Chapter 6

Spherical Aberration

It was discussed in Chapter 4 and illustrated in Figure 4.5 that the field-independent astigmatic aberration comprises defocus and spherical aberration. This means that they are not functions of field angle and are constant aberrations over the entire field of view, and that collectively they are a function of the odd powers of the entrance pupil radius ρ for the transverse aberration form; that is,

$$\underbrace{\rho}_{\text{Defocus}} + \underbrace{\rho^3 + \rho^5 + \rho^7 + \ldots}_{\text{Spherical Aberration}}$$

In this chapter, we will consider both defocus and spherical aberration. In Section 4.3.1, the transverse defocus in an image plane located ξ from the paraxial image plane was shown to be expressed as

$$\begin{aligned} DF(\rho, \xi) &= -\xi \tan \nu'_a \\ &= -\frac{\rho D_{ent.pupil}}{2f}\xi \text{ (when the object is located at infinity)} \\ &= -\frac{\rho}{2\,f\text{-number}}\xi \end{aligned} \tag{6-1}$$

where ν'_a is the angle of the marginal paraxial ray in image space, ρ is the normalized entrance pupil radius, and f is the focal length. Figure 6.1 shows a typical ray plot of a fan of axial meridional rays. The ray intercepts with the defocused image plane and forms a line that is rotated from the abscises. Now, the marginal ray intercept $\varepsilon_y = (DF(1, \xi))$ times $2\,f$-number equals the defocus ξ.

The direct calculation of spherical aberration is a simple matter. A meridional ray is traced from object to image, passing through the desired zone of a lens, and the image distance L' is found. This is compared directly with the l' of a corresponding paraxial ray from the same object point. Then

$$\text{longitudinal spherical aberration} = LA' = L' - l'. \tag{6-2}$$

DOI: 10.1016/B978-0-12-374301-5.00010-3

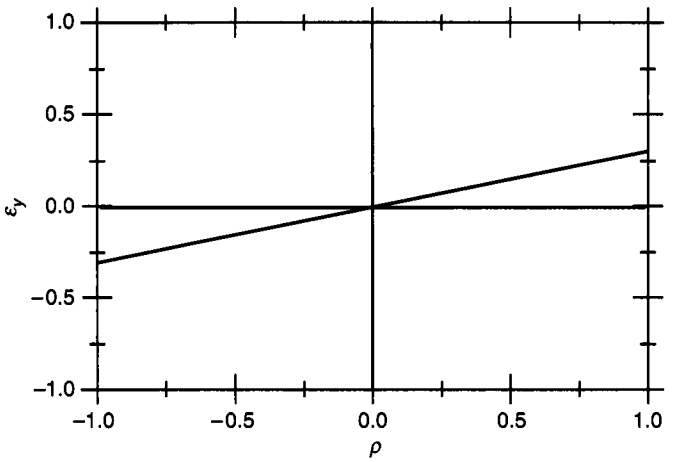

Figure 6.1 Transverse defocus error in the paraxial image plane.

Historically, longitudinal spherical aberration was used for several reasons. First, the intersection of the meridional ray with the optical axis directly provided the length L'. Second, the calculation was reasonably easy using hand computing methods. And finally, the longitudinal spherical aberration is inherently independent of defocus.[1]

Figure 5.3a illustrated a typical presentation of longitudinal spherical aberration for various colors of light, which was mentioned already, spherochromatism. In the following discussions, consider only the D line. The transverse spherical aberration for a desired zone of a lens is measured as the intersection height of a meridional ray with the paraxial image plane ($\varepsilon_y = -LA' \tan U'$); however, the actual image plane may be displaced from the paraxial image plane by ξ so that the ray intersection height comprises both spherical and defocus components. Figure 6.2 shows a ray plot when the image plane is inside the paraxial focus and for positive primary spherical aberration. Notice that the composite aberration (sum of defocus and spherical aberration) ray plot is rotated because of the presence of defocus.

As was explained in Chapter 4, defocus can influence the blur caused by the other astigmatic aberrations, but does nothing to the comatic aberrations. Consider now the problem of locating the position of the defocused image plane to achieve the minimum blur diameter in the presence of primary spherical aberration. Figure 6.3 shows the ray plot for third-order spherical aberration in the paraxial image plane. The aberration is expressed by $\varepsilon_y(\rho, 0, 0, 0) = 0.6\rho^3$ where $\sigma_1 = 0.6$. It is evident that the blur diameter is 1.2 (lens units). The central dashed line in Figure 6.3 represents the defocus aberration. If one then views

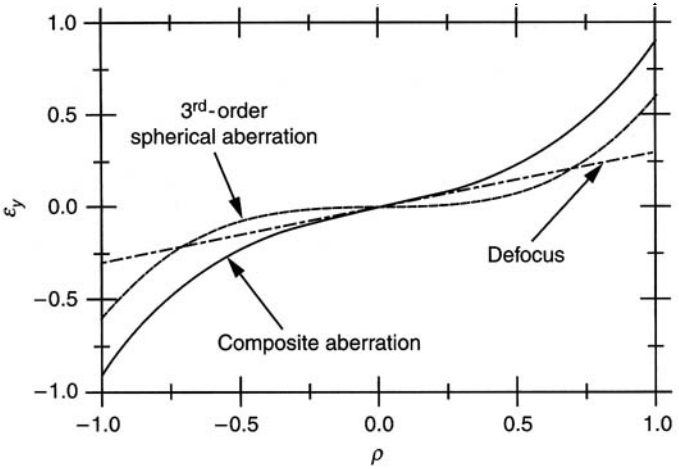

Figure 6.2 Transverse ray plot when image plane is inside the paraxial focus and for positive primary spherical aberration.

this line as the x-axis in the image plane, the primary spherical aberration curve is positioned to give the minimum blur diameter. The parallel lines bound the primary spherical aberration curve and it is easily shown that the minimum defocused blur diameter is $\frac{1}{4}$ the amount in the paraxial image plane, that is, $\frac{1}{2}\sigma_1$. Using Eq. (6-1), the amount of defocus needed is $\frac{3}{4}LA'_{\text{marginal}}$. In the case shown in Figure 6.3, the best focus image plane lies outside of the paraxial image plane.

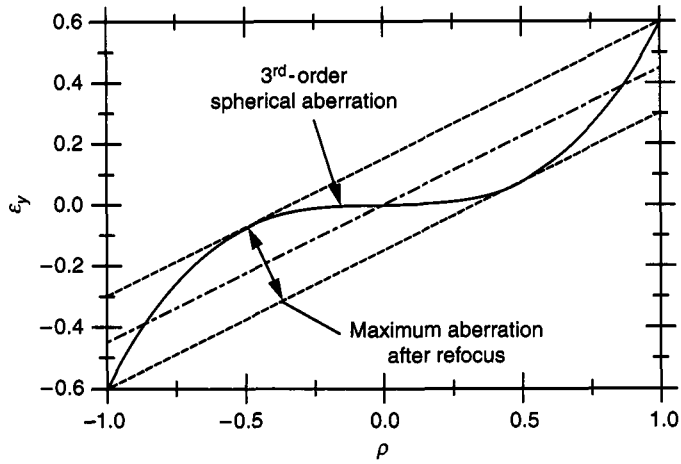

Figure 6.3 Ray plot for third-order spherical aberration in the paraxial image plane.

6.1 SURFACE CONTRIBUTION FORMULAS

The simple relationship given by Eq. (6-2) is often inadequate, both because it gives the aberration as a small difference between two large numbers, and also because it gives no clue as to where the aberration arises. It is therefore much more useful to compute the aberration as the sum of a series of surface contributions. A convenient formula has been given by Delano[2]; the derivation follows from Figure 6.4. Note that these surface contributions are for all orders of spherical aberration, not just for the primary term. In this diagram, entering marginal and paraxial rays are shown at a spherical surface. The length S is the perpendicular drawn from the paraxial object point P onto the marginal ray. The marginal ray is defined by its Q and U, the paraxial ray by its y and u. Then

$$S = Q - l \sin U, \quad \text{hence, } Su = Qu - y \sin U$$

We now replace u on the right by $i - yc$ and $\sin U$ by $\sin I - Qc$, where c is the surface curvature as usual. Multiplying through by n gives

$$Snu = Qni - yn \sin I$$

Doing the same thing for the refracted ray and subtracting plain from prime gives

$$S'n'u' - Snu = (Q' - Q)ni$$

We write this for every surface and add. After extensive cancellation because $(S'n'u')_1 = (Snu)_2$, we get for k surfaces

$$(S'n'u')_k - (Snu)_1 = \sum (Q' - Q)ni \tag{6-3}$$

Inspection of Figure 6.4 shows that

$$LA = -S/\sin U \quad \text{and} \quad LA' = -S'/\sin U'$$

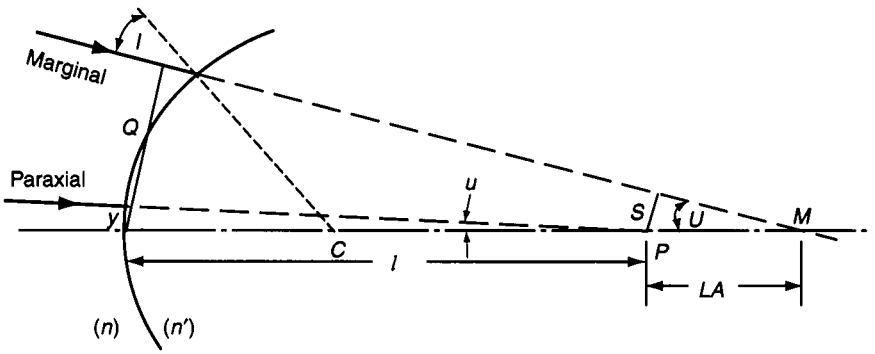

Figure 6.4 Spherical aberration contribution.

