

Ray Vectors and Ray Matrices

QuadriviumPress

Abstract

This comprehensive textbook covers the fundamental principles of optics, from geometrical optics and wave phenomena to advanced topics including lasers, fiber optics, and optical instrumentation. Designed for undergraduate students in physics and electrical engineering, it provides a rigorous yet accessible treatment of topics including ray optics, interference, diffraction, polarization, lasers, and fiber optics.

Keywords optics, physics, education, textbook, light, lenses, lasers, fiber optics, polarization, interference, diffraction

What you should know and be able to do after studying this chapter

- Understand the concept of ray vectors $(n\alpha, y)^T$ and how they describe rays in paraxial optics.
- Know the ray matrices for fundamental optical elements: spherical surfaces, spherical mirrors, and propagation through homogeneous media.
- Apply ray matrix multiplication to analyze optical systems with multiple elements, understanding that matrices are multiplied in reverse order of light propagation.
- Derive and use the lens matrix for both thin and thick lenses, including the concept of optical power $\mathcal{P}$.
- Apply the imaging condition ($C = 0$) to determine when two planes are conjugate (image each other).
- Use the lensmaker formula for lens power and the thin-lens imaging equation for object and image distances.
- Calculate magnification from ray matrix elements and understand the Newtonian form of the lens equation.
- Analyze two-lens systems and calculate combined focal lengths and magnifications.
- Understand principal planes for thick lenses and how they simplify imaging calculations.
- Distinguish the meridional rays in this two-dimensional treatment from skew rays.

Within the first-order, paraxial approximation of Gaussian geometrical optics, a spherical surface maps an object point to a real or virtual image point. A sequence of such surfaces can be treated one surface at a time: the image formed by each surface becomes the object for the next, and intermediate objects or images may be virtual. Real optical systems also have aberrations outside this approximation.

Although this procedure is in principle simple, it is nevertheless convenient in Gaussian geometrical optics to introduce the concept of ray vectors and ray matrices to deal with optical systems consisting of several spherical surfaces. With ray matrices it is easy to derive how the distance of a given ray to the optical axis and its direction change during propagation through an optical system. This in turn can be used to determine the image plane in an optical system for a given object plane.

In any plane perpendicular to the z -axis, a ray is determined by the y -coordinate of the point of intersection of the ray with the plane and the angle α with the optical (z)-axis. This angle has a sign and is defined as follows. Let (y_1, z_1) and (y_2, z_2) be the coordinates of two points on the ray and let the light propagate from point 1 to point 2. Then we define

$$\alpha = \frac{y_2 - y_1}{z_2 - z_1}. \quad (1)$$

Examples of positive and negative α are given in Figure 1. The case $z_2 - z_1 < 0$ occurs when a ray propagates in the negative z -direction after it has been reflected by a mirror. According to the sign convention table in the [Geometrical Optics chapter](#), the refractive index of the ambient medium should after the reflection be taken negative. After a second reflection due to which the ray propagates again in the positive z -direction the refractive index should be chosen positive again.

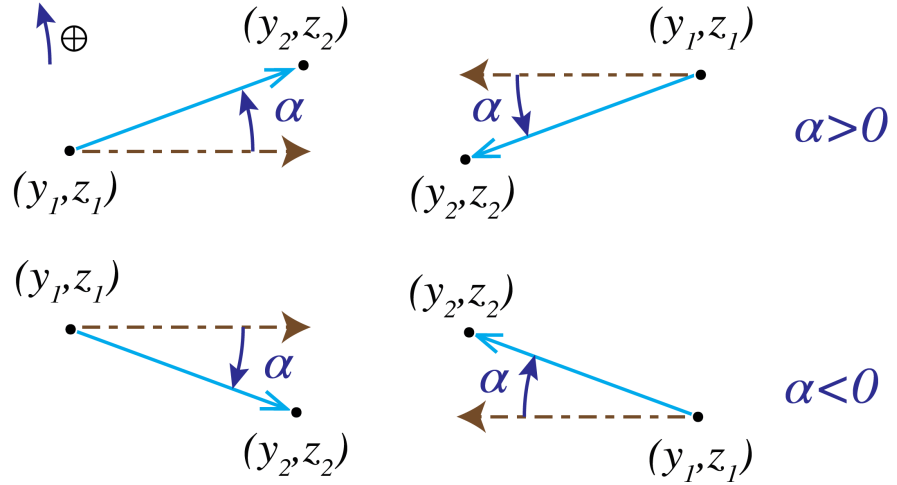


Figure 1: Sign convention for the ray angle. In the upper two figures $\alpha > 0$ while in the lower two figures $\alpha < 0$.

We define the ray vector

$$\begin{pmatrix} n\alpha \\ y \end{pmatrix}, \quad (2)$$

where n is the local refractive index. The definition with the refractive index as factor in the first element of the ray vector turns out to be convenient. The ray vectors of a

ray in any two planes $z = z_1, z = z_2$, with $z_2 > z_1$, are related by a so-called ray matrix:

$$\begin{pmatrix} n_2 \alpha_2 \\ y_2 \end{pmatrix} = \text{cal } M \begin{pmatrix} n_1 \alpha_1 \\ y_1 \end{pmatrix}. \quad (3)$$

where

$$\text{cal } M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}. \quad (4)$$

The elements of matrix $\text{cal } M$ depend on the optical components and materials between the planes $z = z_1$ and $z = z_2$.

As an example consider the ray matrix that relates a ray vector in the plane immediately before a spherical surface (see the [Geometrical Optics chapter](#) for details on spherical interface imaging) to the corresponding ray vector in the plane immediately behind that surface. Using the paraxial Snell's law and small-angle approximations from the [Geometrical Optics chapter](#), it follows

$$n_1 \alpha_1 - n_2 \alpha_2 = \frac{(n_2 - n_1)y_1}{R}, \quad (5)$$

where we have replaced α_2 by $-\alpha_2$ in the paraxial Snell's law equation from the [Geometrical Optics chapter](#), because according to the sign convention, the angle α_2 should be taken negative. Because furthermore $y_2 = y_1$, we conclude

$$\begin{aligned} \begin{pmatrix} n_2 \alpha_2 \\ y_2 \end{pmatrix} &= \begin{pmatrix} n_1 \alpha_1 - \frac{(n_2 - n_1)y_1}{R} \\ y_1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -P \\ 0 & 1 \end{pmatrix} \begin{pmatrix} n_1 \alpha_1 \\ y_1 \end{pmatrix}, \quad \textit{spherical surface}, \end{aligned} \quad (6)$$

where

$$\text{cal } P = \frac{n_2 - n_1}{R}, \quad (7)$$

is as before the **power** of the surface.

