1 Gaussian Optics and Uncertainty Principle

This chapter contains *Gaussian optics* and employs a matrix formalism to describe optical image formation through light rays. In optics, a ray is an idealized model of light. However, in a subsequent chapter (Chapter 3, Section 3.5), we will also see that a matrix formalism can also be used to describe, for example, a Gaussian laser beam under diffraction through the wave optics approach. The advantage of the matrix formalism is that any ray can be tracked during its propagation through the optical system by successive matrix multiplications, which can be easily programmed on a computer. This is a powerful technique and is widely used in the design of optical elements. In the chapter, some of the important concepts in resolution, depth of focus, and depth of field are also considered based on the ray approach.

1.1 Gaussian Optics

Gaussian optics, named after Carl Friedrich Gauss, is a technique in geometrical optics that describes the behavior of light rays in optical systems under the paraxial approximation. We take the *optical axis* to be along the z-axis, which is the general direction in which the rays travel, and our discussion is confined to those rays that lie in the x–z plane and that are close to the optical axis. In other words, only rays whose angular deviation from the optical axis is small are considered. These rays are called *paraxial rays*. Hence, the sine and tangent of the angles may be approximated by the angles themselves, that is, $\sin\theta \approx \tan\theta \approx \theta.$ Indeed, the mathematical treatment is simplified greatly because of the linearization process involved. For example, a linearized form of *Snell's law of refraction*, $n_1 \sin\phi_i = n_2 \sin\phi_t$, is $n_1\phi_i = n_2\phi_t$. Figure 1.1-1 shows ray refraction for Snell's law. ϕ_i and ϕ_t are the angles of incidence and refraction, respectively, which are measured from the normal, ON, to the interface POQ between Media 1 and 2. Media 1 and 2 are characterized by the constant refractive indices, n_1 and n_2, respectively. In the figure, we also illustrate the *law of reflection,* that is, $\phi_i = \phi_r$, where ϕ_r is the *angle of reflection*. Note that the incident ray, the refracted ray, and the reflected ray all lie in the same *plane of incidence*.

Consider the propagation of a paraxial ray through an optical system as shown in Figure 1.1-2. A ray at a given z-plane may be specified by its height x from the optical axis and by its launching angle θ. The convention for the angle is that θ is measured in radians and is anticlockwise positive from the z-axis. The quantities (x, θ) represent the

https://doi.org/10.1017/9781009053204.002 Published online by Cambridge University Press

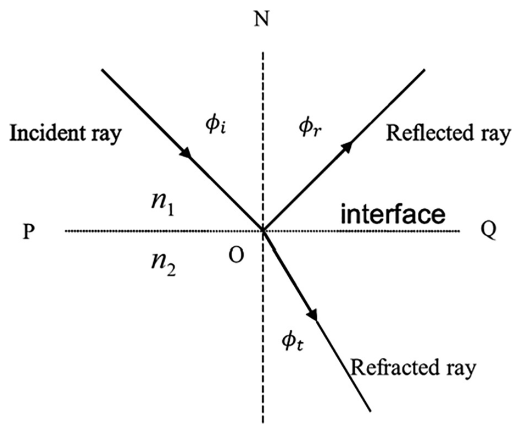


Figure 1.1-1 Geometry for Snell's law and law of reflection

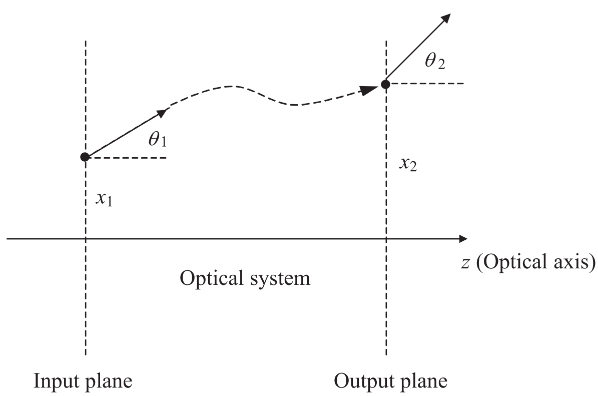


Figure 1.1-2 Ray propagating in an optical system

coordinates of the ray for a given z-plane. However, instead of specifying the angle the ray makes with the z-axis, another convention is used. We replace the angle θ by the corresponding $v = n\theta$, where n is the refractive index of the medium in which the ray is traveling. As we will see later, the use of this convention ensures that all the matrices involved are positive unimodular. A *unimodular matrix* is a real square matrix with determinant +1 or −1, and a positive unimodular matrix has determinant +1.

To clarify the discussion, we let a ray pass through the input plane with the *input ray coordinates* $(x_1, v_1 = n_1\theta_1)$. After the ray passes through the optical system, we denote the *output ray coordinates* $(x_2, v_2 = n_2\theta_2)$ on the output plane. In the paraxial approximation, the corresponding output quantities are linearly dependent on the input quantities. In other words, the output quantities can be expressed in terms of the weighted sum of the input quantities (known as the *principle of superposition*) as follows:

https://doi.org/10.1017/9781009053204.002 Published online by Cambridge University Press

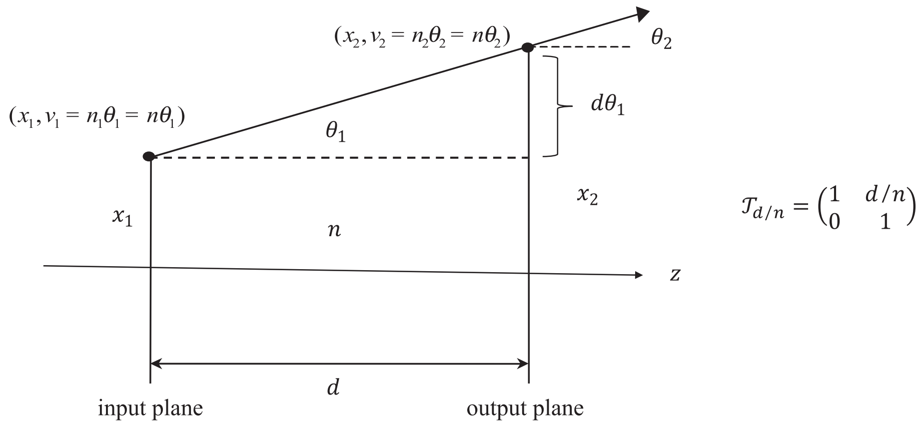


Figure 1.1-3 Ray propagating in a homogeneous medium with the input and output coordinates $(x_1, v_1 = n_1\theta_1)$ and $(x_2, v_2 = n_2\theta_2)$, respectively

$$x_2 = Ax_1 + Bv_1 \text{ and } v_2 = Cx_1 + Dv_1,$$

where A, B, C, and D are the weight factors. We can cast the above equations into a matrix form as

$$\begin{pmatrix} x_2 \\ v_2 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} x_1 \\ v_1 \end{pmatrix}. \tag{1.1-1}$$

The *ABCD* matrix in Eq. (1.1-1) is called the *ray transfer matrix*, or the *system matrix* S, if it is represented by the multiplication of ray transfer matrices. In what follows, we shall derive several important ray transfer matrices.