Hence,

$$LA' = LA\left(\frac{n_1 u_1 \sin U_1}{n'_k u'_k \sin U'_k}\right) - \sum \frac{(Q' - Q)ni}{n'_k u'_k \sin U'_k} \tag{6-4}$$

The quantity under the summation sign is the contribution of each surface to the spherical aberration of this particular ray (see Section 4.4), and the first term is the transfer of the object aberration across the lens to the image space. It may be thought of as the contribution of the object to the final aberration.

As an example of the use of this formula, we will take the lens used in Section 2.5. A marginal ray and a corresponding paraxial ray entering this lens from infinity at height 2.0 has been traced; additional data required for use of Delano's formulas are given in Table 6.1. It will be noted that the sum of the aberration contributions agrees closely with the $L' - l'$ aberration obtained directly from ray tracing:

$$L' = 11.29390$$
$$l' = 11.28586$$
$$LA' = L' - l' = 0.00804$$

However, the L' and l' values are good only to about 1 in the fifth decimal place, when using tables for manual ray tracing whereas the contributions are good to 1 in the seventh place. The contribution method is clearly the more precise of the pair.

Note, too, that the first and third surfaces of this lens contribute undercorrected aberration, the third giving twice as much as the first in spite of its flat

Table 6.1

Surface Contributions to Spherical Aberration

c	0.1353271	−0.1931098	−0.0616427	
d	1.05	0.40		
n	1.517	1.649		
		Paraxial ray data		
u	0 −0.0922401	−0.0554372	−0.1666664	
$yc + u = i$	0.2706542	−0.4597566	−0.1713855	
		Marginal ray data		
Q	2.0	1.9178334	1.9186619	
Q'	2.0171179	1.9398944	1.8814033	
$Q' - Q$	0.0171179	0.0220610	−0.0372586	
ni	0.2706542	−0.6974508	−0.2826147	
$-n'_k u'_k \sin U'_k$	−0.0277643	−0.0277643	−0.0277643	
Spherical contribution	−0.1668701	0.5541815	−0.3792578	$\sum = 0.0080536$

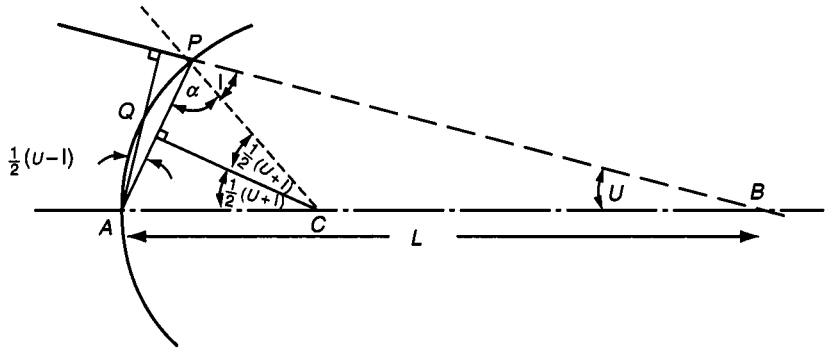

Figure 6.5 Diagram showing that $Q = PA \cos \frac{1}{2}(-U - I)$.

curvature; the second surface contributes more overcorrection than the total undercorrection of the two outer surfaces in spite of the small index difference between the media on each side of it.

An alternative representation of the contribution formula is sometimes useful. Its derivation depends on the relation between Q and the chord PA (Figure 6.5). In triangle APB we have

$$\frac{PA}{\sin U} = \frac{-L}{\sin(\alpha + I)}$$

Therefore,

$$PA = \frac{-L \sin U}{\sin(\alpha + I)} = \frac{-Q}{\sin(\alpha + I)}$$

However,

$$\alpha = 90° - \tfrac{1}{2}(I - U)$$

Therefore,

$$\alpha + I = 90° - \tfrac{1}{2}(I - U) \quad \text{and} \quad Q = PA \cos \tfrac{1}{2}(I - U)$$

Hence,

$$\begin{aligned}(Q - Q') &= PA[\cos \tfrac{1}{2}(I - U) - \cos \tfrac{1}{2}(I' - U')] \\ &= PA[-2 \sin(\tfrac{1}{2}\text{sum}) \sin(\tfrac{1}{2}\text{diff})] \\ &= 2PA \sin \tfrac{1}{2}(I' + U') \sin \tfrac{1}{2}(I' - I)\end{aligned}$$

The spherical aberration contribution formula can therefore be written

$$LA' = LA\left(\frac{n_1 u_1 \sin U_1}{n'_k u'_k \sin U'_k}\right) + \sum \frac{2PA \sin \frac{1}{2}(I' - I) \sin \frac{1}{2}(I' + U) ni}{n'_k u'_k \sin U'_k} \tag{6-5}$$

where

$$I' - I = U' - U, \text{ and } I' + U = I + U'$$

6.1.1 The Three Cases of Zero Aberration at a Surface

In Eq. (6-5), the quantity under the summation sign becomes zero in the following special cases:

(a) if $PA = 0$,
(b) if $I' = I$,
(c) if $i = 0$,
(d) if $I' = -U$

In case (a) the object and image are both at the vertex of the surface. In case (b) the marginal ray suffers no refraction at the surface; this could occur because the object is at the center of curvature of the surface, as also in case (c), but it could occur trivially if the refractive index were the same on both sides of the surface. Case (d) arises if $I' = -U$ or if $I = -U'$. This very important case must be considered further.

By Eq. (2-1) we see that in this case

$$\sin I = Qc + \sin U = \sin U - \left(\frac{L \sin U}{r}\right) = \left(1 - \frac{L}{r}\right) \sin U$$

But $\sin I = (n'/n) \sin I'$, and since, in this special case, $I' = -U$, we find that $(L/r) - 1 = n'/n$, where

$$L = r(n + n')/n$$

and similarly

$$L' = r(n + n')/n'$$

It can also be shown that, for this particular pair of conjugates,

$$Q = Q', \quad nL = n'L', \quad 1/L + 1/L' = 1/r$$

We can understand case (b) better with a numerical example. Consider the aplanatic hemispherical magnifier shown in Figure 6.6, which has a convex surface with air on the right and glass of index 1.5 on the left. We find that

$$L = 2.5r \text{ and } L' = 1.6667r$$

The image is free of all orders of spherical aberration, third-order coma, and axial color.

The aplanatic points B and B' are shown in Figure 6.7. All rays in the object space directed toward B will pass through B' after refraction, no matter at what angle they enter the surface. This pair of conjugates is known as the *aplanatic points* of the surface. Note that the distances of these points from the center of curvature of the surface are respectively equal to $B = r(n'/n)$ and

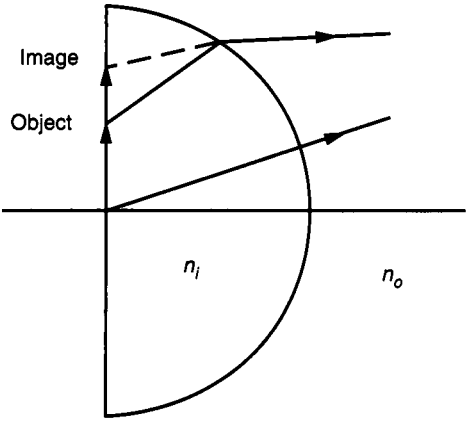

Figure 6.6 Aplanatic hemispherical magnifier with the object and image located at the center of curvature of the spherical surface. This type of magnifier has a magnification of n'/n and can be used as a contact magnifier or as an immersion lens.

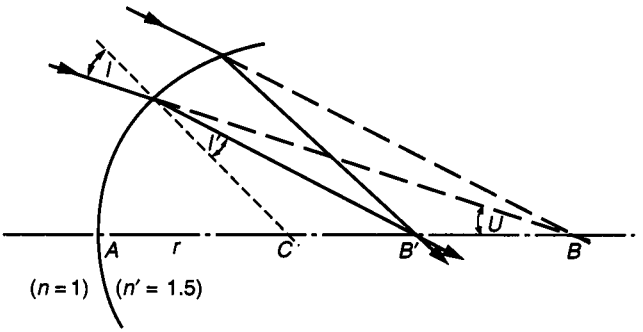

Figure 6.7 The aplanatic points of a surface.

$B' = r(n/n')$. Aplanatic surfaces of this type are used in many types of lens, particularly high-power microscope objectives and immersion lenses which make detectors appear larger.

A similar magnifier can be constructed by using a hyperhemispherical surface and a plano surface as depicted in Figure 6.8. The lateral magnification is $(n'/n)^2$. This lens, called an *Amici lens,* is based on the fourth aplanatic case. The image is free from all orders of spherical aberration, third-order coma, and third-order astigmatism. These magnifiers are often used as desktop magnifiers, having a magnification of about 2.5.

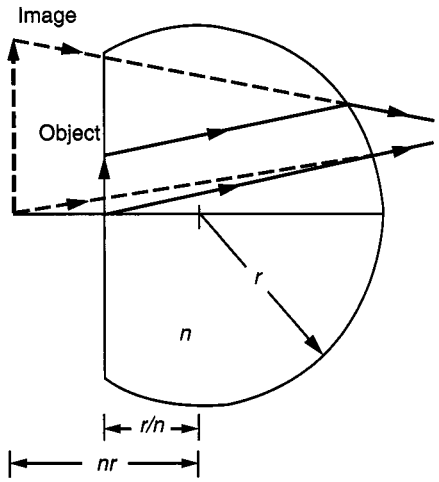

Figure 6.8 Aplanatic hyperhemispherical magnifier or Amici lens has the object located at an aplanatic point.

DESIGNER NOTE

It must be borne in mind that an aplanatic surface is capable only of increasing the convergence of converging light or increasing the divergence of diverging light. The greater the convergence or divergence, the greater will be the effect of an aplanatic surface. For parallel entering light, the aplanatic surface is a plane and produces no change in convergence.