Next we consider a spherical mirror with radius of curvature R . We will show that the ray matrix between the planes just before and after the mirror is given by:

$$\begin{pmatrix} n_2 \alpha_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & -\text{cal } P \\ 0 & 1 \end{pmatrix} \begin{pmatrix} n_1 \alpha_1 \\ y_1 \end{pmatrix}, \quad \textit{spherical reflector}, \quad (8)$$

where

$$\text{cal } P = \frac{2n}{R}, \quad (9)$$

is the power of the mirror, $n_1 = n$ but $n_2 = -n$, because the convention is used that if a ray propagates from **right to left** (i.e. in the negative z -direction), the refractive index in the ray vectors and ray matrices is chosen **negative**. Note that when the mirror is flat: $R = \infty$, the ray matrix of the reflector implies

$$n_2\alpha_2 = n_1\alpha_1, \quad (10)$$

which agrees with the fact that $n_2 = -n_1$ and according to Eq. (1) α_2 and α_1 have opposite sign for a mirror.

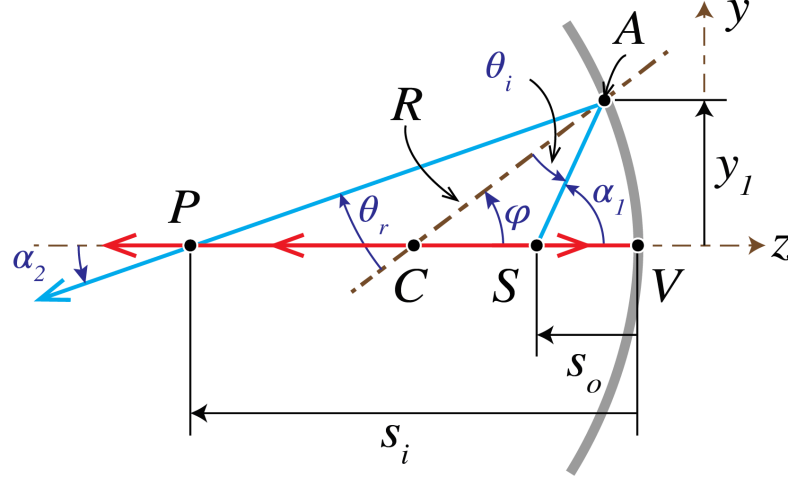


Figure 2: Reflection by a mirror.

With all angles positive for the moment, it follows from Figure 2

$$\alpha_1 = \theta_i + \phi, \quad (11)$$

$$\alpha_2 = \phi - \theta_r = \phi - \theta_i. \quad (12)$$

Hence,

$$\alpha_2 = -\alpha_1 + 2\phi. \quad (13)$$

Now

$$\phi \approx \frac{y_1}{R} \quad (14)$$

In the situation drawn in Figure 2, Eq. (1) implies that both α_2 and α_1 are positive. By choosing the refractive index negative after reflection, we conclude from Eq. (13) and Eq. (14):

$$n_2\alpha_2 = -n\alpha_2 = n\alpha_1 - \frac{2n}{R}y_1 = n_1\alpha_1 - \frac{2n}{R}y_1. \quad (15)$$

This proves Eq. (8).

We now consider the ray matrix when a ray propagates from a plane z_1 to a plane z_2 through a medium with refractive index n . In that case we have $\alpha_2 = \alpha_1$ and $y_2 = y_1 + \alpha_1(z_2 - z_1)$, hence

$$\text{cal } M = \begin{pmatrix} 1 & 0 \\ \frac{z_2 - z_1}{n} & 1 \end{pmatrix}, \quad \text{homogeneous space.} \quad (16)$$

Note that if the light propagates from the left to the right: $z_2 > z_1$ and hence $z_2 - z_1$ in the first column and second row of the matrix is positive, i.e. it is the distance between the planes.

For two planes between which there are a number of optical components, possibly separated by regions with homogeneous material (e.g. air), the ray matrix can be obtained by multiplying the matrices of the individual components and of the homogeneous regions. The order of the multiplication of the matrices is such that the **right-most matrix corresponds to the first component that is encountered while propagating**, and so on.

The matrices here describe **meridional rays** in a plane containing the optical axis, giving a two-dimensional model. **Skew rays** do not intersect the optical axis and require a fuller three-dimensional treatment; a skew ray can still be paraxial.

Remarks.

1. In matrix [Eq. \(16\)](#), z_1 and z_2 are **signed coordinates**.
2. Instead of choosing the refractive index negative in ray vectors of rays that propagate from right to left, one can reverse the direction of the positive z -axis after every reflection. The convention to make the refractive index negative is however more convenient in ray tracing software.
3. The determinants of the ray matrices [Eq. \(6\)](#), [Eq. \(8\)](#) and [Eq. \(16\)](#) are all 1. Since all ray matrices considered below are products of these elementary matrices, the determinant of every ray matrix considered is unity.

1. THE LENS MATRIX

We apply ray matrices to a lens. [Figure 3](#) shows a lens with two spherical surfaces. The refractive index of the lens is n_l and the refractive indices of the media to the left and right of the lens are n_1 and n_2 , respectively. Let the distance between the vertices be d .