1.1.1 Ray Transfer Matrices

Translation Matrix

A ray travels in a homogenous medium of refractive index n in a straight line (see Figure 1.1-3). Let us denote the input and output planes with the ray's coordinates, and then we try to relate the input and output coordinates with a matrix after the traveling of the distance d. Since $n_1 = n_2 = n$ and $\theta_1 = \theta_2, v_2 = n_2\theta_2 = n_1\theta_1 = v_1$. From the geometry, we also find $x_2 = x_1 + d\tan\theta_1 \approx x_1 + d\theta_1 = x_1 + dv_1/n$. Therefore, we can relate the output coordinates of the ray with its input coordinates as follows:

$$\begin{pmatrix} x_2 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 & d/n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ v_1 \end{pmatrix} = \mathcal{T}_{d/n} \begin{pmatrix} x_1 \\ v_1 \end{pmatrix}, \tag{1.1-2}$$

where

$$\mathcal{T}_{d/n} = \begin{pmatrix} 1 & d/n \\ 0 & 1 \end{pmatrix}, \tag{1.1-3a}$$

https://doi.org/10.1017/9781009053204.002 Published online by Cambridge University Press

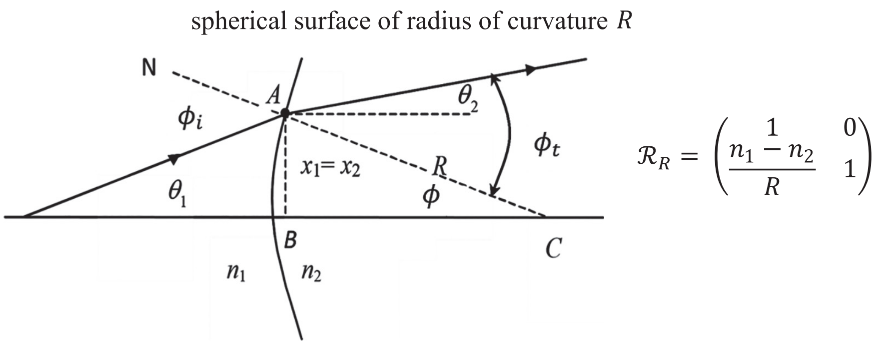


Figure 1.1-4 Ray trajectory during refraction at a spherical surface separating two regions of refractive indices n_1 and n_2

which is called the *translation matrix*. The matrix describes the translation of a ray for a distance d along the optical axis in a homogenous medium of n. The determinant of $\mathcal{T}_{d/n}$ is

$$|\mathcal{T}_{d/n}| = \begin{vmatrix} 1 & d/n \\ 0 & 1 \end{vmatrix} = 1,$$

and hence $\mathcal{T}_{d/n}$ is a positive unimodular matrix. For a translation of the ray in air, we have $n = 1$, and the translation can be represented simply by

$$\mathcal{T}_d = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}. \tag{1.1-3b}$$

Refraction Matrix

We consider a spherical surface separating two regions of refractive indices n_1 and n_2 as shown in Figure 1.1-4. The center of the spherical surface is at C, and its radius of curvature is R. The convention for the *radius of curvature* is as follows. The radius of curvature of the surface is taken to be positive (negative) if the center C of curvature lies to the right (left) of the surface. The ray strikes the surface at the point A and gets refracted into media n_2. Note that the input and output planes are the same. Hence, the height of the ray at A, before and after refraction, is the same, that is, $x_2 = x_1$. ϕ_i and ϕ_t are the angles of incidence and refraction, respectively, which are measured from the normal NAC to the curved surface. Applying Snell's law and using the paraxial approximation, we have

$$n_1\phi_i = n_2\phi_t. \tag{1.1-4}$$

Now, from geometry, we know that $\phi_i = \theta_1 + \phi$ and $\phi_t = \theta_2 + \phi$ (Figure 1.1-4). Hence, the left side of Eq. (1.1-4) becomes

$$n_1\phi_i = n_1(\theta_1 + \phi) = v_1 + n_1 x_1 / R, \tag{1.1-5}$$

https://doi.org/10.1017/9781009053204.002 Published online by Cambridge University Press

where we have used $\sin\phi = x_1 / R \approx \phi$. Now, the right side of Eq. (1.1-4) is

$$n_2\phi_t = n_2(\theta_2 + \phi) = v_2 + n_2 x_2 / R, \tag{1.1-6}$$

where $x_1 = x_2$ as the input and output planes are the same.

Finally, putting Eqs. (1.1-5) and (1.1-6) into Eq. (1.1-4), we have

$$v_1 + n_1 x_1 / R = v_2 + n_2 x_2 / R$$

or

$$v_2 = v_1 + (n_1 - n_2) x_1 / R, \quad \text{as} \quad x_1 = x_2. \tag{1.1-7}$$

We can formulate the above equation in terms of a matrix equation as

$$\begin{pmatrix} x_2 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \dfrac{n_1 - n_2}{R} & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ v_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -p & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ v_1 \end{pmatrix} = \mathcal{R}_R \begin{pmatrix} x_1 \\ v_1 \end{pmatrix}, \tag{1.1-8}$$

where

$$\mathcal{R}_R = \begin{pmatrix} 1 & 0 \\ -p & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \dfrac{n_1 - n_2}{R} & 1 \end{pmatrix}.$$

The determinant of $\mathcal{R}_R$ is

$$|\mathcal{R}_R| = \begin{vmatrix} 1 & 0 \\ -p & 1 \end{vmatrix} = 1.$$

The 2 × 2 ray transfer matrix $\mathcal{R}_R$ is a positive unimodular matrix and is called the *refraction matrix*. The matrix describes refraction for the spherical surface. The quantity p given by

$$p = \frac{n_2 - n_1}{R}$$

is called the *refracting power* of the spherical surface. When R is measured in meters, the unit of p is called *diopters*. If an incident ray is made to converge on (diverge from) by a surface, the power is positive (negative) in sign.

Thick- and Thin-Lens Matrices

A thick lens consists of two spherical surfaces as shown in Figure 1.1-5. We shall find the system matrix that relates the system's input coordinates (x_1, v_1) to system's output ray coordinates (x_2, v_2). Let us first relate (x_1, v_1) to (x_1', v_1') through the spherical surface defined by R_1. (x_1', v_1') are the output coordinates due to the surface R_1. According to Eq. (1.1-8), we have

$$\begin{pmatrix} x_1' \\ v_1' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \dfrac{n_1 - n_2}{R_1} & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ v_1 \end{pmatrix} = \mathcal{R}_{R_1} \begin{pmatrix} x_1 \\ v_1 \end{pmatrix}. \tag{1.1-9}$$

https://doi.org/10.1017/9781009053204.002 Published online by Cambridge University Press