6.1.2 An Aplanatic Single Element

It is possible to make an aplanatic single-element lens for use in a converging light beam by making the front face aplanatic and the rear face perpendicular to the marginal ray. Such a lens increases the convergence of a converging beam, which is useful in certain situations. In parallel light an aplanatic lens is merely a parallel plate. In a diverging beam an aplanatic lens element is a negative meniscus that increases the divergence of the beam without, of course, introducing any spherical aberration.

6.1.3 Effect of Object Distance on the Spherical Aberration Arising at a Surface

We have seen that the contribution of a single surface to the spherical aberration is zero if the (virtual) object is at A, C, or B, as in Figure 6.7. We may

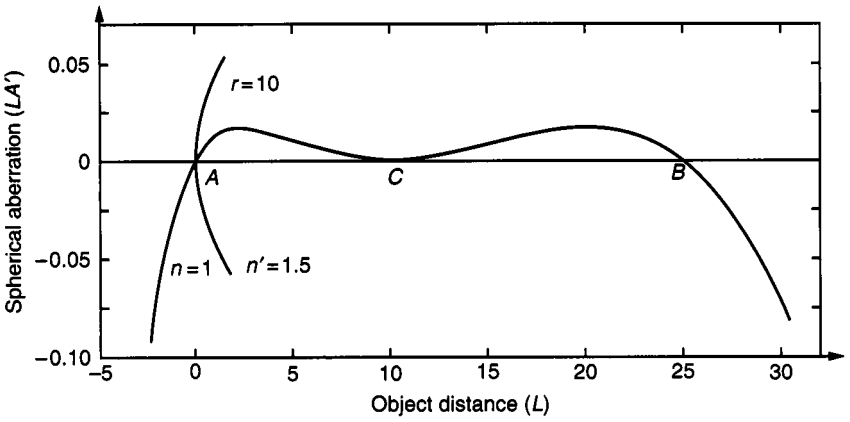

Figure 6.9 Effect of object distance on spherical aberration.

now inquire what will happen if the object lies in any of the regions between these points, the light entering from the left in all cases. As an example, consider the case of a surface of radius 10 with air on the left and glass of index 1.5 to the right. We will let a ray enter this surface at a fixed slope angle of 11.5°, and we calculate the spherical aberration in the image as the object moves along the axis. This is shown in Figure 6.9.

If the object lies between the surface A and the aplanatic point B, a collective surface such as we are considering here contributes *overcorrected* spherical aberration, which is decidedly unexpected and can be quite useful. The peak value of this overcorrection occurs, in our case, when the object distance is about twice the surface radius. As can be seen in Figure 6.9 the peak of overcorrection is close to the aplanatic point, and to achieve it we must use a surface somewhat flatter than the aplanatic radius. There is also a second but much less useful peak close to the surface itself, the value of L in this case being about 0.2 times the surface radius. As a general rule, using ordinary glasses, the maximum overcorrection will be obtained if r is set at a value about 1.2 times the aplanatic radius of $L(1 + n')$, assuming that there is air to the left of the surface as shown in Figure 6.9.

6.1.4 Effect of Lens Bending

One of the best methods for changing the spherical aberration of a lens is to *bend* it (see Section 3.4.5). If the lens is thin, changing all the surface curvatures by the same amount has the effect of changing the lens shape while leaving the

focal length and the chromatic aberration unchanged. Generally spherical aberration varies with bending in a parabolic fashion when plotted against some reasonable shape parameter such as c_1. At extreme bendings either to the left or right, a thin positive lens is decidedly undercorrected, and the aberration reaches a mathematical maximum at some intermediate bending. The aberration of a single thin lens with a distant object is never zero, but in a positive achromat the aberration exhibits a region of overcorrection at and close to the maximum. To bend a thick lens, it is customary to change all the surface curvatures except the last by a chosen value of Δc, the last radius being then solved by the ordinary angle solve procedure (Section 3.1.4) to maintain the paraxial focal length.

DESIGNER NOTE

This procedure, of course, slightly affects the chromatic aberration but it alters the spherical aberration far more. It should be noted, however, that if the aberration is at the maximum, then quite a significant bending will have little or no effect on the spherical aberration. When this condition exists, bending can be used as an effective design tool to vary other aberrations such as coma or field curvature while minimally impacting the spherical aberration.

6.1.5. A Single Lens Having Minimum Spherical Aberration

A single positive lens can be made to have minimum spherical aberration at one wavelength by taking a series of bendings, in each case solving for the last radius to hold focal length. When this is done, it is found that in the minimum-aberration lens each surface contributes about the same amount of aberration, with the front surface (in parallel light) contributing slightly more than the rear surface.

As an example, suppose we wish to design such a lens of focal length 10 and aperture *f*/4, using glass of index 1.523. A suitable thickness is 0.25. The two surface contributions become equal at $c_1 = 0.1648$, for which bending the total aberration is found to be –0.15893. A careful plot shows that the true minimum occurs at $c_1 = 0.1670$ with an aberration of –0.15883, but of course the difference between these two values of the aberration is utterly insignificant. We shall therefore make no noticeable error by aiming at equal contributions for the minimum bending. The error may become much greater, however, in lenses made from high-index materials for the infrared.

The results of ray tracing with $Y_1 = 1.25$ are as shown in the following:

c_1	0.15	0.16	0.17	0.18
Solved c_2	–0.041742	–0.031639	–0.021519	–0.011380
Spherical aberration contribution (1)	–0.05977	–0.07267	–0.08730	–0.10376
Spherical aberration contribution (2)	–0.10434	–0.08705	–0.07174	–0.05828
Total	–0.16411	–0.15972	–0.15904	–0.16205

6.1.6 A Two-Lens Minimum Aberration System

A considerable reduction in spherical aberration can be achieved by taking two identical lenses of twice the desired focal length and mounting them close together. In our case this procedure, after scaling to a focal length of 10, gave a spherical aberration of –0.0788, about half that of the original single element. However, a much greater improvement can be made by bending the second lens so that each of the four refracting surfaces contributes an identical amount of aberration. The required condition is that each surface should have the same value of $(Q' - Q)ni$, since it is this product that determines the aberration contribution of the surface. When computed manually, the curvature of each surface is determined by a few trials, and then, if the resulting focal length is not correct, the whole lens is scaled up or down until it is.

As an example, suppose we add another element to the single minimum-aberration lens of Section 6.1.5. Finding c_3 and c_4 by trial to make all four contributions equal gives the lens shown in Figure 6.10 that has the following prescription and spherical aberration contributions:

c	d	n	Spherical aberration contribution for $Y_1 = 1.25$
0.1648			–0.01703
	0.25	1.523	
–0.02678			–0.01702
	0.05	(air)	
0.3434			–0.01703
	0.25	1.523	
0.1216			–0.01700

The focal length is now 4.6155 and the aperture is *f*/1.85. Scaling the lens to a focal length of 10.0 and tracing a ray at $Y_1 = 1.25$ (*f*/4) gives the total spherical aberration as –0.0310, about one-fifth that of the single element. It is worth noting that the focal lengths of the two elements are now not equal, being respectively 21.7 and 18.4.

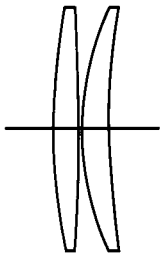

Figure 6.10 A two-lens minimum aberration system.

There is a common misconception regarding this type of two-element lens—namely, that to secure minimum spherical aberration, the marginal ray must be deviated equally at each of the four surfaces. To see how far this is from the truth, these are the surface ray deviations for the last example:

Surface	Angle U (deg)	Angle U' (deg)	Deviation $U' - U$ (deg)
1	0	1.877	1.877
2	1.877	3.318	1.441
3	3.318	6.019	2.701
4	6.019	7.195	1.176

The reason the third surface does so much refracting "work" without the introduction of excessive aberration is its close proximity to the aplanatic condition.

It should be noted that when designing a two-element infrared lens with a material having a refractive index higher than about 2.5, such as silicon or germanium, it will be found that if r_3 is chosen to give the maximum possible overcorrection, it may actually overcompensate the undercorrection of the front minimum-aberrations lens, making it possible to correct the spherical aberration completely. The last radius is then chosen to have its center of curvature at the final image to eliminate any aberration there.

As an example, we will design an *f*/1 lens made of silicon having a refractive index of 3.4. Following the suggested procedure, we come up with the prescription that follows Figure 6.11. This figure shows a longitudinal section of this lens. The strong rear element is highly meniscus, as can be seen. High-index materials such as silicon and germanium appear to behave quite oddly to anyone only familiar with the properties of ordinary glass lenses.

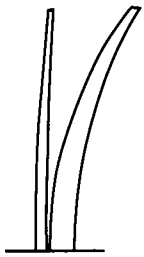

Figure 6.11 An *f*/1 silicon lens.

c	*d*	*n*	Spherical aberration contribution	
0.02790			−0.006017	$f' = 10.283$
	0.25	3.4		
0.01572			−0.006004	$l' = 9.717$
	0.05	(air)		
0.12632			+0.012009	aperture = 10 (*f*/1)
	0.50	3.4		
0.10291			0	

Focal length of front component, 33.99; of rear component, 14.88

6.1.7 A Four-Lens Monochromat Objective

As was stated in Section 6.1.2, a single aplanatic lens element for use in parallel light is nothing but a planoparallel plate and not a lens at all. However, by making use of the small overcorrection that can be obtained from a convex surface slightly weaker than a true aplanat, it is possible to construct an aplanatic system for use with a distant object by placing a minimum-aberration lens first, and following this by a series of overcorrected menisci in the converging beam produced by the first lens.

As an example we may take the single minimum-aberration *f*/4 lens in Section 6.1.5, and follow it with three menisci, the front face of each being chosen to give the maximum of overcorrection, while the rear faces are perpendicular to the marginal ray. Nothing is gained by departing from the strict perpendicular condition for the rear surfaces of the menisci because, being dispersive surfaces, any departure from perpendicularity in either direction would yield spherical undercorrection, which is just what we are trying to avoid.