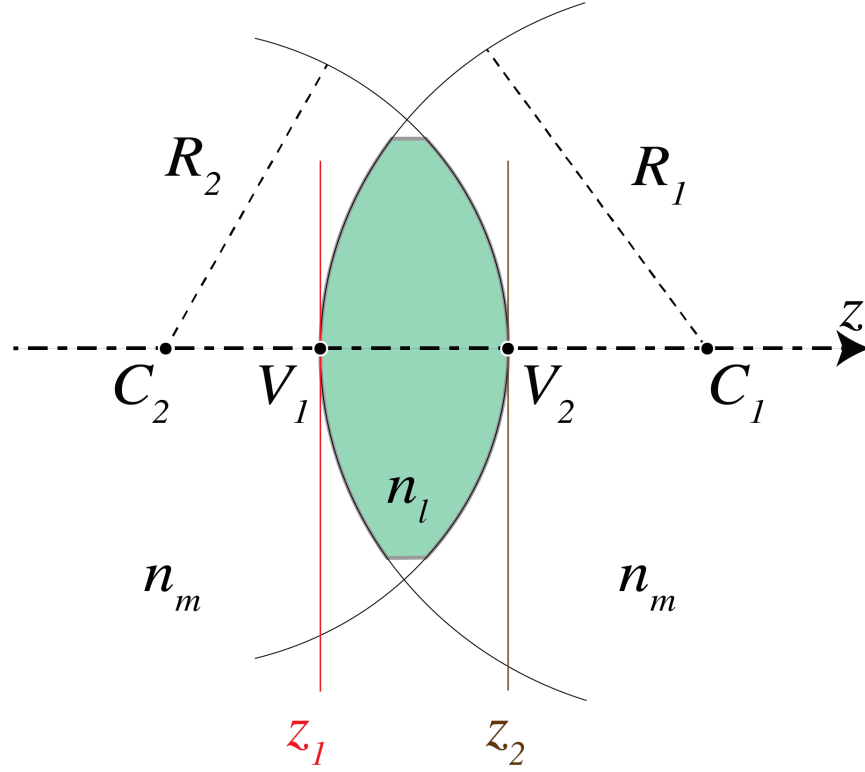


Figure 3: A lens with thickness d . The ray matrix is defined between the planes immediately before and after the lens.

We will first derive the matrix which maps the ray vector in the plane **immediately in front** of the lens to that in the plane **immediately behind** the lens. Let

$$\begin{pmatrix} n_1 \alpha_1 \\ y_1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} n_2 \alpha_2 \\ y_2 \end{pmatrix} \quad (17)$$

be two vectors in the two planes which correspond to the same ray. The ray is first refracted by the spherical surface with radius R_1 and center C_1 . Using [Eq. \(6\)](#) and [Eq. \(7\)](#) it follows that the matrix between the ray vectors just before and just behind the spherical surface with radius R_1 and center C_1 is given by

$$\text{cal } M_1 = \begin{pmatrix} 1 & -\text{cal } P_1 \\ 0 & 1 \end{pmatrix}, \quad (18)$$

where

$$\text{cal } P_1 = \frac{n_l - n_1}{R_1}. \quad (19)$$

The ray then propagates over the distance d through the material of which the lens is made. The matrix that maps ray vectors from the plane inside the lens immediately behind the left spherical surface to a ray vector in the plane immediately before the right spherical surface follows from [Eq. \(16\)](#):

$$\text{cal } M_2 = \begin{pmatrix} 1 & 0 \\ \frac{d}{n_l} & 1 \end{pmatrix}. \quad (20)$$

Finally, the matrix that maps ray vectors from the plane in the lens immediately before the second spherical surface to vectors in the plane immediately behind it is

$$\text{cal } M_3 = \begin{pmatrix} 1 & -\text{cal } P_2 \\ 0 & 1 \end{pmatrix}. \quad (21)$$

with

$$\text{cal } P_2 = \frac{n_2 - n_l}{R_2}. \quad (22)$$

Hence the matrix that maps ray vectors in the plane immediately before the lens to ray vectors in the plane immediately behind the lens is given by the matrix product:

$$\begin{aligned} \text{cal } M &= \text{cal } M_3 \text{cal } M_2 \text{cal } M_1 \\ &= \begin{pmatrix} 1 - \frac{d}{n_l} P_2 & -P_1 - P_2 + \frac{d}{n_l} P_1 P_2 \\ \frac{d}{n_l} & 1 - \frac{d}{n_l} P_1 \end{pmatrix}, \quad \text{ lens.} \end{aligned} \quad (23)$$

The quantity

$$\text{cal } P = \text{cal } P_1 + \text{cal } P_2 - \frac{d}{n_l} \text{cal } P_1 \text{cal } P_2 \quad (24)$$

is called the **power** of the lens. It has dimension 1/length and is given in diopter (cal D), where $1 \text{ cal } D = \text{m}^{-1}$. The power can be positive or negative. The space to the left of the lens is called the **object space** and that to the right of the lens is called the **image space**.

2. FOCUSING WITH A THIN LENS

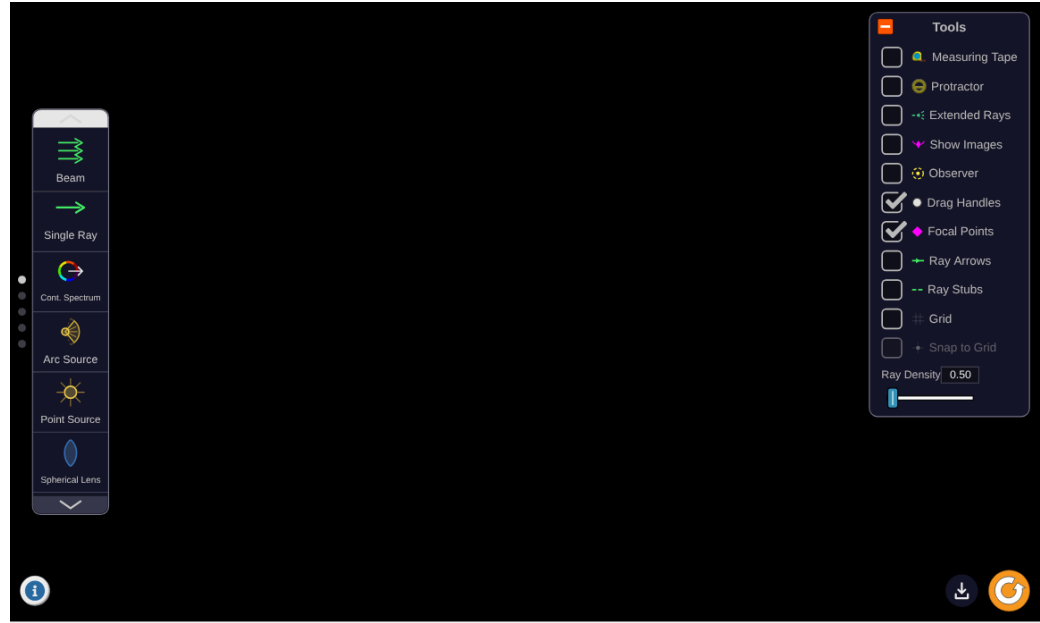


Figure 4: Place a thin lens on the bench, send in a parallel beam, and measure where it crosses the axis — the second focal point of the matrix $\text{cal } M$ derived above. Add an off-axis object ray and recover the imaging relation of the next section.