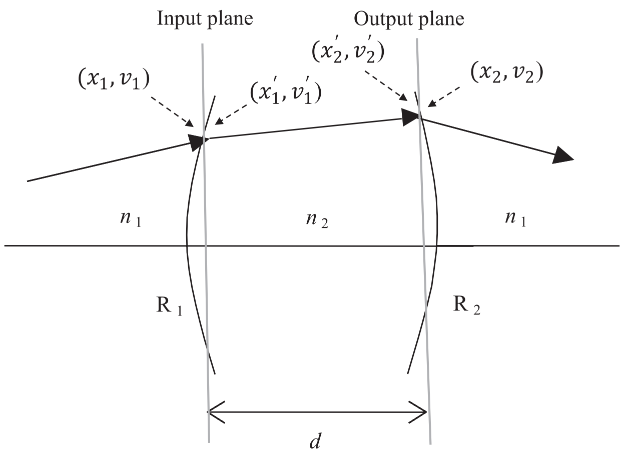


Figure 1.1-5 Thick lens

Now, (x_1', v_1') and (x_2', v_2') are related through a translation matrix as follows:

$$\begin{pmatrix} x_2' \\ v_2' \end{pmatrix} = \begin{pmatrix} 1 & d/n_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1' \\ v_1' \end{pmatrix} = \mathcal{T}_{d/n_2} \begin{pmatrix} x_1' \\ v_1' \end{pmatrix}, \tag{1.1-10}$$

where (x_2', v_2') are the output coordinates after translation, which are also the coordinates of the input coordinates for the surface R_2. Finally, we relate (x_2', v_2') to the system's output coordinates (x_2, v_2) through

$$\begin{pmatrix} x_2 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \dfrac{n_2 - n_1}{R_2} & 1 \end{pmatrix} \begin{pmatrix} x_2' \\ v_2' \end{pmatrix} = \mathcal{R}_{-R_2} \begin{pmatrix} x_2' \\ v_2' \end{pmatrix}. \tag{1.1-11}$$

If we now substitute Eq. (1.1-10) into Eq. (1.1-11), then we have

$$\begin{pmatrix} x_2 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \dfrac{n_2 - n_1}{R_2} & 1 \end{pmatrix} \begin{pmatrix} 1 & d/n_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1' \\ v_1' \end{pmatrix}.$$

Subsequently, substituting Eq. (1.1-9) into the above equation, we have the *system matrix equation* of the entire system as follows:

$$\begin{aligned} \begin{pmatrix} x_2 \\ v_2 \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ \dfrac{n_2 - n_1}{R_2} & 1 \end{pmatrix} \begin{pmatrix} 1 & \dfrac{d}{n_2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \dfrac{n_1 - n_2}{R_1} & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ v_1 \end{pmatrix} \\ &= \mathcal{R}_{-R_2} \mathcal{T}_{d/n_2} \mathcal{R}_{R_1} \begin{pmatrix} x_1 \\ v_1 \end{pmatrix} = \mathcal{S} \begin{pmatrix} x_1 \\ v_1 \end{pmatrix}. \end{aligned} \tag{1.1-12}$$

We have now finally related the system's input coordinates to output coordinates. Note that the system matrix, $\mathcal{S} = \mathcal{R}_{-R_2} \mathcal{T}_{d/n_2} \mathcal{R}_{R_1}$, is a product of three ray transfer

https://doi.org/10.1017/9781009053204.002 Published online by Cambridge University Press

matrices. In general, the system matrix is made up of a collection of ray transfer matrices to account for the effects of a ray passing through the optical system. As the ray goes from left to right in the positive direction of the z-axis, we obtain the system matrix by writing the ray transfer matrices from right to left. This is precisely the advantage of the matrix formalism in that any ray, during its propagation through the optical system, can be tracked by successive matrix multiplications. Let $\mathcal{A}$ and $\mathcal{B}$ be the 2×2 matrices as follows:

$$\mathcal{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{and } \mathcal{B} = \begin{pmatrix} e & f \\ g & h \end{pmatrix}.$$

Then the rule of matrix multiplication is

$$\mathcal{AB} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{pmatrix}.$$

Now, returning to the system matrix in Eq. (1.1-12), the determinant of the system matrix, $\mathcal{S} = \mathcal{R}_{-R_2}\mathcal{T}_{d/n_2}\mathcal{R}_{R_1}$, is

$$|\mathcal{S}| = \left|\mathcal{R}_{-R_2}\mathcal{T}_{d/n_2}\mathcal{R}_{R_1}\right| = \left|\mathcal{R}_{-R_2}\right| \times \left|\mathcal{T}_{d/n_2}\right| \times \left|\mathcal{R}_{R_1}\right| = 1.$$

Note that even the system matrix is also positive unimodular. The condition of a unit determinant is a necessary but not a sufficient condition on the system matrix.

We now derive a matrix of an idea thin lens of focal length f, called the *thin-lens matrix*, $\mathcal{L}_f$. For a thin lens in air, we let $d \to 0$ and $n_1 = 1$ in the configuration of Figure 1.1-5. We also write $n_2 = n$ for notational convenience. Then the system matrix in Eq. (1.1-12) becomes

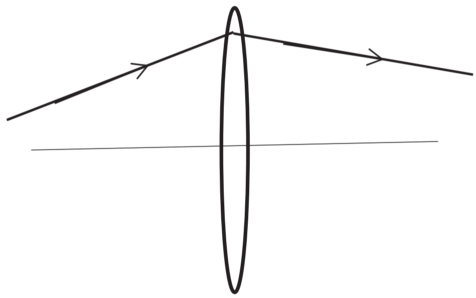
$$\begin{aligned} \mathcal{S} &= \begin{pmatrix} 1 & 0 \\ \frac{n-1}{R_2} & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \frac{1-n}{R_1} & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ \frac{n-1}{R_2} & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \frac{1-n}{R_1} & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} = \mathcal{L}_f, \end{aligned} \tag{1.1-13}$$

where f is the *focal length* of the thin lens and is given by

$$\frac{1}{f} = (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right).$$

For $f > 0$, we have a converging (convex) lens. On the other hand, we have a diverging (concave) lens when $f < 0$. Figure 1.1-6 summarizes the result for the ideal thin lens.

https://doi.org/10.1017/9781009053204.002 Published online by Cambridge University Press



$$\mathcal{L}_f = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix}$$

Convex : $f > 0$; Concave : $f < 0$

Figure 1.1-6 Ideal thin lens of focal length f and its associated ray transfer matrix

Note that the determinant of $\mathcal{L}_f$ is

$$|\mathcal{L}_f| = \begin{vmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{vmatrix} = 1.$$

1.1.2 Ray Tracing through a Thin Lens

As we have seen from Section 1.1.1, when a thin lens of focal length f is involved, then the matrix equation, from Eq. (1.1-13), is

$$\begin{pmatrix} x_2 \\ v_2 \end{pmatrix} = \mathcal{L}_f \begin{pmatrix} x_1 \\ v_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ v_1 \end{pmatrix}. \tag{1.1-14}$$

Input Rays Traveling Parallel to the Optical Axis

From Figure 1.1-7a, we recognize that $x_1 = x_2$ as the heights of the input and output rays are the same for the thin lens. Now, according to Eq. (1.1-14), $v_2 = -x_1/f + v_1$. For $v_1 = 0$, that is, the input rays are parallel to the optical axis, $v_2 = -x_1/f$. For positive $x_1, v_2 < 0$ as $f > 0$ for a converging lens. For negative $x_1, v_2 > 0$. All input rays parallel to the optical axis converge behind the lens to the back focal point (a distance of f away from the lens) of the lens as shown in Figure 1.1-7a. Note that for a thin lens, the front focal point is also a distance of f away from the lens.