After several trials to obtain the greatest possible amount of overcorrection, and finally scaling to $f' = 10.0$ with an aperture of *f*/2, we obtain the following system:

c	d	n	Spherical aberration contribution at $f/2$	
0.066014[a]			−0.020622	−0.041232
	0.3	1.523		
−0.0103636[a]			−0.020610	
	0.05	(air)		
0.082192			+0.002463	
	0.3	1.523		
0.055672			0	
	0.05	(air)		
0.113932			+0.005962	
	0.3	1.523		
0.077543			0	
	0.05	(air)		
0.158867			+0.014476	
	0.3	1.523		
0.109134			0	
		Total	−0.018331	

[a] *Crossed lens* in parallel light (see Section 6.3.2).

The focal length of the first lens alone is now 24.969. It is clear that even with three menisci it is not possible to compensate for the undercorrection of the first lens.

However, we can do much better by starting with the two-lens minimum aberration form given in Section 6.1.6, and following this with only two menisci. By this procedure we can design a four-lens spherically corrected system for use in parallel light with an aperture as high as $f/2$. Scaled to $f' = 10$ this becomes

c	d	n	Spherical aberration contribution at $f/2$
0.041520			−0.005090
	0.3	1.523	
−0.006726			−0.005098
	0.05		
0.084883			−0.005106
	0.3	1.523	
0.029164			−0.005098
	0.05		
0.113764			+0.005966
	0.3	1.523	
0.077891			0
	0.05		
0.159353			+0.014387
	0.3	1.523	
0.109941			−0.000016
		Total	−0.000068

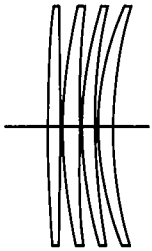

Figure 6.12 A four-lens *f*/2 aplanatic objective.

This lens is shown in Figure 6.12. The focal length of the first two lenses is now 18.380. This system has been used in monochromat microscope objectives made of quartz for use at a single wavelength in the ultraviolet. The design has been discussed by Fulcher.[3]

6.1.8 An Aspheric Planoconvex Lens Free from Spherical Aberration

Two cases arise, the first when the curved aspheric surface faces the distant object, and the other when the plane surface faces the object. In 1637, Descartes described and explained the general properties of utilizing concave and convex lenses, both singularly and in combination. He was the first to create a mathematical formulation to explain spherical aberration. Descartes also made a detailed study of elliptical and hyperbolic surfaces, particularly the plano-hyperbolic lens. Descartes and his colleagues spent substantial money and effort trying to fabricate such a lens since it would be free of spherical aberration.

For all their efforts, not one lens could be fabricated with the tools and methods then available. Fortunately, technology has advanced to allow fabrication of both elliptical and hyperbolic surfaces of high quality. The topic of perfect imaging from one point to another point was first addressed by utilizing Fermat's principle, which states that perfect conjugate imaging occurs when all the rays passing through the conjugate points have the same optical path length. Using Fermat's principle, Luneburg[4] showed in 1944 that the surface can be represented by a fourth-order curve, which is also known as the Cartesian Oval after René Descartes.

Convex to the Front

The left side of the ellipse shown in Figure 6.13a is the portion of the ellipse used as the surface contour for the ellipsoid-plano lens with the image being formed at the rear surface of the lens, which is planar. This lens has no spherical

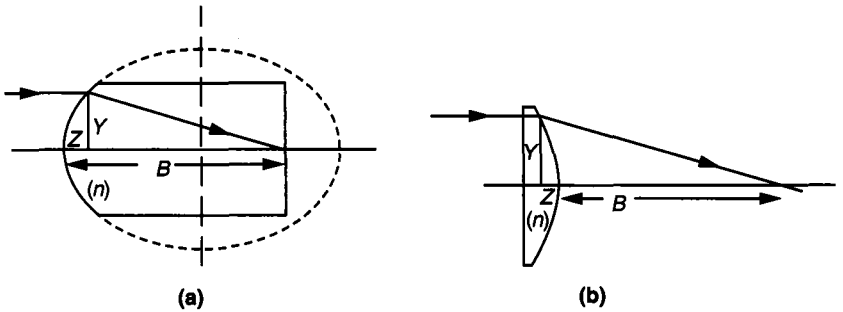

Figure 6.13 Aspheric single lenses corrected for spherical aberration.

aberration at the design wavelength for collimated light input and does suffer from spherochromatism. The coordinate system used for ray tracing has the coordinate system origin located at the left surface vertex of the ellipsoid. The z-axis is the major axis of ellipsoid and the x-axis and y-axis define the vertex tangent plane. The sag or z-coordinate displacement from the tangent plane at the vertex of the surface can be determined from a commonly used mathematical representation of conical surfaces given by

$$z = \frac{y^2 c}{1 + \sqrt{1 - (1 + \kappa)c^2 y^2}}$$

where y is the coordinate of the ray intercept on the surface, c is the radius of curvature (reciprocal of the radius) at the surface vertex, and κ is the conic constant (see Eq. (2-7)). This equation form is typically used in optical design programs with c (or r) and κ as an input description of the surface.

Figure 6.14 presents the basic parameters describing an ellipse.[5] These include the lengths of its major and minor axes, and foci. The distance d can

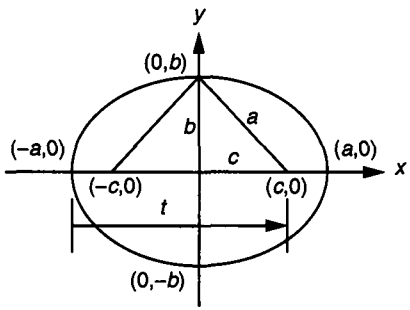

Figure 6.14 Geometrical parameters of an ellipse.

be considered the focal length of the elliptical lens and is the sum of the major axis semi-length a and the foci distance c from the ellipse center.

A conic section can also be represented by

$$\frac{\left(x - d\dfrac{n_1}{n_0 + n_1}\right)^2}{d^2\left(\dfrac{n_1}{n_0 + n_1}\right)^2} + \frac{z^2}{d^2\dfrac{n_1 - n_0}{n_0 + n_1}} = 1$$

and is an ellipse if $n_1 > n_0$ and a hyperbola if $n_0 > n_1$. This surface is called a rotationally symmetric surface. Examination of this equation indicates that the major a and minor b axes' semi-lengths can be computed from

$$a^2 = d^2\left(\frac{n_1}{n_0 + n_1}\right)^2$$

and

$$b^2 = d^2\frac{n_1 - n_0}{n_0 + n_1}.$$

By basic geometry, the foci distances are computed by

$$c^2 = a^2 - b^2$$

and by substituting in the terms for a^2 and b^2, we obtain

$$c^2 = d^2\left(\frac{n_0}{n_0 + n_1}\right)^2.$$

From geometry, the eccentricity $\varepsilon = \frac{c}{a}$ which is determined by using the preceding equations to be $\varepsilon = \frac{n_0}{n_1}$. The conic constant κ is defined as

$$\kappa = -\varepsilon^2 = \frac{-n_0^2}{n_1^2}.$$

Again from geometry (semi-latus rectum), the vertex radius r is given by

$$r = \pm\frac{b^2}{a} = d\left(\frac{n_1 - n_0}{n_1}\right).$$

It should be evident that the same relationship between r, t, n_0, and n_1 can be determined using paraxial optics and a purely spherical surface. This result

should be expected when the aperture is very small. The inclusion of the conic constant to transform the surface into an ellipsoid does not change the first-order properties of this lens, but does mitigate the inherent spherical aberration of a spherical surface.

Consider an example where $n_0 = 1$ and $n_1 = 1.5$, and $d = 20$ mm. It follows that

$$\kappa = \frac{-n_0^2}{n_1^2} = \frac{-1}{1.5^2} = -0.44444$$

$$r = d\left(\frac{n_1 - n_0}{n_1}\right) = 20\left(\frac{0.5}{1.5}\right) = 6.66666$$

A surface of this kind has long been used on highway reflector "buttons."[6] The same surface profile can be used to form a cylindrical lens, which has various applications. An array of such lenses can be created to form a lenticular array commonly used in the printing industry to make prints providing 3D photographic projections or the display of different images as the print is tilted.[7]

Consider a situation when an ellipsoid-plano lens is bonded to one or more materials. An example would be a detector array affixed to the plano side of the lens using an optical glue which can have different refractive index from the lens. The lens thickness would need to be reduced to account for the thickness of the glue. From basic aberration theory, the different refractive indices of the lens and substrate materials can introduce additional spherical aberration since the light beam is converging (see Section 6.4).

To compensate for the different refractive indices, the equation relating r, n, and d can be modified as presented in the following equation to determine the radius r of the lens when the composite optical element comprises two or more different materials.

$$r = (n_1 - 1)\left(d_1/n_1 + d_2/n_2 + \ldots d_n/n_n\right)$$

This equation is readily derived using paraxial ray tracing. Note that the substrates are each assumed to be planar. The summation shown in the parentheses is the effective optical thickness of the assemblage. If the substrate thicknesses are small compared to the lens thickness d_1, then the conic constant can be estimated to be given by $-n_1^{-2}$.

In the event that the substrate thicknesses are a significant fraction of the lens thickness, the conic constant estimation is somewhat more complex in that an estimate of the effective refractive index is required. Consider that a paraxial ray having $n_0 u_0 = 0$ is incident on the refracting surface at a height y_1 where u_0 is the angle the ray makes with the optical axis. As this ray propagates through the assemblage, it intercepts the substrates at heights $y_2, y_3, \ldots, y_n$.

The effective refractive index n_{eff} can then be expressed, following the mean value theorem of integral calculus, as

$$n_{\text{eff}} = \frac{\sum_{i=1}^{n} y_i t_i n_i}{\sum_{i=1}^{n} y_i t_i}$$

and the corresponding conic constant is $-n_{\text{eff}}^{-2}$.

