3 Geometrical optics

Jim Schwiegerling

Contents
3.1 Introduction 27
  3.1.1 What is geometrical optics? 27
  3.1.2 Sign convention 28
  3.1.3 Wavelength, speed of light, and refractive index 28
3.2 Waves, rays
  3.2.1 Vergence 28
  3.2.2 Rays and wavefronts 29
3.3 Laws of refraction and reflection
  3.3.1 Reflection from a planar surface 29
  3.3.2 Snell’s law at an interface 29
  3.3.3 Total internal reflection 30
  3.3.4 Prisms 30
3.4 Refraction and reflection from a spherical surface
  3.4.1 Refraction from a spherical surface 30
  3.4.2 Reflection from a spherical mirror 31
3.5 Gaussian imaging equation
  3.5.1 Thick lenses and Gaussian imaging 32
  3.5.2 Cardinal points 33
  3.5.3 Aperture stop and pupils 34
  3.5.4 Chief and marginal rays 34
3.6 Cylindrical and toric surfaces
  3.6.1 Power and axis of a cylindrical lens 34
  3.6.2 Toric and spherocylindrical surfaces 34
3.7 Visual instruments
  3.7.1 Simple magnifier and magnifying power 37
  3.7.2 Microscopes 37
  3.7.3 Telescopes and angular magnification 38
3.8 Summary 38

3.1 INTRODUCTION

3.1.1 WHAT IS GEOMETRICAL OPTICS?

Geometrical optics is a branch of optics that typically examines the transfer of light from a source to some destination via an optical system. These systems can be composed of multiple lenses, mirrors, prisms, and windows. Consequently, geometrical optics concerns itself with the refraction and reflection of light at interfaces and its propagation through various media. Light exhibits wave phenomena such as interference and diffraction (treated in Chapter 5). However, the description of refraction, reflection, and propagation of waves is mathematically intensive. Geometrical optics makes a series of approximations that vastly simplifies the analysis. These approximations are valid for incoherent light. The main approximation of geometrical optics is to replace the wavefront with a series of rays and analyze how the rays move through space and interact with objects. Rays represent the local normals to wavefronts and illustrate the direction of propagation of the wave at a given point. In homogeneous materials, rays will travel in a straight line until they reach a boundary. At the boundary, the rays will refract and reflect and then continue their straight-line propagation in the ensuing material. This simplification enables optical systems to be designed and analyzed in a straightforward manner leading to systems for illumination or imaging having the desired properties of the designer. This chapter examines the basic description of the properties of rays and their interaction with optical elements.
such as lenses and prisms. In addition, definitions for the various properties of optical systems are provided.

### 3.1.2 SIGN CONVENTION

Prior to developing the foundations of geometrical optics, a coordinate system and a consistent sign convention need to be defined. The axis for an optical system will, in general, be taken as the z-axis. Light will typically travel from left (more negative values of z) to right (more positive values of z) with regard to this axis, unless the light is reflected from a surface. The vertical axis will be denoted as the positive y-axis. The coordinate system will be considered right-handed, meaning that the positive x-axis will be into the plane of the paper in the ensuing figures. Additional quantities such as distances, radii of curvature, and angles and their respective signs are defined with respect to this coordinate system. Distances are measured from a reference point and are signed to be consistent with the coordinate system. For example, if one optical surface is located at \( z = 0 \) and a second optical surface is located at \( z = +10 \), then the distance from the first surface to the second surface is positive. Similarly, distances measured in the \(-z\) direction are negative. The same convention for distances holds for the \( x \) and \( y \) directions. An object with its base on the \( z \)-axis and its top at \( y = 5 \) would have a positive height, while its image may be upside down with its base still on the \( z \)-axis, but its “top” now at \( y = -3 \). In this latter case, the image height is negative. Angles are measured with regard to a reference line such as the \( x \)-axis or a local normal to an optical surface. Counterclockwise angles are considered positive and clockwise angles are negative. For example, a ray that starts on the \( z \)-axis (\( y = 0 \)) at an object may propagate in the positive \( z \) direction to an optical surface and strike that surface at a height of \( y = 2 \). This ray would have a positive angle with respect to the \( z \)-axis. Finally, the radius of curvature of a spherical surface has a positive value if its center of curvature is to the right of the surface. Similarly, surfaces with negative radii have their center of curvatures to the left of the surface. Arrowheads are used in the figures to help illustrate the sign of the quantities depicted. Figure 3.1 summarizes the various sign and coordinate conventions described earlier.

### 3.1.3 WAVELENGTH, SPEED OF LIGHT, AND REFRACTIVE INDEX

The electromagnetic spectrum represents a continuum of radiation that propagates as transverse waves. The distinguishing feature between various elements of the spectrum is the wavelength, \( \lambda \), or the distance between the peaks of the propagating waves. The electromagnetic spectrum describes everything from gamma rays with a wavelength comparable to the size of atomic nuclei to radio waves with wavelengths comparable to the size of a skyscraper. All of these waves have further in common their speed in vacuum, \( c \equiv 3 \times 10^8 \text{ m/s} \). In visual optics, typically only a small subset of the electromagnetic spectrum is considered where the wavelengths interact with the components of the human visual system to form images and/or enable diagnostic and therapeutic treatment of the eye. In the ensuing discussion, the spectrum will be restricted to the ultraviolet, visible, and near-infrared wavelengths. This range corresponds to wavelengths of roughly 0.1 \( \leq \lambda \leq 1.0 \mu \text{m} \). Finally, when these waves enter a material such as water or glass, they slow down. For example, the speed of visible light in glass is typically reduced to about \( 2 \times 10^8 \text{ m/s} \). This speed change in various materials can be captured by defining the material’s index of refraction, \( n \), as

\[
    n = \frac{\text{Speed in vacuum}}{\text{Speed in material}}. \tag{3.1}
\]

For the given glass example, the refractive index would be \( n = 1.5 \). Note, when the electromagnetic waves travel in air, their speed is essentially the same as when they are in vacuum. Consequently, the refractive index of air is typically taken as \( n = 1.0 \).