Input Rays Traveling through the Center of the Lens

For input rays traveling through the center of the lens, their input ray coordinates are $(x_1, v_1) = (0, v_1)$. The output ray coordinates, according to Eq. (1.1.14), are $(x_2, v_2) = (0, v_1)$. Hence, we see as $v_2 = v_1$, all rays traveling through the center of the lens will pass undeviated as shown in Figure 1.1-7b.

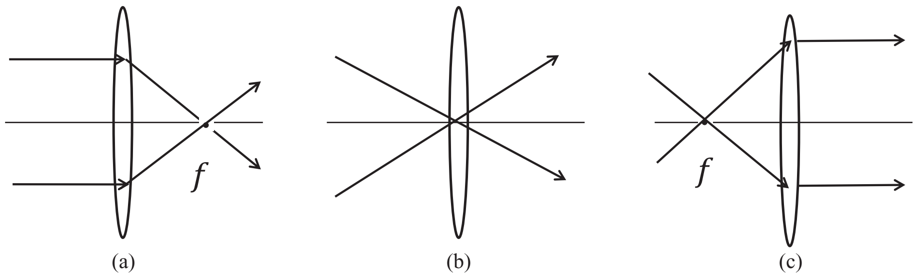


Figure 1.1-7 Ray tracing through a thin convex lens

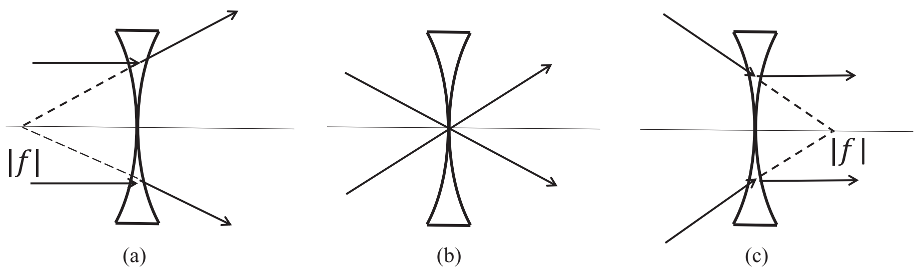


Figure 1.1-8 Ray tracing through a thin concave lens

Input Rays Passing through the Front Focal Point of the Lens

For this case, the input ray coordinates are $(x_1, v_1 = x_1 / f)$, and, according to Eq. (1.1-14), the output ray coordinates are $(x_2 = x_1, v_2 = 0)$, which means that all output rays will be parallel to the optical axis ($v_2 = 0$), as shown in Figure 1.1-7c.

Similarly, in the case of a diverging lens, we can draw conclusions as follows. The ray after refraction diverges away from the axis as if it were coming from a point on the axis a distance $|f|$ in front of the lens, as shown in Figure 1.1-8a. The ray traveling through the center of the lens will pass undeviated, as shown in Figure 1.1-8b. Finally, for an input ray appearing to travel toward the back focus point of a diverging lens, the output ray will be parallel to the optical axis, as shown in Figure 1.1-8c.

Example: Imaging by a Convex Lens

We consider a single-lens imaging as shown in Figure 1.1-9, where we assume the lens is immersed in air. We first construct a *ray diagram* for the imaging system. An object OO′ is located a distance d_0 in front of a thin lens of focal length f. We send two rays from a point O′ towards the lens. Ray 1 from O′ is incident parallel to the optical axis, and from Figure 1.1-7a, the input ray parallel to the optical axis converges behind the lens to the back focal point. A second ray, that is, ray 2 also from O′, is now drawn through the center of the lens without bending and that is the result from Figure 1.1-7b. The interception of the two rays on the other side of the lens forms an image point of O′. The image point of O′ is labeled at I′ in the diagram. The final image is real, inverted, and is called a *real image.*

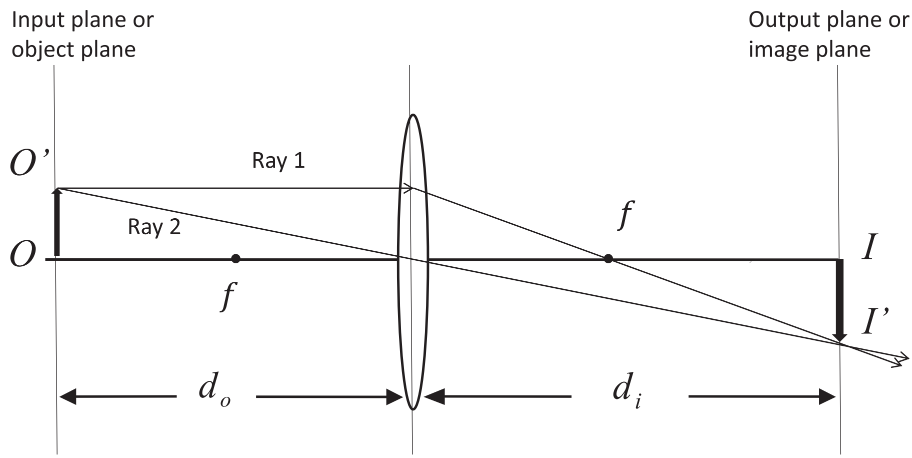


Figure 1.1-9 Ray diagram for single-lens imaging

Now we investigate the imaging properties of the single thin lens using the matrix formalism. The input plane and the output plane of the optical system are shown in Figure 1.1-9. We let (x_0, v_0) and (x_i, v_i) represent the coordinates of the ray at O′ and I′, respectively. We see there are three matrices involved in the problem. The overall system matrix equation becomes

$$\begin{pmatrix} x_i \\ v_i \end{pmatrix} = \mathcal{T}_{d_i}\mathcal{L}_f\mathcal{T}_{d_0}\begin{pmatrix} x_0 \\ v_0 \end{pmatrix} = \mathcal{S}\begin{pmatrix} x_0 \\ v_0 \end{pmatrix}. \tag{1.1-15}$$

The overall system matrix,

$$\mathcal{S} = \mathcal{T}_{d_i}\mathcal{L}_f\mathcal{T}_{d_0} = \begin{pmatrix} 1 & d_i \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -\dfrac{1}{f} & 1 \end{pmatrix}\begin{pmatrix} 1 & d_0 \\ 0 & 1 \end{pmatrix},$$

is expressed in terms of the product of three matrices written in order from right to left as the ray goes from left to right along the optical axis, as explained earlier. According to the rule of matrix multiplication, Eq. (1.1-15) can be simplified to

$$\begin{pmatrix} x_i \\ v_i \end{pmatrix} = \begin{pmatrix} 1 - d_i / f & d_0 + d_i - d_0 d_i / f \\ -\dfrac{1}{f} & 1 - d_0 / f \end{pmatrix}\begin{pmatrix} x_0 \\ v_0 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} x_0 \\ v_0 \end{pmatrix}. \tag{1.1-16}$$

To investigate the conditions for imaging, let us concentrate on the *ABCD* matrix of $\mathcal{S}$ in Eq. (1.1-16). In general, we see that $x_i = Ax_0 + Bv_0 = Ax_0$ if $B = 0$, which means that all rays passing through the input plane at the same object point x_0 will pass through the same image point x_i in the output plane. This is the condition of *imaging*. In addition, for $B = 0$, $A = x_i / x_0$ is defined as the *lateral magnification* of

https://doi.org/10.1017/9781009053204.002 Published online by Cambridge University Press