Figure 6.15 illustrates a lens system comprising a spherical refracting surface and two substrates while Figure 6.16 shows the same lens system with the effective conic constant. The prescriptions of these lens systems are provided in Table 6.2 where the only difference between them is the conic constant. As it is evident by examining Figure 6.15, the lens system with spherical refracting surface suffers from significant spherical aberration. The inclusion of the effective conic constant effectively mitigates the spherical aberration illustrated in Figure 6.16.

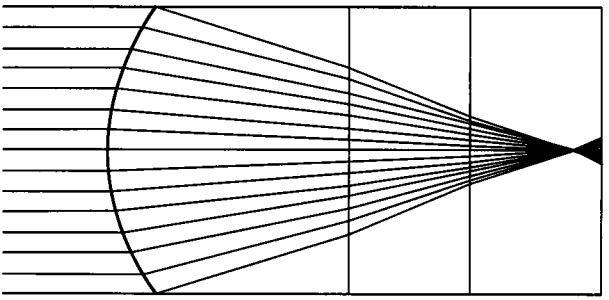

Figure 6.15 Lens system with multiple substrates and spherical refractive surface.

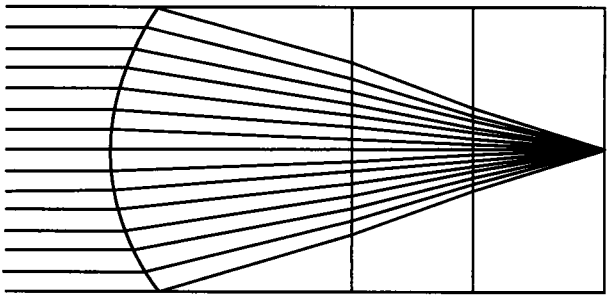

Figure 6.16 Lens system with multiple substrates and elliptical refractive surface.

Table 6.2

Prescription of Lens System Depicted in Figures 6.15 and 6.16

Surface	Radius	Thickness	Refractive index	Marginal ray height	Conic constant
Object	∞	∞	1		
1	10	10	1.8	6	–0.34542
2	∞	5	1.4	3.33333	
3	∞	5.396825	1.6	1.61905	
Image	∞			0	

A solid-optics (containing no air gaps) lens system for making an afocal telescope can be designed by using the foregoing information. Two ellipsoid-plano lenses would be placed with their plano sides facing one another and some optical bonding material of finite thickness placed between them after having accounted for the several thicknesses so that the focal point of the two lenses coincide. The lenses can be of different materials and radii, and the angular magnification is simply the ratio of their effective focal lengths.

Plane Surface in Front

Equating optical paths in the air behind the lens shown in Figure 6.13b gives

$$B + nZ = [Y^2 + (B + Z)^2]^{1/2}$$

where

$$\frac{\{Z + [B/(n+1)]\}^2}{[Bn/(n+1)]^2} - \frac{Y^2}{B^2(n-1)/(n+1)} = 1$$

There is a clear resemblance between these two cases. The plane-in-front lens has a hyperbolic surface with semimajor axis equal to $B/(n + 1)$, and semiminor axis equal to $B[(n - 1)/(n + 1)]^{1/2}$ as before (Figure 6.13b), the eccentricity now being equal to the refractive index n and the conic constant being $\kappa = -n^2$.

Using a *y-nu* ray trace, it is trivial to show that the focal length of a plano-hyperboloid lens is $\frac{-r}{n-1}$ which is shown as B in Figure 6.13b. To create a solid optical element for use at finite magnification, a hyperbolic surface can be applied on both faces of a biconvex lens with the ratio of focal lengths being the ratio of the radii. The light is of course collimated between the faces so there is no fundamental restriction of the lens thickness. The axial image is free of spherical aberration for up to as high an aperture as required; however, the field of this lens is restricted by coma.

6.2 ZONAL SPHERICAL ABERRATION

As we have seen, it is possible by the use of opposing positive and negative elements to design a lens such that the focus of the marginal ray coincides with the paraxial image point. We say that this lens has zero spherical aberration. However, it generally happens that the foci of rays passing through the intermediate zones of the lens fall closer to the lens than the paraxial image-point, and occasionally but rarely fall further from it. Thus we can plot a graph connecting entrance height Y with the spherical aberration, as shown in Figure 6.17. This zonal residual is generally known as *zonal aberration.* It can be expressed as a power series containing only even powers of Y, as

$$LA' = aY^2 + bY^4 + cY^6 + \ldots$$

The successive terms of this series have been called primary, secondary, tertiary ... aberration, but of course they have no separate existence, and the actual aberration of the lens is the sum of all these terms. However, we can plot them separately to see how they vary (Figure 6.17). If Y is small, the secondary and higher terms are very small or negligible, and the primary term represents

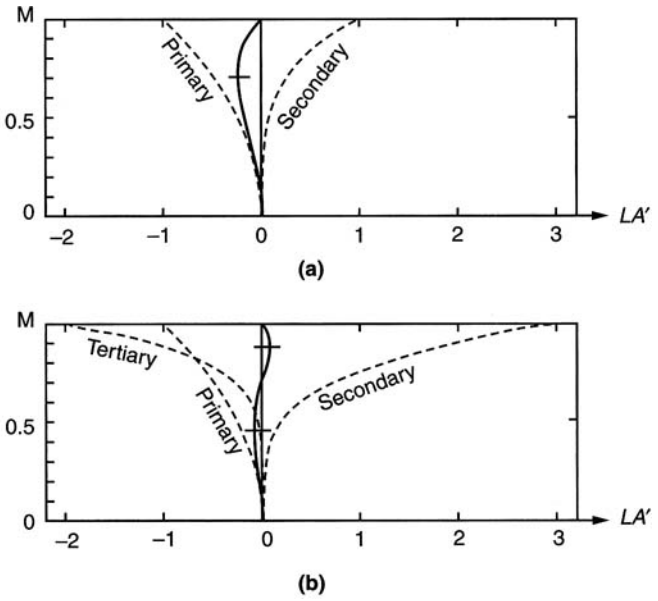

Figure 6.17 Interaction of various orders of spherical aberration: (a) Primary and secondary only; (b) primary, secondary, and tertiary.

the whole aberration. Then, at increasing values of Y, first the secondary and then the tertiary terms begin to increase and finally dominate the situation.

In the example shown in Figure 6.17a, the primary term is negative and the secondary term is positive, and they have equal and opposite values for the marginal ray. Consider now when only the first two terms are present and the marginal spherical aberration is zero, that is, $LA' = a\rho^2 + b\rho^4 = 0$ when $\rho = 1$. This implies that $a = -b$. Now the peak residual can be found by solving $\frac{dLA'}{d\rho} = 2a\rho + 4b\rho^3 = 0$. Now dividing by ρ and substituting $a = -b$, it follows that $2a + 4b\rho^2 = -2b + 4b\rho^2$ which yields that $\rho = 1/\sqrt{2} = 0.707$. In an actual lens system, the peak zonal residual occurs when ρ is equal to the marginal ρ_m multiplied by 0.7071. The magnitude of the zonal residual, in the case of suffering only third- and fifth-order spherical aberration, equals one-quarter of the primary term at the marginal zone of the lens, that is, $LA'_{0.707} = a/4$.

Because tertiary aberration is not greatly different from secondary, it may be positive and add to the secondary; in this case the maximum zonal residual falls higher than the 0.7 zone, and the marginal aberration increases very rapidly. On the other hand, if the tertiary aberration is negative, it tends to oppose the secondary, and it is then possible to eliminate both the marginal and the zonal aberrations, as indicated in Figure 6.17b. It will be noticed that the secondary and tertiary aberrations are now much larger than in the simple case of Figure 6.17a, but the resulting aberration curve is nearly flat, having small equal and opposite residuals above and below the 0.7 zone. An analysis of the situation reveals that the maximum and minimum residuals fall at values of ρ given by

$$\frac{\rho}{\rho_m} = \sqrt{\frac{1 \pm 1/\sqrt{3}}{2}} = 0.8881 \text{ and } 0.4597$$

The locations of the maximum residuals are indicated by short horizontal lines on these diagrams.

DESIGNER NOTE

As a consequence of the nature of the expansion of spherical aberration being in even orders (for longitudinal form), it is virtually always true that the signs of the coefficients a, b, and c must alternate to achieve correction for the marginal ray. This is seen in Figure 6.17a for primary and secondary aberration only and Figure 6.17b when the tertiary aberration is present. Examination of the ray plot for a lens can tell the lens designer what orders of spherical aberration are present and if they have the correct signs to achieve correction.

The effect of refocusing when only primary spherical aberration is present was shown in Figure 6.3. Consider now the effect of refocusing when both

primary and secondary spherical aberration are present, which will be seen to be more complicated. When only primary spherical aberration is present, it is fairly evident what the optimum refocus should be. With the presence of both primary and secondary aberration, the optimum refocus is not so simple.

Consider Figure 6.18 that shows the transverse ray errors versus the normalized entrance pupil radius for the case where the marginal spherical aberration is zero when $\rho = 1$ (curve A) and the case where the marginal spherical aberration is zero when $\rho = 1.12$ (curve B). The first case represents what lens designers often attempt to achieve, that is, having the marginal-ray error equal zero in the paraxial image plane $(\varepsilon_y = \sigma_1\rho^3 + \mu_1\rho^5 = 0$ when $\rho = 1)$. The refocus boundaries are shown and represent the blur diameter that contains 100% of the energy. If a different refocus is used (slope of the boundary lines indicates the amount of refocus) in this case, a brighter core can be obtained. This is illustrated in Figure 6.18. Since these boundary lines intersect curve A at about $\rho = \pm 0.9$, this bright core region contains about 80% of the energy and would provide an improvement in resolution.