### 3.2 WAVES, RAYS

#### 3.2.1 VERGENCE

If a stone is dropped into a calm pool of water, circular ripples will propagate outward from the point where the stone entered the water. In a similar fashion, a point source of light will radiate spherical wavefronts in all directions. These wavefronts will remain perfectly spherical until they interact with some object in the environment. The properties of this spherical wavefront at any point in space can be defined by its vergence. Vergence is defined as

\[
    \text{Vergence} = \frac{n}{z}, \tag{3.2}
\]

where
- \( n \) is the refractive index of the material in which the spherical wave is propagating
- \( z \) is the distance between the measurement point in space and the location of the point source
In visual optics, it is common to measure this distance in units of meters. The units of vergence, therefore, are reciprocal meters or diopters (D). As the distance between the point source and the measurement point increases, the radius of the spherical wavefront will continue to become larger. At extreme distances, \( z \) will approach infinity and the vergence will become zero. The preceding description of vergence considers the case of diverging spherical wavefronts. By convention, diverging spherical waves have a negative vergence, which requires that \( z \) be negative. Converging spherical waves, where the spherical wavefronts collapse to a single point, can also be described with vergence. In this case, the distance \( z \) is taken as a positive value. To summarize, a negative vergence describes a diverging spherical wave, a positive vergence describes a converging spherical wave, and a flat or plane wave is equivalent to a vergence of zero.

### 3.2.2 Rays and wavefronts

Rays are a mathematical simplification that describes the local direction a wavefront is propagating. Consequently, rays are always perpendicular to the wavefront. For the spherical waves described in the preceding section, a wave with positive vergence is converging. The rays associated with this wave all point to the single point to which the spherical wave collapses. A wave with negative vergence is diverging. The rays associated with this wave all appear to be emanating from the same point. Finally, the rays associated with a plane wave are all parallel. Figure 3.2 illustrates the various spherical wavefronts, their vergences, and associated rays.

### 3.3 Laws of refraction and reflection

#### 3.3.1 Reflection from a planar surface

The law of reflection governs how a ray reflects from a surface. In much the same fashion as an elastic collision, a ray incident on a planar surface will reflect from the surface at the same angle. Figure 3.3a shows the reflective surface and its normal. A ray incident at an angle \( i \) with respect to this normal will reflect from the surface at an angle \( i' \). The law of reflection, in conjunction with the sign convention, states that \( i = -i' \). As drawn in the figure, the angle \( i \) is positive and the angle \( i' \) is negative. A further consequence of the law of reflection is that the incident and emerging rays and the surface normal must all lie in the same plane.

#### 3.3.2 Snell’s law at an interface

When an incident ray reaches a boundary between two transparent media, the transmitted ray is bent or refracted. Figure 3.3b illustrates the refracting of a ray at the boundary between two regions of indices \( n \) and \( n' \), respectively. Snell’s law, as shown in the following equation, describes the degree of refraction:

\[
 n \sin i = n' \sin i',
\]

where \( i \) and \( i' \) are the angles the incident and emerging rays form with respect to the surface normal. The typical convention of using unprimed variables for values before the interface and primed variables for values after the interface will be used in the ensuing discussion. A ray in a medium of refractive index \( n \) that is incident on a surface at an angle \( i \) will refract to leave the surface at angle \( i' \). The newly refracted ray will then continue in a straight line until it intercepts a new boundary. As with the law of reflection, the incident and emerging rays and the surface normal must all lie in the same plane.

As seen in the figure, the incident and refracted angles are measured counterclockwise from the surface normal and are consequently both positive. Finally, it is convenient to combine Snell’s law and the law of reflection into a single mathematical expression. This can be easily accomplished in the reflective...
Geometrical optics

3.3.3 TOTAL INTERNAL REFLECTION

Snell’s law can also be rewritten in the form
\[
sin i = \frac{n'}{n} \sin i'.
\]
(3.5)

The maximum value of the sine function is one. In cases where \( n' > n \), meaning the incident ray is in a higher refractive index material than the refracted ray, the incident angle \( i \) cannot exceed to so-called critical angle, \( i_c \). Otherwise, the maximum value of the sine function is exceeded. The critical angle occurs when \( \sin i' = 1 \) or
\[
i_c = \sin^{-1} \left( \frac{n'}{n} \right).
\]
(3.6)

For angles of incidence less than the critical angle, the incident ray will refract as governed by Snell’s law. If the incident ray meets or exceeds the critical angle, the ray will reflect as governed by the law of reflection. This reflective process is known as total internal reflection and is often exploited to provide reflection without the need to mirror coat a surface.

3.3.4 PRISMS

Prisms are optical elements, often triangular in shape, which are used to invert, revert, or deviate the direction of light beams within an optical system. Prisms can combine the effects of total internal reflection, conventional reflection (with a mirrored surface), and refraction to cause these changes in beam direction. A common prism used in visual optics is a thin triangular prism used to treat strabismus or misalignment between the two eyes. The prism is placed in front of the misaligned eye so that the line of sight through that eye is deviated such that it becomes parallel with the line of sight of the fellow eye. In this manner, the two eyes are looking at the same point in the scene and the images from each eye can be fused. Figure 3.4 shows a typical thin triangular prism. The magnitude of the deviation is dictated by the apex angle of the triangular prism and is usually given in units of prism diopters (\( \Delta \)). The definition of one prism diopter is the prismatic deviation of a beam of light by 1 cm at a distance of 1 m.

