CHAPTER 1
Optical Aberrations

1.1 INTRODUCTION

This chapter starts with the concepts of *aperture stop* and *entrance* and *exit pupils* of an optical imaging system. Certain special rays, such as the *chief* and the *marginal*, are defined. The *wave aberration* associated with a ray is defined and its relationship to the corresponding transverse *ray aberration* is given. Representations of wavefront *defocus* and *tilt* aberrations are given. We introduce different forms of the *primary aberration function* of a *rotationally symmetric system.* How this function changes as the aperture stop of the system is moved from one position to another is discussed. The primary aberration function for the simplest imaging system, namely, a single spherical refracting surface, is given for an arbitrary position of the aperture stop. Finally, we outline a procedure by which the aberration function of a multielement system may be calculated. This procedure is utilized in later chapters, for example, to calculate the aberration of a *thin lens* (Chapter 2) and *a plane-parallel plate* (Chapter 3). This chapter forms the basis of Part I on *geometrical optics.*

1.2 OPTICAL IMAGING

An optical imaging system consists of a series of refracting and/or reflecting surfaces. The surfaces refract or reflect light rays from an object to form its image. The image obtained according to geometrical optics in the *Gaussian approximation*, i.e., according to Snell's law in which the sines of the angles are replaced by the angles, is called the *Gaussian image.* The Gaussian approximation and the Gaussian image are often referred to as the *paraxial approximation* and the *paraxial image*, respectively. We assume that the surfaces are rotationally symmetric about a common axis called the *optical axis* (*OA*). Figure 1-1 illustrates the imaging of an on-axis point object P_0 and an off-axis point object P, respectively, by an optical system consisting of two thin lenses. (For definition of a thin lens, see Section 2.2.) P' and P_0' are the corresponding Gaussian image points. An object and its image are called *conjugates* of each other, i.e., if one of the two conjugates is an object, the other is its image.

An aperture in the system that physically limits the solid angle of the rays from a point object the most is called the *aperture stop* (*AS*). For an extended (i.e., a nonpoint) object, it is customary to consider the aperture stop as the limiting aperture for the axial point object, and to determine vignetting, or blocking of some rays, by this stop for off-axis object points. The object is assumed to be placed to the left of the system so that initially light travels from left to right. The image of the stop by surfaces that precede it in the sense of light propagation, i.e., by surfaces that lie between it and the object, is called the *entrance pupil* (*EnP*). When observed from the object side, the entrance pupil appears to limit the rays entering the system to form the image of the object. Similarly, the image of the aperture stop by surfaces that follow it, i.e., by surfaces that lie between it and the