The remaining 20% of the energy would be spread around this bright core to form a dim flare having a diameter about five times larger than the core. Now consider the second case which is illustrated as curve B in Figure 6.18. It can be shown that if $\varepsilon_y = \sigma_1\rho^3 + \mu_1\rho^5 = 0$ when $\rho = 1.12$, then the smallest 100% blur diameter is obtained when refocused and it is about 50% larger than the bright core for

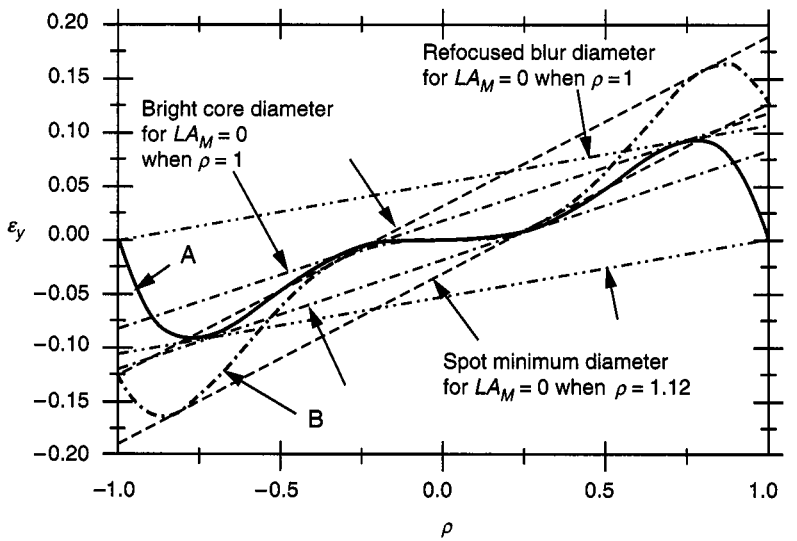

Figure 6.18 Geometric blur for third- and fifth-order spherical aberration for zero marginal spherical aberration when (curve A) $\rho = 1$ and (curve B) $\rho = 1.12$.

curve A. Which of the preceding cases and amount of refocus are best for a given application must be determined by the lens designer. It is important to note that achieving a marginal-ray error equal to zero in the paraxial image plane is not always appropriate.

6.3 PRIMARY SPHERICAL ABERRATION

6.3.1 At a Single Surface

To isolate the primary term, we would have to make Y of infinitesimal magnitude, and then we cannot use the formula in Eq. (6-4) to compute the aberration for the same reason that we cannot trace a paraxial ray by the ordinary ray-tracing formulas. However, the primary term can be determined as a limit:

$$LA'_{\text{primary}} = \lim_{y \to 0}(LA'_y)$$

To find this limit, we use paraxial ray data to fill in the numbers in the accurate version of Eq. (6-5). Making this substitution gives the primary aberration equation:

$$LA'_p = LA_p\left(\frac{n_1 u_1^2}{n'_k u'^2_k}\right) + \sum \frac{2y \cdot \frac{1}{2}(i' - i) \cdot \frac{1}{2}(i' + u)ni}{n'_k u'^2_k}$$

Here LA_p is the primary aberration of the object, if any; it is transferred to the final image by the ordinary longitudinal magnification rule. The quantity under the summation sign is the primary aberration arising at each surface.

These *surface contributions* (SC) can be written

$$SC = yni(u' - u)(i + u')/2n'_k u'^2_k \tag{6-6}$$

Only paraxial ray data are required to evaluate this formula. To interpret it, we note that for pure primary aberration,

$$LA'_p = aY^2$$

and that the radius of curvature of the spherical-aberration graphs in Figure 6.17, at the point where the graph crosses the axis, has the value

$$\rho = Y^2/2LA'_p = 1/2a$$

Therefore, the coefficient of primary aberration a is an inverse measure of twice the radius of curvature of the spherical-aberration graph at the point where it crosses the lens axis. Hence, by tracing one paraxial ray, we not only discover the location of the image point, but we also ascertain the shape of the aberration curve as it crosses the axis at that point. It is remarkable how much information can be obtained from so very little ray-tracing effort.

Table 6.3

Calculation of Primary Spherical Aberration

y	2	1.9031479	1.8809730		
n	1	1.517	1.649		
$yc + u = i$	0.2706542	−0.4597566	−0.1713855		
u	0	−0.0922401	−0.0554372		−0.1666664
y	2	1.9031479	1.8809730		
ni	0.2706542	−0.6974508	−0.2826147		
$u' - u$	−0.0922401	0.0368029	−0.1112292		
$i + u'$	0.1784141	−0.5151938	−0.3380519		
$1/2u_k'^2$	18	18	18		
Product = SC	−0.160349	0.453014	−0.359792	Σ =	−0.067127

As an example of the use of this formula, we will calculate the primary spherical aberration contributions of the three surfaces of the cemented doublet shown in Section 2.5 that we have already used several times (see Table 6.3). It is interesting to compare these primary aberration contributions with the exact contributions given in Section 6.1. The contributions are as follows:

Surface	1	2	3	Sum
Exact contribution	−0.16687	+0.55418	−0.37926	0.00805
Primary contribution	−0.16035	+0.45301	−0.35979	−0.06713
Difference (contribution of higher orders)	−0.00652	+0.10117	−0.01947	

At each surface the true and primary contributions are similar in magnitude and have the same sign, but the cemented interface shows the greatest difference. This is due to the presence of a significant amount of secondary and higher-order aberrations there, while the outer surfaces show very little sign of higher-order aberrations. It is the presence of the considerable amount of higher-order aberrations at the cemented interface that is the cause of the large zonal aberration of this lens. Examination of curve *D* in Figure 5.3a indicates that the spherical aberration is comprised of almost totally primary and secondary contributions.

6.3.2 Primary Spherical Aberration of a Thin Lens

By combining the *SC* values for the two surfaces of a thin lens element, we find that a thin lens, or a thin group of lenses in close contact, within a system

contributes the following amount to the primary spherical aberration at the final image:

$$SC = -\frac{y^4}{n_0' u_0'^2} \sum (G_1 c^3 - G_2 c^2 c_1 + G_3 c^2 v_1 + G_4 c c_1^2 - G_5 c c_1 v_1 + G_6 c v_1^2) \quad (6\text{-}7)$$

where the terms with suffix 0 refer to the final image, the other terms applying to each single element. Here c and c_1 have their usual meanings, namely, $c_1 = 1/r_1$ and $c = 1/f'(n-1)$. The symbol v_1 is the reciprocal of the object distance of the element, and the Gs are functions of the refractive index, namely,

$$G_1 = \tfrac{1}{2}n^2(n-1), \qquad G_2 = \tfrac{1}{2}(2n+1)(n-1),$$

$$G_3 = \tfrac{1}{2}(3n+1)(n-1), \qquad G_4 = (n+2)(n-1)/n,$$

$$G_5 = 2(n^2-1)/n, \qquad G_6 = \tfrac{1}{2}(3n+2)(n-1)/n$$

The details of the derivation of this formula have been given by Conrady.[8] The summation sign in Eq. (6-7) is used only if the thin lens contains more than one element, for example, if it is a thin doublet or triplet; otherwise it may be omitted. If there is more than one element we must assume a very thin layer of air to exist between the elements in place of cement, c_1 being the curvature of the first surface of each element and v_1 being the reciprocal of the object distance in air. Thus for the second lens of a cemented doublet we take

$$(c_1)_b = (c_1)_a - c_a \quad \text{and} \quad (v_1)_b = (v_1)_a + c_a(n_a - 1)$$

In the case of an isolated thin element or thin system in air, not forming part of a more complex system, $n_0' = 1$ and $u_0' = y/l'$. Also the aberration of the object (if any) must be transferred to the image and added to the new aberration arising at the lens. Thus in such a case we have

$$LA_p' = LA_p\left(\frac{l'}{l}\right)^2 - y^2 l'^2 \sum (G \text{ sum})$$

the (G sum) referring to the six-term expression in parentheses in Eq. (6-7).

By use of the G-sum formula we can plot a graph showing how the primary spherical aberration of a thin lens varies with the bending (Figure 6.19). For a single thin positive element, this graph is a vertical parabola, the vertex of which nearly but not quite reaches the zero aberration line.

The thin single lens having the minimum primary spherical aberration is called a *crossed lens*. Its shape can be found by differentiating the G-sum expression with respect to c_1 yielding

$$c_1 = \frac{\frac{1}{2}n(2n+1)c + 2(n+1)v_1}{n+2}$$

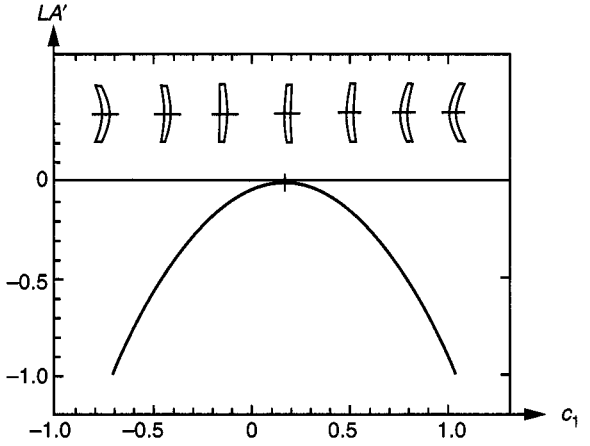

Figure 6.19 Effect of bending on spherical aberration.

For the special case of a very-distant object, $v_1 = 0$, we find that $c_1/c = n(2n + 1)/2(n + 2)$ and $c_2/c_1 = (2n^2 - n - 4)/n(2n + 1) = r_1/r_2$. For glass having an index of 1.6861, the crossed lens is exactly planoconvex; however, for other glass indices, the departure from the planconvex form is slight. The very high-refractive indices of infrared materials cause the crossed lens to be a deeply curved thin meniscus.

It is well-known that the spherical aberration of a lens is a function of its shape factor or bending. Although several definitions for the shape factor have been suggested (see Section 3.4.5), a useful formulation is

$$\chi = \frac{c_1}{c_1 - c_2} \tag{6-8}$$

where c_1 and c_2 are the curvatures of the lens, with the first surface facing the object. By adjusting the lens bending, the spherical aberration can be seen to have a minimum value.