The orientation of a prism is defined by its base (the wide end of the prism). Base up and base down orientations cause vertical deviations, while base out and base in cause horizontal deviations. By rotating the prism about the line of sight, deviations in different directions can be achieved. Base out orients the prism base oriented toward the temple, while base in orients the prism nasally. Figure 3.5 shows the orientation of the prism with respect to the misaligned eye for different types of strabismus.

3.4 REFRACTION AND REFLECTION FROM A SPHERICAL SURFACE

3.4.1 REFRACTION FROM A SPHERICAL SURFACE

Prisms and other optical elements with flat surfaces only deviate the direction of a beam passing through them, but do not change the beam’s vergence. To change vergence, a surface must have curvature. Snell’s law, as defined in Equation 3.3, holds for spherical surfaces as well. The angle of incidence and refraction are measured with respect to the surface normal. However, since the surface is curved, the orientation of the normal now depends upon where the ray strikes the surface. The spherical refracting surface shown in Figure 3.6 separates two optical spaces with refractive indices of \( n \) and \( n' \), respectively. The optical axis is the line passing through the center of curvature of the spherical surface. The intersection of the optical axis with the spherical surface is called the surface vertex. This surface has a radius of curvature \( R \), which is the distance from a point on the surface to the center of curvature. This radius defines the shape of the spherical surface. From the previously defined sign convention, \( R \) is a positive value in this case since the center of curvature lies...
3.4 Refraction and reflection from a spherical surface

### Fundamentals

#### 3.4.1 Refraction from a Spherical Surface

The spherical surface can be described by its radius \( R \), which is measured in units of meters. The power \( \phi \) of a spherical surface is defined as

\[
\phi = \frac{1}{n} \left( \frac{1}{n} - 1 \right) \frac{1}{R}.
\]

where \( n \) is the index of refraction of the surface and \( R \) is the radius of curvature.

#### 3.4.2 Reflection from a Spherical Mirror

The spherical mirror can be described by its radius \( R \) and index of refraction \( n \). The power \( \phi \) of a spherical mirror is defined as

\[
\phi = \frac{n}{n'} - 1.
\]

where \( n' \) is the index of refraction of the medium on the opposite side of the surface.

#### 3.4.3 Focal Lengths

The focal length \( f_R \) of a spherical surface is given by

\[
f_R = -\frac{1}{R}.
\]

Similarly, the distance from the surface to the focal point \( F \) is called the focal length \( f_F \), and is defined as

\[
f_F = -\frac{1}{n} \frac{1}{R}.
\]

where \( n \) is the index of refraction of the surface.

#### Example

For example, the average cornea has a radius of curvature of 1.376 m. In air (\( n = 1.0 \)), the average cornea would have a power of

\[
\phi = \frac{1}{R} - 1 = \frac{1}{1.376} - 1 = 0.022.
\]

This value is in units of diopters.

#### Thin Lenses

When the thickness of lenses is much less than the radius of curvature, lenses can be approximated as "thin" lenses. For thin lenses, the powers of two thin lenses in contact add to the net power.

\[
\phi_{net} = \phi_1 + \phi_2.
\]

where \( \phi_1 \) and \( \phi_2 \) are the powers of the individual thin lenses.

#### Thick Lenses

A thick lens is an optical element made of a material with different refractive indices. The spherical refracting surface can be extended to a thick lens. A thick lens is made of a material with different refractive indices, and is treated as a set of thin lenses in contact.

The focal length of a thick lens is given by

\[
f_{net} = \frac{1}{\phi_{net}} = \frac{1}{\phi_1 + \phi_2}.
\]

where \( \phi_{net} \) is the net power of the thick lens.

### Figure 3.6

**Spherical refractive surface of radius \( R \), separating spaces with refractive index \( n \) and \( n' \).**

- **Optical axis**: The line along which rays are incident and pass through the surface.
- **Vertex**: The point on the surface where the rays intersect.
- **Focal lengths**: The distances from the surface to the focal points, defined as
  - **Front focal length**: \( f_R = -\frac{1}{R} \)
  - **Rear focal length**: \( f_F = -\frac{1}{n} \frac{1}{R} \)

**Equations**

- \((3.10)\): \( f = -\frac{1}{R} \)
- \((3.11)\): \( f = -\frac{1}{n} \frac{1}{R} \)
- \((3.12)\): \( \phi = \frac{n}{n'} - 1 \)
- \((3.13)\): \( \phi = \phi_1 + \phi_2 \)
- \((3.14)\): \( \phi_{net} = \frac{1}{\phi_{net}} \)

**Table**

<table>
<thead>
<tr>
<th>Index</th>
<th>n'</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

**Example**

- If the two thin lenses are placed in contact, then \( t = 0 \) and the net power is given by
  \[
  \phi_{net} = \phi_1 + \phi_2.
  \]

**Notes**

- The power of a spherical refracting surface can also be defined by its curvature. The power is defined as the reciprocal of the radius of curvature.
- The power is measured in units of meters, the power of the surface is in units of meters, and the surface in air is measured in units of meters. In cases such as that shown in Figure 3.6, when \( n > n' \), rays parallel to the optical axis will converge as they pass through the surface. This situation is associated with a negative power. When \( n' > n \), the same incident rays would diverge after passing through the surface.
Geometrical optics

Figure 3.7 Collimated light focuses to a point halfway to the center of curvature of the mirror.

and the front and rear focal lengths are

\[ f_f = f'_f = -\frac{n}{\phi} = -nf = \frac{R}{2} = \frac{1}{2C}. \quad (3.18) \]

This last result means that the front and rear focal points for a mirror lie halfway between the center of curvature and the vertex of the surface. Figure 3.7 shows a concave mirror. In visual optics, concave mirrors are often used in illumination systems for surgical microscopes to concentrate light onto the surgical field. The properties of convex mirrors are used in keratometry to measure the radius of curvature of the anterior corneal surface.