Interactive simulation: [Optics Lab](#)

For a thin lens the vertices V_1 and V_2 coincide and $d = 0$, hence Eq. (23) becomes

$$\text{cal } M = \begin{pmatrix} 1 & -P \\ 0 & 1 \end{pmatrix}, \quad \text{thinlens}, \quad (25)$$

where

$$P = P_1 + P_2 = \frac{n_l - n_1}{R_1} - \frac{n_2 - n_l}{R_2}, \quad (26)$$

This power–curvature relation is the **lensmaker formula** for a thin lens. The origin of the coordinate system is chosen at the common vertex $V_1 = V_2$.

By considering a ray in medium 1 which is parallel to the optical axis ($\alpha_1 = 0$) and at height y_1 , we get $n_2\alpha_2 = -Py_1$ and $y_2 = y_1$. Hence, when $P > 0$, the angle α_2 of the ray has sign opposite to y_2 and therefore the ray in image space is bent back to the optical axis, yielding a **second focal point** or **image focal point** F_i . Its z -coordinate f_i is:

$$f_i = -\frac{y_2}{\alpha_2} = \frac{n_2}{\text{cal } P}. \quad (27)$$

For a ray emerging in image space at height y_2 and parallel to the optical axis: $\alpha_2 = 0$, we have $y_1 = y_2$ and

$$n_1\alpha_1 = Py_1. \quad (28)$$

If the power is positive: $\text{cal } P > 0$, the angle α_1 has the same sign as y_1 , which implies that the ray in object space has intersected the optical axis in a point F_o with z -coordinate: $z = f_o$

$$f_o = -\frac{y_1}{\alpha_1} = -\frac{n_1}{\text{cal } P}. \quad (29)$$

The point F_o is called the **first focal point** or **object focal point**.

We conclude that when the power $\text{cal } P$ of the lens is positive, $f_i > 0$ and $-f_o > 0$, which means that the image and object focal points are in the image and object space, respectively, hence they are both real. A lens with positive power is called **convergent** or **positive**. It makes incident bundles of rays convergent or less divergent.

A lens with negative power is called divergent and has $f_i < 0$, $-f_o < 0$. It makes incident rays more divergent or less convergent. Incident rays which are parallel to the optical axis are refracted away from the optical axis and seem to come from a point in front of the lens with z -coordinate $f_i < 0$. Hence the image focal point does not correspond to a location where there is an actual concentration of light intensity, i.e. it is virtual. The object focal point is a virtual object point, because only a bundle of incident rays that are converging to a certain point behind the negative lens can be turned into a bundle of rays parallel to the optical axis.

With the results obtained for the focal coordinates we can rewrite the lens matrix of a thin lens as

$$\text{cal } M = \begin{pmatrix} 1 & -\frac{n_2}{f_i} \\ 0 & 1 \end{pmatrix}, \quad \text{thinlens}. \quad (30)$$

3. IMAGING WITH A THIN LENS

We first consider a general ray matrix [Eq. \(3\)](#), [Eq. \(4\)](#) between two planes $z = z_1$ and $z = z_2$ and ask the following question: what are the properties of the ray matrix such that the two planes are images of each other, or (as this is also called) are each other's conjugate? Clearly for these planes to be each other's image, we should have that for every point coordinate y_1 in the plane $z = z_1$ there is a point with some coordinate y_2 in the plane $z = z_2$ such that any ray through (y_1, z_1) (within some cone of rays) will pass through the point (y_2, z_2) . Hence for any angle α_1 (in some interval of angles) there is an angle α_2 such that [Eq. \(3\)](#) is valid. This means that for any y_1 there is a y_2 such that for all angles α_1 :

$$y_2 = Cn_1\alpha_1 + Dy_1, \quad (31)$$

This requires that

$$C = 0, \quad \text{condition for imaging}. \quad (32)$$

The ratio of y_2 and y_1 is the magnification M . Hence,

$$M = \frac{y_2}{y_1} = D, \quad (33)$$

is the **magnification** of the image (this quantity has sign).

To determine the image by a thin lens we first derive the ray matrix between two planes $z = z_1 < 0$ and $z = z_2 > 0$ on either side of the thin lens. The origin of the coordinate system is again at the vertex of the thin lens. This ray matrix is the product of the matrix for propagation from $z = z_1$ to the plane immediately in front of the lens, the matrix of the thin lens and the matrix for propagation from the plane immediately behind the lens to the plane $z = z_2$:

$$\begin{aligned} \text{cal } M &= \begin{pmatrix} 1 & 0 \\ \frac{z_2}{n_2} & 1 \end{pmatrix} \begin{pmatrix} 1 & -\text{cal } P \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \frac{-z_1}{n_1} & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 + \frac{z_1}{n_1} \text{cal } P & -\text{cal } P \\ -\frac{z_1}{n_1} + \frac{z_2}{n_2} + \frac{z_1 z_2}{n_1 n_2} \text{cal } P & 1 - \frac{z_2}{n_2} \text{cal } P \end{pmatrix} \end{aligned} \quad (34)$$

The imaging condition Eq. (32) implies:

$$-\frac{n_1}{s_o} + \frac{n_2}{s_i} = \text{cal } P, \quad \text{thin -lens imaging equation}, \quad (35)$$

where we have written $s_o = z_1$ and $s_i = z_2$ for the z -coordinates of the object and the image. Because for the thin lens matrix Eq. (34): $D = 1 - z_2/f_i$, it follows by using Eq. (35) that the magnification Eq. (33) is given by

$$M = \frac{y_i}{y_o} = 1 - \frac{s_i}{f_i} = \frac{n_1 s_i}{n_2 s_o}, \quad (36)$$

where we have written now y_o and y_i instead of y_1 and y_2 , respectively.

Remark. The thin-lens imaging equation can alternatively be derived by using the single-surface imaging formula from the [Geometrical Optics chapter](#) for the two spherical surfaces of the lens. We first image a given point S by the left spherical surface using that imaging formula as if the second surface were absent. The obtained intermediate image P' is then imaged by the second spherical surface as if the first surface were absent. P' can be a real or virtual object for the second surface. The derivation is carried out in Problem 2.5.

