, and an object, image pair.
We consider again the equiconvex lens of Example 1.3.1, and redraw the diagram for this lens in Figure 1.44. We list its properties below the diagram with $-\ell$ and ℓ' put in the t column to represent the object and image distances as t quantities.

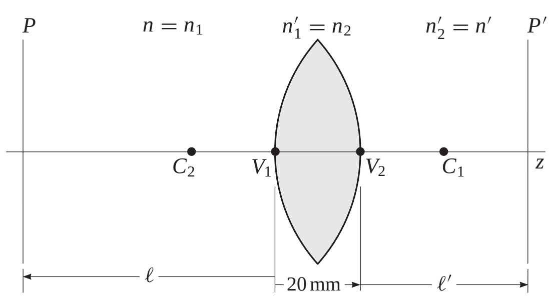


Figure 1.44

r (mm)	n	t (mm)
	1.000	$-\ell$
40.00		
	1.500	20.00
−40.00		
	1.000	ℓ'

We have already calculated the system matrix and the Gaussian constants a, b, c, d in Example 1.4.1; for convenience, we list these constants again:

$$a = 0.0229167$$
$$b = 0.833333$$
$$c = 0.833333$$
$$d = -13.3333$$

Now suppose an object is positioned so that the object distance is $\ell = -60.00$ mm. To find the corresponding image distance ℓ', we substitute the above Gaussian constants and the indices of refraction $n = n' = 1$ into the ℓ, ℓ' relation of Equation 1.73:

$$\begin{aligned} \ell' &= n' \left(\frac{d + \dfrac{c\ell}{n}}{b + \dfrac{a\ell}{n}} \right) \\ &= (1) \left(\frac{-13.3333 + \dfrac{(0.833333)(-60)}{1}}{0.833333 + \dfrac{(0.0229167)(-60)}{1}} \right) \\ &= 116.92\,\text{mm} \end{aligned} \tag{1.82}$$

With Equation 1.76 we calculate the transverse magnification

$$\begin{aligned} m_T &= c - \frac{a\ell'}{n'} = 0.833333 - \frac{(0.0229167)(116.92)}{1} \\ &= -1.846 \end{aligned} \tag{1.83}$$

If we had used $m_T = 1/(b + a\ell/n)$ in Equation 1.77, we would have obtained the same result. Because m_T is negative, the image is inverted relative to the object.

If the object height is chosen as

$$y = 12.00\,\text{mm} \tag{1.84a}$$

then the image height y' is calculated with the help of Equation 1.75 to be

$$y' = m_T y = (-1.8462)(12.00) = -22.15\,\text{mm} \tag{1.84b}$$

In Figure 1.45, we make a scale drawing of the object and the corresponding image produced by the lens using the information gleaned from our calculations; we also trace five rays from the object arrowhead through the lens to the image plane using the paraxial methods for ray tracing we illustrated in Examples 1.3.1 to 1.3.5. We note that all the rays intersect at the image arrowhead; in fact, any ray leaving the arrowhead of the object has this property. Rays drawn from any of the

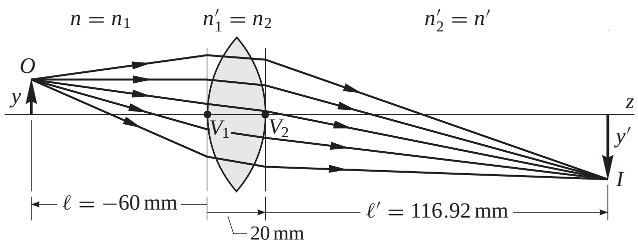


Figure 1.45 The object height is $y = 12.00$ mm; the image height is $y' = -22.15$ mm.

other points making up the continuous object would intersect at corresponding image points to produce a continuous image; we don't show this property for simplicity, but understand that it happens.

Observe in Figure 1.45 that instead of drawing the rays to the lens surfaces, we drew them to vertical planes positioned at V_1 and V_2 (see Figure 1.27). This practice was followed because many of the rays drawn do not satisfy the paraxial approximation as given in Section 1.3.2, which said that paraxial rays should travel near the symmetry axis; also angles, such as the angle of inclination δ, should be small. The main problem is that in the translation matrix the thickness of the lens measured from V_1 to V_2 is used as the translation distance of the ray through the lens, even though some of the rays are quite far from the symmetry axis. For such rays it is more correct to draw them to vertical planes that are spaced by the thickness of the lens, as we first discussed in Example 1.3.5. However, to obtain simpler diagrams, we usually do not draw these vertical planes.

It is important to note that the rules of paraxial mathematics that we developed in Section 1.3 give consistent results even for rays that are no longer paraxial; for example, such rays in Figure 1.45 still pass through the image arrowhead just like the paraxial rays do. Sometimes, it is advantageous to draw rays that are no longer truly paraxial to obtain a diagram that better shows the property being illustrated—in Figure 1.45, it is the object-image property. Again, it should be emphasized that in drawing these rays, we used the matrix ray-tracing method of Examples 1.3.1 to 1.3.5 with the help of *Mathematica* to make all the calculations; we don't show all the steps for lack of space.

Finally, we would like to illustrate what the object-image matrix M_{IO} in Equation 1.80 can do. To trace a ray through the system in Figure 1.45, we can write

$$V_I = T_{I2} R_2 T_{21} R_1 T_{1O} V_O \tag{1.85}$$

which says that to transform the elements of V_O into those of V_I, five square matrices $T_{I2} \ldots T_{1O}$ must be multiplied together. The product of these five square matrices is simply the matrix M_{IO} of the four elements given in Equation 1.80:

$$M_{IO} = \begin{pmatrix} 1/m_T & -a \\ 0 & m_T \end{pmatrix} \tag{1.86}$$

As a numerical example, let's consider the ray in Figure 1.45 that leaves the object arrowhead parallel to the symmetry axis; that is, $\delta = 0$ rad and $y = 12.00$ mm. Then using the m_T and a values we got earlier in this example, we have

$$V_I = M_{IO} V_O$$

$$\begin{aligned} \begin{pmatrix} n'\delta' \\ y' \end{pmatrix} &= \begin{pmatrix} 1/m_T & -a \\ 0 & m_T \end{pmatrix} \begin{pmatrix} n\delta \\ y \end{pmatrix} \\ &= \begin{pmatrix} 1/(-1.8462) & -0.0229167 \\ 0 & -1.8462 \end{pmatrix} \begin{pmatrix} (1)(0) \\ 12 \end{pmatrix} \\ &= \begin{pmatrix} -0.2750 \\ -22.15 \end{pmatrix} \end{aligned} \tag{1.87}$$

We read off $y' = -22.15$ mm, which agrees with the image height we obtained before in Equation 1.84b. We also see that $n'\delta' = -0.2750$; since $n' = 1$, we easily obtain $\delta' = -0.2750$ rad ≈ -16 deg. We check this value for δ' by looking at Figure 1.45 and then concentrate on the incident ray that travels parallel to the symmetry axis. We see that it emerges from the lens with a downhill slope as it travels to the image arrowhead, which appears consistent with the -16 deg value; in fact, a small protractor can be used to check this value.

1.4.5 A paradox and its resolution

Because the $ABCD$ matrix

$$\begin{pmatrix} B & A \\ D & C \end{pmatrix}$$

which we described in Equation 1.63 is composed of translation and refraction matrices, its determinant must equal one:

$$BC - AD = 1 \tag{1.88}$$

This requirement would appear to mean that certain quantities can't be zero at the same time; for example, C and D should not both be zero simultaneously. However, consider an on-axis point object O that has the corresponding image I, as shown in Figure 1.46. As we stated in Equation 1.72, the

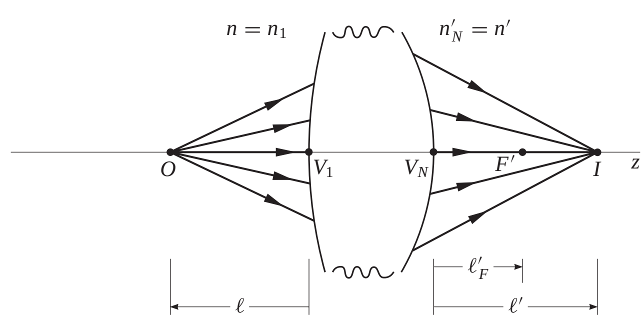


Figure 1.46

quantities ℓ, ℓ' in an object-image relationship must obey the object-image condition, namely,

$$D = \frac{\ell'}{n'}\left(b + \frac{a\ell}{n}\right) - \left(d + \frac{c\ell}{n}\right) = 0 \tag{1.89}$$

which has the solution given in Equation 1.73 as

$$\ell' = n'\left(\frac{d + \frac{c\ell}{n}}{b + \frac{a\ell}{n}}\right) \tag{1.90}$$

where we called it the ℓ, ℓ' relation.

Now suppose we maintain the object-image relationship, but move the object to the left towards $-\infty$; that is, we let $\ell \to -\infty$. To determine what happens to the image distance ℓ', we divide numerator and denominator by ℓ inside the parentheses in Equation 1.90, and then let $\ell \to -\infty$:

$$\ell' = n' \lim_{\ell \to -\infty}\left(\frac{\frac{d}{\ell} + \frac{c}{n}}{\frac{b}{\ell} + \frac{a}{n}}\right) = n'\frac{c}{a} \tag{1.91}$$

agreeing with Equation 1.70, which says that $\ell'_F = n'c/a$; thus, the image point I approaches the primed focal point F' as the object point O moves to $-\infty$, as we diagram in Figure 1.47. This diagram also shows that as O goes to $-\infty$, the rays that are incident on the first surface approach a direction that is parallel to the symmetry axis, which is consistent with the optical system focusing these rays at F'. However, incident parallel rays being focused to F' is a Case 3 situation

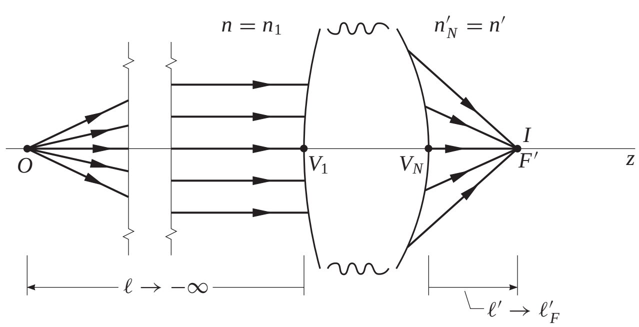


Figure 1.47

for which $C = 0$, as our earlier derivation of Equation 1.70 showed (see also Figure 1.37). Now we have reached our paradox where C and D are both zero simultaneously, which appears to violate Equation 1.88.

In drawing the above diagrams in Figures 1.46 and 1.47, we always maintained the object-image relationship, so certainly $D = 0$, and therefore Equation 1.88 becomes

$$BC = 1 \tag{1.92}$$

Let's take a closer look at B and C. According to Equations 1.64b and 1.64c, we have $B = b + a\ell/n$ and $C = c - a\ell'/n'$. First, in the B equation, a is nonzero because we are not considering an afocal system (see Case 1 and Figure 1.33); also, b has some fixed value that depends on the properties of the optical system. Therefore, as $\ell \to -\infty$, we must have $B \to \pm\infty$, where the $\pm$ sign depends on the sign of a. Second, as we showed in Equation 1.91, when $\ell \to -\infty$, we have $\ell' \to n'c/a$, and when we substitute this result for ℓ' into the C equation, we obtain $C \to 0$. We have reached the interesting situation where $BC \to \pm\infty \cdot 0$, which is indeterminate; that is, this product might give a finite result. In fact, if we multiply BC together before we let ℓ go to $-\infty$, and also substitute the expression for ℓ' given in Equation 1.90, we obtain

$$\begin{aligned} BC &= \left(b + \frac{a\ell}{n}\right)\left(c - \frac{a\ell'}{n'}\right) \\ &= \left(b + \frac{a\ell}{n}\right)\left(c - a\,\frac{d + \frac{c\ell}{n}}{b + \frac{a\ell}{n}}\right) \\ &= bc - ad = 1 \end{aligned} \tag{1.93}$$

That is, when we perform the algebraic multiplication in Equation 1.93 and simplify, we find that all the terms containing ℓ cancel leaving the expression $bc - ad$, which has the value of one by Equation 1.60; thus, the product BC always has the desired value of one no matter what the value of ℓ is as we move from small values of ℓ to the large negative ones of $-\infty$, as diagrammed in Figures 1.46 and 1.47.

A similar argument can be made when the Case 2 condition of $B = 0$ holds, and $C \to \pm\infty$ as $\ell' \to \infty$. In this situation, the object point O moves to the right toward the unprimed focal point F (assumed to be on the left of the optical system shown in Figures 1.46 and 1.47) and the image I moves toward $+\infty$.

1.4.6 The longitudinal magnification m_L

We have already defined the transverse or lateral magnification m_T, which is a comparison of the image and object heights normal to the symmetry axis. The longitudinal magnification m_L compares image and object lengths on or parallel to the symmetry axis. Because of the nonlinear nature of m_L, the definition is given in terms of a derivative:

$$m_L = \frac{d\ell'}{d\ell} \tag{1.94}$$

To obtain a more useful expression for m_L, we start with the ℓ, ℓ' relation of Equation 1.90 (first defined in Equation 1.73):

$$\ell' = n'\left(\frac{d + \frac{c\ell}{n}}{b + \frac{a\ell}{n}}\right) \tag{1.95}$$

All quantities in Equation 1.95 are fixed for a given optical system except ℓ and ℓ'. Therefore, we can view Equation 1.95 as a function of ℓ' in terms of ℓ, and we can differentiate ℓ' with respect to ℓ to obtain an expression for m_L. We apply the quotient rule, manipulate algebraically, and finally using the $bc - ad = 1$ relation of Equation 1.60 and the expression for m_T in Equation 1.77, we obtain

$$\begin{aligned} m_L &= \frac{d\ell'}{d\ell} \\ &= n' \, \frac{(b + a\ell/n)(c/n) - (d + c\ell/n)(a/n)}{(b + a\ell/n)^2} \\ &= \frac{n'}{n} \, \frac{1}{(b + a\ell/n)^2} \\ &= \frac{n'}{n} \, m_T^2 \end{aligned} \tag{1.96}$$

Equation 1.96 shows that m_L is always positive so that $d\ell$ and $d\ell'$ always point in the same direction. Assuming that $d\ell$ and $d\ell'$ are both positive, we draw the diagram in Figure 1.48 to illustrate the meaning of m_L. If we think of $d\ell$ as representing an object point moving from the tail to the head of $d\ell$, we see that the corresponding image point moves from the tail to the head of $d\ell'$; that is, object and image always move in the same direction.

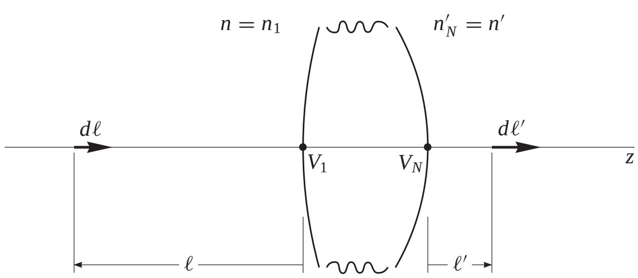


Figure 1.48

1.5 The Cardinal Points and Planes

In the previous section, we have seen how the system matrix composed of four numbers, the Gaussian constants, described the properties of an optical system. The optical system could be quite simple, consisting of a single refracting surface or a single lens, or it could be quite elaborate, made of many lenses, close together or far apart. Another way to describe an optical system is in terms of six points, called the cardinal points. These points, and their corresponding vertical planes, completely characterize the paraxial properties of an optical system, and also allow the system to be described in a way that is similar to the thin lens description usually presented in a beginning physics course. These six points come in three pairs: the focal points, the unit or principal points, and the nodal points. They are quite easily described in terms of the Gaussian constants a, b, c, d, and the first and last indices of refraction n, n'.

1.5.1 The focal points, F and F'

We have already described these two important points as Cases 2 and 3 in Section 1.4.2. We will simply summarize the results here. In Equation 1.68, we found that the distance of the unprimed focal point F from the first vertex V_1 was

$$\ell_F = -n\frac{b}{a} \tag{1.97}$$

as shown in Figure 1.49, where a and b are Gaussian constants of the optical system. The perpendicular plane through F is called the unprimed focal plane. As the diagram indicates, the property of F is that rays traveling from this point move through the optical system in such a way as to emerge traveling parallel to the symmetry axis. Actually, the situation is more general than Figure 1.49 indicates, where the rays emanate from (or pass through) F, for the focal point F can be positioned to the right of the optical system so that the forward extension of the incident rays pass through F—we will illustrate this feature in later examples.

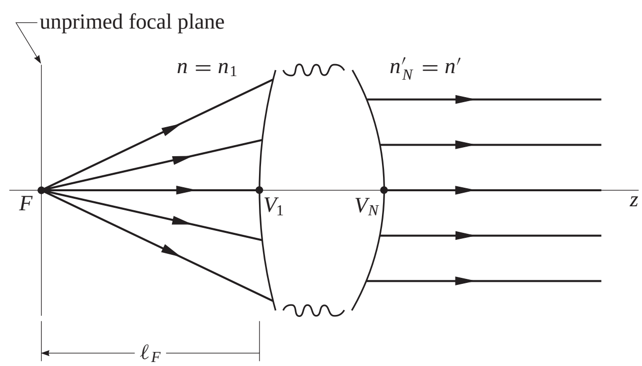


Figure 1.49

Similar to the definition we gave for the unprimed focal point F, we summarize the definition of the primed focal point F'. According to Equation 1.70,

$$\ell'_F = n'\frac{c}{a} \tag{1.98}$$

The F' point has the property that incident rays parallel to the symmetry axis emerge from the optical system to pass through it, as Figure 1.50 shows.

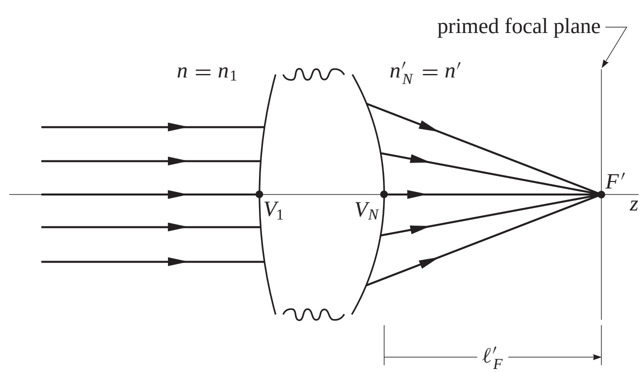


Figure 1.50

1.5.2 The unit or principal points, H and H'

The unit or principal points are a pair of points, H and H', on the symmetry axis such that an object located at point H produces an image located at point H' with the property that the transverse magnification m_T is unity—that is, the object and image have the same height. To find the object distance ℓ_H that locates H, we use the expression for m_T given in terms of the Gaussian constants in Equation 1.77 and set it equal to one to obtain

$$m_T = \frac{1}{b + \dfrac{a\ell_H}{n}} = 1 \tag{1.99}$$

where we have replaced the general ℓ with the particular ℓ_H. Solving Equation 1.99 for ℓ_H gives

$$\ell_H = -n\frac{b-1}{a} \tag{1.100}$$

In a similar manner, to find the image distance ℓ'_H that locates H', we set Equation 1.76 equal to one and replace the general ℓ' by the particular ℓ'_H:

$$m_T = c - \frac{a\ell'_H}{n'} = 1 \tag{1.101}$$

From this equation, we obtain

$$\ell'_H = n'\,\frac{c-1}{a} \tag{1.102}$$

To illustrate the property of the H, H' principal points, we draw the diagram in Figure 1.51; the vertical planes through these points are called the principal planes. The diagram shows that an object of height y on the H principal plane is reproduced by the optical system as an image of the same height on the H' principal plane.