### 3.5 GAUSSIAN IMAGING EQUATION

The Gaussian imaging equation allows the determination of the object and image locations for a surface of power \( \phi \). The distance from the surface vertex to the object will be denoted by \( z \) and the distance from the object to the image plane is denoted by \( z' \). In keeping with the sign convention, these distances are negative if the distance is measured to the left of the surface and positive if it is to the right of the surface. The Gaussian imaging equation relates these distances to the surface power as

\[ \frac{n'}{z'} = \phi + \frac{n}{z}. \quad (3.19) \]

The planes located at \( z \) and \( z' \) are said to be conjugate since they satisfy the Gaussian imaging equation. Figure 3.8 shows an object point a distance \( z \) from the refracting surface. This point is imaged to a point a distance \( z' \) from the single refracting surface. In examining Equation 3.19 and comparing it to the definition of vergence in Equation 3.2, it is evident that \( n/z \) is the vergence of the object point as measured from the plane of the refracting surface. Similarly, \( n'/z' \) is the vergence of the image point as measured from the plane of the refracting surface. Consequently, the image vergence is the sum of the object vergence and the power of the refracting surface. The refracting surface modifies the object vergence to produce the image vergence. The Gaussian imaging equation holds for mirror surfaces as well, again with the assumption that \( n' = -n \).

The transverse (or lateral) magnification, \( m \), of the system for an extended object can also be defined by the object and image distances. Figure 3.8 shows an extended object of height \( y \), located a distance \( z \) to the left of the surface. The image of this object has a height \( y' \). The sign convention says that positive heights are above the optical axis and negative heights are below the optical axis. The transverse magnification is the ratio of the image height to the object height and is related to the object and image distances by

\[ m = \frac{y'}{y} = \frac{nz'}{nz}. \quad (3.20) \]

Figure 3.8 shows a case where \( m < 0 \), which means that the image is inverted relative to the object.

#### 3.5.1 THICK LENSES AND GAUSSIAN IMAGING

In the previous section, only a single refracting surface was considered. In a similar manner, the imaging equation will hold for a thin lens with its power defined as the sum of the surface powers as in Equation 3.13. In applying the Gaussian imaging equation to thick lenses with power \( \Phi \) as defined in Equation 3.12, a slight modification is required. For the thick lens, the object and image distances \( z \) and \( z' \) need to be measured relative to the front and rear principal planes of the lens, \( P \) and \( P' \). The principal planes can be considered the planes of effective refraction for the lens. The object wave with its object vergence entering the front principal plane is mapped to the rear principal plane, where it is then converted to the image vergence by adding the power of the lens. In the cases of a single refracting (or reflecting) surface and a thin lens, both the principal planes were located at the surface. For a thick lens, the principal planes are in general displaced from the surface vertices and are separated from one another. The front principal plane is located at a distance \( d \) from the first surface vertex, where this distance is given by

\[ d = \frac{\Phi}{\Phi - \frac{1}{n_{lens}}}. \quad (3.21) \]
Similarly, the rear principal plane is located a distance $d'$ from the second surface vertex, where this distance is given by

$$d' = \frac{\phi_1 t - f_1 F}{\phi_2 n_{\text{lens}}},$$  \hspace{1cm} (3.22)

where
- $\phi_1$ and $\phi_2$ are the powers of the individual refracting surfaces of the thick lens
- $t$ is the thickness of the lens
- $n_{\text{lens}}$ is the refractive index of the lens material

The total power of the thick lens $\Phi$ is defined in Equation 3.12. Figure 3.9 shows the locations of the principal planes for a thick lens.

### 3.5.2 CARDINAL POINTS

In general, optical systems will consist of multiple thick lens elements whose surface shapes and materials are carefully chosen to provide the imaging requirements and quality desired for the application. These complex systems can be reduced to a series of six “special” points known as the cardinal points that fully define the Gaussian imaging and magnification properties of the system. Four of these six points have already been encountered in the preceding discussion. The front and rear focal points, $F$ and $F'$, arise when collimated light enters or leaves the optical system. For collimated incident rays, the optical system converts the vergence of the bundle so that the exiting vergence appears to emanate from the rear focal point. Similarly, an object point located on the front focal point will result in a beam emerging from the optical system that has zero vergence or, in other words, is collimated. The second set of points associated with the optical system is the front and rear principal points. The principal points are located where their respective principal planes cross the optical axis. For a thick lens, the locations of the principle planes (points) were defined by Equations 3.21 and 3.22. For a general optical system, the locations can be identified by ray tracing the system. Once located, the object and image distance are measured with respect to the principal points and the Gaussian imaging formula holds. The principal planes are said to be planes of unit magnification. A ray striking the front principal plane at a given location is mapped to an identical location on the rear principal plane and appears to emerge from that point. In general, the emerging ray will be traveling in a different direction than the incident ray.