Analogous to the case of a single spherical surface, an image is called a **real image** if it is to the right of the lens ($s_i > 0$) and is called a **virtual image** when it seems to be to the left of the lens ($s_i < 0$). An object is called a **real object** if it is to the left of the lens ($s_o < 0$) and is a **virtual object** if it seems to be to the right of the lens ($s_o > 0$). For a positive lens: $\text{cal } P > 0$ and hence Eq. (35) implies that $s_i > 0$ provided $|s_o| > |f_o|$, which means that the image by a convergent lens is real if the object is further from the lens than the object focal point F_o . The case $s_o > 0$ corresponds to a virtual object, i.e. to the case of a converging bundle of incident rays, which for an observer in object space seems to converge to a point at distance s_o behind the lens. A convergent lens ($f_i > 0$) will then make an image between the lens and the second

focal point. In contrast, a diverging lens ($f_i < 0$) can turn the incident converging bundle into a real image only if the virtual object point is between the lens and the focal point. If the virtual object point has larger distance to the lens, the convergence of the incident bundle is too weak and the diverging lens then refracts this bundle into a diverging bundle of rays which seem to come from a virtual image point in front of the lens ($s_i < 0$).

Instead of using ray matrices, one can construct the image with a ruler. Consider the imaging of a finite object S_1S_2 as shown in Figure 5 for the case that the media to the left and right of the lens are the same. Let y_o be the y-coordinate of S_2 . We have $y_o > 0$ when the object is above the optical axis.

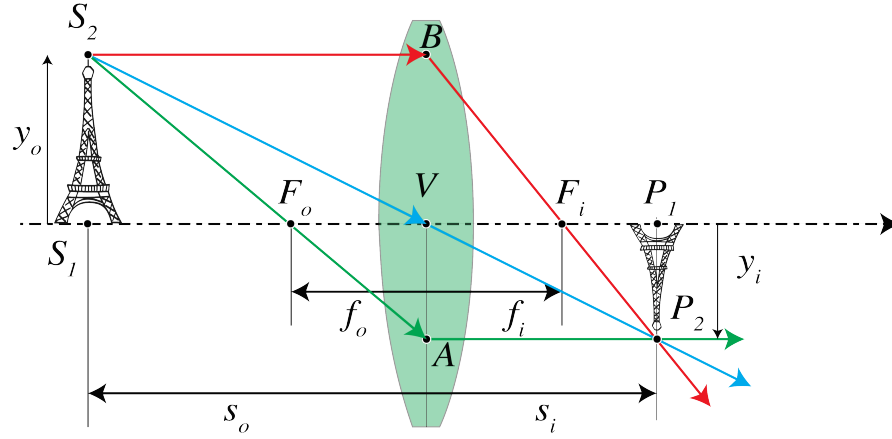


Figure 5: Object and image for a thin lens.

Draw the ray through the focal point F_o in object space and the ray through the center V of the lens. The first ray becomes parallel in image space. The latter intersects both surfaces of the lens almost in their (almost coinciding) vertices and therefore the refraction is opposite at both surfaces and the ray exits the lens parallel to its direction of incidence. Furthermore, its lateral displacement can be neglected because the lens is thin. (Of course, this is not correct when the refractive indices to the left and right of the lens are different). Hence, **the ray through the center of a thin lens is not refracted**. The intersection in image space of the two rays gives the location of the image point P_2 of S_2 . The image is real if the intersection occurs in image space and is virtual otherwise. For the case of a convergent lens with a real object with $y_o > 0$ as shown in Figure 5, it follows from the similar triangles $\triangle BVF_i$ and $\triangle P_2P_1F_i$ that

$$\frac{y_o}{|y_i|} = \frac{f_i}{s_i - f_i}. \quad (37)$$

From the similar triangles $\triangle S_2S_1F_o$ and $\triangle AVF_o$:

$$\frac{|y_i|}{y_o} = \frac{f_i}{f_o - s_o}. \quad (38)$$

Here we used $|f_o| = f_i$. (The absolute value of y_i is taken because according to our sign convention y_i in Figure 5 is negative whereas Eq. (38) is a ratio of lengths). By multiplying these two equations we get the **Newtonian form** of the lens equation (valid when $n_2 = n_1$):

$$x_o x_i = -f_i^2 = -f_o^2, \quad (39)$$

where x_o and x_i are the z -coordinates of the object and image relative to those of the first and second focal point, respectively:

$$x_o = s_o - f_o, \quad x_i = s_i - f_i. \quad (40)$$

Hence x_o is negative if the object is to the left of F_o and x_i is positive if the image is to the right of F_i .

The **transverse magnification** is

$$M = \frac{y_i}{y_o} = \frac{s_i}{s_o} = -\frac{x_i}{f_i}, \quad (41)$$

where the second identity follows from considering the similar triangles $\Delta P_2 P_1 F_i$ and $\Delta B V F_i$ in Figure 5. A positive M means that the image is erect, a negative M means that the image is inverted.

All equations are also valid for a thin negative lens and for virtual objects and images. Examples of real and virtual object and image points for a positive and a negative lens are shown in Figure 6 and Figure 7.

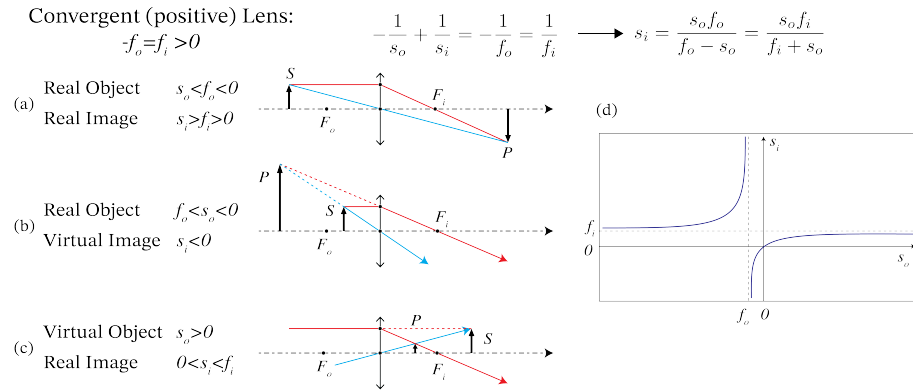


Figure 6: Real and virtual objects and images for a convergent thin lens in the same medium on both sides, so $-f_o = f_i > 0$. (a) A real object beyond the front focal point ($s_o < f_o$) gives a real image ($s_i > 0$). (b) A real object between that focal point and the lens ($f_o < s_o < 0$) gives a virtual image ($s_i < 0$). (c) Incident rays converging toward a point behind the lens define a virtual object ($s_o > 0$) and give a real image. (d) The graph shows s_i versus s_o from the thin-lens imaging equation.