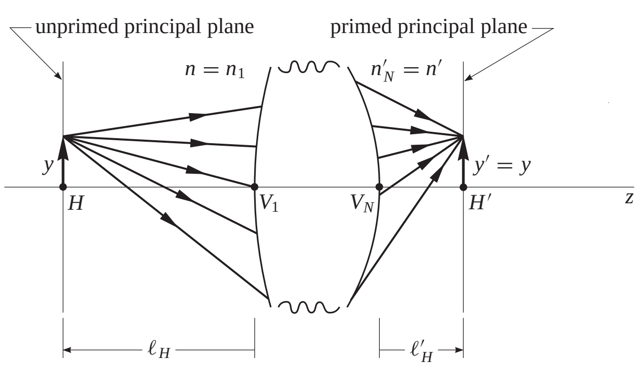


Figure 1.51

The concept of rays traveling from one principal plane to another is more general than we might first think. As the diagram in Figure 1.51 indicates, all rays leaving an object point pass through the corresponding image point at the same height above the symmetry axis. This idea means that any incident ray passing through a point P on the H principal plane must emerge from the optical system to pass through a point P' on the H' principal plane at the same height above the symmetry axis (see Figure 1.52). What is confusing sometimes is that for this property to be true in the more general cases, we must work with the forward or backward extensions of the incident and emergent rays, as we shall show in examples.

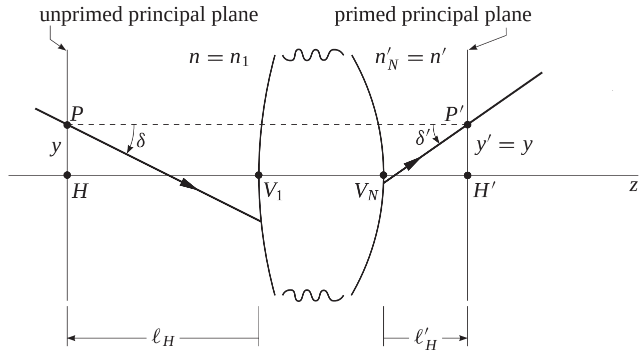


Figure 1.52

We obtain another important relationship by using the object-image matrix. For an object-image pair on the principal planes, the equation that governs the relationship between the ray vector V_O and the corresponding ray vector V_I is given by Equation 1.81 as

$$V_I = M_{IO} V_O \tag{1.103a}$$

Filling in these matrices, as we did at the beginning of Equation 1.87, we have

$$\begin{pmatrix} n'\delta' \\ y' \end{pmatrix} = \begin{pmatrix} 1/m_T & -a \\ 0 & m_T \end{pmatrix} \begin{pmatrix} n\delta \\ y \end{pmatrix} \tag{1.103b}$$

or, because $m_T = 1$ for an object-image pair on principal planes,

$$\begin{pmatrix} n'\delta' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & -a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} n\delta \\ y \end{pmatrix} \tag{1.103c}$$

which yields

$$\delta' = \frac{n}{n'}\delta - \frac{a}{n'}y \tag{1.104a}$$

$$y' = y \tag{1.104b}$$

Equation 1.104b simply says the the object and image points have the same y coordinates on principal planes. But we get new information from Equation 1.104a; namely, that when we set $y = 0$, we have

$$\delta' = \frac{n}{n'}\,\delta \tag{1.105}$$

which says that δ' will have the same sign as δ for rays passing through the H, H' points, as we illustrate in Figure 1.53.

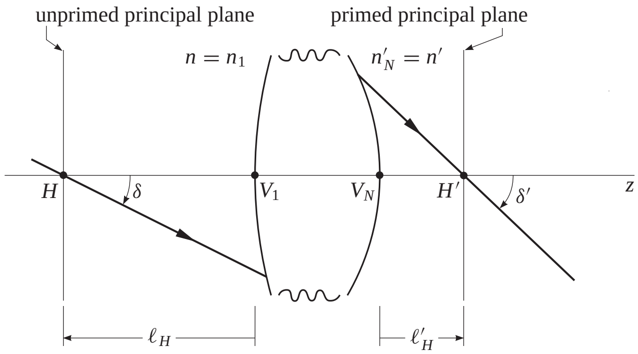


Figure 1.53

Even more interesting is when the same medium is on either side of the optical system, for then $n' = n$ and Equation 1.105 becomes

$$\delta' = \delta \tag{1.106}$$

and we have the diagram in Figure 1.54. Since we usually work with such optical systems (most often, in air) the behavior of rays passing through H and H' with $\delta' = \delta$, as illustrated in Figure 1.54, is very useful for ray tracing.

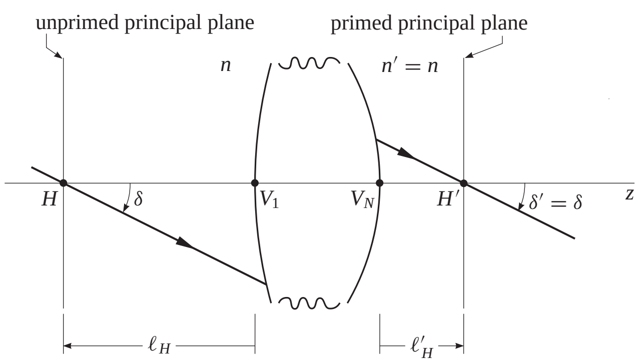


Figure 1.54

1.5.3 The nodal points, N and N'

Derivation of equations. Consider a ray passing through an object point O on the symmetry axis, traversing an optical system, and then passing through the corresponding image point I, as shown in Figure 1.55. The properties of such a ray are described by the same Equations 1.103 we used in Section 1.5.2, except m_T is no longer equal to one. Thus, we have

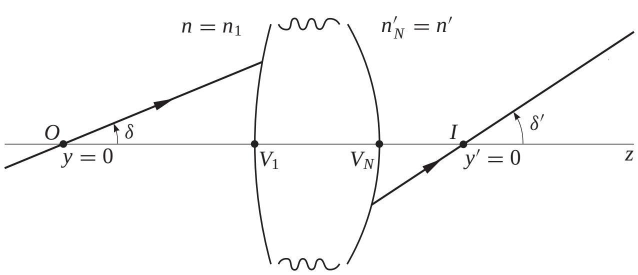


Figure 1.55

$$V_I = M_{IO} V_O \tag{1.107a}$$

$$\begin{pmatrix} n'\delta' \\ y' \end{pmatrix} = \begin{pmatrix} 1/m_T & -a \\ 0 & m_T \end{pmatrix} \begin{pmatrix} n\delta \\ y \end{pmatrix} \tag{1.107b}$$

which yields

$$\delta' = \frac{n}{n'\, m_T}\delta - \frac{a}{n'} y \tag{1.108a}$$

$$y' = m_T\, y \tag{1.108b}$$

Because $y = 0$ in Figure 1.55, Equations 1.108 become

$$\delta' = \frac{n}{n'\, m_T}\delta \tag{1.109a}$$

$$y' = 0 \tag{1.109b}$$

We define the Nodal points N, N' as the object-image points on the symmetry axis when (see Figure 1.56)

$$\delta' = \delta \tag{1.110}$$

which happens, according to Equation 1.109a, when

$$\frac{n}{n'\, m_T} = 1 \tag{1.111}$$

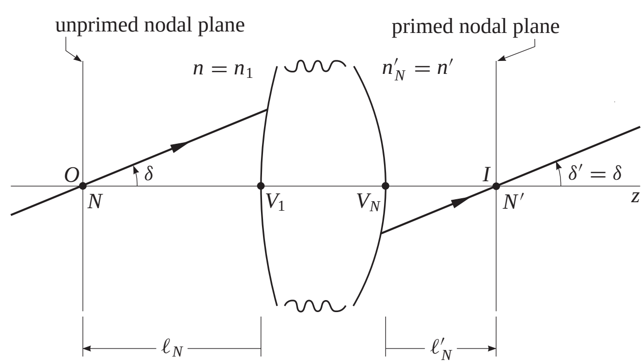


Figure 1.56

We can find the distance ℓ_N (see Figure 1.56) of the nodal point N from the vertex V_1 by substituting into Equation 1.111 the expression for m_T given in Equation 1.77, and then solve for ℓ_N:

$$\frac{n}{n'}\left(b + \frac{a\ell_N}{n}\right) = 1$$

$$\ell_N = -n\,\frac{b - n'/n}{a} \tag{1.112}$$

In a similar way, we find ℓ'_N (see Figure 1.56) by substituting Equation 1.76 into Equation 1.111:

$$\frac{n}{n'\left(c - \dfrac{a\ell'_N}{n'}\right)} = 1$$

$$\ell'_N = n' \frac{c - n/n'}{a} \tag{1.113}$$

To compare the emergent angle of inclination δ' with the incident angle δ, we define the m_A, namely,

$$m_A = \frac{\delta'}{\delta} \tag{1.114a}$$

By Equation 1.109a, we see that

$$m_A = \frac{n}{n' m_T} \tag{1.114b}$$

The nodal points and the principal points are closely related. In Figure 1.57, we position both H and H' the same distance to the right of N and N', respectively. We show that the separation between the H, N points is the same as that between the H', N' points by first using Equations 1.100 and 1.112 to find

$$\begin{aligned}\ell_H - \ell_N &= -n\frac{b-1}{a} + n\,\frac{b - n'/n}{a}\\ &= \frac{n - n'}{a}\end{aligned} \tag{1.115a}$$

and then using Equations 1.102 and 1.113 to get

$$\begin{aligned}\ell'_H - \ell'_N &= n'\,\frac{c-1}{a} - n'\,\frac{c - n/n'}{a}\\ &= \frac{n - n'}{a}\end{aligned} \tag{1.115b}$$

Each difference gives the same result; therefore, these spacings are the indeed the same. Also, by looking at the results, we see that if both $a > 0$ and $n > n'$ (or $a < 0$ and $n < n'$), then the principal points are to the right of their corresponding nodal points, as we have drawn the diagram in Figure 1.57; if the inequalities are opposite, then the reverse is true.

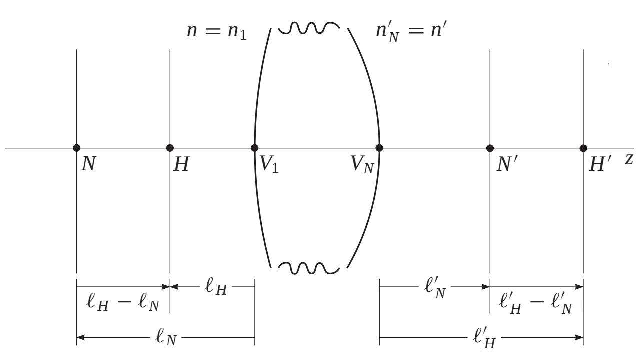


Figure 1.57

Because the separations $\ell_H - \ell_N$ and $\ell'_H - \ell'_N$ are equal, the distance $\ell_{HH'}$ between the principal points equals the

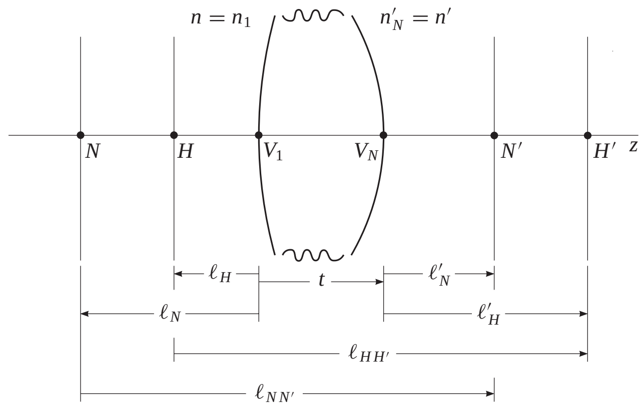


Figure 1.58

distance $\ell_{NN'}$ between the nodal points. With the help of Figure 1.58, we find the separation of the principal points using Equations 1.100 and 1.102

$$\begin{aligned}\ell_{HH'} &= -\ell_H + t + \ell'_H\\ &= t + \frac{n(b-1) + n'(c-1)}{a}\end{aligned} \tag{1.116a}$$

and the separation of the nodal points with Equations 1.112 and 1.113

$$\begin{aligned}\ell_{NN'} &= -\ell_N + t + \ell'_N\\ &= t + \frac{n(b-1) + n'(c-1)}{a}\end{aligned} \tag{1.116b}$$

The results are clearly the same; therefore, these distances must be equal.

Finally, we observe that if the same medium is on either side of the optical system—that is, if $n = n'$—then the nodal point distances in Equations 1.112 and 1.113 are

$$\ell_N = -n\,\frac{b-1}{a} \tag{1.117a}$$

and

$$\ell'_N = n'\,\frac{c-1}{a} \tag{1.117b}$$

which are the same as the principal point distances ℓ_H, ℓ'_H in Equations 1.100 and 1.102. Therefore, when the same medium is on either side of the optical system, the nodal points (and planes) become identical with the principal points (and planes). Also, Equation 1.114b becomes simply

$$m_A = \frac{1}{m_T} \tag{1.118}$$

that is, the ray-angle magnification is the reciprocal of the transverse magnification when $n = n'$.

An important property. The $\delta' = \delta$ property of nodal points N, N' (see Equation 1.110 and Figure 1.56) makes it easy to determine the position of N, N' experimentally. We first investigate geometrically how this property is used by drawing Figures 1.59 and 1.60. In Figure 1.59 we represent a set of parallel incident rays by drawing three rays (numbered 1, 2, and 3): The important ray is number 1, the ray aimed at the nodal point N; the other rays, 2 and 3, are chosen to pass through F and H for convenience in drawing the diagram. The parallel set of rays is inclined at angle δ to the horizontal, the symmetry axis in Figure 1.59. Because these rays are parallel, they are all focused by the optical system at the image point P' on the primed focal plane. Since we imagine an opaque screen at this focal plane, we would see a bright dot at P'. The important ray is ray 1, which is aimed at the nodal point N; by the $\delta' = \delta$ property of nodal points, this ray 1 emerges from the optical system (also called ray 1 for convenience) to pass through the nodal point N' making the same angle δ with the horizontal. What is important is that the two rays numbered 1 are parallel to each other—this fact is emphasized by the dashed line a. (Note that the N in V_N has nothing to do with the nodal points N or N'.)

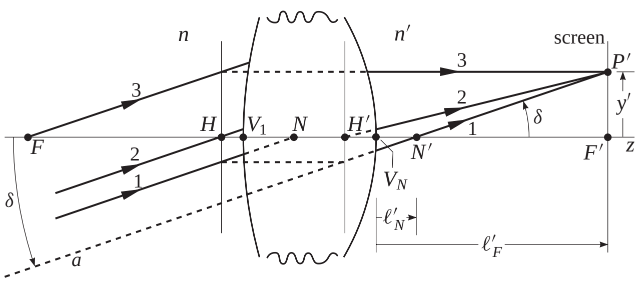


Figure 1.59

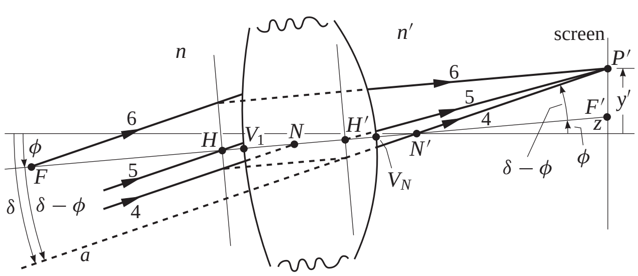


Figure 1.60

Now suppose we rotate the optical system (and its symmetry axis z) through a small angle ϕ about the nodal point N', as shown in Figure 1.60; the set of parallel incident rays, the dashed line a, and the screen remain fixed. Another combination of rays (numbered 4, 5, and 6) in the parallel set now pass through (or are aimed at or aimed from) the points N, H, F and N', H', F'. However, the nodal rays 4 and the dashed line a make the same angle δ with the horizontal and follow the same paths in space as rays 1 and the dashed line a in Figure 1.59. Therefore the emergent ray 4 passes through the same point P' on the screen as before in Figure 1.59. The image point may be shifted slightly to the left or right of the P' on the screen, but for small ϕ, this change will not be noticeable. Thus, small rotations ϕ (counterclockwise or clockwise) about N' produce no change in the position of the P'. In conclusion, the nodal point N' is found by searching for a center of rotation that produces no change in the position of P' on the screen for small rotations ϕ. The nodal slide is a physical device that makes this process more convenient to implement. By rotating the optical system 180 deg, the same procedure is used to find the other nodal point N.

When the same medium is on either side of the optical system, usually air ($n = n' = 1$), then the nodal points N, N' coincide with the respective principal points H, H'. In this case, the usual one, the method outlined gives the positions of H, H' on the symmetry axis.