The last pair of cardinal points is the nodal points, $N$ and $N'$. A ray traveling at an angle to the optical axis and passing through the front nodal point emerges from rear nodal point at the same angle. The nodal points can be found relative to the principal points. In general, the nodal points of a system are shifted relative to the corresponding principal points such that

$$PN = PN' = f + f'' = (n' - n)f,$$ \hspace{1cm} (3.23)

where
- $n$ is the object space index
- $n'$ is the image space index

In cases where the object and image space indices are the same (e.g., a camera lens in air), the nodal points are located at the principal planes. In cases where the object and image refractive indices differ (e.g., the eye where $n = 1.0$ and the image space contains vitreous humor with refractive index $n' = 1.336$), the nodal points are shifted relative to the principal points.

With knowledge of these six cardinal points, a multielement optical system can be reduced to a “black box” and properties of the rays entering and leaving the system can be easily determined without knowledge of all of the surfaces, spacing, and materials within the box. Figure 3.10 shows the properties of the cardinal points of a generalized optical system.
3.5.3 APERTURE STOP AND PUPILS

The aperture stop is a mask within the system that limits the size of the bundle of rays that passes through an optical system. This mask can be a separate mask or the clear opening of one of the optical elements. The aperture stop is typically circular, but this is not a requirement. Each point source on an object radiates light in all directions. The aperture stop in effect blocks most of these rays and only allows a limited cone of light from the point source to pass through the optical system. The object and image locations, as well as the system magnification, are independent of the aperture stop location. However, the aperture stop dimensions and location affects the amount of light that reaches the image, as well as how well the rays come to a focus. A judicious choice of the location of the aperture stop can optimize the quality of an optical system.

The entrance pupil is the image of the aperture stop formed by all of the lens (optical) surfaces preceding the stop, and the exit pupil is the image of the stop formed by all of the surfaces following the stop. If the entrance pupil is the object for the entire system, its image is formed at the exit pupil. For a well-corrected system, the exit pupil is a 1:1 mapping of the entrance pupil, which is to say that the scaling of rays passing through the entrance pupil will be mapped to the same relative position in the exit pupil. The entrance pupil can be thought of as a port that captures light from the object scene. The light that gets into the entrance pupil makes it to the exit pupil (assuming no loss due to reflections and/or vignetting) and contributes to the image. In the eye, for example, the iris acts as the aperture stop. The cornea acts as lens and forms an image of the iris located approximately 3 mm posterior to the corneal vertex. This image is the entrance pupil of the eye and it is approximately 10% larger than the physical size of the opening in the iris.

3.5.4 CHIEF AND MARGINAL RAYS

There are several special rays that provide useful information regarding the properties of an optical system. One set of rays are called marginal rays. These rays start at the object on the optical axis and pass through the edge of the aperture stop. In an ideal system, the marginal rays will propagate through the optical system and ultimately converge to the optical axis at the image plane. Furthermore, if the marginal rays intersect the optical axis at some location between the object and image, then an intermediate image is formed. These intermediate image planes are useful in that a mask or reticle can be placed at the location, having the effect of superimposing the reticle pattern onto the final image. This technique is often used in microscope eyepieces, for example, to allow scaled rulings to be placed over the image to measure features.

A second special ray defined with regard to the aperture stop is the chief ray (principal ray). The chief ray is defined as the ray that starts at the edge of the object and passes through the center of the aperture stop. When the chief ray crosses the optical axis, it is called a pupil plane. The size of this pupil is defined by the height of the marginal ray in the pupil plane.

The positions of the entrance and exit pupil are determined by where the chief ray appears to cross the optical axis in object and image space. The chief ray incident on the first surface is projected to determine the point it crosses the optical axis. This point defines the location of the entrance pupil. The size of the entrance pupil is determined by projecting the incident marginal ray onto the plane of the entrance pupil. The marginal ray will define the boundary of the entrance pupil. Similar definitions hold for the exit pupil, where the emerging chief and marginal rays are used to define the location and size of the exit pupil.

3.6 CYLINDRICAL AND TORIC SURFACES

Most optical systems are rotationally symmetric. The spherical surfaces of the lenses are chosen to provide the imaging properties and magnification required for a specific task. Multiple elements may be used to optimize image quality. Occasionally, rotationally symmetric aspheric surfaces are used to further improve image quality or reduce the required number of elements. However, some optical systems lack this rotational symmetry. The human eye, for example, often suffers from astigmatism since its optical surfaces are not necessarily rotationally symmetric. Consequently, astigmatic lenses often appear in visual optics since they are used to compensate for astigmatism in the eye. The simplest astigmatic lens is a cylindrical lens.

3.6.1 POWER AND AXIS OF A CYLINDRICAL LENS

A cylindrical lens has one or more surfaces that are a section of a cylinder. Figure 3.11a shows an example of a positive powered cylindrical lens. A cross section through one meridian of the cylindrical lens shows a circular surface of radius $R_s$, meaning the power in this meridian is given by Equation 3.7. However, in a perpendicular meridian, the cross section is flat, meaning there is no power in this direction. Consequently, a cylindrical lens can also be rotated about the optical axis. To describe this orientation, the cylinder axis $\Theta_j$ is defined as the angle of the zero-power meridian, measured counterclockwise from the horizontal axis.