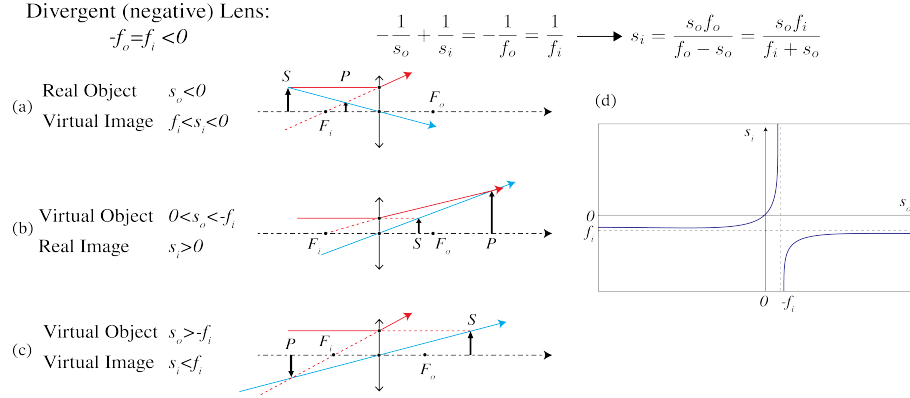


Figure 7: Real and virtual objects and images for a divergent thin lens in the same medium on both sides, so $-f_o = f_i < 0$. (a) A real object ($s_o < 0$) gives a virtual image ($s_i < 0$). For a virtual object, incident rays would converge at S behind the lens if the lens were absent. (b) If $0 < s_o < -f_i$, the image is real. (c) If $s_o > -f_i$, the image is virtual. (d) The graph shows s_i versus s_o from the thin-lens imaging equation.

4. TWO THIN LENSES

The ray matrix is a suitable method to study the imaging of a system consisting of several thin lenses. For two lenses however, the imaging can still easily be obtained by construction. We simply construct the image obtained by the first lens as if the second lens were not present and use this image as a (possibly virtual) object for the second lens. In Figure 8 an example is shown where the distance between the lenses is larger than the sum of their focal lengths. First the image P' of S is constructed as obtained by L_1 as if L_2 were not present. We construct the intermediate image P' due to lens L_1 using rays 2 and 3. P' is a real image for lens L_1 and also a real object for lens L_2 . Ray 3 is parallel to the optical axis between the two lenses and is thus refracted by lens L_2 through its back focal point F_{i2} . Ray 4 is the ray from P' through the center of lens L_2 . The image point P is the intersection of rays 3 and 4.

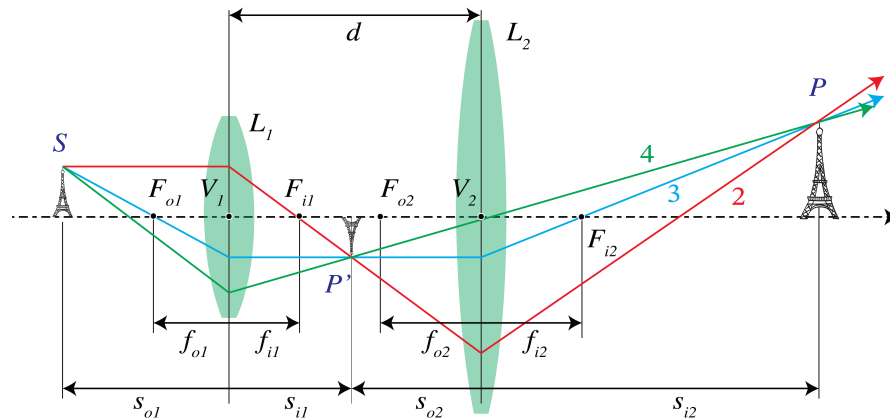


Figure 8: Two thin lenses separated by a distance that is larger than the sum of their focal lengths.

In the case of Figure 9 the distance d between the two positive lenses is smaller than their focal lengths. The intermediate image P' is a real image for L_1 obtained as the intersection of rays 2 and 4 passing through the object and image focal points F_{o1} and F_{i1} of lens L_1 . P' is now a virtual object for lens L_2 . To find its image by L_2 , draw ray 3 from P' through the center of lens L_2 back to S (this ray is refracted by lens L_1 but not by L_2) and draw ray 4 as refracted by lens L_2 . Since ray 4 is parallel to the optical axis between the lenses, it passes through the back focal point F_{2i} of lens L_2 . The intersection point of rays 3 and 4 is the final image point P .

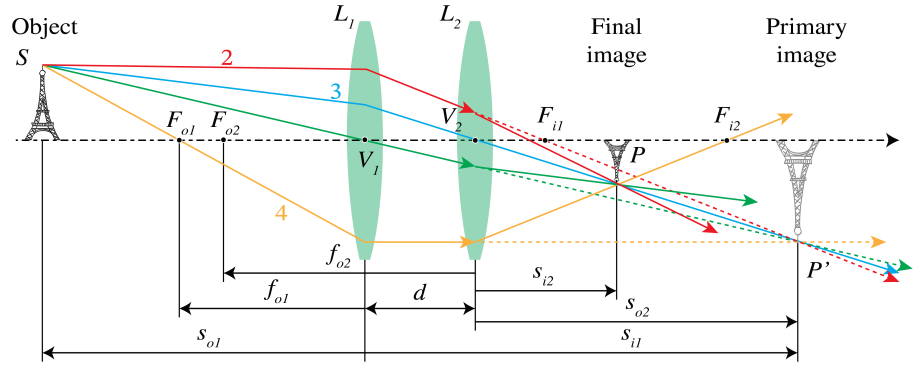


Figure 9: Two thin lenses at a distance smaller than their focal lengths.