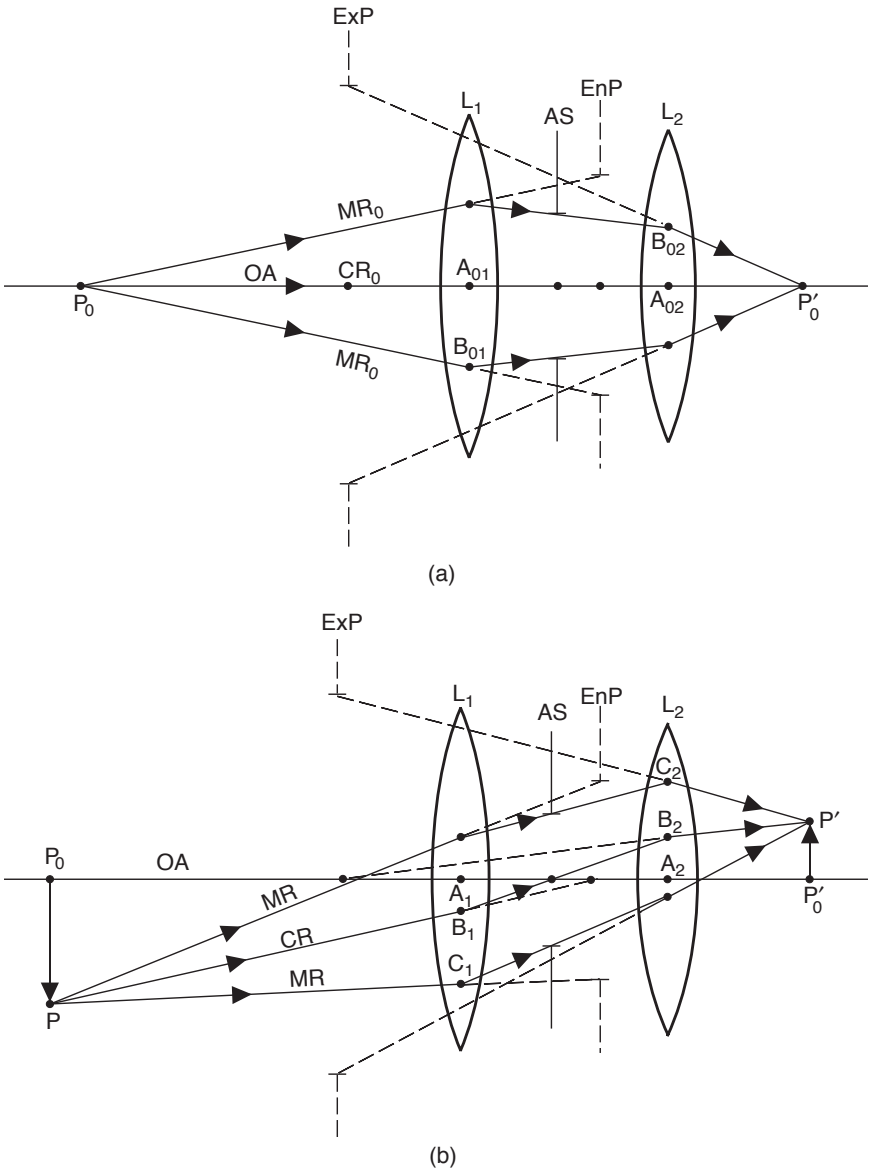

Figure 1-1. (a) Imaging of an on-axis point object P_0 by an optical imaging system consisting of two lenses L_1 and L_2. *OA* is the optical axis. The Gaussian image is at P_0'. *AS* is the aperture stop; its image by L_1 is the entrance pupil *EnP,* and its image by L_2 is the exit pupil *ExP*. CR_0 is the axial chief ray, and MR_0 is the axial marginal ray. (b) Imaging of an off-axis point object *P.* The Gaussian image is at *P′*. *CR* is the off-axis chief ray, *MR* is the off-axis marginal ray.

image, is called the *exit pupil* (*ExP*). The object rays reaching its image appear to be limited by the exit pupil. Since the entrance and exit pupils are images of the stop by the surfaces that precede and follow it, respectively, the two pupils are conjugates of each other for the whole system; i.e., if one pupil is considered as the object, the other is its image formed by the system.

An object ray passing through the center of the aperture stop and appearing to pass through the centers of the entrance and exit pupils is called the *chief* (or the *principal) ray* (*CR*). An object ray passing through the edge of the aperture stop is called a *marginal ray* (*MR*). The rays lying between the center and the edge of the aperture, and, therefore, appearing to lie between the center and edge of the entrance and exit pupils, are called *zonal rays.*

It is possible that the stop of a system may also be its entrance and/or exit pupil. For example, a stop placed to the left of a lens is also its entrance pupil. Similarly, a stop placed to the right of a lens is also its exit pupil. Finally, a stop placed at a single thin lens is both its entrance and exit pupils.

1.3 WAVE AND RAY ABERRATIONS

In this section, we define the wave aberration associated with a ray and relate it to its transverse ray aberration in an image plane. The *optical path length* of a ray in a medium of refractive index n is equal to n times its geometrical path length. If rays from a point object are traced through the system and up to the exit pupil such that each one travels an optical path length equal to that of the chief ray, the surface passing through their end points is called the system *wavefront* for the point object under consideration. If the wavefront is spherical with its center of curvature at the Gaussian image point, we say that the Gaussian image is perfect. If, however, the wavefront deviates from this *Gaussian spherical wavefront*, we say that the Gaussian image is aberrated. The optical deviation (i.e., geometrical deviations times the refractive index n of the image space) of the wavefront along a certain ray from the Gaussian spherical wavefront is called the *wave aberration* of that ray. It represents the difference between the optical path lengths of the ray under consideration and the chief ray in traveling from the point object to the reference sphere. Accordingly, the wave aberration associated with the chief ray is zero. The wave aberration associated with a ray is positive if it has to travel an extra optical path length, compared to the chief ray, in order to reach the Gaussian spherical wavefront. The Gaussian spherical wavefront is also called the *Gaussian reference sphere.*

Figures 1-2a and 1-2b illustrate the reference sphere S and the aberrated wavefront W for on- and off-axis point objects whose Gaussian images lie at P_0' and P', respectively. The coordinate system is also illustrated in these figures. We choose a right-hand coordinate system such that the optical axis lies along the z axis. The object, entrance pupil, exit pupil, and the Gaussian image lie in mutually parallel planes that are perpendicular to this axis, with their origins lying along the axis. We assume that a point object such as P lies along the x axis. The zx plane containing the point object and the optical axis is called the *tangential* or the *meridional plane.* The Gaussian image P' lying in the Gaussian image plane along its x axis also lies in the tangential plane. This may be seen by a consideration of a tangential object ray and Snell's law according to which the incident and refracted or reflected rays at a surface lie in the same plane. The chief ray