The power and orientation of a cylindrical lens can be described by $\phi_s \times \Theta_j$, where $\phi_s = (n - n)R_s$ is the cylindrical lens power. Note that due to the shape of the cylindrical lens, there is a redundancy in the definition of the cylinder axis. If the zero-power axis of cylindrical lens is oriented along the horizontal meridian, then $\Theta_j = 0^\circ$ or $\Theta_j = 180^\circ$. By convention, the cylinder axis is uniquely defined as being in the range $0^\circ < \Theta_j < 180^\circ$. With this definition, the horizontally oriented cylindrical lens has a cylinder axis $\Theta_j = 180^\circ$.

3.6.2 TORIC AND SPHEROCYLINDRICAL SURFACES

A more complicated astigmatic lens introduces both spherical and cylindrical power simultaneously. Examples of refractive surfaces that exhibit power variation are toric and spherocylindrical surfaces. These surfaces have a short radius of curvature $R_s$ along one meridian (called the steep meridian) and a longer radius of curvature $R_f$ along the orthogonal meridian (called the flat meridian). The shape of toric and spherocylindrical surfaces match along these principal meridia, but in general the surface shapes are slightly different away from these meridia. The powers
3.6 Cylindrical and toric surfaces

determine the magnification imparted by the anterior cornea. From the magnification \( m \), the corneal radius can be determined by combining Equations 3.18 through 3.20, such that

\[
R = \frac{2nz}{m-1}
\]  

(3.25)

where \( z \) is the distance between the keratometer target and the eye. Note that in the sign convention described previously, \( z \) is a negative quantity. In cases of corneal astigmatism, the ring image will be elliptical with the long axis of the ellipse corresponding to the flat meridian and the short axis of the ellipse corresponding to the steep meridian. Thus, the orientation of the ellipse determines the orientation of the principal meridia of the cornea. The magnification along both the short and long axes of the ellipse can be measured separately and Equation 3.25 is used to determine the radii \( R_f \) and \( R_s \) along their respective axis. Finally, Equation 3.24 is used to determine corneal power along each direction. If \( n' = 1.376 \) and \( n = 1.0 \) in these equations, then the true optical power of the anterior cornea is determined. In keratometry, an artificial refractive index called the “keratometric refractive index” \( n' = n_k \) is used in these expressions. The value of \( n_k \) has been chosen to reduce the anterior corneal power to account for the negative power imparted by the posterior corneal surface. So keratometry is an estimate of the total corneal power (combined anterior and posterior powers) based solely on the anterior corneal shape. Typical values of the keratometric refractive index are \( n_k = 1.3375 \) and \( n_k = 1.3315 \), depending on the device used to measure the cornea.

As an example, suppose the cornea has a radius \( R = 7.8 \) mm along a meridian oriented at \( 30^\circ \) and a radius \( R_f = 8.0 \) mm along a meridian oriented at \( 120^\circ \). Based on Equation 3.23 and using \( n' = 1.376 \), the true anterior corneal powers are \( \phi_f = 48.21 \) D and \( \phi_s = 47.00 \) D. The keratometry values associated with this cornea are given by

\[
K_f = \frac{1.3375 - 1}{7.8} = 43.27 \text{ D} \quad \text{and} \quad K_s = \frac{1.3375 - 1}{8.0} = 42.19 \text{ D},
\]  

(3.26)

where \( K_f \) and \( K_s \) have been used to distinguish between the keratometric power and the true anterior corneal power. Note, \( K_f \) is typically referred to as the “steep-K value” and \( K_s \) is typically referred to as the “flat-K value.” Finally, the orientation of the steep and flat meridians are specified along with these keratometric powers and the “@” symbol. In the preceding example, the full keratometry measurement would be described as \( 43.27 \) D @ \( 30^\circ \)/42.19 D @ \( 120^\circ \). The absolute difference between the keratometric powers suggests that the example cornea has slightly more than 1 D or corneal astigmatism.

Lenses for correcting refractive error also incorporate astigmatic surfaces. From a thin lens standpoint, these types of lenses can be considered as a combination of a spherical lens and a cylindrical lens. In Equation 3.15, the powers of two thin lenses simply added when the separation between them was negligible. For astigmatic thin lenses, the same effect holds, but each of the principal meridia needs to be considered independently. Combining a thin spherical lens with a thin cylindrical
Geometrical optics

A lens results in a thin spherocylindrical lens. This resultant spherocylindrical lens has the spherical lens power along one principal meridian and the combined spherical and cylindrical power along the other principal meridian. For example, the combination of a −3.00 D spherical lens with a +1.50 D cylindrical lens with cylinder axis at 180° leads to a spherocylindrical lens with power of −3.00 D along the horizontal axis since the cylindrical lens has no power in this direction. Along the vertical axis, the resultant lens has a power of −1.50 D since along this meridian the powers of the spherical and cylindrical lens add. The prescription for this spherocylindrical lens is typically written as

$$-3.00/+1.50 \times 180°,$$  \hspace{1cm} (3.27)

where the first component is the spherical lens power and the second component is the cylindrical lens power and its cylinder axis. This prescription is said to be in “plus cylinder form” since the cylindrical lens power is positive.

As a second example, consider a spherical lens with a power of −1.50 D combined with a cylindrical lens of power −1.50 D with cylinder axis at 90°. In this case, the resultant spherocylindrical lens has a power of −1.50 D along the vertical axis since only the spherical lens contributes in this direction. Along the horizontal axis, the power is −3.00 D since both the spherical and cylindrical lens combines in this direction. The prescription for this lens is typically written as

$$-1.50/-1.50 \times 90°,$$ \hspace{1cm} (3.28)

where again the first component is the spherical lens power and the second component is the cylindrical lens power and its cylinder axis. This prescription is said to be in “minus cylinder form” since the cylindrical lens power is negative. However, comparing the properties of the resultant spherocylindrical lenses from the two examples shows that both lenses have identical power distributions. There are always two ways to achieve a given spherocylindrical lens. One prescription is in plus cylinder form and the other is in minus cylinder form.