We have finished our geometrical investigation. Now we would like to examine this same property of nodal points, but from the analytical viewpoint. Basically, the problem is one of the translation and rotation of axes, a topic usually treated in analytic geometry. To help us set up the problem, we draw Figure 1.61. The coordinates of the image point P' in the z'', y'' rotated coordinate system are simply (z'', y''); as we see in the diagram, the origin of this coordinate system is O'—this point, we imagine, moves until it coincides with N'

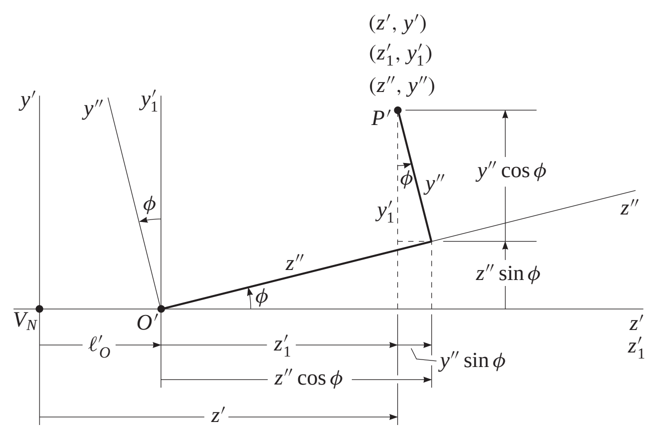


Figure 1.61

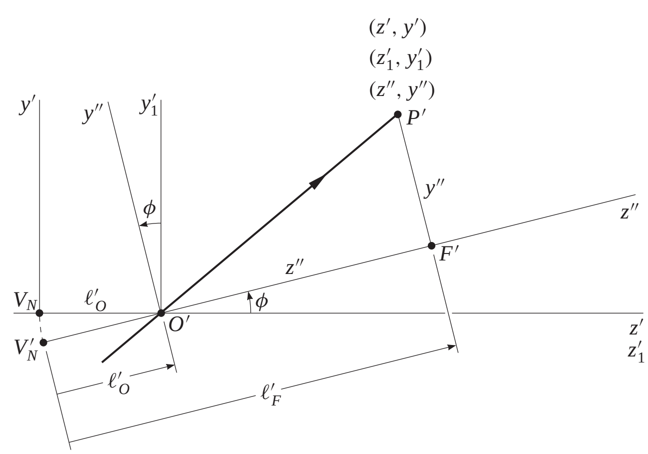


Figure 1.62

in our discussion. It is convenient to choose V_N as the fixed origin. In our diagrams, z'' is the symmetry axis of the rotated optical system, and z', z'_1 are the horizontal axes. Because notation is a problem in the analytical work, we think it is better to stay away from the z and y symbols. We have also omitted drawing the screen in the diagrams for simplicity.

By inspection of the diagram in Figure 1.61, we see that in the z'_1, y'_1 coordinate system the coordinates of P' are

$$z'_1 = z'' \cos\phi - y'' \sin\phi \tag{1.119a}$$

$$y'_1 = z'' \sin\phi + y'' \cos\phi \tag{1.119b}$$

and in the z', y' coordinate system

$$\begin{aligned} z' &= \ell'_O + z'_1 \\ &= \ell'_O + z'' \cos\phi - y'' \sin\phi \end{aligned} \tag{1.120a}$$

$$\begin{aligned} y' &= y'_1 \\ &= z'' \sin\phi + y'' \cos\phi \end{aligned} \tag{1.120b}$$

If angle of rotation ϕ is small, then $\sin\phi \approx \phi$, $\cos\phi \approx 1$, and Equations 1.120 become more simply

$$z' = \ell'_O + z'' - y''\phi \tag{1.121a}$$

$$y' = z''\phi + y'' \tag{1.121b}$$

We work first with the y' expression in Equation 1.121b, which we see says that y' is a function of the angle ϕ. We want to show that by choosing O' appropriately y' becomes independent of ϕ, which in turn means that the image point P' does not move vertically as we change ϕ, the goal of our analytical discussion. To see what z'' equals, we draw Figure 1.62, which is a simplified version of Figure 1.61, to show a ray traveling through O' to the image point P'. We have also extended the z'' axis backwards to the rotated position of V_N, naming it V'_N, and we see that the distance between it and O' is ℓ'_O. The image point P' must be on the focal plane through F'; therefore, the distance from V_N to F' is ℓ'_F, as indicated. By inspection of Figure 1.62, we obtain

$$z'' = \ell'_F - \ell'_O \tag{1.122a}$$

and substituting Equation 1.98 for ℓ'_F, we have

$$z'' = n'\frac{c}{a} - \ell'_O \tag{1.122b}$$

Substituting Equation 1.122b into Equation 1.121b, we get

$$y' = \left(n'\frac{c}{a} - \ell'_O\right)\phi + y'' \tag{1.123}$$

Next we need an expression for y''. We discussed earlier the situation of a parallel set of rays that are incident on an optical system as Case 3 ($C = 0$), where we derived Equations 1.71 and drew Figure 1.36. The diagram for our current discussion that corresponds to Figure 1.36 we draw in Figure 1.63. In Figure 1.36, we note that δ is the angle of inclination to z, the symmetry axis; in Figure 1.63, it is $\delta - \phi$ that is the angle of inclination to the symmetry axis z''. We now apply Equation 1.71b, which says that $y' = (n/a)\delta$; for our diagram in Figure 1.63, this equation becomes

$$y'' = \frac{n}{a}(\delta - \phi) \tag{1.124}$$

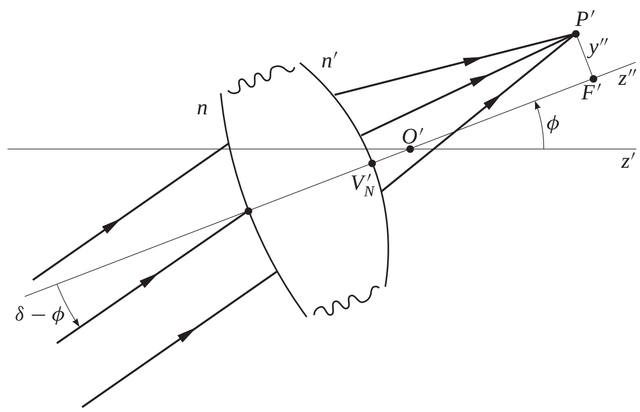


Figure 1.63

We now substitute Equation 1.124 into Equation 1.123 to get the important equation

$$\begin{aligned} y' &= \left(n'\frac{c}{a} - \ell'_O\right)\phi + \frac{n}{a}(\delta - \phi) \\ &= \left(n'\frac{c}{a} - \ell'_O - \frac{n}{a}\right)\phi + \frac{n}{a}\delta \end{aligned} \tag{1.125}$$

To remove the dependence on ϕ, all we have to do is set the coefficient of ϕ to zero, and then solve for ℓ'_O; we get

$$\begin{aligned} \ell'_O &= n'\frac{c}{a} - \frac{n}{a} \\ &= n'\frac{c - n/n'}{a} = \ell'_N \end{aligned} \tag{1.126}$$

where we have set the last expression equal to ℓ'_N by using Equation 1.113—we have obtained the desired result.

Finally, we get the z' dependence on ϕ by substituting Equations 1.122a and 1.124 into Equation 1.121a to obtain

$$\begin{aligned} z' &= \ell'_O + \left(\ell'_F - \ell'_O\right) - \frac{n}{a}(\delta - \phi)\phi \\ &= \ell'_F - \frac{n}{a}\delta\phi + \frac{n}{a}\phi^2 \end{aligned} \tag{1.127}$$

Remembering that the screen is positioned to pass through F' in the unrotated system (see Figure 1.59), Equation 1.127 says that the image point is shifted horizontally from the screen position; but as long as ϕ is small, the shift is small.

1.5.4 The focal lengths, f and f'

There are two more important quantities that are useful in the description of an optical system, the unprimed and primed focal lengths, f and f'—also called the effective focal lengths, the efl. These quantities are related to the cardinal points by the definitions:

$$f = \ell_F - \ell_H \tag{1.128a}$$

and

$$f' = \ell'_F - \ell'_H \tag{1.128b}$$

as illustrated in Figure 1.64. The dimension arrows in the diagram show that f is negative when F is to the left of H; f' is positive when F' is to the right of H'.

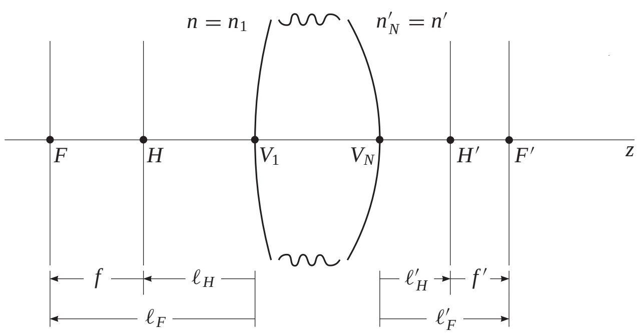


Figure 1.64

We express the focal lengths f, f' in terms of the Gaussian constants by substituting Equations 1.97 and 1.100 into Equation 1.128a to obtain

$$f = -n\frac{b}{a} - \left(-n\frac{b-1}{a}\right) = -\frac{n}{a} \tag{1.129a}$$

and by substituting Equations 1.98 and 1.102 into Equation 1.128b to obtain in a similar way

$$f' = n'\frac{c}{a} - n'\frac{c-1}{a} = \frac{n'}{a} \tag{1.129b}$$

We note that f and f' always have opposite signs. When $f' > 0$, we define the optical system to be a converging system; when $f' < 0$, the system is a diverging one.

The dissimilar signs of f and f' result from using analytic geometry as the basis of our sign rules; this dissimilarity can be somewhat awkward. However, we avoid most of the difficulty by normally using f' in our equations. Lens designers try to avoid the problem by defining f to be equal to f', but this definition destroys the symmetry in our diagram of Figure 1.64.

Example 1.5.1 An equiconvex lens: the cardinal points and planes, and the focal lengths.
It is convenient to work again with the equiconvex lens of Example 1.3.1, which we redraw in Figure 1.65 with its properties listed below the diagram.

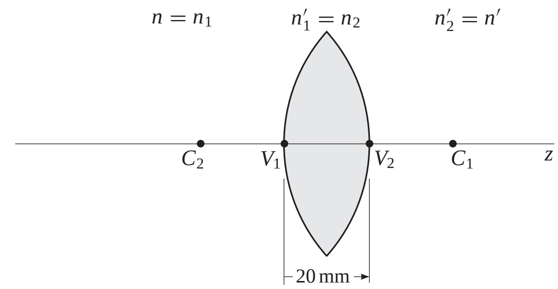


Figure 1.65

r (mm)	n	t (mm)
	1.000	
40.00		
	1.500	20.00
−40.00		
	1.000	

We copy the Gaussian constants for this lens from Example 1.4.1 as

$$\begin{aligned} a &= 0.0229167 \\ b &= 0.833333 \\ c &= 0.833333 \\ d &= -13.3333 \end{aligned}$$

and then use Equations 1.97 and 1.98 to calculate

$$\ell_F = -n\frac{b}{a} = -(1)\frac{0.833333}{0.0229167} = -36.36\,\text{mm}$$

$$\ell'_F = n'\frac{c}{a} = (1)\frac{0.833333}{0.0229167} = 36.36\,\text{mm}$$

Equations 1.100 and 1.102 are used to calculate ℓ_H and ℓ'_H (since the same medium is on either side of the lens, ℓ_N and ℓ'_N are equal to ℓ_H and ℓ'_H, respectively):

$$\ell_H = -n\frac{b-1}{a} = -(1)\frac{0.833333-1}{0.0229167} = 7.27\,\text{mm}$$

$$\ell'_H = n'\frac{c-1}{a} = (1)\frac{0.833333-1}{0.0229167} = -7.27\,\text{mm}$$

Finally, Equations 1.129a and 1.129b give f and f':

$$f = -\frac{n}{a} = -\frac{1}{0.0229167} = -43.64\,\text{mm}$$

$$f' = \frac{n'}{a} = \frac{1}{0.0229167} = 43.64\,\text{mm}$$

We employ Figure 1.64 as a guide to use the values we have just calculated to draw the focal points F, F', the principal points H, H', and their respective planes in Figure 1.66; we locate the focal lengths f, f', as well. The nodal points and planes are identical with H, H' because the same medium is on either side of the lens, namely air ($n = n' = 1$).

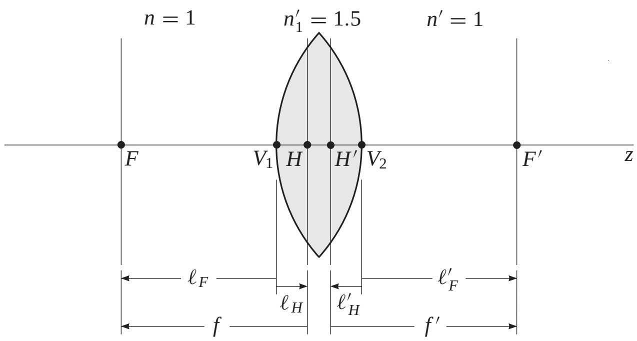


Figure 1.66

The three-ray method. These cardinal points and planes are utilized to determine the location and height of the image by simply drawing several rays from the object, as shown in Figure 1.67. In this diagram, we start with the same object that we used in Example 1.4.2, namely one with the object distance $\ell = -60.00$ mm and height $y = 12.00$ mm. It's usually easiest to draw ray 1 first, the ray drawn from the head O of the object and parallel to the symmetry axis until it hits the first surface. To review the properties of this ray, we look at Figure 1.50, where we see that a ray traveling parallel to the symmetry axis must emerge from the optical system to pass through the primed focal point F'; furthermore, according to Figure 1.52, these rays (or their forward or backward extensions) must pass through points on the H, H' planes at the same height above (or below) the symmetry axis. Thus, we extend the incident ray 1 forward to the H and H' principal planes, and the ray that emerges we draw as ray 1 passing through F' with its backward extension intersecting at the same point on the H' plane as the forward extension of the incident ray. We draw these extensions as dashed lines to indicate that the rays do not actually follow these paths inside the lens; this solid-line, dashed-line scenario is followed because the H, H' principal planes are inside the lens.

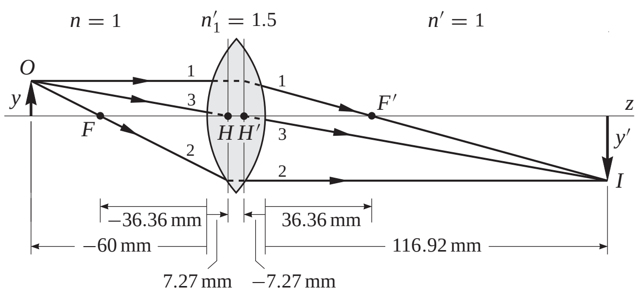


Figure 1.67 The three-ray method illustrated.

The next ray we draw in Figure 1.67 is ray 2 passing through the unprimed focal point F. As illustrated in Figure 1.49, this ray must emerge from the lens parallel to the symmetry axis. Just as with ray 1, the forward and backward extensions of ray 2 inside the lens are dashed to intersect at the H, H' planes to satisfy the requirement of Figure 1.52. The interesting feature now is that when this emergent ray is extended it intersects ray 1, this intersection point marks the head I of the image, as shown in Figure 1.67.

As a check on the image location, we draw ray 3 to implement the property shown in Figure 1.54; namely, when the same medium is on both sides of the optical system, a ray (or its extension) passing through the H point at an angle δ with the symmetry axis must emerge from the optical system so that it (or its extension) passes through H' at the same angle δ. In our current example, the H, H' points are inside the lens, so it is the extensions of the incident and emergent ray 3 that pass through these points, as shown by the dashed lines. The emergent ray 3 when drawn forward passes through the image point I, as we see in Figure 1.67.

It is interesting to check whether the δ's on either side of the lens for ray 3 are indeed the same. We first calculate δ for the incident ray with the help of Figures 1.67 and 1.68. To correctly calculate δ with its appropriate sign, it's best for δ to be the angle formed by the incident ray and the horizontal at O, as shown. By inspection of Figure 1.68, we see that the horizontal projection of this ray is 67.27 mm, and the vertical projection is -12.00 mm. Then

$$\delta = \frac{-12.00}{60.00 + 7.27} = -0.178\,\text{rad}$$

We might be inclined to use the tangent function to compute δ, but we must remember that in paraxial mathematics, $\tan\delta \approx \delta$, as we discussed with Equations 1.28 in Section 1.3.3, the section on the translation matrix.

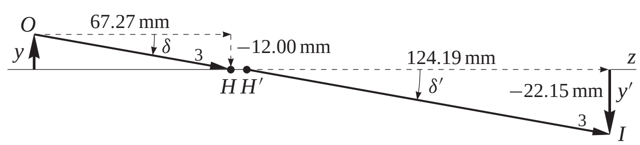


Figure 1.68

We next calculate the angle δ' for the emergent ray 3 in Figure 1.68, where the horizontal projection is 124.19 mm, and the vertical projection is the image height -22.15 mm (see Example 1.4.2 for the calculation of y'). We obtain

$$\delta' = \frac{-22.15}{7.27 + 116.92} = -0.178\,\text{rad}$$

and it is clearly the same as the δ for the incident ray 3.

We next look at an example of a concave lens, where we find that the focal points F, F' have their locations interchanged when compared to a convex lens.

Example 1.5.2 An equiconcave lens: the cardinal points and planes, and the focal lengths.
As a second example of the location and use of the cardinal points and planes, we look at the equiconcave lens that we traced a ray through in Example 1.3.3. We redraw the diagram of this lens in Figure 1.69 and list its properties in the table below the diagram.

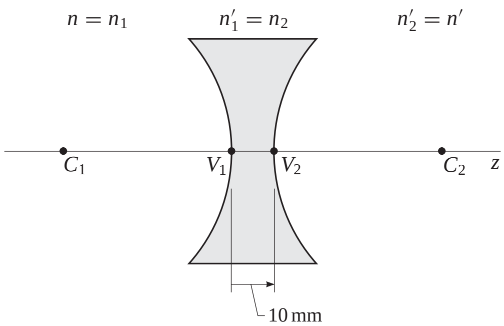


Figure 1.69

r (mm)	n	t (mm)
	1.000	
−40.00		
	1.500	10.00
40.00		
	1.000	

As we illustrated in Example 1.4.1, we first calculate the system matrix

$$
\begin{aligned}
S_{21} &= R_2 T_{21} R_1 \\
&= \begin{pmatrix} 1 & -p_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \dfrac{t'_1}{n'_1} & 1 \end{pmatrix} \begin{pmatrix} 1 & -p_1 \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1.08333 & 0.0260417 \\ 6.66667 & 1.08333 \end{pmatrix} = \begin{pmatrix} b & -a \\ -d & c \end{pmatrix}
\end{aligned}
$$

where

$$
\begin{aligned}
p_1 &= \frac{n'_1 - n_1}{r_1} = \frac{1.5 - 1}{-40} = -0.0125\,\text{mm}^{-1} \\
p_2 &= \frac{n'_2 - n_2}{r_2} = \frac{1 - 1.5}{40} = -0.0125\,\text{mm}^{-1}
\end{aligned}
$$

Next we read off the Gaussian constants by inspection as

$$
\begin{aligned}
a &= -0.0260417 \\
b &= 1.08333 \\
c &= 1.08333 \\
d &= -6.66667
\end{aligned}
$$

We then use these Gaussian constants to calculate the positions of the cardinal points and planes, and the focal lengths:

$$
\begin{aligned}
\ell_F &= -n\frac{b}{a} = -(1)\frac{1.08333}{-0.0260417} = 41.60\,\text{mm} \\
\ell'_F &= n'\frac{c}{a} = (1)\frac{1.08333}{-0.0260417} = -41.60\,\text{mm} \\
\ell_H &= -n\frac{b-1}{a} = -(1)\frac{1.08333 - 1}{-0.0260417} = 3.20\,\text{mm} \\
\ell'_H &= n'\frac{c-1}{a} = (1)\frac{1.08333 - 1}{-0.0260417} = -3.20\,\text{mm} \\
f &= -\frac{n}{a} = -\frac{1}{-0.0260417} = 38.40\,\text{mm} \\
f' &= \frac{n'}{a} = \frac{1}{-0.0260417} = -38.40\,\text{mm}
\end{aligned}
$$

We use these values to construct the cardinal points and planes, along with the focal lengths, in Figure 1.70. Again, because the same medium is on either side of the lens, the nodal points and planes are identical with the corresponding principal points and planes.