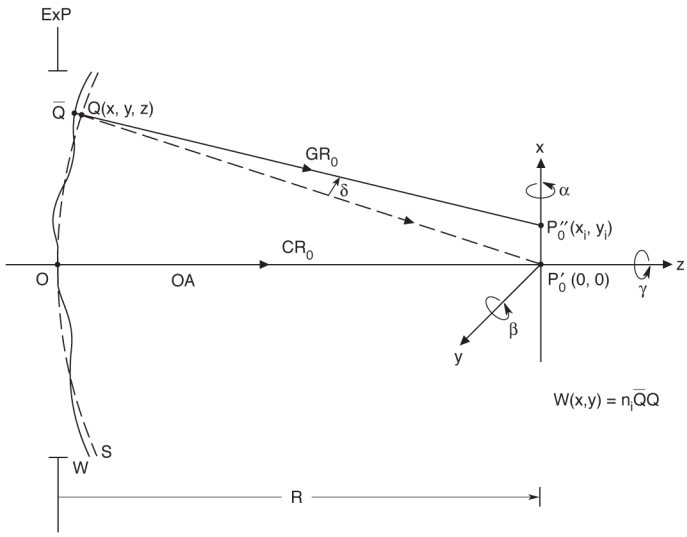

Figure 1-2a. Aberrated wavefront for an on-axis point object P_0. The reference sphere S of radius of curvature R is centered at the Gaussian image point P_0'. The wavefront W and reference sphere pass through the center O of the exit pupil ExP. A right-hand Cartesian coordinate system showing x, y, and z axes is illustrated, where the z axis is along the optical axis of the imaging system. Angular rotations α, β, and γ about the three axes are also indicated. CR_0 is the chief ray, and a general ray GR_0 is shown intersecting the Gaussian image plane at P_0''.

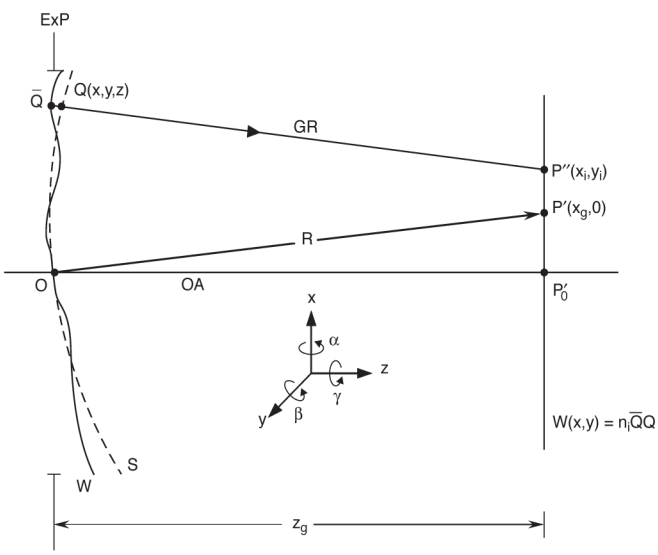

Figure 1-2b. Aberrated wavefront for an off-axis point object P. The reference sphere S of radius of curvature R is centered at the Gaussian image point P'. The value of R in this figure is slightly larger than its value in Figure 1-2a. GR is a general ray intersecting the Gaussian image plane at the point P''. By definition, the chief ray (not shown) passes through O, but it may or may not pass through P'.

always lies in the tangential plane. The plane normal to the tangential plane but containing the chief ray is called the *sagittal plane.* As the chief ray bends when it is refracted or reflected by an optical surface, so does the sagittal plane.

Consider an image ray such as GR in Figure 1-2b passing through a point Q with coordinates (x, y, z) on the reference sphere of radius of curvature R centered at the image point. We let $W(x, y)$ represent its wave aberration $n\overline{Q}Q$, since z is related to x and y by virtue of Q being on the reference sphere. It can be shown that the ray intersects the Gaussian image plane at a point P'' whose coordinates with respect to the Gaussian image point P' are approximately given by

$$(x_i, y_i) = \frac{R}{n}\left(\frac{\partial W}{\partial x}, \frac{\partial W}{\partial y}\right) \quad . \tag{1-1}$$

[Equation (1-1) has been derived by Mahajan, Born and Wolf, and Welford. However, Welford uses a sign convention for the wave aberration that is opposite to ours.]