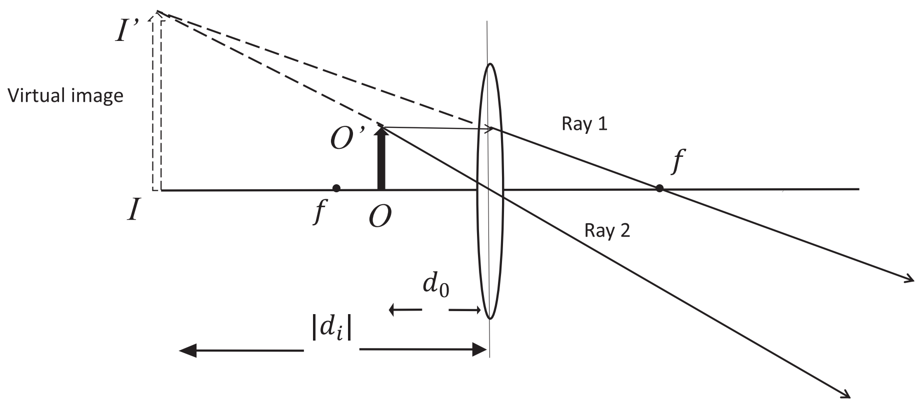


Figure 1.1-10 Imaging of a converging lens with object inside the focal length

the imaging system. Now in our case of thin-lens imaging, $B = 0$ in Eq. (1.1-16) leads to $d_0 + d_i - d_0 d_i / f = 0$, which gives the well-known *thin-lens formula*:

$$\frac{1}{d_0} + \frac{1}{d_i} = \frac{1}{f}. \tag{1.1-17}$$

The sign convention for Eq. (1.1-17) is that the object distance d_0 is positive (negative) if the object is to the left (right) of the lens. If the image distance d_i is positive (negative), the image is to the right (left) of the lens and it is real (virtual). In Figure 1.1-9, we have $d_0 > 0$, $d_i > 0$, and the image is therefore real, which means physically that light rays actually converge to the formed image I′. Hence, for imaging, Eq. (1.1-16) becomes

$$\begin{pmatrix} x_i \\ v_i \end{pmatrix} = \begin{pmatrix} 1 - d_i / f & 0 \\ -\dfrac{1}{f} & 1 - d_0 / f \end{pmatrix} \begin{pmatrix} x_0 \\ v_0 \end{pmatrix}, \tag{1.1-18}$$

which relates the input ray and output ray coordinates in the imaging system. Using Eq. (1.1-17), the *lateral magnification* M of the imaging system is

$$M = A = \frac{x_i}{x_0} = 1 - \frac{d_i}{f} = -\frac{d_i}{d_0}. \tag{1.1-19}$$

The sign convention is that if $M > 0$, the image is erect, and if $M < 0$, the image is inverted. As shown in Figure 1.1-9, we have an inverted image as both d_i and d_0 are positive.

If the object lies within the focal length, as shown in Figure 1.1-10, we follow the rules as given in Figure 1.1-8 to construct a ray diagram. However, now rays 1 and 2, after refraction by the lens, are divergent and do not intercept on the right side of the lens. They seem to come from a point I′. In this case, since $d_0 > 0$ and $d_i < 0$, $M > 0$. The final image is virtual, erect, and is called a *virtual image*.

https://doi.org/10.1017/9781009053204.002 Published online by Cambridge University Press

1.2 Resolution, Depth of Focus, and Depth of Field

1.2.1 Circular Aperture

The *numerical aperture* (NA) of a lens is defined for an object or image located infinitely far away. Figure 1.2-1 shows an object located at infinity, which sends rays parallel to the lens with a circular aperture. The angle θ_{im} used to define the NA on the image side is

$$\text{NA}_i = n_i \sin(\theta_{\text{im}} / 2), \tag{1.2-1}$$

where n_i is refractive index in the image space. Note that the *aperture stop* limits the angle of rays passing through the lens, which affects the achievable NA. Let us now find the lateral resolution, Δr.

Since we treat light as particles in geometrical optics, each particle then can be characterized by its momentum, p_0. According to the *uncertainty principle* in quantum mechanics, we relate the minimum uncertainty in a position of quantum, Δr, to the uncertainty of its momentum, Δp_r, according to the relationship

$$\Delta r \Delta p_r \geq h, \tag{1.2-2}$$

where h is *Planck's constant*, and Δp_r is the momentum spread in the r-component of the photons. The momentum of the FB ray (chief ray, a ray passes through the center of the lens) along the r-axis is zero, while the momentum of the FA ray (marginal ray, a ray passes through the edge of the lens) along the r-axis is $\Delta p_{AB} = p_0 \sin(\theta_{\text{im}} / 2)$, where $p_0 = h / \lambda_0$ with λ_0 being the wavelength in the medium, that is, in the image space. Hence, Δp_r is $\Delta p_r = 2\Delta p_{AB} = 2 p_0 \sin(\theta_{\text{im}} / 2)$ to accommodate a maximum variation (or spread) of the momentum direction by an angle θ_{im}. By substituting this into Eq. (1.2-2), we have *lateral resolution*

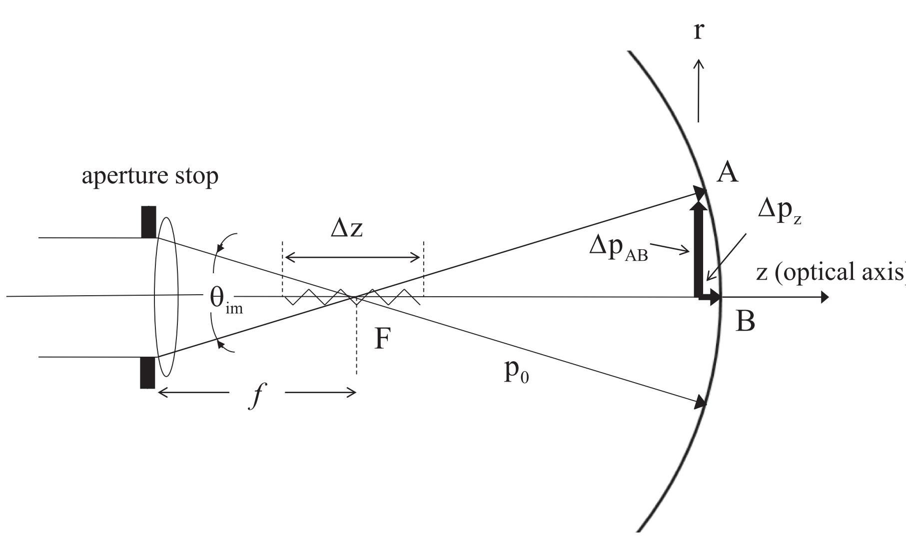


Figure 1.2-1 Uncertainty principle used in finding lateral resolution and depth resolution

https://doi.org/10.1017/9781009053204.002 Published online by Cambridge University Press

$$\Delta r \geq \frac{h}{\Delta p_r} = \frac{h}{2p_0 \sin(\theta_{\text{im}}/2)} = \frac{\lambda_0}{2\sin(\theta_{\text{im}}/2)}. \tag{1.2-3}$$

Note that this equation reveals that in order to achieve high resolution, we can increase $\theta_{\text{im}}/2$. For example, when $\theta_{\text{im}}/2 = 90°$, Δr is half the wavelength, that is, $\lambda_0/2$, which is the theoretical maximum lateral resolution.