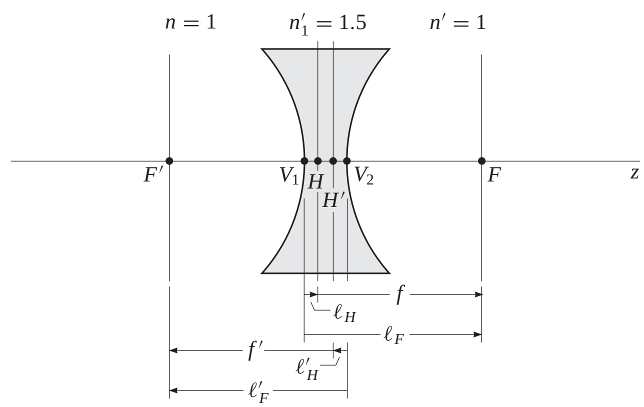


Figure 1.70

It is important to compare this diagram for a diverging lens with that for the converging one in Figure 1.66. The principal points are shifted in the same relative directions from the vertices in both cases, but the focal points have interchanged positions making the focal length f' negative—a diverging optical system always has negative f' (and a positive f).

Finally, we draw a ray diagram similar to the one shown in Figure 1.67 by choosing an object location and height to give a appropriate diagram: we select $\ell = -40.00$ mm, and $y = 16.00$ mm. Then with the ℓ, ℓ' relation of Equation 1.73, we obtain $\ell' = -23.53$ mm. To get y', we first calculate $m_T = 0.4706$ using either Equation 1.76 or 1.77, and then obtain $y' = m_T y = 7.53$ mm. Actually, we can determine the image location and height without calculating the values for ℓ' and y' by simply drawing the ray diagram following the procedure we outlined in the previous example with rays 1, 2, 3—the calculated values provide a check of the diagram. The ray diagram we obtain is shown in Figure 1.71. We see that the rays do not pass through the image, only the backward extensions do; therefore, the image is virtual, as we indicate by dashing the image. This virtual-image property also follows from the negative value for ℓ', as listed in the table of Figure 1.43.

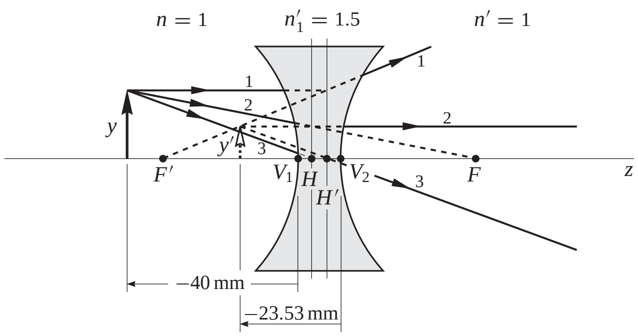


Figure 1.71 The object height is $y = 16.00$ mm; the image height is $y' = 7.53$ mm. The image is virtual/erect.

Example 1.5.3 An optical system of two lenses: the cardinal points and planes, and the focal lengths.
We next look at an optical system of two converging lenses which are planoconvex, as shown in Figure 1.72, and list the properties of this system below the diagram. This particular system is called a Ramsden eyepiece; it is frequently used in microscopes as the lens system that is close to the eye. Each of the two lenses has its own set of cardinal points and planes, but there is another set for the entire optical system—it is the one we want to obtain.

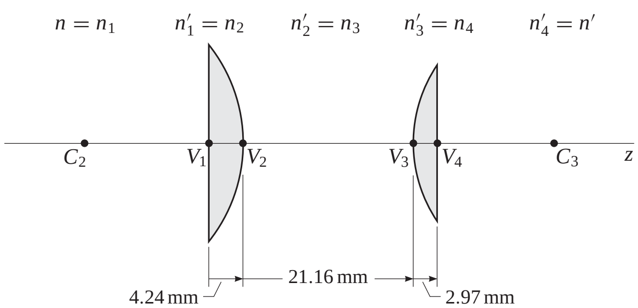


Figure 1.72

r (mm)	n	t (mm)
	1.000	
∞		
	1.517	4.24
−19.70		
	1.000	21.16
17.51		
	1.517	2.97
∞		
	1.000	

The system matrix for this optical system is

$$S_{41} = R_4 T_{43} R_3 T_{32} R_2 T_{21} R_1$$
$$= \begin{pmatrix} b & -a \\ -d & c \end{pmatrix}$$

Using a calculator or computer software, these R and T matrices can be multiplied to give the Gaussian constants:

$$a = 0.0393734$$
$$b = 0.265182$$
$$c = 0.367599$$
$$d = -22.9221$$

With these values for the Gaussian constants, we determine the cardinal point and plane locations as

$$\ell_F = -n\frac{b}{a} = -(1)\frac{0.265182}{0.0393734} = -6.74\,\text{mm}$$
$$\ell'_F = n'\frac{c}{a} = (1)\frac{0.367599}{0.0393734} = 9.34\,\text{mm}$$
$$\ell_H = -n\frac{b-1}{a} = -(1)\frac{0.265182-1}{0.0393734} = 18.66\,\text{mm}$$
$$\ell'_H = n'\frac{c-1}{a} = (1)\frac{0.367599-1}{0.0393734} = -16.06\,\text{mm}$$
$$f = -\frac{n}{a} = -\frac{1}{0.0393734} = -25.40\,\text{mm}$$
$$f' = \frac{n'}{a} = \frac{1}{0.0393734} = 25.40\,\text{mm}$$

and diagram their positions in Figure 1.73. Because f' is positive, we have a convergent system; also, since the same medium (air) is on either side of the optical system, the nodal

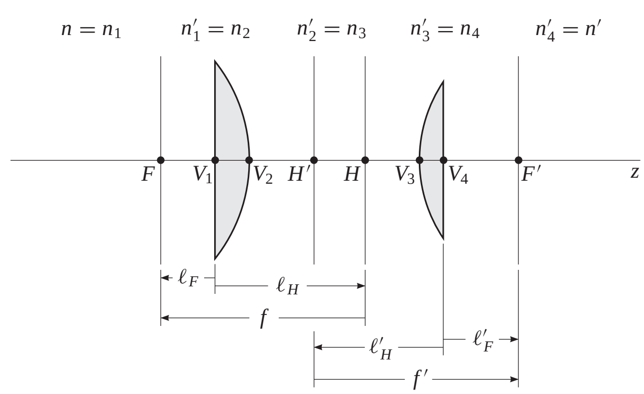


Figure 1.73

points N and N' are identical to the principal points H and H'. The interesting feature to note in this diagram is that H and H' have interchanged positions (or have crossed positions) when compared to the diagram in Figure 1.66, even though both are convergent systems.

Choosing a convenient object location and height to obtain a good illustration for this system, we draw the ray diagram shown in Figure 1.74 using the rays 1, 2, 3 as in the previous examples.

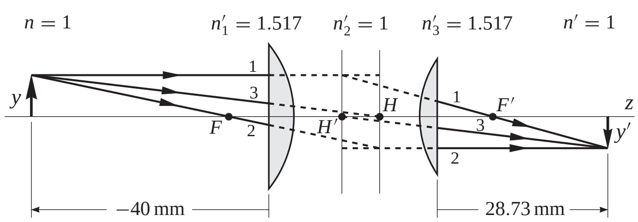


Figure 1.74 The object height is $y = 7.00$ mm; the image height is $y' = -5.34$ mm. The image is real/inverted.

Example 1.5.4 An optical system of one surface: the important relationships.
As a last example, we consider the simplest optical system of all; namely, a single refracting surface of radius r, as shown in Figure 1.75. We assume a positive r in the diagram, but the equations we derive hold for a negative r as well. Furthermore, we assume that $n \neq n'$.

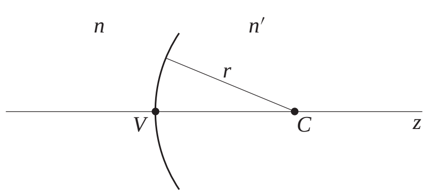


Figure 1.75

Following our usual procedure, we first calculate the system matrix:

$$S = R = \begin{pmatrix} 1 & -p \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -\dfrac{n'-n}{r} \\ 0 & 1 \end{pmatrix}$$

Then we read off the Gaussian constants by inspection as

$$\left.\begin{aligned} a &= \frac{n'-n}{r} \\ b &= 1 \\ c &= 1 \\ d &= 0 \end{aligned}\right\} \tag{1.130}$$

Because $b = c = 1$, we note that f and ℓ_F are equal, and so are f' and ℓ'_F:

$$f = \ell_F = -n\frac{b}{a} = -\frac{nr}{n'-n} \tag{1.131a}$$

$$f' = \ell'_F = n'\frac{c}{a} = \frac{n'r}{n'-n} \tag{1.131b}$$

Also because $b = c = 1$, the principal points are located at the vertex V:

$$\ell_H = -n\frac{b-1}{a} = 0 \tag{1.132a}$$

$$\ell'_H = n'\frac{c-1}{a} = 0 \tag{1.132b}$$

Finally, because $n \neq n'$, the nodal point locations N, N' are different from those occupied by the principal points at the vertex; we obtain (with the help of Equations 1.130)

$$\ell_N = -n\,\frac{b - n'/n}{a} = \frac{n'-n}{a} = r \tag{1.133a}$$

$$\ell'_N = n'\,\frac{c - n/n'}{a} = \frac{n'-n}{a} = r \tag{1.133b}$$

Therefore, both the unprimed and primed nodal points are located at the center of curvature C in Figure 1.75.

If we imagine an object located a distance ℓ from V in Figure 1.75, then we can calculate the location of the image with the ℓ, ℓ' relation of Equation 1.73 as

$$\ell' = n'\left(\frac{d + \dfrac{c\ell}{n}}{b + \dfrac{a\ell}{n}}\right) = n'\left(\frac{\dfrac{\ell}{n}}{1 + \dfrac{n'-n}{r}\dfrac{\ell}{n}}\right) \tag{1.134}$$

where we substituted for a, b, c, d using Equations 1.130. We rework this equation into a simple object-image relationship:

$$-\frac{n}{\ell} + \frac{n'}{\ell'} = \frac{n'-n}{r} \tag{1.135a}$$

The transverse magnification is found from Equation 1.76 with the help of Equations 1.130 and 1.135a:

$$m_T = \frac{y'}{y} = c - \frac{a\ell'}{n'} = \frac{n\ell'}{n'\ell} \tag{1.135b}$$

We have completed the last of the examples for this section. In the next section, we examine other ways to describe the object and image locations, as well as the transverse magnification m_T.

1.6 The Gaussian and Newtonian Formulations

So far, we have worked with only one equation to relate an object to its image: the ℓ, ℓ' relation of Equation 1.73, which used the Gaussian constants a, b, c, d. The object position was specified by its distance ℓ from the left vertex V_1 of the optical system, and the image position by its distance ℓ' from the right vertex V_N. The transverse magnification m_T also made use of the Gaussian constants and ℓ, ℓ'. But there are two other convenient ways of relating the object and image positions, and specifying m_T. One way is to measure the object and image distances from the principal points (or planes)—the Gaussian method—and the other way is to measure the distances from the focal points (or planes)—the Newtonian method. All three formulations are useful in working out the properties of optical systems; it depends on the particular optical system and what we want to know as to which formulation is easier to apply. The diagram that we will use to aid in the derivation of the Gaussian and Newtonian equations is shown in Figure 1.76.

1.6.1 The Gaussian equations

We first rearrange slightly the ℓ, ℓ' relation of Equation 1.73 to get

$$\frac{\ell'}{n'} = \frac{d + \dfrac{c\ell}{n}}{b + \dfrac{a\ell}{n}} \tag{1.136}$$

Then, for convenience, we list and renumber Equations 1.100 and 1.102, which give the positions of the principal points and planes:

$$\ell_H = -n\,\frac{b-1}{a} \qquad \text{and} \qquad \ell'_H = n'\,\frac{c-1}{a} \tag{1.137}$$

The first step in the derivation is to read off from Figure 1.76 that

$$s = \ell - \ell_H \qquad \text{and} \qquad s' = \ell' - \ell'_H \tag{1.138}$$

Substituting Equations 1.137 into Equations 1.138, and solving for ℓ/n in the first equation and for ℓ'/n' in the second, we obtain

$$\frac{\ell}{n} = \frac{s}{n} - \frac{b-1}{a} \tag{1.139a}$$

and

$$\frac{\ell'}{n'} = \frac{s'}{n'} + \frac{c-1}{a} \tag{1.139b}$$

Now substitute Equations 1.139a and 1.139b for ℓ/n and ℓ'/n', respectively, in Equation 1.136 to get

$$\frac{s'}{n'} + \frac{c-1}{a} = \frac{d + c\left(\dfrac{s}{n} - \dfrac{b-1}{a}\right)}{b + a\left(\dfrac{s}{n} - \dfrac{b-1}{a}\right)} \tag{1.140}$$

The Gaussian constants have the property that $bc - ad = 1$ (see Equation 1.60); thus, with some effort, we can manipulate Equation 1.140 into the simple relationship

$$-\frac{n}{s} + \frac{n'}{s'} = a \tag{1.141a}$$

Equations 1.129 say that $f = -n/a$ and $f' = n'/a$, so that we can expand Equation 1.141a to read

$$-\frac{n}{s} + \frac{n'}{s'} = a = -\frac{n}{f} = \frac{n'}{f'} \tag{1.141b}$$

or if the same medium is on either side of the optical system, then $n = n'$ and

$$-\frac{1}{s} + \frac{1}{s'} = \frac{a}{n} = -\frac{1}{f} = \frac{1}{f'} \tag{1.141c}$$

Equations 1.141 constitute the Gaussian lens equations. The last equation, Equation 1.141c, is close to the one normally found in a beginning physics course, except the minus signs are usually plus signs. However, in a beginning course, this

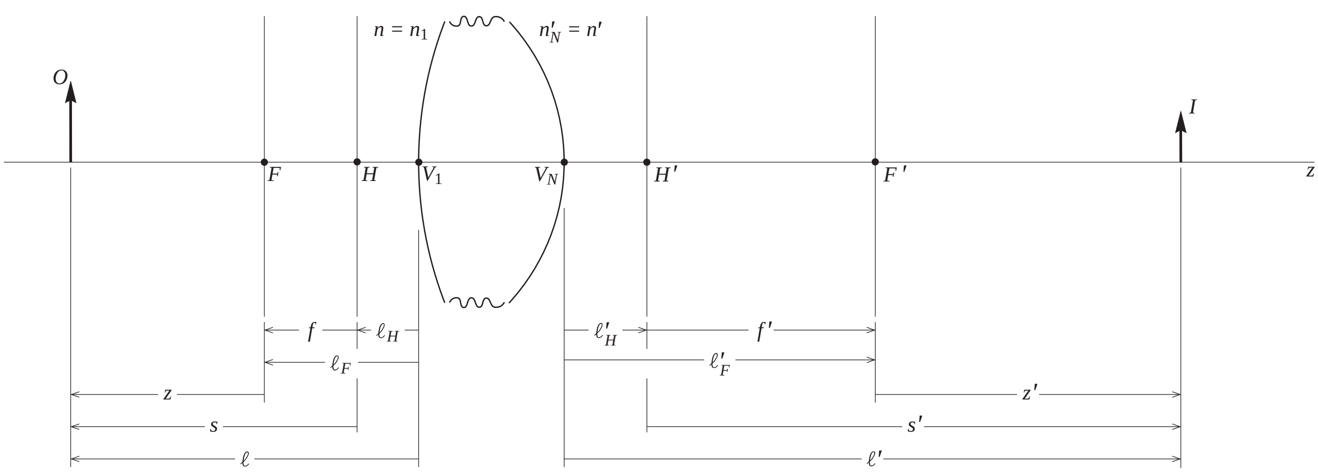


Figure 1.76 The reference diagram for (ℓ_F, ℓ'_F); (ℓ_H, ℓ'_H); $(f,\ f')$; (ℓ, ℓ'); and the new quantities (s, s'); (z, z').

equation is applied to a single thin lens (the subject of the next section), whereas here it governs a quite general optical system composed of any number of thick or thin lenses by measuring s and s' from the principal planes—a much more general treatment.

We also get a simple equation that relates the transverse magnification m_T to s and s'. We start with Equation 1.76,

$$m_T = c - \frac{a\ell'}{n'} \tag{1.142}$$

and substitute Equation 1.139b for ℓ'/n' to obtain

$$m_T = c - a\left(\frac{s'}{n'} + \frac{c-1}{a}\right) = 1 - \frac{as'}{n'} \tag{1.143}$$

Substituting Equation 1.141a for a, and simplifying, we get the Gaussian form of the transverse magnification to be

$$m_T = \frac{n\,s'}{n'\,s} \tag{1.144a}$$

If the same medium is on either side of the optical system, then Equation 1.144a becomes

$$m_T = \frac{s'}{s} \tag{1.144b}$$

which, except for the absence of a minus sign, is the formula usually given in beginning physics texts for the transverse magnification.

The Equations 1.141 and 1.144 are collectively called the Gaussian equations; we see that these equations describe in a rather pleasing style the object-image properties in terms of s and s', the distances from the respective principal planes.

1.6.2 The Newtonian equations

Instead of using s and s', the Newtonian equations describe the object-image properties in terms of z and z', the distances from their respective focal planes, as shown in Figure 1.76.

To get these equations, we follow the same procedural pattern that we used to get the Gaussian equations. First, we look back at Equations 1.68 and 1.70 to obtain

$$\ell_F = -n\frac{b}{a} \qquad \text{and} \qquad \ell'_F = n'\frac{c}{a} \tag{1.145}$$

Next, we read from Figure 1.76 that

$$z = \ell - \ell_F \qquad \text{and} \qquad z' = \ell' - \ell'_F \tag{1.146}$$

Substituting Equations 1.145 into Equations 1.146, and solving for ℓ/n in the first equation and for ℓ'/n' in the second, we get

$$\frac{\ell}{n} = \frac{z}{n} - \frac{b}{a} \tag{1.147a}$$

and

$$\frac{\ell'}{n'} = \frac{z'}{n'} + \frac{c}{a} \tag{1.147b}$$

Then, we substitute Equations 1.147 into Equation 1.136 for ℓ/n and ℓ'/n', respectively, to obtain

$$\frac{z'}{n'} + \frac{c}{a} = \frac{d + c\left(\dfrac{z}{n} - \dfrac{b}{a}\right)}{b + a\left(\dfrac{z}{n} - \dfrac{b}{a}\right)} \tag{1.148}$$

which simplifies with the help of $bc - ad = 1$ to

$$zz' = -\frac{nn'}{a^2} \tag{1.149a}$$

Because $f = -n/a$ and $f' = n'/a$, we rewrite Equation 1.149a as

$$zz' = ff' \tag{1.149b}$$

Equations 1.149 make up the Newtonian lens equations.