The displacement $P_0'P_0''$ in Figure 1-2a (or $P'P''$ in Figure 1-2b) of a ray from the Gaussian image point is called its *geometrical* or *transverse ray aberration*, and its coordinates (x_i, y_i) in the Gaussian image plane relative to the Gaussian image point are called its ray aberration components. Since a ray is normal to a wavefront, the ray aberration depends on the shape of the wavefront and, therefore, on its geometrical path difference from the reference sphere. The division of W by n in Eq. (1-1) converts the optical path length difference into geometrical path length difference. When an image is formed in free space, as is often the case in practice, then $n = 1$. The angle $\delta \simeq P_0'P_0''/R$ between the ideal ray QP_0' and the actual ray QP_0'' is called the *angular ray aberration.*

The distribution of rays from a point object in an image plane is called the *ray spot.* (Such diagrams are discussed in Chapter 7.) When the wavefront is spherical with its center of curvature at the Gaussian image point, then the wave and ray aberrations are zero. In that case, all of the object rays transmitted by the system pass through the Gaussian image point, and the image is *perfect.* We shall refer to $W(x, y)$ as the wave at a projected point (x, y) in the plane of the exit pupil. If (r, θ) represent the corresponding polar coordinates, they are related to the rectangular coordinates according to

$$(x, y) = r(\cos\theta,\ \sin\theta) \quad . \tag{1-2}$$

1.4 DEFOCUS ABERRATION

We now discuss defocus wave aberration of a system and relate it to its longitudinal defocus. Consider an imaging system for which the Gaussian image of a point object is located at P_1. As indicated in Figure 1-3, let the wavefront for this point object be spherical with a center of curvature at P_2 (due to field curvature discussed in Section 1.6) such that P_2 lies on the line OP, joining the center 0 of the exit pupil and the Gaussian image point P_1'. The aberration of the wavefront with respect to the Gaussian reference

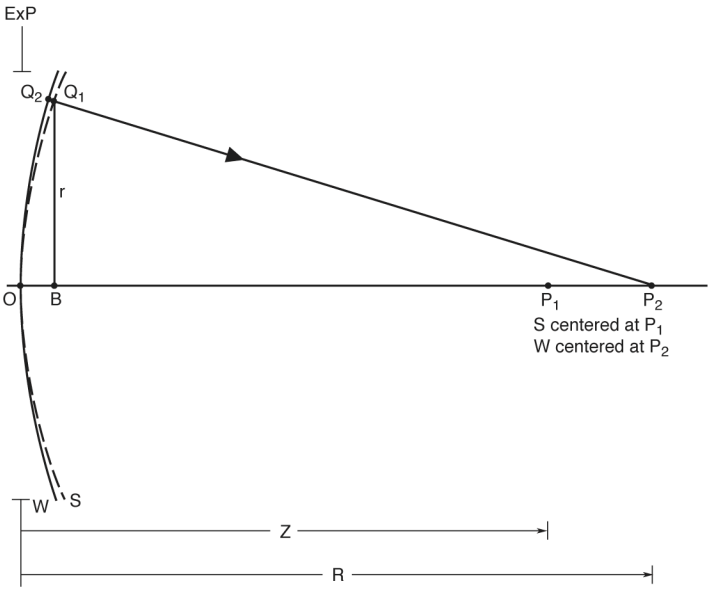

Figure 1-3. Wavefront defocus. Defocused wavefront *W* is spherical with a radius of curvature *R* centered at P_2. The reference sphere *S* with a radius of curvature *z* is centered at P_1. Both *W* and *S* pass through the center *O* of the exit pupil *ExP*. The ray Q_2P_2 is normal to the wavefront at Q_2. *OB* represents the sag of Q_1.

sphere is its optical deviation from it. This deviation is given by nQ_2Q_1, where *n* is the refractive index of the image space and Q_2Q_1, as indicated in the figure, is approximately equal to the difference in the sags of the reference sphere and the wavefront at a height *r*. (The *sag* of a surface at a certain point on it represents its deviation at that point along its axis of symmetry from a plane surface that is tangent to it at its vertex). Thus, the *defocus wave aberration* at a point Q_1 at a distance *r* from the optical axis is given by

$$W(r) = \frac{n}{2}\left(\frac{1}{z} - \frac{1}{R}\right)r^2 \quad , \tag{1-3a}$$

z and R are the radii of curvature of the reference sphere *S* and the spherical wavefront *W* centered at P_1 and P_2, respectively, passing through the center *O* of the exit pupil, and *r* is the distance of Q_1 from the optical axis. We note that the defocus wave aberration is proportional to r^2. If $z \simeq R$, then Eq. (1-3a) may be written

$$W(r) \simeq -\frac{n}{2}\frac{\Delta}{R^2}r^2 \quad , \tag{1-3b}$$

where $\Delta = z - R$ is called the *longitudinal defocus*. We note that the defocus wave aberration and the longitudinal defocus have numerically opposite signs. The ray aberrations corresponding to a defocus wave aberration are discussed in Chapter 7.

A defocus aberration is also introduced if the image is observed in a plane other than the Gaussian image plane. Consider, for example, an imaging system forming an