It is easy to express the z -coordinate s_i with respect to the coordinate system with origin at the vertex V_2 of the final image point, in terms of the z -coordinate s_o with respect to the origin at the vertex of lens L_1 of the object point. We use the thin-lens imaging equation for each lens while taking care that the proper local coordinate systems are used. The intermediate image P' due to lens L_1 has z -coordinate s_{1i} with respect to the coordinate system with origin at the vertex V_1 , which satisfies:

$$-\frac{1}{s_o} + \frac{1}{s_{1i}} = \frac{1}{f_{1i}}. \quad (42)$$

As object for lens L_2 , P' has z -coordinate with respect to the coordinate system with origin at V_2 given by: $s_{2o} = s_{1i} - d$, where d is the distance between the lenses.

Hence, with $s_i = s_{2i}$ the imaging equation for lens L_2 implies:

$$-\frac{1}{s_{1i} - d} + \frac{1}{s_i} = \frac{1}{f_{2i}}. \quad (43)$$

By solving Eq. (42) for s_{1i} and substituting the result into Eq. (43), we find

$$s_i = \frac{-df_{1i}f_{2i} + f_{2i}(f_{1i} - d)s_o}{f_{1i}(f_{2i} - d) + (f_{1i} + f_{2i} - d)s_o}, \quad \text{two thin lenses.} \quad (44)$$

By taking the limit $s_o \rightarrow -\infty$, we obtain the z -coordinate f_i of the image focal point of the two lenses, while $s_i \rightarrow \infty$ gives the z -coordinate f_o of the object focal point:

$$f_i = \frac{(f_{1i} - d)f_{2i}}{f_{1i} + f_{2i} - d}, \quad (45)$$

$$f_o = -\frac{(f_{2i} - d)f_{1i}}{f_{1i} + f_{2i} - d}, \quad (46)$$

We found in [Focusing with a Thin Lens](#) that when the refractive indices of the media before and after the lens are the same, the signed object and image focal coordinates of a thin lens have equal magnitudes and opposite signs. As [Eq. \(45\)](#) and [Eq. \(46\)](#) show, their distances from the first and last vertices can differ for separated lenses.

By construction using the intermediate image, it is clear that the magnification of the two-lens system is the product of the magnifications of the two lenses:

$$M = M_1 M_2. \quad (47)$$

Remarks.

1. When $f_{1i} + f_{2i} = d$, parallel incident rays emerge parallel: the system is **afocal**. Telecentricity additionally depends on the aperture-stop placement.
2. In the limit where the lenses are very close together: $d \rightarrow 0$, [Eq. \(44\)](#) becomes

$$-\frac{1}{s_o} + \frac{1}{s_i} = \frac{1}{f_{1i}} + \frac{1}{f_{2i}}. \quad (48)$$

The focal length f_i of the system of two lenses in contact thus satisfies:

$$\frac{1}{f_i} = \frac{1}{f_{1i}} + \frac{1}{f_{2i}}. \quad (49)$$

In particular, by using two identical lenses in contact, the focal length is halved.

3. Although for two lenses the image coordinate can still be expressed relatively easily in the object distance, for systems with more lenses finding the overall ray matrix and then using the image condition [Eq. \(32\)](#) is a much better strategy.

5. THE THICK LENS

At the left of [Figure 10](#) a thick lens is shown. The object focal point is defined as the point whose rays are refracted such that the emerging rays are parallel to the optical axis. By extending the incident and emerging rays by straight segments, the points of intersection are found to be on a curved surface which, close to the optical axis (in the paraxial approximation), is approximately a plane perpendicular to the optical axis. This plane is called the **primary principal plane** and its intersection with the optical axis is called the primary principal point H_1 .

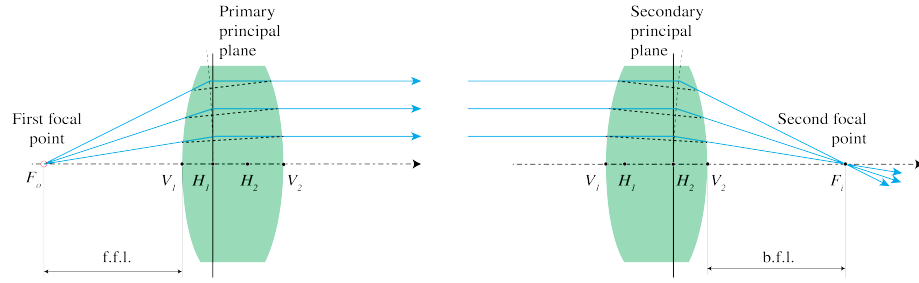


Figure 10: Principal planes of a thick lens, with front and back focal lengths: f.f.l and b.f.l.

By considering incident rays which are parallel to the optical axis and therefore focused in the image focal point, the **secondary principal plane** and secondary principal point H_2 are defined in a similar way (see the drawing at the right in [Figure 10](#)). The principal planes can be outside the lens. For meniscus lenses, this is usually the case as shown in [Figure 11](#). It can be seen from [Figure 10](#) that the principal planes are images of each other, with unit magnification. Hence, if an object is placed in the primary principal plane (hypothetically if this plane is inside the lens), its image is in the secondary principal plane. The image is erect and has unit magnification.

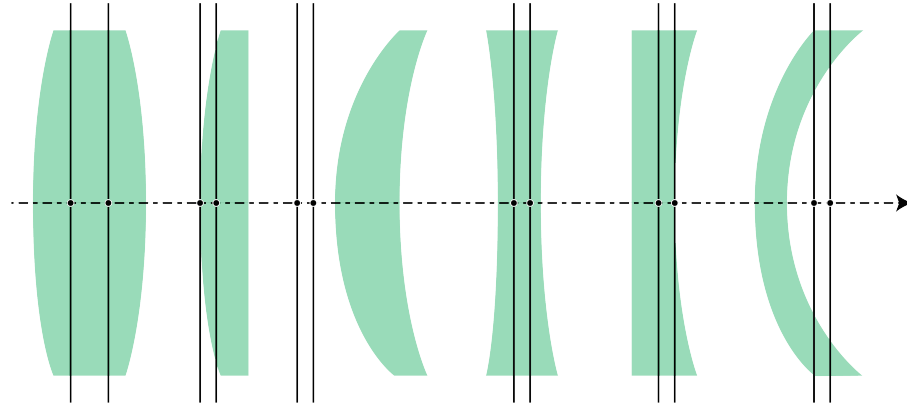


Figure 11: Position of the principal planes for several lenses.