An example where any of the three Newtonian lens equations quickly give a useful result is to observe that if an object is located on the F focal plane ($z = 0$), then the image must be located at $\mp$infinity ($z' = \mp\infty$). To see that it is $\mp$, let z approach zero from the right of F, and then from the left (remember $ff' < 0$ because f and f' have opposite signs).

To get the corresponding Newtonian form of the transverse magnification m_T, we again start with Equation 1.76

$$m_T = c - \frac{a\ell'}{n'} \tag{1.150}$$

but replace ℓ/n' with Equation 1.147b to get

$$m_T = c - a\left(\frac{z'}{n'} + \frac{c}{a}\right) = -a\frac{z'}{n'} \tag{1.151}$$

Equation 1.129b says that $f' = n'/a$, or $a = n'/f'$. Substituting this expression for a into Equation 1.151, and simplifying, we get

$$m_T = -\frac{z'}{f'} \tag{1.152a}$$

Solving Equation 1.149b for z'/f' and substituting into Equation 1.152a gives

$$m_T = -\frac{f}{z} \tag{1.152b}$$

The Equations 1.149 and 1.152 are collectively called the Newtonian equations, and they describe the object-image properties of the optical system in terms of the distances, z and z', measured from their respective focal planes.

Comparing the Gaussian and Newtonian equations, we see that the Newtonian equations have the simplest form, which makes their use appear attractive. However, the s and s' in the Gaussian equations are measured from the principal planes which, in simple optical systems, are close to the first and last surfaces of the optical system—this fact can be an advantage. One such simple optical system is the thin lens, the topic of the next section.

1.7 The Thin Lens Approximation

Although the paraxial approximation permits a great simplification in describing the properties of an optical system, there is one more approximation that makes the work even simpler; this approximation postulates that each of the lenses comprising the optical system is thin—that is, each lens has a thickness that approaches zero.

1.7.1 Thick lens analysis

To see what this new approximation means, we first analyze in detail a thick lens (see Figure 1.77). The properties of this lens are shown in the usual tabular form below the diagram. The lens shape can be quite general: it can be converging or diverging, symmetric or asymmetric.

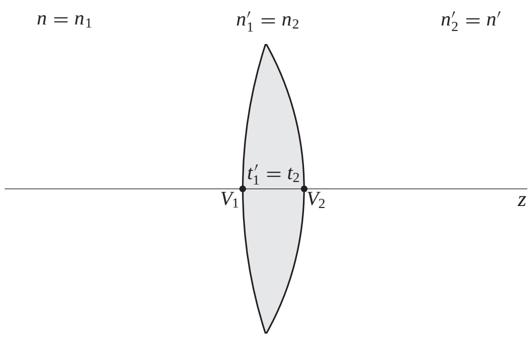


Figure 1.77

r	n	t
	n	
r_1		
	n_1'	t_1'
r_2		
	n'	

Our first step is to obtain the Gaussian constants for the thick lens by calculating its system matrix:

$$\begin{aligned} S_{21} &= R_2 T_{21} R_1 \\ &= \begin{pmatrix} 1 & -p_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \dfrac{t_1'}{n_1'} & 1 \end{pmatrix} \begin{pmatrix} 1 & -p_1 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 - p_2 \dfrac{t_1'}{n_1'} & -p_1 - p_2 + p_1 p_2 \dfrac{t_1'}{n_1'} \\ \dfrac{t_1'}{n_1'} & 1 - p_1 \dfrac{t_1'}{n_1'} \end{pmatrix} \\ &= \begin{pmatrix} b & -a \\ -d & c \end{pmatrix} \end{aligned} \tag{1.153}$$

where

$$p_1 = \frac{n_1' - n}{r_1} \qquad \text{and} \qquad p_2 = \frac{n' - n_1'}{r_2} \tag{1.154}$$

By inspection of Equation 1.153, we read that

$$a = p_1 + p_2 - p_1 p_2 \frac{t_1'}{n_1'} \tag{1.155a}$$

$$b = 1 - p_2 \frac{t_1'}{n_1'} \tag{1.155b}$$

$$c = 1 - p_1 \frac{t_1'}{n_1'} \tag{1.155c}$$

$$d = -\frac{t_1'}{n_1'} \tag{1.155d}$$

Next, we use the above Gaussian constants to obtain the positions of the cardinal points and planes, and the focal lengths:

$$\ell_F = -n\frac{b}{a} = -n\,\frac{1 - p_2 \dfrac{t_1'}{n_1'}}{a} \tag{1.156a}$$

$$\ell_F' = n'\frac{c}{a} = n'\,\frac{1 - p_1 \dfrac{t_1'}{n_1'}}{a} \tag{1.156b}$$

$$\ell_H = -n\frac{b-1}{a} = \frac{np_2}{a}\frac{t_1'}{n_1'} \tag{1.156c}$$

$$\ell_H' = n'\frac{c-1}{a} = -\frac{n'p_1}{a}\frac{t_1'}{n_1'} \tag{1.156d}$$

$$\begin{aligned} \ell_N &= -n\frac{b - \dfrac{n'}{n}}{a} \\ &= -n\frac{1 - \dfrac{n'}{n} - p_2\dfrac{t_1'}{n_1'}}{a} \end{aligned} \tag{1.156e}$$

$$\begin{aligned} \ell_N' &= n'\frac{c - \dfrac{n}{n'}}{a} \\ &= n'\frac{1 - \dfrac{n}{n'} - p_1\dfrac{t_1'}{n_1'}}{a} \end{aligned} \tag{1.156f}$$

$$f = -\frac{n}{a} \tag{1.156g}$$

$$f' = \frac{n'}{a} \tag{1.156h}$$

where a is given by Equation 1.155a. These equations look rather intimidating, but we need to see where the lens thickness t_1' appears. In the next section we shall simplify these equations greatly. But first we want to obtain an equation for the focal length f' of a thick lens in air ($n = n' = 1$). To obtain a simpler equation we work with the reciprocal of f' and f. Using Equations 1.156h, 1.156g, 1.155a, and 1.154, we obtain

$$\begin{aligned}\frac{1}{f'} &= -\frac{1}{f} = a \\ &= (n_1' - 1)\left[\frac{1}{r_1} - \frac{1}{r_2} + \frac{(n_1'-1)t_1'}{n_1' r_1 r_2}\right] \end{aligned} \quad (1.157)$$

This equation is called the lensmaker's equation for a thick lens because it expresses the focal length in terms of its geometry and lens material.

1.7.2 The thin lens in air

To obtain the properties of a thin lens, we allow the thickness t_1' tend to zero in the preceding equations (namely, Equations 1.155, 1.156, and 1.157). Also, because we want to simplify our equations as much as possible, we assume that air surrounds any thin lens system so $n = n' = 1$.

It is easiest to work backwards; thus, we start with Equation 1.157 and obtain

$$\frac{1}{f'} = -\frac{1}{f} = a = (n_1' - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) \quad (1.158)$$

which is the lensmaker's equation for a thin lens. Equations 1.156 for the cardinal points and planes, and the focal lengths, can be summarized as

$$\ell_F = f = -1/a \quad (1.159a)$$

$$\ell_F' = f' = 1/a \quad (1.159b)$$

$$\ell_H = \ell_H' = \ell_N = \ell_N' = 0 \quad (1.159c)$$

The Gaussian constants in Equations 1.155 become

$$a = 1/f' \quad (1.160a)$$

$$b = 1 \quad (1.160b)$$

$$c = 1 \quad (1.160c)$$

$$d = 0 \quad (1.160d)$$

so the system matrix for a thin lens is simply

$$Z = \begin{pmatrix} b & -a \\ -d & c \end{pmatrix} = \begin{pmatrix} 1 & -1/f' \\ 0 & 1 \end{pmatrix} \quad (1.161)$$

where we use the symbol Z to represent the system matrix of a thin lens. We have achieved quite a simplification; whereas, before we had to multiply three matrices to obtain the system matrix for a lens (see Equation 1.153), now we need only the simple one in Equation 1.161.

Because $\ell_H = \ell_H' = 0$ (see Equation 1.159c), the relations between s, s' and ℓ, ℓ' simplify by Equation 1.138 to

$$s = \ell \qquad \text{and} \qquad s' = \ell' \quad (1.162)$$

Remembering that $n = n' = 1$, the object and image positions are given by Equation 1.141c as

$$-\frac{1}{s} + \frac{1}{s'} = -\frac{1}{f} = \frac{1}{f'} \quad (1.163)$$

and the transverse magnification m_T is, by Equations 1.75 and 1.144b,

$$m_T = \frac{y'}{y} = \frac{s'}{s} \quad (1.164)$$

The thin converging lens. The meaning of the quantities in Equations 1.162, 1.163, and 1.164 is shown for a thin converging (or positive) lens in Figure 1.78; the three-ray method is used to locate the image—both object and image are real.

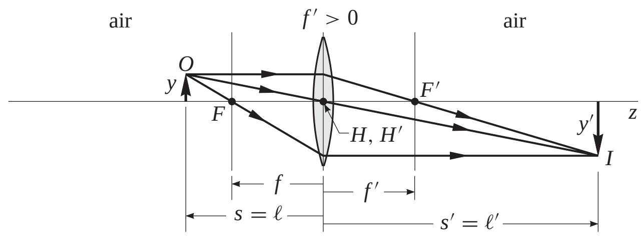


Figure 1.78 A thin positive lens and the three-ray method.

For a thin converging lens $f' > 0$, and so we solve Equation 1.163 for s' in terms of f' to get

$$s' = \frac{f's}{f' + s} \quad (1.165)$$

and graph this equation in Figure 1.79. We show values of m_T at selected points on this graph marked by the solid dots.

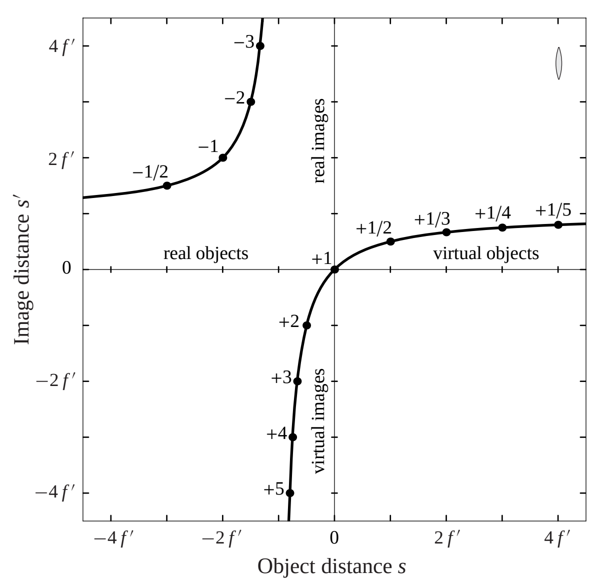


Figure 1.79 A graph of s' versus s for a thin converging lens in air. The values of m_T are given for selected values of s and s' at the solid dots.

Finally, we draw object-image diagrams in Figure 1.80 at several of the "dot" values shown in Figure 1.79 for an object that moves from left to right along the z axis. We assume an object of height +1 unit so that the image height equals m_T. In the diagrams, dashed lines indicate a virtual object or image.

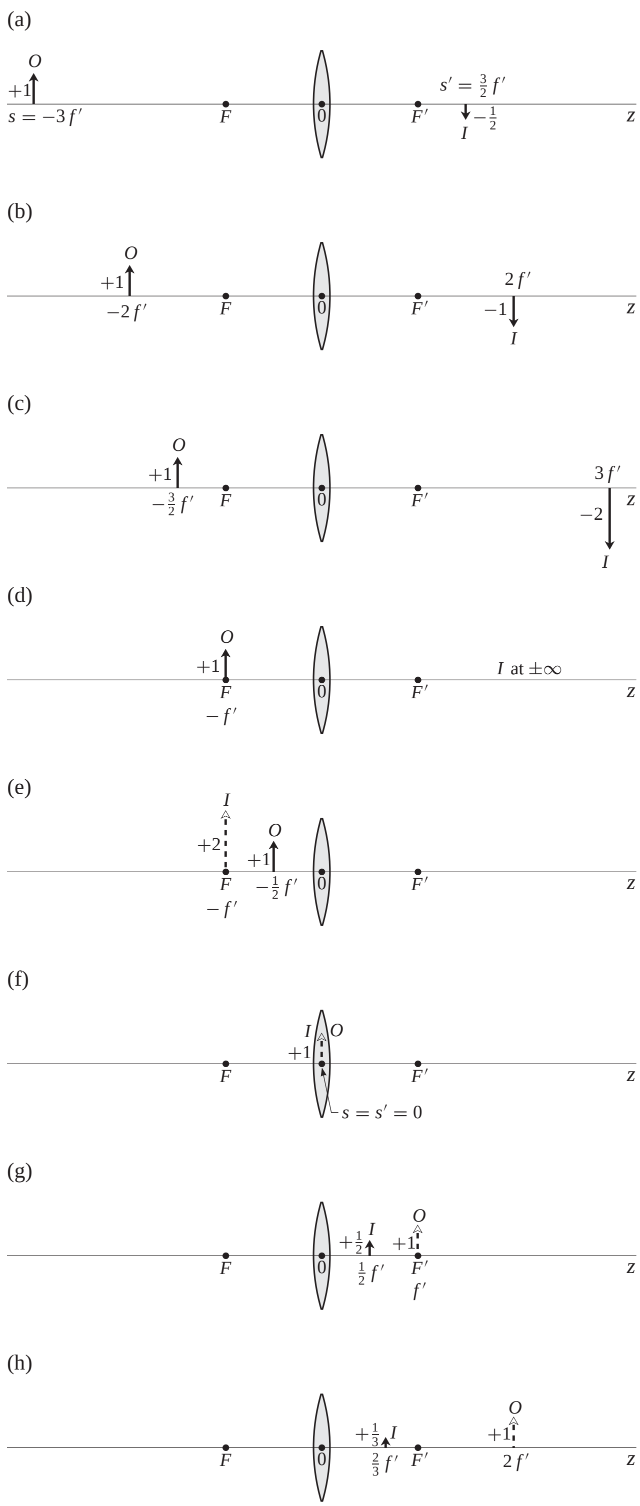


Figure 1.80

The thin diverging lens. It is instructive to carry out the same procedure for a thin diverging (or negative) lens. We show such a lens in Figure 1.81. The three-ray method is used to locate the image for a virtual object—the image is also virtual.

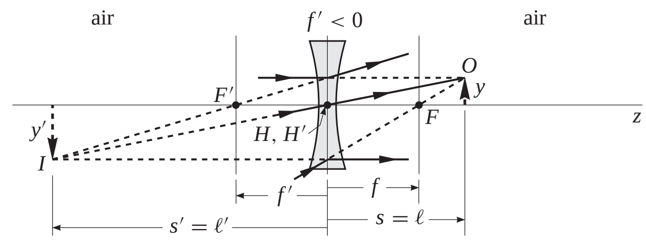


Figure 1.81 A thin negative lens and the three-ray method.

Of course, Equations 1.162, 1.163, and 1.164 still hold; however, we must remember that focal lengths f, f' for a diverging lens have interchanged signs relative to a converging lens. In fact, because $f > 0$, it is easier to draw a graph of s' versus s in terms of f, not f', as we did before for the thin converging lens; thus, the form of Equation 1.163 that we use is

$$-\frac{1}{s} + \frac{1}{s'} = -\frac{1}{f} \tag{1.166}$$

Solving for s', we get

$$s' = \frac{fs}{f - s} \tag{1.167}$$

We graph this equation in Figure 1.82, where s and s' are expressed in units of f. We also show values of m_T at the solid dots.

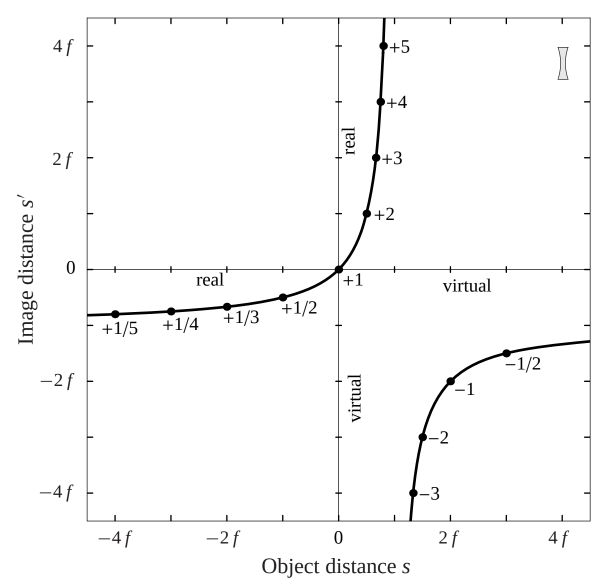


Figure 1.82 A graph of s' versus s for a thin diverging lens in air. The values of m_T are given for selected values of s and s' at the solid dots.

Just as we did before with a converging lens, we draw object-image diagrams at some of the "dot" values in Figure 1.82 for a thin diverging lens in air—we show these diagrams in Figure 1.83. As before, an object of height $+1$ moves from left to right along the z axis; dashed lines indicate a virtual object or image.