Since the wavelength in the image space, λ_0, is equal to λ_v/n_i, where λ_v is the wavelength in air or in vacuum, Eq. (1.2-3) becomes, using Eq. (1.2-1),

$$\Delta r \geq \frac{\lambda_v}{2n_i \sin(\theta_{\text{im}}/2)} = \frac{\lambda_v}{2\text{NA}_i}. \tag{1.2-4}$$

Similarly, we can calculate the *depth of focus*, Δz. Depth of focus, also called *longitudinal resolution* in the image space, is the axial distance over which the image can be moved without loss of sharpness in the object. To find the depth of focus, we use

$$\Delta z \Delta p_z \geq h, \tag{1.2-5}$$

where Δp_z is the momentum difference between rays FB and FA along the z-direction, as shown in Figure 1.2-1, which is given by

$$\Delta p_z = p_0 - p_0 \cos(\theta_{\text{im}}/2). \tag{1.2-6}$$

By substituting this expression into Eq. (1.2-5), we have

$$\Delta z \geq \frac{h}{\Delta p_z} = \frac{h}{p_0\left[1-\cos(\theta_{\text{im}}/2)\right]} = \frac{\lambda_0}{\left[1-\cos(\theta_{\text{im}}/2)\right]}. \tag{1.2-7}$$

This equation reveals that in order to achieve small depth of focus, we can increase $\theta_{\text{im}}/2$. For example, when $\theta_{\text{im}}/2 = 90°$, Δz is one wavelength, that is, λ_0, which is the theoretical maximum longitudinal resolution. Equation (1.2-7) can be written in terms of the numerical aperture and is given by

$$\Delta z \geq \frac{\lambda_0}{1-\sqrt{1-\sin^2(\theta_{\text{im}}/2)}} = \frac{\lambda_v}{n_i - \sqrt{n_i^2 - \text{NA}_i^{\,2}}}. \tag{1.2-8}$$

For small angles, that is, $\theta_{\text{im}} << 1$, one can use the approximation $\sqrt{1-\sin^2\beta} \approx 1-(\sin^2\beta)/2$ to get

$$\Delta z \geq \frac{2n_i\lambda_v}{\text{NA}_i^{\,2}}. \tag{1.2-9}$$

Taking the equality in Eq. (1.2-9) and combining with Eq. (1.2-4), we have

$$\frac{(\Delta r)^2}{\Delta z} \approx \frac{\lambda_0}{8} \tag{1.2-10}$$

https://doi.org/10.1017/9781009053204.002 Published online by Cambridge University Press

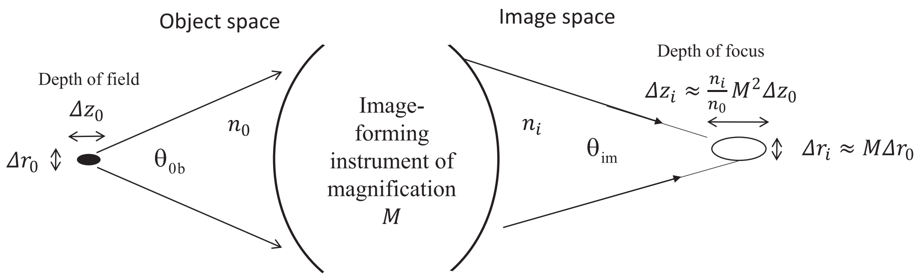


Figure 1.2-2 Image-forming instrument illustrating resolutions in the object space and image space

under the small-NA approximation. This equality is derived from uncertainty relationship, which says that during imaging the higher the lateral resolution is, the shorter the depth of focus. For example, by increasing the lateral resolution by a factor of 2 will result in the depth of focus decreased by a factor of 4.

Figure 1.2-2 shows an image-forming instrument with θ_{ob} and θ_{im} denoting the ray of maximum divergent angle and maximum convergent angle from the object side and the image side, respectively.

If the resolutions in the object space are given by $\Delta r_0 \approx \lambda_v / 2\mathrm{NA}_0$ and $\Delta z_0 \approx 2n_0\lambda_v / \mathrm{NA}_0{}^2$, where $\mathrm{NA}_0 = n_0 \sin(\theta_{ob}/2)$ and n_0 is the refractive index in the object space, and if the lateral magnification of the instrument is M, the resolution on the image space is then $\Delta r_i \approx M\Delta r_0$. Let us establish the relationship between the *depth of field* Δz_0 and the *depth of focus* Δz_i, where $\Delta z_i \approx 2n_i\lambda_v / \mathrm{NA}_i^2$ with $\mathrm{NA}_i = n_i \sin(\theta_{im}/2)$. Since $\Delta r_i \approx \lambda_v / 2\mathrm{NA}_i = M\Delta r_0 = M\lambda_v / 2\mathrm{NA}_0$, we have $\mathrm{NA}_0 = M \times \mathrm{NA}_i$. Hence,

$$\begin{aligned}\Delta z_i \approx 2n_i\lambda_v / \mathrm{NA}_i^2 &= 2n_i\lambda_v / \left(\mathrm{NA}_0 / M\right)^2 \\ &= M^2 (2n_i\lambda_v / \mathrm{NA}_0^2) \\ &= \frac{n_i}{n_0} M^2 \Delta z_0. \end{aligned} \tag{1.2-11}$$

This result indicates that the longitudinal resolutions in the object space and image space are related by a factor of M^2. Take a 40×, $\mathrm{NA}_0 \approx 0.6$ microscope objective as an example, we have $\Delta r_0 \approx \lambda_v / 2\mathrm{NA}_0 \approx 0.5\mu\mathrm{m}$ for red light with wavelength of 632 nm and the depth of field, $\Delta z_0 \approx 2n_0\lambda_v / \mathrm{NA}_0^2 \approx 3.5\mu\mathrm{m}$ for $n_0 = 1$ in air. In the image space, the lateral resolution is $\Delta r_i \approx M\Delta r_0 = 40 \times 0.5\mu\mathrm{m} = 20\mu\mathrm{m}$ and the depth of focus is $\Delta z_i \approx M^2 \Delta z_0 = 40^2 \times 3.5\mu\mathrm{m} = 0.56\mathrm{mm}$ for $n_i = 1$.