Now, if the object coordinates and object focal point are defined with respect to the origin at H_1 and the image coordinates and image focal point are defined with respect to the origin at H_2 , the imaging equation [Eq. \(35\)](#) can also be used for a thick lens.

Proof

We recall the result [Eq. \(23\)](#) for the ray matrix between the planes through the front and back vertices V_1, V_2 of a thick lens with refractive index n_l and thickness d :

$$\text{cal } M_{V_1 V_2} = \begin{pmatrix} 1 - \frac{d}{n_l} P_2 & -P \\ \frac{d}{n_l} & 1 - \frac{d}{n_l} P_1 \end{pmatrix}, \quad \text{thick lens}, \quad (50)$$

where

$$P_1 = \frac{n_l - n_1}{R_1}, \quad P_2 = \frac{n_2 - n_l}{R_2}, \quad (51)$$

and n_1, n_2 are the refractive indices to the left and the right of the lens, respectively, and where

$$P = P_1 + P_2 - \frac{d}{n_l} P_1 P_2. \quad (52)$$

If h_1 is the z -coordinate of the first principal point H_1 with respect to the coordinate system with origin at vertex V_1 , we have according to Eq. (16) for the ray matrix between the primary principal plane and the plane through vertex V_1

$$\text{cal } M_1 = \begin{pmatrix} 1 & 0 \\ -\frac{h_1}{n_1} & 1 \end{pmatrix}. \quad (53)$$

Similarly, if h_2 is the coordinate of the secondary principal point H_2 with respect to the coordinate system with V_2 as origin, the ray matrix between the plane through vertex V_2 and the secondary principal plane is

$$\text{cal } M_2 = \begin{pmatrix} 1 & 0 \\ \frac{h_2}{n_2} & 1 \end{pmatrix}. \quad (54)$$

The ray matrix between the two principal planes is then

$$\text{cal } M_{H_1 H_2} = \text{cal } M_2 \text{ cal } M_{V_1 V_2} \text{ cal } M_1. \quad (55)$$

The coordinates h_1 and h_2 can be found by imposing the imaging condition on the resulting matrix Eq. (32): $C = 0$ and the condition that the magnification should be unity: $D = 1$, which follows from Eq. (33). We omit the details and only give the resulting expressions here:

$$h_1 = \frac{n_1}{n_l} \frac{P_2}{P} d, \quad (56)$$

$$h_2 = -\frac{n_2}{n_l} \frac{P_1}{P} d. \quad (57)$$

With these results, Eq. (55) becomes

$$\text{cal } M_{H_1 H_2} = \begin{pmatrix} 1 & -P \\ 0 & 1 \end{pmatrix}. \quad (58)$$

We see that **the ray matrix between the principal planes is identical to the ray matrix of a thin lens** Eq. (25). We therefore conclude that if the coordinates in object space are chosen with respect to the origin in the primary principal point H_1 , and the coordinates in image space are chosen with respect to the origin in the secondary principal point H_2 , the expressions for the first and second focal points and for the coordinates of the image point in terms of those of the object point are identical to those for a thin lens. An example of imaging by a thick lens is shown in Figure 12.

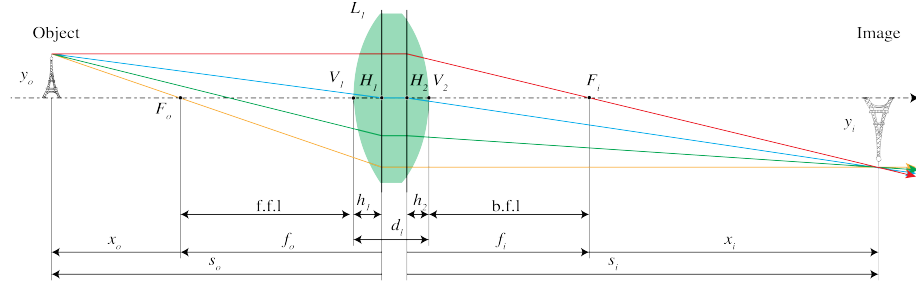


Figure 12: Thick-lens ray construction through the principal planes H_1 and H_2 . When the external media have the same refractive index, the signed focal coordinates obey $f_i = -f_o$; object and image coordinates are measured from H_1 and H_2 , respectively.

6. CHAPTER SUMMARY

- **Ray vectors** $(n\alpha, y)^T$ describe a ray by its angle with the axis (weighted by refractive index) and its height.
- **Ray matrices** (ABCD matrices) relate ray vectors before and after optical elements: $\begin{pmatrix} n_2\alpha_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} n_1\alpha_1 \\ y_1 \end{pmatrix}$.
- **System matrix**: For multiple elements, multiply matrices in reverse order of light propagation.
- **Propagation matrix**: $\begin{pmatrix} 1 & 0 \\ d/n & 1 \end{pmatrix}$ for distance d in medium with index n .
- **Spherical surface matrix**: Power $\mathcal{P} = (n_2 - n_1)/R$; matrix has $-\mathcal{P}$ in upper right.
- **Thin lens matrix**: $\begin{pmatrix} 1 & -\mathcal{P} \\ 0 & 1 \end{pmatrix}$; its signed focal coordinates are $f_i = n_2/\mathcal{P}$ and $f_o = -n_1/\mathcal{P}$.
- **Imaging condition**: $C = 0$ means two planes are conjugate (image each other).
- **Lensmaker formula**: The thin-lens power follows from its surface curvatures; the **imaging equation** $-n_1/s_o + n_2/s_i = \mathcal{P}$ relates object and image coordinates.
- **Magnification**: $M = D$ when $C = 0$; for equal external indices, **Newton's equation** is $x_o x_i = f_o f_i = -f_i^2$.
- **Two-lens systems**: In air, the effective power is $\mathcal{P}_{\text{upright eff}} = \mathcal{P}_1 + \mathcal{P}_2 - d\mathcal{P}_1\mathcal{P}_2$; focal coordinates measured from the first and last lenses also depend on d .
- **Principal planes** for thick lenses: The ray matrix between principal planes equals that of a thin lens.