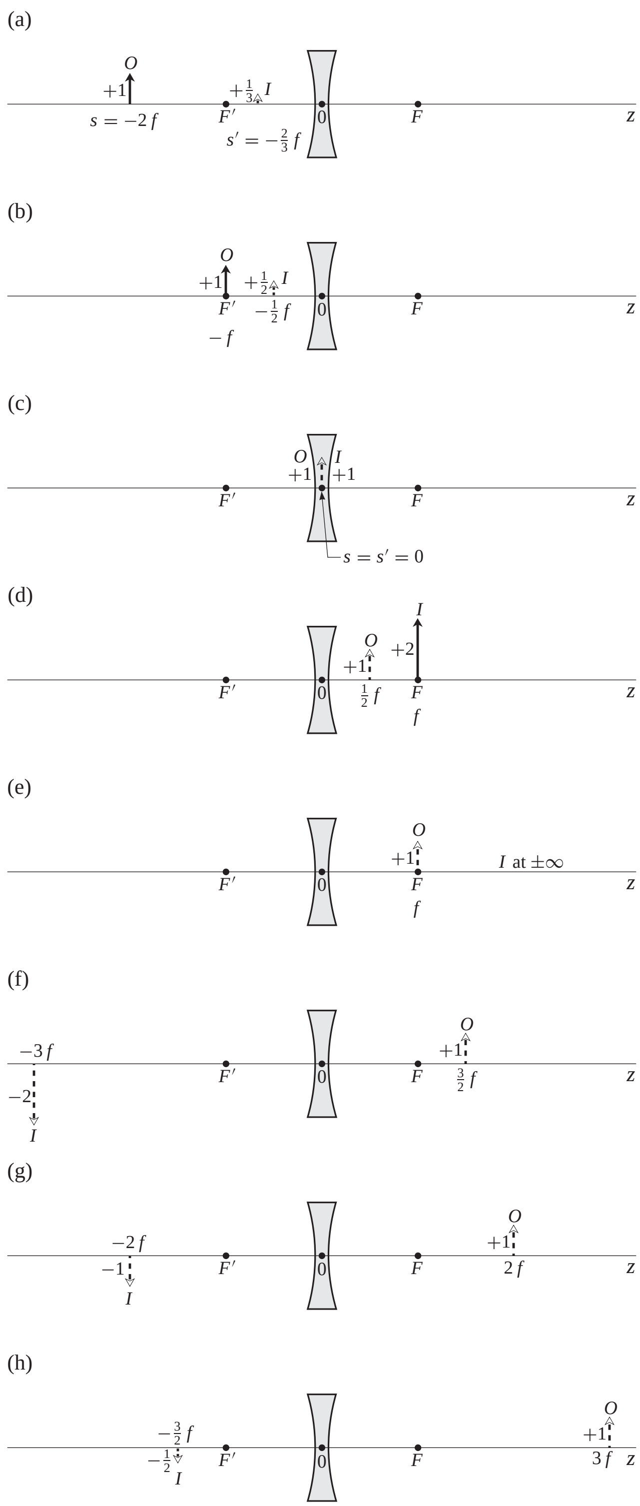


Figure 1.83

We now analyze several optical systems in air using the thin-lens approximation. In the diagrams, we use $L_1, L_2, \ldots$ to mark a thin lens position, and place them at the H, H' principal planes of each lens, as indicated by the dots (see Figure 1.84). Even though the lenses are in air, it is still advantageous to have an index of refraction notation so that we can write the matrix equations in symbolic form—the notation we use is shown in Figure 1.84. There is now a conflict with how we employed this notation before with thick lenses; for example, n'_1 and n_2 represented the index of refraction of the first lens—now we shall use n_{L_1} to represent the index of refraction of the first lens L_1. This notation is duplicated for other lenses in the system.

Example 1.7.1 The telephoto lens.

A simple telephoto lens, shown in Figure 1.84, consists of two thin lenses: one converging, the other diverging. One goal of this combination is to obtain an image that is larger than what a single converging lens would yield of a distant object. We give the numeric properties of a representative telephoto lens system in the table below the diagram.

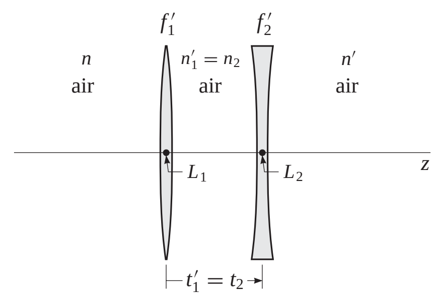


Figure 1.84

f' (mm)	n	t (mm)
	1.000	
60.00		
	1.000	40.00
−60.00		
	1.000	

We calculate the system matrix for the optical system using the system matrix for a thin lens given in Equation 1.161:

$$S_{21} = Z_2 T_{21} Z_1$$

$$= \begin{pmatrix} 1 & -1/f'_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ t'_1/n'_1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1/f'_1 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & -1/(-60) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 40/1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1/60 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1.66667 & -0.0111111 \\ 40 & 0.333333 \end{pmatrix} = \begin{pmatrix} b & -a \\ -d & c \end{pmatrix}$$

By inspection we read the Gaussian constants as

$$\left.\begin{aligned} a &= 0.0111111 \\ b &= 1.66667 \\ c &= 0.333333 \\ d &= -40 \end{aligned}\right\} \tag{1.168}$$

Next, we calculate the positions of the focal and principal points, and the focal lengths (see Section 1.5 for the equations—also, the nodal points and principal points are the same):

$$\ell_F = -n\frac{b}{a} = -(1)\frac{1.66667}{0.0111111} = -150.00\,\text{mm}$$

$$\ell'_F = n'\frac{c}{a} = (1)\frac{0.333333}{0.0111111} = 30.00\,\text{mm}$$

$$\ell_H = -n\frac{b-1}{a} = -(1)\frac{1.66667-1}{0.0111111} = -60.00\,\text{mm}$$

$$\ell'_H = n'\frac{c-1}{a} = (1)\frac{0.333333-1}{0.0111111} = -60.00\,\text{mm}$$

$$f = -\frac{n}{a} = -\frac{1}{0.0111111} = -90.00\,\text{mm}$$

$$f' = \frac{n'}{a} = \frac{1}{0.0111111} = 90.00\,\text{mm}$$

We draw the diagram that corresponds to the above values in Figure 1.85. To create a real and somewhat distant object, we choose $\ell = -350.00$ mm. Then, we calculate $\ell' = 70.50$ mm and $m_T = -0.45$, where we have used the ℓ, ℓ' relation (see Equation 1.73) to obtain ℓ' and Equations 1.76 or 1.77 to obtain m_T. Since ℓ' is positive, the image is real; since m_T is negative, the image is inverted. The ray diagram for this object-image combination is shown in Figure 1.86, where we choose an object height of $y = 50$ mm to get a convenient ray diagram—the corresponding image height is $y' = -22.5$ mm. We note that both the F' plane and the image are located to the right of the diverging lens, and fairly close to it; these locations are appropriate when the telephoto lens is used with a camera, so that the image can conveniently fall upon the camera film plane. We draw the

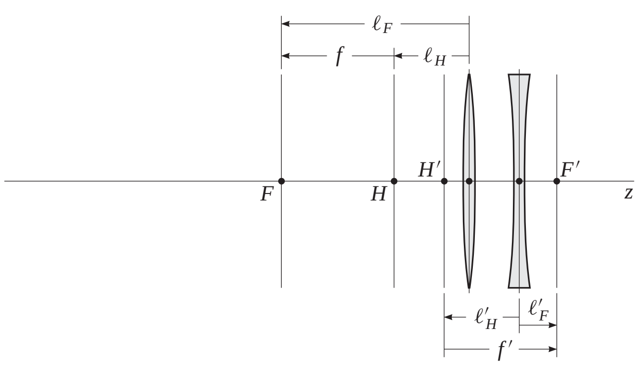


Figure 1.85

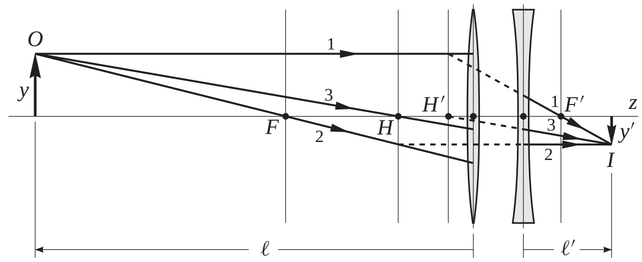


Figure 1.86

rays 1, 2, 3 from the object O through the lens system to locate the image I just as we did before in Section 1.5 with Examples 1.5.1 to 1.5.4. Again we must remember that rays 1, 2, 3 are simply three rays that allow quick determination of the image; there are many other rays that travel through the lens system from object to image.

Example 1.7.2 The Ramsden eyepiece.
One of the popular eyepieces found in microscopes and telescopes is the Ramsden eyepiece. It is usually composed of two identical planoconvex lenses with the curved surfaces facing each other, and with a lens separation equal to two-thirds the focal length of either lens, as shown in Figure 1.87. As in the previous example, we use L_1 and L_2 to name our thin lenses and to also mark their positions, as denoted by the dots; we place the dots on the plane surfaces, rather than in the center, for convenience in drawing the diagrams (doesn't really matter because the lenses are thin). The numerical values we choose for our system are shown in the table below the diagram.

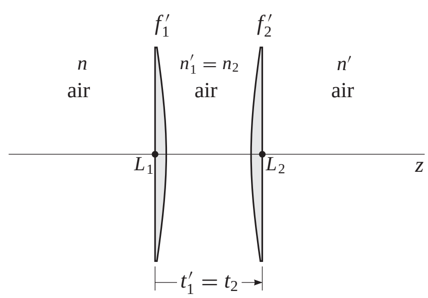


Figure 1.87

f' (mm)	n	t (mm)
	1.000	
+30.00		
	1.000	20.00
+30.00		
	1.000	

We make calculations and draw diagrams that are similar to those in the previous example. Again we use the matrix Z (see Equation 1.161) to represent the refraction properties of a thin lens. The system matrix for the Ramsden eyepiece of this example is

$$S_{21} = Z_2 T_{21} Z_1$$
$$= \begin{pmatrix} 1 & -1/f_2' \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ t_1'/n_1' & 1 \end{pmatrix} \begin{pmatrix} 1 & -1/f_1' \\ 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 0.333333 & -0.0444444 \\ 20 & 0.333333 \end{pmatrix} = \begin{pmatrix} b & -a \\ -d & c \end{pmatrix}$$

from which we read the Gaussian constants to be

$$\left.\begin{aligned} a &= 0.0444444 \\ b &= 0.333333 \\ c &= 0.333333 \\ d &= -20 \end{aligned}\right\} \tag{1.169}$$

Using the same formulas that we did in the previous example, we calculate

$$\begin{array}{lll} \ell_F = -7.50\,\text{mm} & \ell_H = 15.00\,\text{mm} & f = -22.50\,\text{mm} \\ \ell_F' = 7.50 & \ell_H' = -15.00 & f' = 22.50 \end{array}$$

With these values, we draw the focal points and planes, the principal points and planes, and the focal lengths, in Figure 1.88. We observe that the principal points and planes are crossed; that is, H' is to the left of H. In Figure 1.89, we draw an object-image diagram for this system. In actual use,

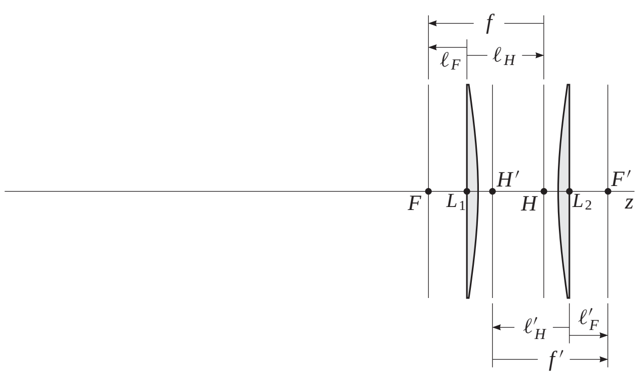


Figure 1.88

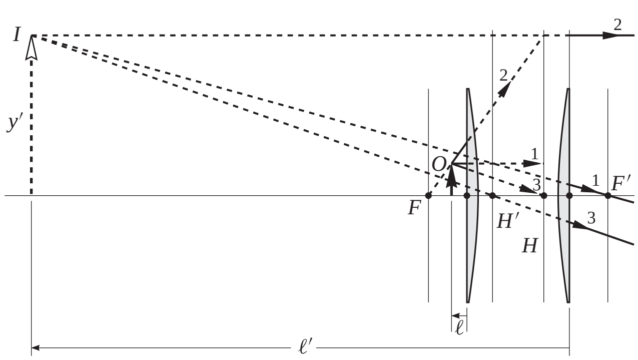


Figure 1.89

the object for this system is typically placed at or just to the right of the focal point F; but this position produces an image that is too far away to draw conveniently. So we place the object closer to the lens by choosing $\ell = -3.00\,\text{mm}$, and obtain a virtual and erect image at $\ell' = -105.00\,\text{mm}$ with lateral magnification $m_T = 5.00$. We draw the object height as 6.25 mm, and then obtain an image height $y' = 31.25\,\text{mm}$.

One other interesting aspect to the ray diagram in Figure 1.89 is that ray 2 never passes through lens L_2. We solve this problem by simply extending the appropriate principal planes until the ray intersects them—the ray would actually behave in this manner if the lens L_2 was bigger. Following this procedure still gives correct results.

Example 1.7.3 The Huygens' eyepiece.
Another popular eyepiece is the Huygens' eyepiece; it is also a two-lens system composed of planoconvex lenses. However, as shown in Figure 1.90, the curved surfaces of both lenses face the incident light, the focal lengths are different, and the separation t_1' is $(f_1' + f_2')/2$—see table below the diagram for the values we shall use.

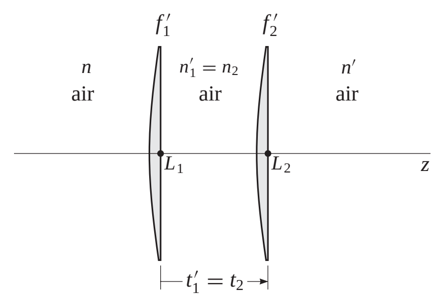


Figure 1.90

f' (mm)	n	t (mm)
	1.000	
+30.00		
	1.000	20.00
+10.00		
	1.000	

We calculate the system matrix for this lens system following the procedure illustrated in the previous two examples, again using the matrix Z to represent the thin lens properties:

$$S_{21} = Z_2 T_{21} Z_1$$
$$= \begin{pmatrix} 1 & -1/f_2' \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ t_1'/n_1' & 1 \end{pmatrix} \begin{pmatrix} 1 & -1/f_1' \\ 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} -1 & -0.0666667 \\ 20 & 0.333333 \end{pmatrix} = \begin{pmatrix} b & -a \\ -d & c \end{pmatrix}$$

By inspection, we easily read the values of the Gaussian constants as

$$\left.\begin{aligned} a &= 0.0666667 \\ b &= -1 \\ c &= 0.333333 \\ d &= -20 \end{aligned}\right\} \qquad (1.170)$$

In the usual way, we calculate

$$\begin{array}{lll} \ell_F = 15.00\,\text{mm} & \ell_H = 30.00\,\text{mm} & f = -15.00\,\text{mm} \\ \ell'_F = 5.00 & \ell'_H = -10.00 & f' = 15.00 \end{array}$$

We now use these values to draw the corresponding points and planes shown in Figure 1.91. Compared to the Ramsden eyepiece of the previous example, we see a more complicated arrangement; not only are the H, H' planes crossed, they also lie outside the focal planes F, F'. Furthermore, H', F lie between the two lenses. Just as in the case of the Ramsden eyepiece, the object is typically placed at or near the focal plane F; in this example, we will place the object O at F. The image location is then at infinity, a result that we can obtain with the usual ℓ, ℓ' relation in Equation 1.73 by substituting $\ell = \ell_F = -nb/a$. In fact, by using this same equation, we find that as the object moves toward F from the left, the image moves toward $+\infty$; when the object moves toward F from the right, the image moves toward $-\infty$. Thus, there is a discontinuity in the image location when the object is in the focal plane F; that is, we can regard the image as at either $+\infty$ or $-\infty$. This property is illustrated in Figure 1.92 where the emergent rays 1 and 3 travel parallel to each other to $+\infty$; the backward extensions of rays 1 and 3, are also parallel and extend to $-\infty$.

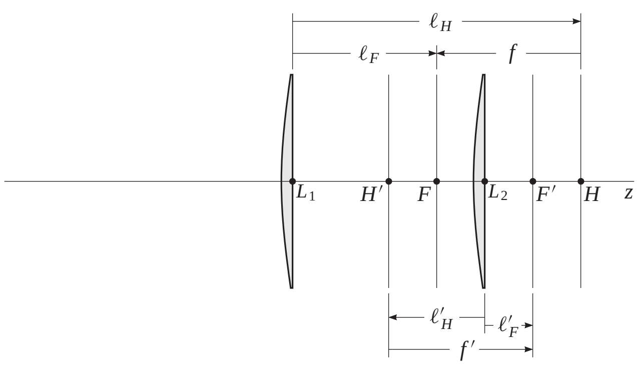


Figure 1.91

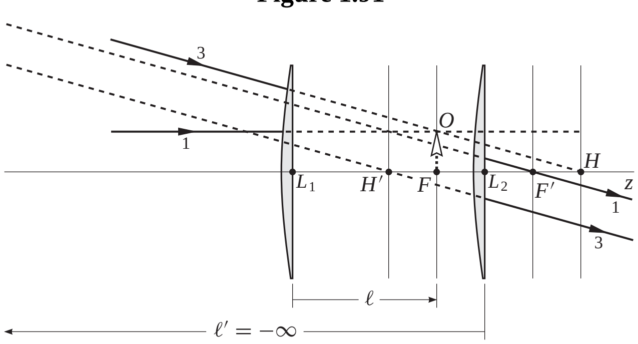


Figure 1.92

There is an additional complication. Because the object is to the right of the first lens L_1, it must be a virtual object; that is, the rays that form this object never get to it, as the incident rays 1 and 3 show in Figure 1.92. In fact, these rays come from another lens system (not shown) to the left of the eyepiece called the objective in a microscope or telescope. Consistent with the object being virtual is the positive value of ℓ (see the table in Figure 1.43); in this example, $\ell = 15.00$ mm, equal to ℓ_F.

To describe in more detail the paths the rays follow, we use Figure 1.92 and start with incident ray 1 which is parallel to the symmetry axis. It travels toward the object arrowhead until it hits the lens L_1. To determine how this ray emerges from lens L_2, we make use of the fact that all rays parallel to the symmetry axis must emerge traveling through the focal point F'. Furthermore, the backward extension of this emergent ray 1 must pass through the same point on the H' axis as the forward extension of the incident ray 1.

Because the object is in the F focal plane, drawing ray 2 is not helpful; it would be a vertical line passing through both the F focal point and the object arrowhead.

Ray 3 begins as an incident ray that is directed toward both the object arrowhead and the H point, as shown by the dashed line in the diagram. As we described in Figure 1.54, when the same medium is on either side of an optical system, an incident ray that is directed toward a principal point H must emerge parallel to the incident ray, but be directed toward the principal point H' (here, it is the backward extension that passes through H')—thus, we get the emergent ray 3 in the diagram. We see that rays 1 and 3 emerge parallel to each other, indicating that a real image is at $+\infty$; or since the backward extensions are also parallel to each other, we can say there is a virtual image at $-\infty$.

Since the eye is placed to the right of lens L_2 to focus the emergent rays 1 and 3 on the retina, it is convenient to think of the image as at $-\infty$; that is, $\ell' = -\infty$, as we have indicated in Figure 1.92.

Example 1.7.4 A zoom lens.