The power of a thin lens or the reciprocal of its focal length is given by

$$\phi = \frac{(n-1)c_1}{\chi} \tag{6-9}$$

When the object is located at infinity, the shape factor for minimum spherical aberration can be represented by

$$\chi = \frac{n(2n+1)}{2(n+2)} \tag{6-10}$$

and

$$\frac{c_2}{c_1} = \frac{2n^2 - n - 4}{n(2n+1)}.$$

The resultant third-order spherical aberration of the marginal ray in angular units is

$$SA3 = \frac{n^2 - (2n+1)k + (1+2/n)\chi^2}{16(n-1)^2(f\text{-number})^3} \tag{6-11}$$

or after some algebraic manipulations,

$$SA3 = \frac{n(4n-1)}{64(n+2)(n-1)^2(f\text{-number})^3} \tag{6-12}$$

When the object is located at a finite distance s_0, the equations for the shape factor and residual spherical aberration are more complex. Recalling that the magnification m is the ratio of the object distance to the image distance and that the object distance is negative if the object lies to the left of the lens, the relationship between the object distance and the magnification is

$$\frac{1}{s_0\phi} = \frac{m}{1-m} \tag{6-13}$$

where m is negative if the object distance and the lens power have opposite signs. The term $1/s_0\phi$ represents the reduced or ϕ-normalized reciprocal object distance v_1; that is, s_0 is measured in units of focal length, ϕ^{-1}. The shape factor for minimum spherical aberration is given by

$$\chi = \frac{n(2n+1)}{2(n+2)} + \frac{2(n^2-1)}{n+2}\left(\frac{m}{1-m}\right) \tag{6-14}$$

and the resultant third-order spherical aberration of the marginal ray in angular units is

$$SA3 = \frac{1}{16(n-1)^2(f\text{-number})^3}\left[n^2 - (2n+1)\chi + \frac{n+2}{n}\chi^2 + (3n+1)(n-1)\left(\frac{m}{1-m}\right) - \frac{4(n^2-1)}{n}\left(\frac{m}{1-m}\right)\chi + \frac{(3n+2)(n-1)^2}{n}\left(\frac{m}{1-m}\right)^2\right] \tag{6-15}$$

and

$$c_1 = \frac{\frac{1}{2}n(2n+1)c + 2(n+1)v_1}{n+2}. \tag{6-16}$$

When the object is located at infinity, the magnification becomes zero and the above equations reduce to those previously given.

Figure 6.20 illustrates the variation in shape factor as a function of reciprocal object distance for refractive indices of 1.5 to 4 for an f-number = 1.[9] Notice that lenses have a shape factor of 0.5 regardless of the refractive index when

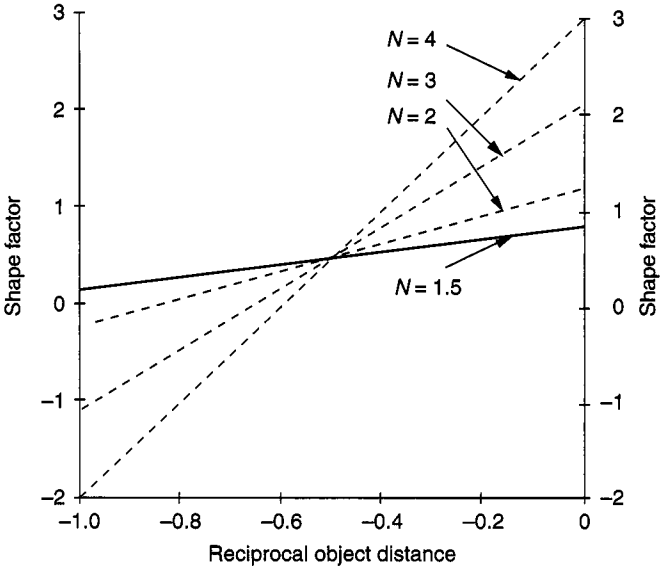

Figure 6.20 The shape factor for a single lens is shown for several refractive indexes as a function of reciprocal object distance v_1 where the distance is measured in units of focal length.

the magnification is –1 or $v_1 = -0.5$. For this shape factor, all lenses have biconvex surfaces with equal radii. When the object is at infinity, a lens having a refractive index of 1.5 has a somewhat biconvex shape with the second surface having a radius about six times greater than the first surface radius.

Since the minimum-spherical lens shape is selected for a specific magnification, the spherical aberration will vary as the object-image conjugates are adjusted. For example, a lens having a refractive index of 1.5 and configured for $m = 0$ ($v_1 = 0$ and image at f) exhibits a substantial increase in spherical aberration when the lens is used at a magnification of –1. Figure 6.21 illustrates the variation in the angular spherical aberration as both a function of refractive index and reciprocal object distance when the lens bending is for minimum spherical aberration with the object located at infinity. As can be observed from Figure 6.21, the ratio of the spherical aberration, when $m = -0.5$ and $m = 0$, increases as n increases.

Figure 6.22 shows the variation in angular spherical aberration when the lens bending is for minimum spherical aberration at a magnification of –1. In a like manner, Figure 6.23 presents the variation in angular spherical aberration for a convex-plano lens with the plano side facing the image. The figure can also be used when the lens is reversed by simply replacing the object distance with the image distance. For these plots, the actual aberration value is determined by dividing the aberration value shown by (f-number)^3.

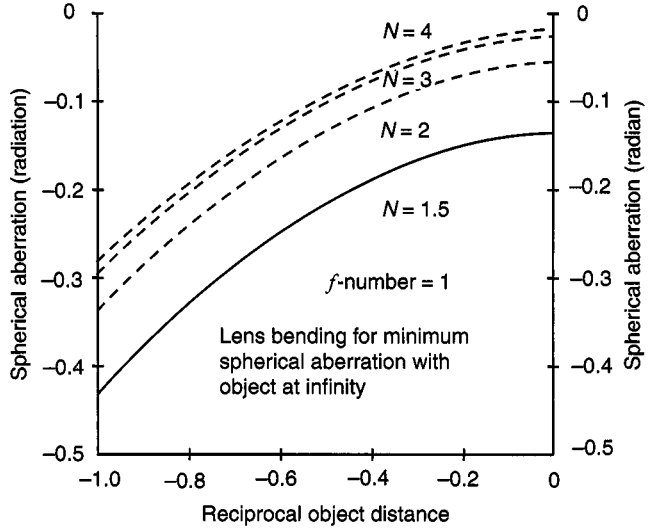

Figure 6.21 Variation of angular spherical aberration as a function of reciprocal object distance ν_1 for various refractive indices when the lens is shaped for minimum spherical aberration with the object at infinity. Spherical aberration for a specific f-number is determined by dividing the aberration value shown by (f-number)^3.

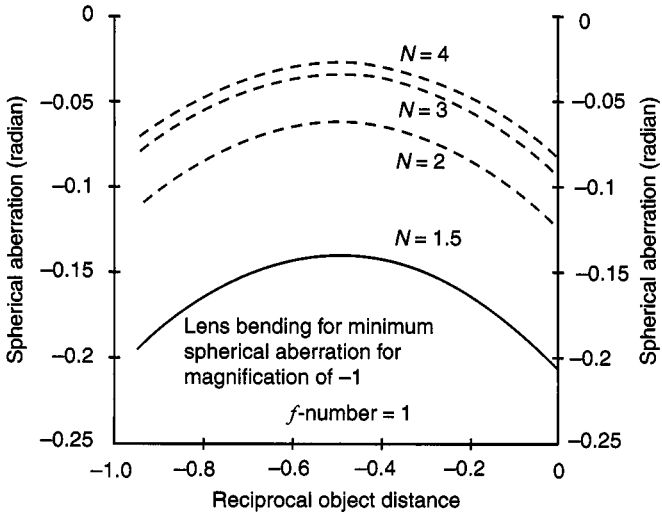

Figure 6.22 Variation of angular spherical aberration as a function of reciprocal object distance ν_1 for various refractive indices when the lens is shaped for minimum spherical aberration for a magnification of −1. Spherical aberration for a specific f-number is determined by dividing the aberration value shown by (f-number)^3.

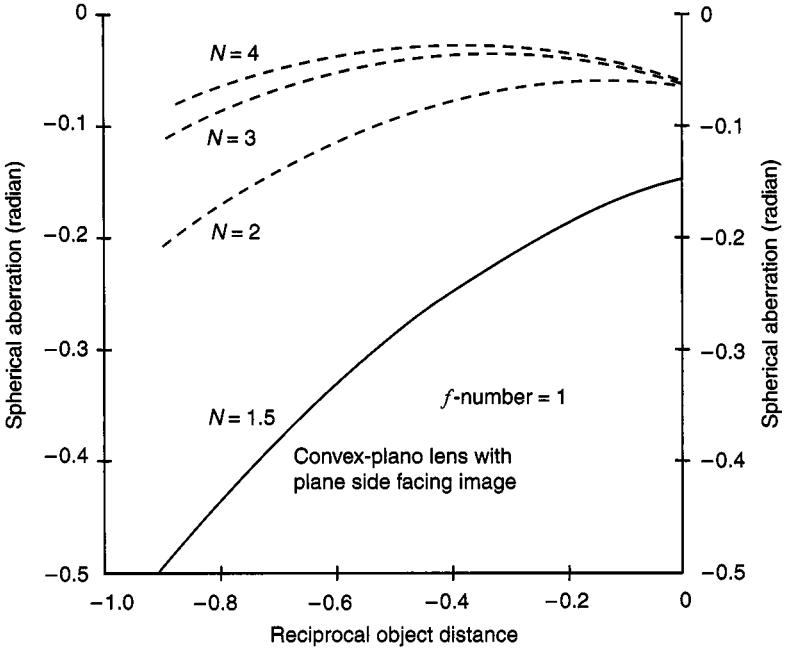

Figure 6.23 Variation of angular spherical aberration as a function of reciprocal object distance v_1 for various refractive indices when the lens has a convex-plano shape with the plano side facing the object. Spherical aberration for a specific f-number is determined by dividing the aberration value shown by (f-number)^3.