1.2.2 Annular Aperture

Three-dimensional imaging in microscopy aims to develop techniques that can provide high lateral resolution, and at the same time maintain a large depth of focus in order to observe a thick specimen. However, calculations have shown that the depth of focus may be increased by reducing the numerical aperture of the lens, but this is achieved at the expense of a decrease in lateral resolution. In what follows, we

https://doi.org/10.1017/9781009053204.002 Published online by Cambridge University Press

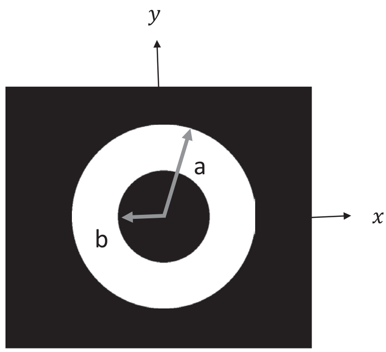


Figure 1.2-3 Annular aperture

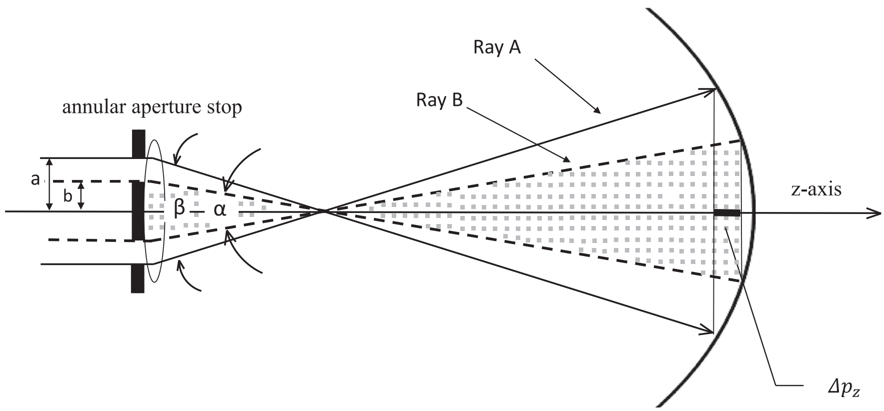


Figure 1.2-4 Uncertainty principle used in finding depth of focus for a lens with annular aperture stop

consider an annular aperture which has the property of increasing the depth of focus and at the same time maintaining the lateral resolution. An *annular aperture* is defined as a clear circular aperture with a central obstruction, as shown in Figure 1.2-3. If the aperture is an annulus with outer radius a and inner radius b, we define a *central obscuration ratio* $\varepsilon = b / a$. For $\varepsilon = 0$, we have a clear circular aperture.

Since $\beta = \theta_{\mathrm{im}}$ as shown in Figure 1.2-1, the possible spread in the r-component of momentum Δp_r is the same in Figure 1.2-4 as it is in Figure 1.2-1. Hence, the lateral resolution remains the same as that in the case of the clear aperture, that is, $\varepsilon = 0$, given by Eq. (1.2-4). However, the spread or uncertainty in the momentum in the z-direction is different due to the central obscuration of the annulus. Δp_r in this case is the momentum difference between the ray passing the upper part of the annulus (Ray A) and the ray passing the lower part of the annulus (Ray B) and is given by

$$\Delta p_z = \Delta p_{\text{Ray A}} - \Delta p_{\text{Ray B}}, \tag{1.2-12}$$

where $\Delta p_{\text{Ray A}}$ and $\Delta p_{\text{Ray B}}$ are given by Eq. (1.2-6) with β and α substituted into the argument of cosine, respectively. Hence Eq. (1.2-12) becomes

https://doi.org/10.1017/9781009053204.002 Published online by Cambridge University Press

$$\Delta p_z = p_0\left[\cos\left(\frac{\alpha}{2}\right) - \cos\left(\frac{\beta}{2}\right)\right],$$

and the depth of focus is

$$\Delta z_{\text{ann}} = \frac{h}{\Delta p_z} = \frac{\lambda_0}{\cos\left(\frac{\alpha}{2}\right) - \cos\left(\frac{\beta}{2}\right)}, \tag{1.2-13}$$

which can be shown to have the form

$$\Delta z_{\text{ann}} = \frac{\lambda_v}{\sqrt{\frac{1-\text{NA}^2}{1+\text{NA}^2\left(\varepsilon^2-1\right)}} - \sqrt{1-\text{NA}^2}}, \tag{1.2-14}$$

where $\text{NA} = \sin(\beta/2)$, that is, assuming the lens is being immersed in air. For small NAs, it can be shown that Eq. (1.2-14) is given by

$$\Delta z_{\text{ann}} = \frac{2\lambda_v}{\text{NA}^2\left(1-\varepsilon^2\right)} = \frac{\Delta z}{\left(1-\varepsilon^2\right)}. \tag{1.2-15}$$

The above equation is consistent with Eq. (1.2-9) for a clear aperture, that is, for the case $\varepsilon = 0$. With 95% obstruction, that is, $\varepsilon = 0.95$, we can increase the depth of focus of a clear lens by more than a factor of 10.

1.3 Illustrative Examples

1.3.1 Three-Dimensional Imaging through a Single-Lens Example

Figure 1.3-1a shows the imaging of two objects in front of the lens, where both of the object lie beyond the focal length of the lens. We note that magnification is different, depending on the object distance to the lens. Let us consider longitudinal magnification, M_z, in addition to lateral magnification, M, considered earlier. Longitudinal magnification M_z is the ratio of an image displacement along the axial direction, δd_i, to the corresponding object displacement, δd_0, that is, $M_z = \delta d_i / \delta d_0$.

Using Eq. (1.1-17) and treating d_i and d_0 as variables, we take the derivative of d_i with respect to d_0 to obtain

$$M_z = \delta d_i / \delta d_0 = -M^2. \tag{1.3-1}$$

This equation is consistent with Eq. (1.2-11) and states that the longitudinal magnification is equal to the square of the lateral magnification. The minus sign in front of the equation signifies that a decrease in the distance of the object from the lens, $|d_0|$, will result in an increase in the image distance, $|d_i|$, and vice versa. The situation of a magnified volume is shown in Figure 1.3-1b, where a cube volume (abcd plus the dimension into the paper) is imaged into a truncated pyramid with a–b imaged into a′–b′ and c–d imaged into c′–d′.

https://doi.org/10.1017/9781009053204.002 Published online by Cambridge University Press

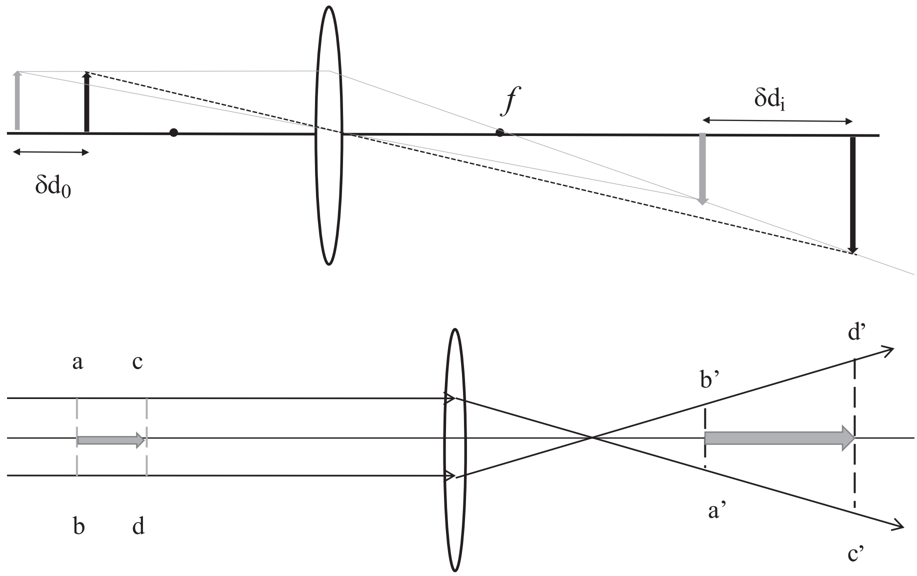


Figure 1.3-1 (a) Illustration of different magnifications and (b) volumetric imaging