A contemporary zoom lens system is usually made of three or more lenses; two (or more) of these lenses are moved along the symmetry axis to change the magnification of the image and at the same time keep the image plane fixed. Such systems have become very popular. The design of practical zoom lenses began in the 1930s driven by the demands of the motion-picture market to change the size of the picture while the camera was running; it was given additional impetus later by the TV industry. However, for cameras the implementation took longer. Designers thought that photographers would be satisfied to change lenses manually to obtain different magnifications. To carry a bag full of interchangeable lenses probably sufficed for the professional, but the ease of use of the zoom lens soon became apparent, and a demand for such lenses was created. Since the 1950s, hundreds of zoom lenses have been designed.

For our example, we shall study a zoom lens that might be used in a camera. The lens system is in air, and is mechanically-compensated; that is, the lenses move physically to maintain the image on the image plane. For our system, we take three thin lenses and arrange them as shown in Figure 1.93; the values we shall use for this system are displayed below the diagram.

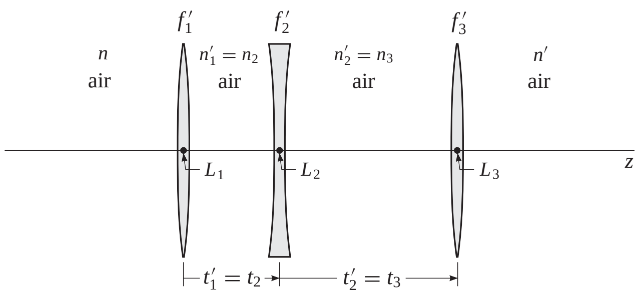


Figure 1.93

f' (mm)	n	t (mm)
	1.000	
+90.00		
	1.000	t_1'
−32.00		
	1.000	t_2'
+29.00		
	1.000	

To get a feeling for this system, we assume representative values for t_1', t_2' and calculate the values of F, H, f and F', H', f'. Then, just as we have in the previous examples, we draw diagrams to illustrate. We choose $t_1' = 10$ mm and $t_1' = 50$ mm; then with the values in the above table for the other quantities, we calculate

$$S_{31} = Z_3 T_{32} Z_2 T_{21} Z_1$$

$$= \begin{pmatrix} 1 & -1/f_3' \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ t_2'/n_2' & 1 \end{pmatrix} \begin{pmatrix} 1 & -1/f_2' \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ t_1'/n_1' & 1 \end{pmatrix} \begin{pmatrix} 1 & -1/f_1' \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} -1.29526 & -0.0427203 \\ 75.625 & 1.72222 \end{pmatrix} = \begin{pmatrix} b & -a \\ -d & c \end{pmatrix}$$

and we read the Gaussian constants to be

$$a = 0.0427203$$
$$b = -1.29526$$
$$c = 1.72222$$
$$d = -75.625$$

Then, as we did in the previous examples, we calculate the cardinal point locations and the focal lengths (again, because the system is in air, the nodal points coincide with the principal points):

$$\ell_F = 30.32\,\text{mm} \quad \ell_H = 53.73\,\text{mm} \quad f = -23.41\,\text{mm}$$
$$\ell_F' = 40.31 \quad \ell_H' = 16.91 \quad f' = 23.41$$

Because $f' > 0$, this lens system is a converging system, as it should be for a camera. With these values we draw the diagram of Figure 1.94. Finally, we choose an object of height $y = 25.00$ mm and distance $\ell = -50.00$ mm; we then calculate with Equations 1.77 and 1.73, the image height $y' = -7.29$ mm and the image distance $\ell' = 47.14$ mm. We draw the ray diagram for this object position in Figure 1.95.

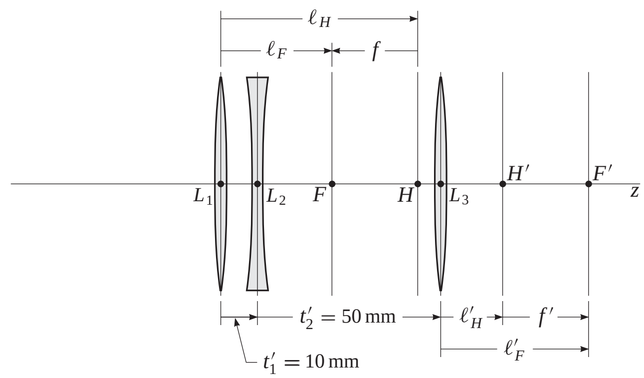


Figure 1.94

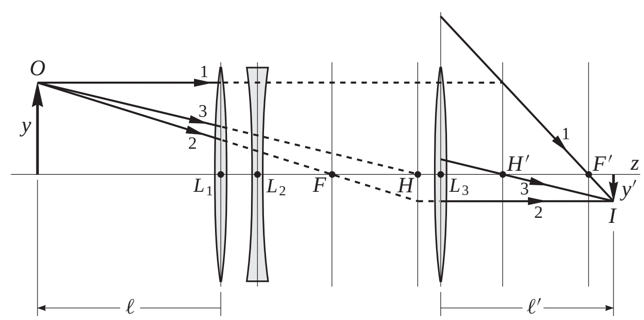


Figure 1.95

For our camera system, we shall move the object much farther from lens L_1 than in Figure 1.95. We choose an object position of $\ell = -1000$ mm, which makes the absolute value of ℓ many times greater than f'; therefore, the image is just to the right of F'. One of the interesting properties of this lens system is that the position of of F' does not change very much from its position in Figure 1.94 (or Figure 1.95) when we move lens L_2 between lenses L_1 and L_3; that is, vary t_1' but keep $t_1' + t_2' = 60$ mm. Thus, the position of the image does not move much either, because it stays near F'. However, the other cardinal point locations F, H, H' move much more, and so the focal lengths change as well as the lateral magnification m_T and the image size y'—remember, the goal is to change the image size. However, not every lens combination of three lenses has this property of F' moving little as the middle lens is moved; only certain combinations do. How to select lens combinations that work is beyond the scope of this text, but one way is to write a program that uses the power of a computer to search for combinations that meet your requirements.

We now want to investigate the properties of our zoom lens system in more detail. In our study, we assume that lens L_1 is fixed in position, and so it is available as a reference for the positions of the object, image, and the lenses L_2, L_3. We look at the behavior of our system in two steps.

Step 1: Only lens L_2 moves. To set the stage, we redraw the system in Figure 1.96 to show the variables we use in the study. We hold the lens L_3 fixed in position (L_1 is also fixed in this example) and require $t_1' + t_2' = 60$ mm; thus $t_2' = 60 - t_1'$. Then, we move lens L_2 by varying t_1' from 0 to 60 mm. As t_1' changes, we calculate D to determine the position of the image, and display the results graphically.

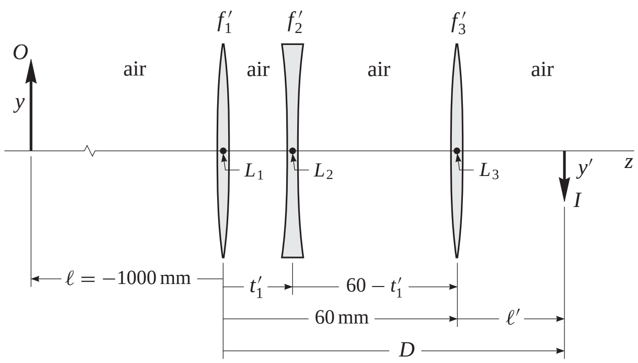


Figure 1.96

We start by writing the system matrix S_{31} for this combination of three thin lenses:

$$
\begin{aligned}
S_{31} &= Z_3 T_{32} Z_2 T_{21} Z_1 \\
&= \begin{pmatrix} 1 & -1/29 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 60 - t_1' & 1 \end{pmatrix} \begin{pmatrix} 1 & -1/(-32) \\ 0 & 1 \end{pmatrix} \\
&\qquad \begin{pmatrix} 1 & 0 \\ t_1' & 1 \end{pmatrix} \begin{pmatrix} 1 & -1/90 \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} b & -a \\ -d & c \end{pmatrix}
\end{aligned}
\qquad (1.171)
$$

Once the Gaussian constants a, b, c, d are determined, we calculate D with the help of the ℓ, ℓ' relation Equation 1.73:

$$
D = 60 + \ell' = 60 + \frac{d + c(-1000)}{b + a(-1000)} \qquad (1.172)
$$

For every new value of t_1', the Gaussian constants must be recalculated by using Equation 1.171 before a new value of D can be found with Equation 1.172. This numerical work is very tedious to perform by hand, but by writing a program in *Mathematica* (or in the programming language of your choice), we can quickly perform the numerical and graphical work to obtain Figure 1.97, which shows how D varies with t_1'.

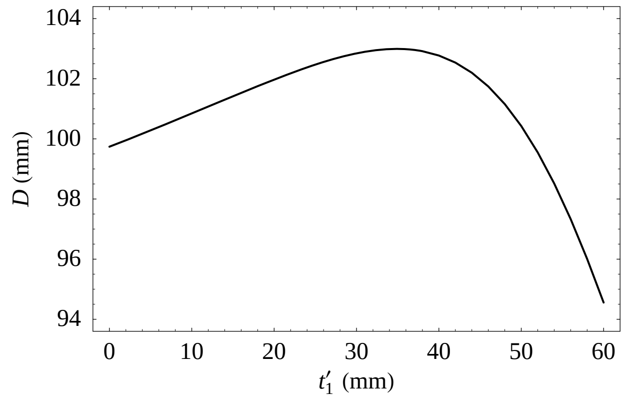


Figure 1.97

We now have a feeling for how the image plane moves as we change the position of L_2 with the other lenses fixed. In our second step, we want to fix the image plane position by moving L_3 as well as L_2.

Step 2: Both L_2 and L_3 move. To set the stage, we alter the diagram in Figure 1.96 so that we can move L_3 independent of L_2—the new diagram is shown in Figure 1.98.

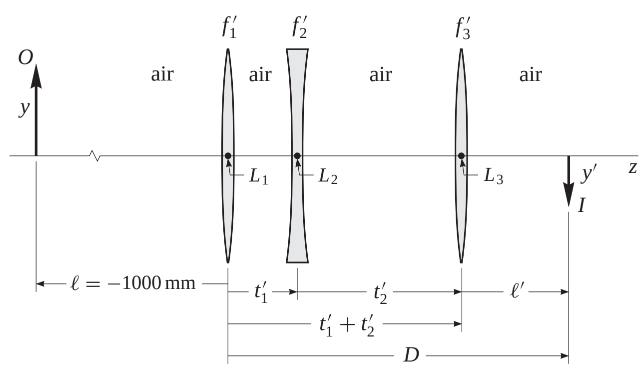


Figure 1.98

We must also change the calculation of the system matrix because both t_1' and t_2' are independent variables; we have

$$
\begin{aligned}
S_{31} &= Z_3 T_{32} Z_2 T_{21} Z_1 \\
&= \begin{pmatrix} 1 & -1/29 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ t_2' & 1 \end{pmatrix} \begin{pmatrix} 1 & -1/(-32) \\ 0 & 1 \end{pmatrix} \\
&\qquad \begin{pmatrix} 1 & 0 \\ t_1' & 1 \end{pmatrix} \begin{pmatrix} 1 & -1/90 \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} b & -a \\ -d & c \end{pmatrix}
\end{aligned}
\qquad (1.173)
$$

The calculation of D, the distance from L_1 to the image, in Equation 1.172 changes to

$$D = t_1' + t_2' + \ell' \tag{1.174}$$

and because we also want to determine the transverse magnification m_T of the image, we have from Equation 1.76

$$m_T = c - a\ell' \tag{1.175}$$

where in both equations

$$\ell' = \frac{d + c(-1000)}{b + a(-1000)} \tag{1.176}$$

Our goal is to adjust L_2 and L_3 so that the image does not move; that is, D should have some fixed value.

To help us determine the fixed value, we go back to Step 1 and look at the graph in Figure 1.97, we see that D varied between approximately 94 and 103 mm. Suppose we set our desired value of D to 100.00 mm. Because our work is numerical, we must specify how close we should come to this value, we choose ± 0.01 mm. The flow chart in Figure 1.99 outlines the method we use to obtain data to draw graphs of $t_1 + t_2$ vs t_1' and m_T vs t_1'.

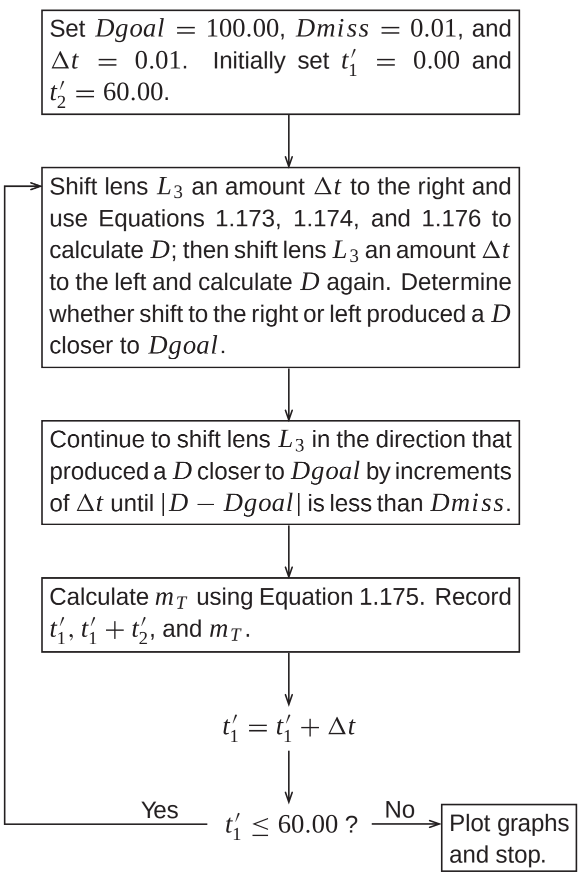


Figure 1.99 Flow chart for moving the lenses L_2 and L_3 of a zoom lens system to provide a changing magnification m_T on a fixed image plane at $D = 100.00 \pm 0.01$ mm.

The graph in Figure 1.100 shows the plot of the L_3 position, $t_1' + t_2'$, versus the L_2 position, t_1'. It shows that as lens L_2 moves from the lens L_1 position ($t_1' = 0$ mm) to the 60 mm position, lens L_3 moves toward lens L_1 and then away.

In Figure 1.101, the lenses are moved just like they are in Figure 1.100, but we graph the lateral magnification m_T. This graph shows that the image is inverted, and varies from a magnification of approximately -0.02 to -0.08, a factor of four increase in size.

A diagram that shows how the lenses move is shown in Figure 1.102. The positions of the lenses are drawn to scale, but not the positions and heights of the objects and images.

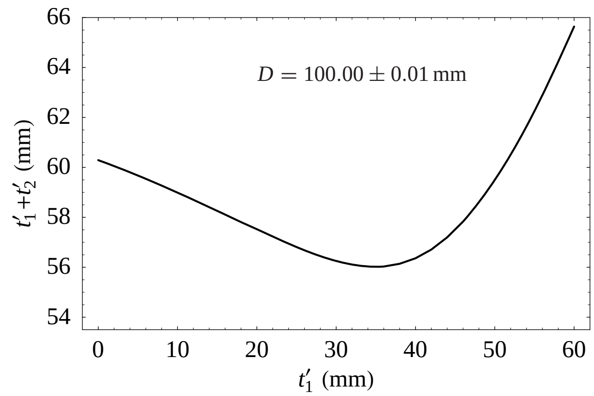


Figure 1.100

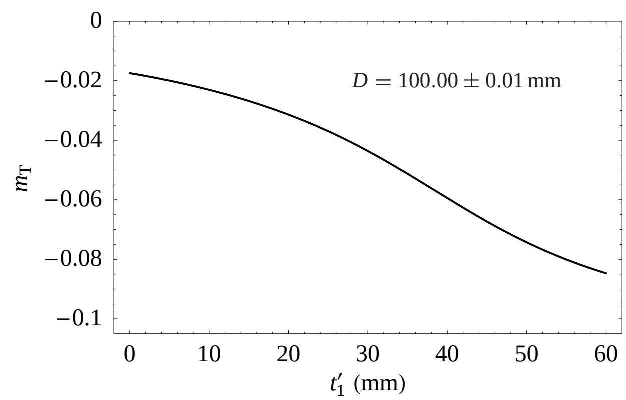


Figure 1.101

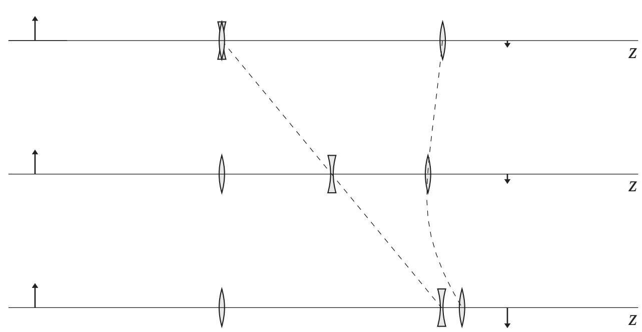


Figure 1.102

Problems

1.1 Suppose

$$A = \begin{pmatrix} 4 & 6 \\ 8 & 10 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 5 \\ 7 & 9 \end{pmatrix}$$

Calculate AB and BA. Is $AB = BA$?

1.2 Suppose

$$A = \begin{pmatrix} 2 & 6 \\ 9 & 8 \end{pmatrix} \qquad X = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

Calculate AX.

1.3 Suppose

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad B = \begin{pmatrix} 9 & 3 \\ 2 & 1 \end{pmatrix} \quad C = \begin{pmatrix} 8 & 1 \\ 7 & 3 \end{pmatrix}$$

Calculate $(AB)C$ and $A(BC)$. Is $(AB)C = A(BC)$?

1.4 Using the matrices of the previous problem, calculate
(a) $|A|\,|B|\,|C|$
(b) $|ABC|$
(c) $|CBA|$
(d) $|ACB|$

1.5 Suppose

$$A = \begin{pmatrix} 3 & -2 \\ 4 & 5 \end{pmatrix}$$

(a) Calculate A^{-1}. (b) Calculate AA^{-1} and $A^{-1}A$.

In the Problems 1.6 to 1.9, use Example 1.3.1 as a guide.

1.6 Trace a ray through a converging lens using the data listed. This ray emerges from the lens and travels to the primed focal point F'.

$\delta = 0.000$ rad $= 0.00$ deg
$y = 10.00$ mm

r (mm)	n	t (mm)
	1.000	20.00
40.00		
	1.517	4.00
−120.00		
	1.000	56.53

Sketch the ray diagram.

1.7 Trace a ray through a converging lens using the data listed. This ray begins at the unprimed focal point F.

$\delta = 0.200$ rad $= 11.46$ deg
$y = 0.00$ mm

r (mm)	n	t (mm)
	1.000	35.45
30.00		
	1.500	20.00
−50.00		
	1.000	10.00

Sketch the ray diagram.