6.4 THE IMAGE DISPLACEMENT CAUSED BY A PLANOPARALLEL PLATE

From Figure 6.24 it is clear that the longitudinal image displacement caused by the insertion of a thick planoparallel plate into the path of a ray having a convergence angle U is $S = BB'$, given by

$$S = \frac{Y}{\tan U'} - \frac{Y}{\tan U}$$
$$= \frac{Y}{\tan U'}\left(1 - \frac{\tan U'}{\tan U}\right)$$

But $Y/\tan U'$ is equal to t, the thickness of the plate. Therefore,

$$S = t\left(1 - \frac{\tan U'}{\tan U}\right) = \frac{t}{N}\left(N - \frac{\cos U}{\cos U'}\right)$$

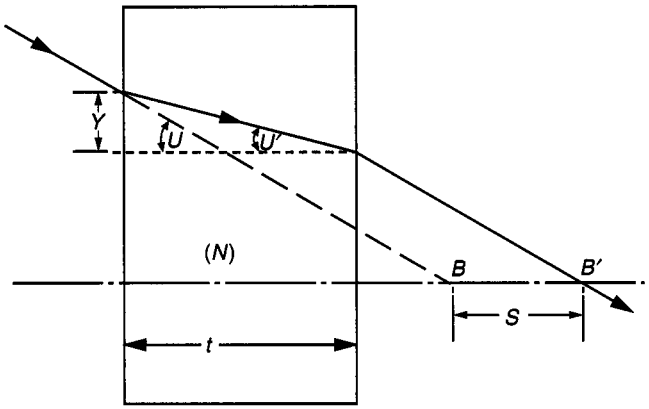

Figure 6.24 Image displacement caused by the insertion of a parallel plate.

where N is the refractive index of the plate. For a paraxial ray this reduces to

$$s = \frac{t}{N}(N - 1)$$

Since $\sin U = n \sin U'$ and $\cos^2 U' + \sin^2 U' = 1$, it follows that

$$\frac{\cos U}{\cos U'} = \frac{n \cos U}{\sqrt{n^2 - \sin^2 U}}.$$

The exact spherical aberration is

$$S - s = \frac{t}{n}\left(1 - \frac{n \cos U}{\sqrt{n^2 - \sin^2 U}}\right).$$

The plate of glass occupies more space than its "air equivalent," which is defined as that thickness of air in which a paraxial ray drops or rises by the same amount as in the glass plate. Thus, the useful relationship,

$$\text{air equivalent} = \text{glass thickness/refractive index}$$

DESIGNER NOTE

The inclusion of a flat glass plate in an optical system can impact the ultimate image quality. Microscope cover glasses and dewar windows for infrared detectors are examples where the aberrations induced by a flat glass plate should be accounted for. Consider a flat glass plate placed between the plano-hyperboloid lens in Section 6.1.8 "Plane Surface in Front" and the image it forms. In this case, the added spherical aberration can be effectively mitigated by slightly weakening the conic constant of the lens.

DESIGNER NOTE

To the lens designer a reflecting prism in a system behaves as though it were a very thick parallel plate. In a converging beam a prism has the effect of overcorrecting the three astigmatic aberrations (spherical, chromatic, and astigmatism) while it undercorrects the comatic aberrations (coma and distortion), and lateral color.

6.5 SPHERICAL ABERRATION TOLERANCES

6.5.1 Primary Aberration

Conrady has shown[10] that if a lens suffers from a small amount of pure primary spherical aberration, the best-fitting reference sphere will touch the emerging wavefront at the center and edge, and the plane of best focus will lie midway between the marginal and paraxial image points. If the aberration is large compared with the Rayleigh limit, geometrical considerations dominate, and the geometrical circle of least confusion becomes the "best" focus. However, there is also a secondary best focus close to the paraxial focus; this has been amply verified by experiment.[11]

In the case of pure primary aberration, the magnitude of the maximum residual *OPD* at this best focus is equal to the Rayleigh quarter-wave limit when

$$LA' = 4\lambda/\sin^2 U'_m = 16\lambda(f\text{-number})^2 \tag{6-17}$$

where *f*-number = focal length/diameter of aperture.

This aberration tolerance is surprisingly large, being four times the extent of the focal range. Some typical values for $\lambda = 0.0005$ mm are given in the following tabulation:

f-number	4.5	6	8	11	16	22
Primary aberration tolerance (mm)	0.2	0.3	0.5	1.0	2.0	3.9

6.5.2 Zonal Aberration

Conrady has also shown[12] that if a lens is spherically corrected for the marginal ray, the residual zonal aberration will reach the Rayleigh limit if its magnitude is

$$LZA = 6\lambda/\sin^2 U'_m \tag{6-18}$$

or 1.5 times the tolerance for pure primary spherical aberration.

DESIGNER NOTE

For telescopes, microscopes, projection lenses, and other visual systems, it is best not to allow any overcorrection of the marginal spherical aberration, even though this would reduce the zonal residual. This is because overcorrection leads to an unpleasant haziness of the image, and the zonal tolerance is so large that it is not likely to be exceeded. Indeed, many projection lenses are deliberately undercorrected, even for the marginal ray, to give the cleanest possible image with maximum contrast. Photographic objectives, on the other hand, are generally given an amount of spherical overcorrection equal to two or three times the zonal undercorrection. The overcorrected haze is often too faint to be recorded on film, especially if the exposure is on the short side, and in any case the lens will generally be stopped down somewhat, which cuts off the marginal overcorrection, leaving a small and often quite insignificant zonal residual.

In this connection, it may well be pointed out that focusing a camera by unscrewing the front element has the effect of rapidly undercorrecting the spherical aberration. This leads to a loss of definition and some degree of focus shift at small apertures, but its convenience to the camera designer outweighs these objections. If the lens is known to be intended for this type of focusing, then it should be designed with a large amount of spherical overcorrection. If possible, the aberration should be well-corrected at a focus distance of about 15 to 20 feet.

6.5.3 Conrady's OPD'_m Formula

Probably the best way to ascertain if a lens is adequately corrected for zonal aberration is to calculate the optical path difference between the emerging wavefront and a reference sphere centered about the marginal image point. Conrady[13] has given a formula by which the contribution of each lens surface to this *OPD* can be found:

$$OPD'_m = \frac{Yn \sin I \sin \frac{1}{2}(U - U') \sin \frac{1}{2}(I - U')}{2 \cos \frac{1}{2}(U + I) \cos \frac{1}{2} U \cos \frac{1}{2} I \cos \frac{1}{2} U' \cos \frac{1}{2} I'} \tag{6-19a}$$

Referring back to Eq. (6-5), we see that by using the Q method of ray tracing, Conrady's expression can be greatly simplified to

$$OPD'_m = \frac{(Q - Q')n \sin I}{4 \cos \frac{1}{2} U \cos \frac{1}{2} I \cos \frac{1}{2} U' \cos \frac{1}{2} I'} \tag{6-19b}$$

This *OPD* term has the same sign as the spherical aberration contribution at any surface. If the lens is spherically corrected for the marginal ray, the magnitude of this sum is a measure of the zonal aberration, the sum being positive for a negative zone. The advantage of using the *OPD* formula is that the tolerance of the sum is known to be two wavelengths. Hence we have an immediate

assessment of the significance of the zonal residual; this is much more accurate than the simple zonal tolerance given in Section 6.5.2, which is valid only for a mixture of primary and secondary aberrations.

If the spherical aberration is zero at both margin and 0.7 zone, as in the diagram of Figure 6-17b, then we can determine the seriousness of the two remaining small zones by calculating the *OPD* sum along the marginal ray (which should be zero) and also along the 0.7 zonal ray.

ENDNOTES

[1] Transverse field-independent astigmatism is a function of odd orders of ρ while longitudinal field-independent astigmatism is a function of even orders of ρ. Consequently, longitudinal field-independent astigmatism has no defocus component and is purely spherical aberration of the form $\rho^2 + \rho^4 + \ldots$.

[2] E. Delano, "A general contribution formula for tangential rays," *J. Opt. Soc. Am.*, 42:631 (1952).

[3] G. S. Fulcher, "Telescope objective without spherical aberration for large apertures, consisting of four crown glass lenses," *J. Opt. Soc. Am.*, 37:47 (1947).

[4] R. K. Luneburg, *Mathematical Theory of Optics*, pp. 129–133, University of California Press, Berkeley, [reproduced by permission from mimeographed notes issued by Brown University in 1944] (1966).

[5] Robert C. Fisher and Allen D. Ziebur, *Calculus and Analytical Geometry*, pp. 167–201, Prentice-Hall, Englewood Cliffs (1963).

[6] This should not be confused with the so-called cat-eye sign reflector that utilizes spherical glass beads; however, the equation for r that precedes this endnote can be used letting $d = 2r$, which yields that $n = 2$ for retroreflection. However, if a material having $n = 2$ was available, spherical aberration would spoil the return beam. Since it is desirable to have a sign appear illuminated over some reasonably significant viewing angle, using a glass having $n = 1.75$, for example, allows the beam to diverge.

[7] R. Barry Johnson and Gary A. Jacobsen, "Advances in lenticular lens arrays for visual display (Invited Paper)," *Proc. SPIE*, 5874:06-1–11 (2005).

[8] A. E. Conrady, p. 95.

[9] Figures 6.20 to 6.23 after R. Barry Johnson, "Lenses," Section 1.10 in *Handbook of Optics, Second Edition*, Chapter 1, Vol. II, McGraw-Hill, New York (1995).

[10] A. E. Conrady, p. 628.

[11] H. G. Conrady, "An experimental study of the effects of varying amounts of primary spherical aberration on the location and quality of optical images," *Phot. J.*, 66:9 (1926).

[12] A. E. Conrady, p. 631.

[13] A. E. Conrady, p. 616.