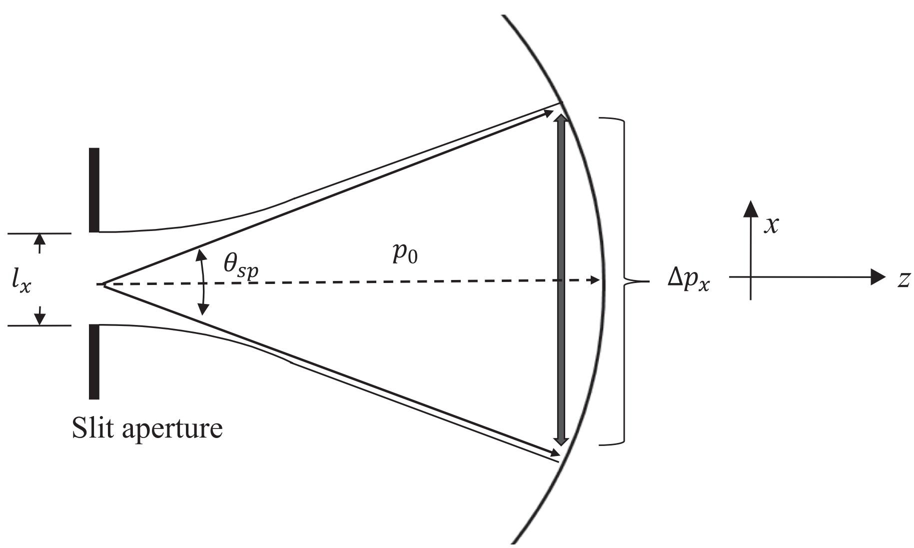


Figure 1.3-2 Spreading from a slit

1.3.2 Angle of Spread from a Slit Example

Consider light emanating from a slit aperture of width l_x, as shown in Figure 1.3-2.

We relate the minimum uncertainty in position Δx of a quantum to the uncertainty in its momentum Δp_x according to

https://doi.org/10.1017/9781009053204.002 Published online by Cambridge University Press

$$\Delta x \Delta p_x \geq h.$$

Because the quantum of light can emerge from any point in the aperture, we have $\Delta x = l_x$. Therefore,

$$\Delta p_x \sim \frac{h}{l_x}.$$

We define the angle of spread as

$$\theta_{\mathrm{sp}} \sim \frac{\Delta p_x}{p_0} = \frac{\dfrac{h}{l_x}}{\dfrac{h}{\lambda_0}} = \frac{\lambda_0}{l_x}, \tag{1.3-2}$$

which is basically the result in Eq. (1.2-3) for small angles. Note that the spread angle is inversely proportional to the width of the aperture.

Problems

1.1 Find the transfer matrix for Snell's law.

1.2 An object 4 cm tall is 10 cm in front of a convex lens of focal length 20 cm. Using the matrix formalism, find the location and magnification of the image. Draw a ray diagram from the object to the image.

1.3 A slide 5 cm tall is located 110 cm from a screen (see Figure P.1.3). What is the focal length of the positive lens, which will project a real image measuring 50 cm on the screen? Use the matrix formalism to solve the problem.

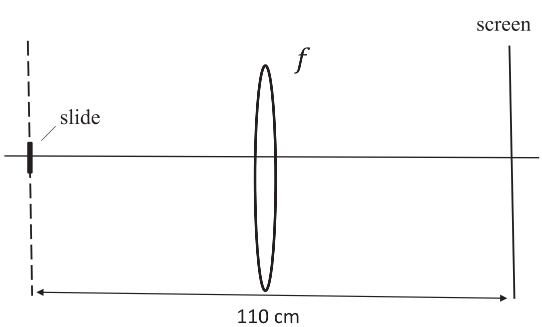


Figure P.1.3 Single-lens system

1.4 If an object 3 cm tall is located on the optical axial 24 cm to the left of the convex lens as shown in Figure P.1.4, find the position and size of the image using the matrix formalism. Also, draw the ray diagram from the object to the image.

https://doi.org/10.1017/9781009053204.002 Published online by Cambridge University Press

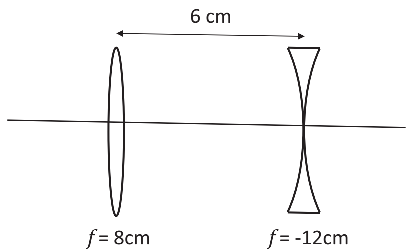


Figure P.1.4 Two-lens system

1.5 In 3-D imaging through a single lens example, it is stated that the longitudinal magnification is equal to the square of the lateral magnification, that is, $M_z = \delta d_i / \delta d_0 = -M^2$. Verify this statement.

1.6 Fill in the blanks below for several microscope objective lenses (assuming the lenses are immersed in air with wavelength of operation using red light of $\lambda_v = 632\,\text{nm}$).

Magnification, NA	Resolution in image space	Depth of field (μm)	Depth of focus (mm)
10×, 0.1			
20×, 0.4			
40×, 0.6			
60×, 0.8			
100×, 0.95			

1.7 Starting from Eq. (1.2-13) and with reference to Figure 1.2-4, show that the depth of focus of an annular aperture with a thin lens is given by

$$\Delta z_{\text{ann}} = \frac{\lambda_v}{\sqrt{\dfrac{1-\text{NA}^2}{1+\text{NA}^2\left(\varepsilon^2-1\right)}} - \sqrt{1-\text{NA}^2}},$$

where $\text{NA} = \sin(\beta/2)$ and for small NAs, it is approximately given by

$$\Delta z_{\text{ann}} = \frac{2\lambda_v}{\text{NA}^2\left(1-\varepsilon^2\right)} = \frac{\Delta z}{\left(1-\varepsilon^2\right)},$$

where Δz is the depth of focus for a clear lens.

Bibliography

1. Banerjee, P. P. and T.-C. Poon (1991). *Principles of Applied Optics*. Irwin, Illinois.
2. Gerard, A. and J. M. Burch (1975). *Introduction to Matrix Methods in Optics*. Wiley, New York.
3. Hecht, E. (2002). *Optics,* 4th ed., Addison Wesley, California.
4. Korpel, A. (1970). *United State Patent (#3,614,310)* Electrooptical Apparatus Employing a Hollow Beam for Translating an Image of an Object.

5. Poon, T.-C. (2007). *Optical Scanning Holography with MATLAB®*, Springer, New York.
6. Poon, T.-C. and M. Motamedi (1987). "Optical/digital incoherent image processing for extended depth of field," *Applied Optics* 26, pp. 4612–4615.
7. Poon, T.-C. and T. Kim (2018). *Engineering Optics with MATLAB®*, 2nd ed., World Scientific, New Jersey.