It is useful to convert between the two cylinder forms. The following steps perform this conversion:

1. The new spherical component is the sum of the spherical and cylindrical powers of the old form.
2. The new cylindrical component is the negative of the old cylinder component.
3. The new cylinder axis is 90° from the old cylinder axis.
4. If the new axis does not fall within the 1°–180° range, then add or subtract 180° from the new axis to place it in this range. This technique works for converting from plus to minus cylinder form, as well as from minus to plus cylinder form.

The previous examples are used to illustrate this conversion technique. To convert the plus cylinder form prescription −3.00/ +1.50 × 180° to minus cylinder form, the first step is to add the spherical and cylindrical powers. The new spherical power is therefore −3.00 + 1.50 = −1.50 D. The second step is to negate the cylindrical power, so the new cylindrical power is −(−1.50) = +1.50 D. The third step is to rotate the axis by 90°, leading to a new cylinder axis of 180° + 90° = 270°. Finally, the new cylinder axis should be in the range of 1°–180°, so it is observed that the 270° is the same as the 90° axis and the new cylinder axis adjusted accordingly. The final prescription in minus cylinder form based on these steps is given in Equation 3.28. Note, steps 3 and 4 can be combined by simply rotating the axis 90° in the direction that puts it in the desired range. In the preceding example, at step 3, the new cylinder axis can easily be obtained by calculating instead 180°−90° = 90°.

A final concept that often arises with astigmatic lenses is spherical equivalent power (SEP). The SEP is the average power of a spherocylindrical lens. As described previously, the power of a spherocylindrical lens varies sinusoidally from a minimum power along its flat meridian to a maximum power along its steep meridian. The SEP is given by the spherical power plus half the cylindrical power. This definition holds for both the plus and minus cylinder forms. From the previous spherocylindrical examples, the SEP is given by

$$\text{SEP} = -3.00 + \frac{+1.50}{2} = -1.50 + \frac{-1.50}{2} = -2.25 \text{D}. \hspace{1cm} (3.29)$$

Finally, the Jackson crossed cylinder is a specialty lens often used in ophthalmic optics. It is a spherocylindrical lens with a SEP of zero. It has a power $\theta$ in one meridian and a power $-\theta$ in the orthogonal meridian. These lenses introduce pure astigmatism and no spherical component. The prescription for such a lens is given by

$$\theta = 2\theta \times 0, \hspace{1cm} (3.30)$$

where $\theta$ is the cylinder axis. The Jackson crossed cylinder is formed by combining a plano-convex positive cylindrical lens with a plano-concave negative cylindrical lens. The powers of each lens have equal magnitude, but opposite sign. The axes of the lenses are set to be orthogonal. Figure 3.12 demonstrates a Jackson crossed cylinder.

![Figure 3.12](image-url) The Jackson crossed cylinder can be formed by combining conventional positive and negative cylindrical lenses.
3.7 VISUAL INSTRUMENTS

There are a variety of instruments that are designed to use the eye as the final detector. In many cases, these systems do not form an image per se but, instead, have the emerging light collimated or slightly diverging so that the eye can ultimately form the final image on the retina. In such cases, the lateral magnification of the optical system often becomes ill-defined because the eye needs to be considered as part of the overall system. Visual instruments typically use the angular subtense of the light emerging from the systems relative to an unaided view as a measure of magnification. The systems described in this section follow these specifications.

3.7.1 SIMPLE MAGNIFIER AND MAGNIFYING POWER

Perhaps, the simplest visual instrument is the simple magnifier. This device can be as simple as a single positive lens. The lens forms a virtual image of a given object where the image distance is a comfortable viewing distance from the eye. This virtual image also subtends a larger angle than the original object so that it appears larger to the viewer. One way of increasing the angular subtense of an object is to bring it close to the eye. Figure 3.13a shows an object with height \( h \) close to the eye that subtends an angle \( u_0 \). The problem with this arrangement is that the eye must accommodate in order to focus on such a near object. This accommodation can produce excessive strain or may not even be possible since people gradually lose their ability to accommodate with age.

Moving the object further away can alleviate this strain by placing the object at a comfortable viewing distance. This distance is typically taken as 250 mm. Figure 3.13b shows the object at this viewing distance. Moving the object away from the eye reduces its angular subtense to \( u_1 \). A simple magnifier is used to recover the original angular subtense, but place the image of the object at a comfortable viewing distance. If the object is placed within the focal length of the magnifier, then a virtual image as shown in Figure 3.13c is formed. This virtual image now has an angular subtense of \( u_0 \). The magnifying power (MP) is defined as

\[
MP = \frac{\text{Angular subtense of the virtual image with magnifier}}{\text{Angular subtense of the object at a comfortable viewing distance}} = \frac{u_0}{u_1}.
\]

The MP is a measure of how much larger the virtual image is relative to the object at some finite distance. This concept will appear again when discussing microscopes in Section 3.7.2.