1.8 Trace two rays through a converging lens using the data listed. This problem illustrates that in paraxial optics there exist conjugate points; that is, object-image points.

r (mm)	n	t (mm)
	1.000	204.55
100.00		
	1.500	20.00
−200.00		
	1.000	382.96

ray 1: $\delta = 0.100$ rad $= 5.73$ deg, $y = 10.00$ mm
ray 2: $\delta = -0.100$ rad $= -5.73$ deg, $y = 10.00$ mm

Sketch the ray diagram.

1.9 Trace a ray through a planoconvex lens using the data listed. This ray travels to the primed focal point F' on the symmetry axis after emerging from the lens.

$\delta = 0.000$ rad $= 0.00$ deg
$y = 10.00$ mm

r (mm)	n	t (mm)
	1.000	20.00
60.00		
	1.666	25.00
∞		
	1.000	75.08

Sketch the ray diagram.

In the Problems 1.10 to 1.11, use Example 1.3.4 as a guide.

1.10 Trace a ray through a biconcave lens using the data shown. In this problem, we finish with the backward extension of the emergent ray ending at a point on the symmetry axis; this point is called the primed focal point F'.

$\delta = 0.000$ rad $= 0.00$ deg
$y = 10.00$ mm

r (mm)	n	t (mm)
	1.000	20.00
65.00		
	1.560	28.00
32.00		
	1.000	−136.82

Sketch the ray diagram.

1.11 Trace a ray through a planoconcave lens using the data listed. The backward extension of the emergent ray passes through the primed focal point F' to the left of the lens.

$\delta = 0.000$ rad $= 0.00$ deg
$y = 10.00$ mm

r (mm)	n	t (mm)
	1.000	20.00
−60.00		
	1.666	25.00
∞		
	1.000	−105.10

Sketch the ray diagram.

1.12 Trace a ray through a plane-parallel plate of glass using the data below. Note that the ray emerges from the plate with same angle of inclination it had when it entered.

	r (mm)	n	t (mm)
$\delta = 0.300$ rad		1.000	20.00
$= 17.19$ deg	∞		
		1.574	35.00
$y = 5.00$ mm	∞		
		1.000	20.00

Sketch the ray diagram.

1.13 An equiconvex lens in air has radii of 52.00 mm in magnitude, an index of refraction of 1.680, and a thickness of 35.00 mm. (a) Calculate the Gaussian constants. (b) Check that $bc - ad = 1$. (c) Calculate $\ell_F, \ell'_F, \ell_H, \ell'_H$, f, and f'. (d) Sketch a diagram of this system.

1.14 A glass lens with an index of refraction 1.600 has radii $r_1 = +30.00$ mm and $r_2 = +30.00$ mm. The thickness of the lens is 30.00 mm. The lens is placed so that air touches the surface of radius r_1, and oil of index of refraction 1.300 touches the surface of radius r_2. (a) Calculate the Gaussian constants. (b) Check that the Gaussian constants satisfy $bc - ad = 1$. (c) Find the cardinal point locations by computing $\ell_F, \ell'_F, \ell_H, \ell'_H, \ell_N, \ell'_N$. Also calculate the focal lengths f and f'. (d) Is this lens a converging or diverging one? How do you know? (e) Sketch a diagram of this system.

In the Problems 1.15 to 1.18, use Figure 1.67 as a guide to draw the ray diagram.

1.15 A glass sphere in air of 20.00 mm radius has an index of refraction 1.700. (a) Calculate the Gaussian constants. (b) Check that the Gaussian constants satisfy the relation $bc - ad = 1$. (c) Calculate $\ell_F, \ell'_F, \ell_H, \ell'_H$. Also calculate the focal lengths f and f'. (d) Suppose an object 2.00 mm tall is located 20.00 mm to the left of the sphere surface. Calculate the position and size of the image. (e) Draw a ray diagram to locate the image. (f) Is the object real or virtual? Is the image real or virtual? Is the image erect or inverted?

1.16 A meniscus-convex lens in air has an index of refraction 1.500; it is 40.00 mm thick and has the radii of $r_1 = 100.00$ mm, $r_2 = 150.00$ mm. (a) Calculate the Gaussian constants. (b) Check that the Gaussian constants satisfy the relation $bc - ad = 1$. (c) Calculate $\ell_F, \ell'_F, \ell_H, \ell'_H, f, f'$. (d) Suppose an object 50.00 mm tall is located 250.00 mm to the left of the lens. Calculate the position and size of the image. (e) Draw a ray diagram to locate the image. (f) Is the object real or virtual? Is the image real or virtual? Is the image erect or inverted?

1.17 A meniscus-concave lens in air has an index of refraction 1.500; it is 30.00 mm thick and has the radii of $r_1 = 150.00$ mm, $r_2 = 100.00$ mm. (a) Calculate the Gaussian constants. (b) Check that the Gaussian constants satisfy the relation $bc - ad = 1$. (c) Calculate $\ell_F, \ell'_F, \ell_H, \ell'_H, f, f'$. (d) Suppose an object 50.00 mm tall is located 250.00 mm to the left of the lens. Calculate the position and size of the image. (e) Draw a ray diagram to locate the image. (f) Is the object real or virtual? Is the image real or virtual? Is the image erect or inverted?

1.18 An equiconvex lens in air of index of refraction 1.626 is 15.00 mm thick and has radii of $r_1 = 50.00$ mm, $r_2 = -50.00$ mm. (a) Calculate the Gaussian constants. (b) Check that the Gaussian constants satisfy the relation $bc - ad = 1$. (c) Calculate $\ell_F, \ell'_F, \ell_H, \ell'_H, f, f'$. (d) Suppose an object 30.00 mm tall is located a distance 200.00 mm to the right of the lens. Calculate the position and size of the image. (e) Draw a ray diagram to locate the image. (f) Is the object real or virtual? Is the image real or virtual? Is the image erect or inverted?

1.19 To obtain Equation 1.71b, the relation $D = 1/a$ was used. Obtain this relation by substituting Equation 1.70 into Equation 1.64d.

1.20 Show the steps for obtaining Equation 1.77.

1.21 Rework Equation 1.134 into Equation 1.135a.

1.22 Obtain Equation 1.135b.

1.23 Obtain Equation 1.141a from Equation 1.140.

1.24 Show that for a thin lens in air, $s' = 0$ and $m_T = +1$ when $s = 0$ (see Figures 1.79 and 1.82).

1.25 A thin converging lens of focal length f' in air forms a real image on a screen a fixed distance D from a real object. (a) Determine the object and image distances: denote one object, image distance pair as (s_1, s'_1), the other as (s_2, s'_2), such that $s'_1 > s'_2$. (b) By inspection, write a relationship between s'_1 and s_2; also, between s'_2 and s_1. (c) If d is the distance between the two lens positions that create images on the screen, show that

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

Observe that $D \geq 4f'$ for the formation of a real image. (d) Show that

$$f' = \frac{D^2 - d^2}{4D}$$

This formula is useful for determining the f' of a converging lens. (e) Show that the ratio of the two lateral magnifications m_T can be written as

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

1.26 Suppose in the previous problem, $f' = 100.00$ mm and $D = 800.00$ mm. (a) Calculate (s_1, s_1') and (s_2, s_2'). (b) Calculate d. (c) Calculate m_T for the bigger image, and m_T for the smaller image. (d) Check to see if the formula in (d) of the previous problem gives f'. (e) Compare the ratio of your m_T values with that given by the formula in the previous problem, part (e).

1.27 A thin converging lens in air forms a real image on a screen located a distance D from a real object. (a) Show that we can write

$$D = f'\left(2 - m_T - \frac{1}{m_T}\right)$$

(b) Graph D in units of f' from $m_T = -4.00$ to -0.25. This graph shows how the image shifts as the lens is moved. Observe that for $D > 4f'$, two images are brought to a focus for a given distance D by moving the lens: one image large, the other small—a simple type of zoom system. (c) Show that the equation in part (a) can be solved for f' to give

$$f' = -\frac{D\,m_T}{(m_T - 1)^2}$$

1.28 As in Problem 1.27, a thin converging lens forms a real image on a screen located a distance D from the object. (a) Show that it is possible to write

$$D = -\frac{s^2}{s + f'}$$

(b) Use calculus to find s for which D has a minimum, and determine the value of D at this minimum. (c) Calculate the value of the lateral magnification m_T at this minimum.

1.29 Two thin lenses are in air, the first has a focal length of 80.00 mm, the second 100.00 mm. They are separated by a distance of 30.00 mm. (a) Calculate the Gaussian constants. (b) Check that $bc - ad = 1$. (c) Locate the four cardinal points by calculating $\ell_F, \ell_F', \ell_H, \ell_H'$. Also calculate the focal lengths, f and f'. (d) Sketch the system.

1.30 An object is placed 130.00 mm to the left of a thin lens. The image is formed on the right 440.00 mm from the lens. (a) Calculate the f and f' of the lens. (b) Calculate the transverse magnification m_T of the lens.

1.31 An upright object 20.0 mm tall is positioned 130.0 mm to the left of a thin lens of focal length $f' = 28.0$ mm. Calculate (a) the image distance s', and (b) the image height y'. (c) Is the image real or virtual. (d) Is the image erect or inverted. (e) Draw a ray diagram.

1.32 The two surfaces of a thin lens made of glass have the radii $r_1 = 200.0$ mm, and $r_2 = -150.0$ mm. The index of refraction is 1.5725. Calculate the focal length f' of this lens.

1.33 A thin planoconvex lens made from extra-light flint glass of $n_1' = 1.5585$ is designed to have the focal length of $f' = 200.0$ mm. Calculate the radius of the spherical surface of this lens.

1.34 A thin equiconvex lens is designed from barium-crown glass of $n_1' = 1.5744$ to have a focal length of $f' = 50.00$ mm. Calculate the radii of the surfaces of this lens.

1.35 A thin equiconcave lens is designed from crown glass of $n_1' = 1.5180$ to have a focal length of $f' = -50.0$ mm. Calculate the radii of the surfaces of the lens.

1.36 The picture on a slide has a height of 25.0 mm. The slide is a distance 4.25 m from a projection screen. (a) Calculate the focal length f' of the thin lens that is required to project an image of height 1.20 m on the screen. Determine also (b) the object distance s, and (c) the image distance s'.

1.37 An object is located 2.80 m from a wall. (a) Calculate the focal length f' of the thin lens that is required to form an image of transverse magnification $m_T = -10.0$. Determine also (b) the object distance s, and (c) the image distance s'.

1.38 A pair of thin lenses of focal lengths f_1', f_2' are spaced a distance t apart in air. (a) Use the system matrix for this optical system to show that the focal length f' of the combination is

$$\frac{1}{f'} = \frac{1}{f_1'} + \frac{1}{f_2'} - \frac{t}{f_1' f_2'}$$

(b) When the thin lenses are so close together that $t = 0$, what simple formula to we obtain?

1.39 Two thin lenses have focal lengths of $f_1' = 90.0$ mm and $f_2' = -180.0$ mm, and are spaced 30.0 mm apart in air. An object 25.0 mm tall is located 200.0 mm to the left of the first lens. Use the Gaussian constants to calculate (a) the position of the image, and (b) the height of the image. (c) Is the image real or virtual? (d) Is the image erect or inverted?

1.40 Use the Gaussian constants determined in the previous problem to calculate (a) $\ell_F, \ell_F', \ell_H, \ell_H', f$, and f'. (b) Use Figure 1.86 as a guide to draw a ray diagram for the object of the previous problem.

1.41 Write a program to obtain the graph in Figure 1.97.

1.42 Use the flow chart in Figure 1.99 as a guide to write programs to obtain (a) the graph in Figure 1.100, and (b) the graph in Figure 1.101.

Chapter 1: Answers to Problems

1.1 $\begin{pmatrix} 54 & 74 \\ 94 & 130 \end{pmatrix}$, $\begin{pmatrix} 52 & 68 \\ 100 & 132 \end{pmatrix}$

1.2 $\begin{pmatrix} 18 \\ 43 \end{pmatrix}$

1.3 $\begin{pmatrix} 139 & 28 \\ 371 & 74 \end{pmatrix}$, $\begin{pmatrix} 139 & 28 \\ 371 & 74 \end{pmatrix}$

1.4 −102, −102, −102, −102

1.5 $\begin{pmatrix} \frac{5}{23} & \frac{2}{23} \\ \frac{-4}{23} & \frac{3}{23} \end{pmatrix}$, $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

1.6

	δ (rad)	δ (deg)	y (mm)
P	0.000	0.00	10.00
1	0.000 −0.0852	0.00 −4.88	10.00 10.00
2	−0.0852 −0.171	−4.88 −9.79	9.66 9.66
P'	−0.171	−9.79	0.00

1.7

	δ (rad)	δ (deg)	y (mm)
P	0.200	11.46	0.00
1	0.200 0.0546	11.46 3.13	7.09 7.09
2	0.0546 0.000	3.13 0.00	8.18 8.18
P'	0.000	0.00	8.18

1.8

ray 1	δ (rad)	δ (deg)	y (mm)
P	0.100	5.73	10.00
1	0.100 −0.0349	5.73 −2.00	30.46 30.46
2	−0.0349 −0.127	−2.00 −7.26	29.76 29.76
P'	−0.127	−7.26	−18.75

ray 2	δ (rad)	δ (deg)	y (mm)
P	−0.100	−5.73	10.00
1	−0.100 −0.0318	−5.73 −1.82	−10.46 −10.46
2	−0.0318 −0.200	−1.82 −1.15	−11.09 −11.09
P'	−0.200	−1.15	−18.75

1.9

	δ (rad)	δ (deg)	y (mm)
P	0.000	0.00	10.00
1	0.000 −0.0666	0.00 −3.82	10.00 10.00
2	−0.0666 −0.111	−3.82 −6.36	8.33 8.33
P'	−0.111	−6.36	0.00

1.10

	δ (rad)	δ (deg)	y (mm)
P	0.000	0.00	10.00
1	0.000 −0.0552	0.00 −3.16	10.00 10.00
2	−0.0552 0.0618	−3.16 3.54	8.45 8.45
P'	0.0618	3.54	0.00

1.11

	δ (rad)	δ (deg)	y (mm)
P	0.000	0.00	10.00
1	0.000 0.0666	0.00 3.82	10.00 10.00
2	0.0666 0.111	3.82 6.36	11.67 11.67
P'	0.111	6.36	0.00

1.12

	δ (rad)	δ (deg)	y (mm)
P	0.300	17.19	5.00
1	0.300 0.191	17.19 10.92	11.00 11.00
2	0.191 0.300	10.92 17.19	17.67 17.67
P'	0.300	17.19	23.67

1.13 (a) 0.0225912, 0.727564, 0.727564, −20.8333
(b) −32.21 mm   32.21 mm
12.06   −12.06
−44.27   44.27

1.14 (a) 0.01375, 1.1875, 0.625, −18.75
(b) −86.36 mm   59.09 mm
−13.64   −35.45
8.18   −13.64
−72.73   94.55

1.15 (a) 0.0411765, 0.176471, 0.176471, −23.5294
(c) −4.29 mm   4.29 mm
20.00   −20.00
−24.29   24.29
(d) 41.82 mm, −3.09 mm; (f) real, real, inverted

1.16 (a) 0.00211111, 1.08889, 0.866667, -26.6667

(c)

-515.79 mm	410.53 mm
-42.11	-63.16
-473.68	473.68

(d) -433.66 mm, 89.11 mm; (f) real, virtual, erect

1.17 (a) -0.00133333, 1.1, 0.933333, $-20.$

(c)

825.00 mm	-700.00 mm
75.00	50.00
750.00	-750.00

(d) -176.74 mm, 34.88 mm; (f) real, virtual, erect

1.18 (a) 0.023594, 0.884502, 0.884502, -9.22509

(c)

-37.49 mm	37.49 mm
4.90	-4.90
-42.38	42.38

(d) 29.92 mm, 5.35 mm; (f) virtual, real, erect

1.25 (a) $s_1 = \left(-D + \sqrt{D(D - 4f')}\right)/2$

$s_1' = \left(\ \ D + \sqrt{D(D - 4f')}\right)/2$

$s_2 = \left(-D - \sqrt{D(D - 4f')}\right)/2$

$s_2' = \left(\ \ D - \sqrt{D(D - 4f')}\right)/2$

(b) $s_1' = -s_2,\ s_2' = -s_1$

1.26 (a) (-117.16 mm, 682.84 mm)
(-682.84 mm, 117.16 mm)
(b) 565.69 mm, (c) -5.83, -0.172

1.28 (b) -2 f', 4 f' (c) -1

1.29 (a) 0.01875, 0.7, 0.625, $-30.$

(b)

-37.33 mm	33.33 mm
16.00	-20.00
-53.33	53.33

1.30 (a) -100.35 mm, 100.35 mm; (b) -3.38

1.31 (a) 35.7 mm; (b) -5.5 mm; (c) real; (d) inverted

1.32 149.7 mm

1.33 $r_1 = 111.7$ mm or $r_2 = -111.7$ mm

1.34 $r_1 = 57.44$ mm, $r_2 = -57.44$ mm

1.35 $r_1 = -51.8$ mm, $r_2 = 51.8$ mm

1.36 (a) 85.0 mm, (b) -86.7 mm, (c) 4160 mm

1.37 (a) 231 mm, (b) -255 mm, (c) 2550 mm

1.38 (b) $1/f' = 1/f_1' + 1/f_2'$

1.39 (a) 518.8 mm; (b) -79.4 mm; (c) real; (d) inverted

References

1. K. Halbach, "Matrix Representation of Gaussian Optics," American Journal of Physics **32**, 90 (1964).

An article on the application of matrices to paraxial optics.

2. E. T. Bell, *Men of Mathematics*, (Simon and Schuster, New York, 1965), pp. 400-402.

A short history of Cayley, the inventor of the theory of matrices.

3. George B. Arfken and Hans J. Weber, *Mathematical Methods for Physicists*, (Academic Press, New York, 1995).

A description of matrix properties, and a good exposition of one of the algorithms for matrix inversion.

4. Allen Nussbaum and Richard A. Phillips, *Contemporary Optics for Scientists and Engineers*, (Prentice-Hall, Englewood Cliffs, 1976).

Also uses sign conventions consistent with analytic geometry.

5. Miles V. Klein, *Optics*, (John Wiley, New York, 1970).

A well done mathematical treatment of many topics.

6. Eugene Hecht, *Optics*, 2nd edition, (Addison-Wesley, Reading, 1987).

A somewhat encyclopedic textbook; many topics covered, usually quite readable. Few examples.

7. Francis A. Jenkins and Harvey E. White, *Fundamentals of Optics*, 4th edition, (McGraw-Hill, New York, 1976).

A time-honored text of optics.

8. Warren J. Smith, *Modern Optical Engineering*, 3rd edition, (McGraw-Hill, New York, 2000).

Recognized as one of the basic textbooks of lens design.