3.7.2 MICROSCOPES

Microscopes extend the capabilities of the simple magnifier, enabling much higher magnifications to be achieved. Most modern microscopes are based on infinity-corrected objective lenses. Figure 3.14 illustrates such a system. The microscope consists of three lens groups: an infinity-corrected objective with power \( \Phi_{\text{obj}} \), a tube lens with power \( \Phi_{\text{tube}} \), and an eyepiece with power \( \Phi_{\text{eye}} \). The corresponding focal lengths are \( f_{\text{obj}} = 1/\Phi_{\text{obj}}, f_{\text{tube}} = 1/\Phi_{\text{tube}} \), and \( f_{\text{eye}} = 1/\Phi_{\text{eye}} \) respectively. The object is placed in the front focal plane of the objective lens. The objective lens collimates light from points on the object. The tube lens collects these collimated beams and creates a magnified intermediate image. The eyepiece then acts much like a simple magnifier and reimages the intermediate image to a comfortable viewing distance with magnifying power \( MP_{\text{eye}} \). The magnification of a microscope with an infinity-corrected objective is given by

\[
\text{Magnification} = \frac{\Phi_{\text{obj}}}{\Phi_{\text{tube}}} MP_{\text{eye}} = \frac{f_{\text{tube}}}{f_{\text{obj}}} MP_{\text{eye}}.
\]

The collimated space between the objective and tube lenses is advantageous over more traditional designs where the objective lens forms the intermediate image directly at a standardized location. In the infinity-corrected objective case, additional
optical elements such as filters, polarization elements, and beam splitters can be placed in the collimated space without introducing aberrations.

### 3.7.3 Telescopes and Angular Magnification

Telescopes are afocal optical systems, or systems with a net power of zero. The net power of a system with two separated thin lenses is given by Equation 3.14. The first thin lens is called the objective lens and has a power \( \Phi_{\text{obj}} \). The second lens is called the eyepiece and has a power \( \Phi_{\text{eye}} \). The corresponding focal lengths are \( f_{\text{obj}} = \frac{1}{\Phi_{\text{obj}}} \) and \( f_{\text{eye}} = \frac{1}{\Phi_{\text{eye}}} \), respectively. From Equation 3.14, the net power of a telescope is

\[
\Phi_{\text{net}} = \Phi_{\text{obj}} + \Phi_{\text{eye}} - \Phi_{\text{obj}} \Phi_{\text{eye}}. \tag{3.33}
\]

For the telescope to have a net power of zero, the separation between the two thin lenses must satisfy

\[
t = \frac{\Phi_{\text{obj}} + \Phi_{\text{eye}}}{\Phi_{\text{obj}} \Phi_{\text{eye}}} = f_{\text{obj}} + f_{\text{eye}}. \tag{3.34}
\]

Since telescopic systems do not form an image, the lateral magnification is again ill-defined. For telescopes, angular magnification is used instead, where the angular magnification, \( m_{\text{a}} \), is defined as the ratio of the angular subtense of the object as viewed from the objective to the angular subtense of the object as viewed by the eye through the telescope.

Two common types of telescopes are the Galilean and the Keplerian telescope. These telescopes are shown in Figure 3.15. The Galilean telescope shown in Figure 3.15a consists of a low power positive lens (objective) and a high power negative lens (eyepiece). Since the focal length of the eyepiece is negative, the separation between the object and the eyepiece is less than the focal length of the objective, leading to a compact system. The angular magnification of the Galilean telescope is given by

\[
m_{\text{a}} = \frac{\Phi_{\text{eye}}}{\Phi_{\text{obj}}} = \frac{f_{\text{eye}}}{f_{\text{obj}}}. \tag{3.35}
\]

Since the powers of the two lenses have opposite sign, the angular magnification is positive, meaning the image viewed through the Galilean telescope is upright. The primary drawback to the Galilean telescope is that the exit pupil is located in between the two lenses. This means that the eye cannot be placed at the position of the exit pupil and consequently the field of view of the telescope is reduced.

A second common telescope type is the Keplerian telescope shown in Figure 3.15b. This telescope consists of two positive powered lenses. The separation between the two thin lenses is again given by the sum of the focal lengths as in Equation 3.34. In this case, both focal lengths are positive, so the separation between the lenses is larger than the equivalent Galilean telescope. The angular magnification for the Keplerian telescope is again given by Equation 3.35. However, since both lens powers are positive, the angular magnification is negative, leading to an inverted image. One advantage of the Keplerian telescope over the Galilean is that the exit pupil of the Keplerian telescope lies outside the two thin lenses. The eye can be placed at the location of the exit pupil, and the field of view of the system is consequently larger compared to the Galilean system.

Both telescopic systems are routinely used in visual optics. A common application is to provide increased angular magnification to people with low vision. Both systems can be used, but the Keplerian telescope requires additional prisms or lenses to make the perceived image upright. The trade-off between the two systems is a compact Galilean telescope with a small field of view or a longer Keplerian telescope with a wider field of view.

### 3.8 Summary

Geometrical optics studies how light beams travel through optical systems. For imaging systems, the main goal is to relay the object plane to the image plane. In addition, the size of the image is typically important. The Gaussian imaging equation provides the connection between the object and image, as well as the system magnification. While the techniques demonstrated in this chapter examine simplified cases of thin lenses and systems of a few lenses, these examples are sufficient to understanding the basic mechanisms of the image-forming process. Optical system designers will typically use many more elements to optimize the quality of the image in terms of brightness and sharpness. This chapter also showed that these more complex systems can be reduced to a “black box,” where only the cardinal points need to be known to determine the imaging properties. In this manner, the designer can be concerned with the elements within the box, and the user can implement the system without full knowledge of every component within the box. Visual optics employs a wide array of optical systems, from systems for presenting visual stimuli, to heads-up display systems, to diagnostic imaging devices, to low vision aids. The workings of the rich and varied systems can often be understood with the fundamentals presented here.