ABERRATION THEORY II

1. Primary Aberrations

If we expand the expression (12.24) for the wave aberration of a general ray from a general object point as a power series in r, τ (aperture and field variables) then it is seen that every term has an even-number power dependence on the combined powers of r and τ. Thus the terms can be grouped as zero-order, second-order, fourth-order, sixth-order, and so on. The zero and second-order terms are

$$ {}_0W_{00} $$

$$ {}_2W_{00}\tau^2 $$

$$ {}_1W_{11}\tau r\cos\phi $$

$$ {}_0W_{20}r^2 $$

and these can all be brought identically to zero by choosing the reference sphere in the exit pupil to be centred on the intersection of the pupil ray with the Gaussian image plane.

There are six fourth-order terms of which one (${}_4W_{00}\tau^4$) is identically zero for the same reason as above. The remaining five are known as the monochromatic primary aberrations, or the fourth-order aberrations, or the Seidel aberrations, or (when discussing the associated transverse ray aberrations) the third-order aberrations. Terms of the sixth and higher orders are often collectively known as "higher-order aberrations".

The importance of primary aberrations lies in the fact that it is normally the case that these aberrations dominate in optical systems of small aperture and field (i.e. the higher-order aberrations are small). As the aperture and field are increased so the higher-order aberrations typically increase, but it is still generally the case that for good image quality in a lens its primary aberrations must not be too large. The optical design process usually consists of a trade-off or balance between similar types of primary and higher-order aberrations, but the early stages of design concentrate on reducing the primary aberrations alone.

The five monochromatic primary aberrations are

${}_0W_{40}r^4$	spherical aberration	
${}_1W_{31}\tau r^3\cos\phi$	coma	
${}_2W_{22}\tau^2r^2\cos^2\phi$	astigmatism	(13.1)
${}_2W_{20}\tau^2r^2$	field curvature	
${}_3W_{11}\tau^3r\cos\phi$	distortion	

2. Wave Aberration Plots ASsociated with Primary Aberrations

The wave aberration of a system as a function of its exit pupil coordinates is conventionally plotted with W as abscissa and the pupil coordinate in a chosen azimuth as ordinate. This inversion of the normal choice of axes (independent variable as ordinate) is so that, in the special case of an infinite image (plane exit pupil sphere), the plot can be thought of as representing a section through the aberrated geometric wavefront at the exit pupil.

Figure 13.1 shows the characteristic plots of the wave aberration curves associated with primary spherical aberration, coma, astigmatism and field curvature in the $\phi = 0$ and $\phi = \pi/2$ azimuths (known as the meridian (or tangential) and sagittal azimuths).

3. Ray Intersection Patterns Associated with Primary Aberrations

Using the expressions (12.6) relating transverse ray aberration (TRA) and wave aberration (WA) for any ray, or using their normalised-coordinate forms (12.15), we can easily obtain the form of the TRA as a function of exit pupil coordinates given the wave aberration as such a function. Because the TRA is essentially the aperture derivative of the wave aberration the field dependence of a given aberration type is the same in both TRA and WA form. For this reason the explicit field dependence is omitted in the discussion below.

It is instructive to take the primary aberration terms one by one and to consider a family comprising an annular fan of rays from an object point intersecting the exit pupil in a circle ($x^2 + y^2 = r^2$) and to enquire what geometrical figure is produced by the intersection of this family of rays with the Gaussian image plane in the presence of each aberration type. This figure (or the family of such figures for ray circles of different radius in the aperture) is known as the ray intersection pattern. The ray intersection pattern associated with an aberration-free system is a single point. Figure 13.2 illustrates the ideas of this section.

The dimensions given below are in terms of normalised or reduced image coordinates G and H, calculated from (12.).

3.1 $W = W_{40}r^4$; Spherical Aberration

The ray intersection pattern for rays lying on the circle $x^2 + y^2 = r^2$ in the exit pupil is another circle of radius $4W_{40}r^3$ centred on the geometrical image point. Thus rays in successive zones of the pupil form a set of concentric circles in the image plane (Figure 13.3).

3.2 $W = W_{31}r^3\cos\phi$; Coma

The pattern is a circle of radius $W_{31}r^2$ whose centre is displaced a distance $-2W_{31}r^2$ from the geometrical image point. Thus rays in concentric zones in the exit pupil give rise to a set of circles whose displacements are proportional to their radii. It is possible to draw a pair of common tangents to these circles (Figure 13.4).

3.3 $W = W_{20}r^2 + W_{22}r^2\cos^2\phi + \delta W_{20}r^2$; Astigmatism, Field-Curvature and Defocus

Three terms are grouped here, the first two representing field curvature and astigmatism and the last representing a longitudinal shift of focal plane. This is included because the pattern changes in an interesting manner as we go through focus.

In general the pattern corresponding to a circular zone of rays in the pupil is an ellipse with semi-axes of lengths $2W_{2_0}r$ and $2(W_{20} + W_{22} + \delta W_{2_0})r$ along the G and H directions, respectively. At the particular focal plane location given by $\delta W = -\left(W_{20} + \frac{W_{22}}{2}\right)$ the ellipse reduces to a circle of radius W_{22}. At the focal plane given by $\delta W_{20} = -W_{20}$ the ellipse becomes a straight line in the H direction of semi-length $2W_{22}$, while at the focal plane given by $\delta W_{20} = -(W_{20} + W_{22})$ the pattern is another line of the same length but at right angles to the first. These orthogonal focal lines (known as the sagittal and tangential focal lines) interspaced by a circle (known as the disc of least confusion) are characteristic of astigmatism. The progression is illustrated in Figure 13.5.

3.4 $W = W_{11} r \cos\phi$; Distortion

The pattern is a single point since the TRA is the same for all rays from a given object point. The nett result is a simple image displacement of an amount $-W_{11}$. Distortion is often described as a field-imaging aberration rather than a point-imaging aberration.

Because the coefficient W_{11} depends on some odd power of τ (primary distortion depends on τ^3) a small square object is not imaged as a square but as shown in Figure 13.6. Depending on the sign of the coefficient either pincushion or barrel distortion is produced.

4. The Sine Condition

In the two drawings of Figure 13.7 both of the optical systems are free from spherical aberration; one of them is also free from coma when the object point is displaced a small distance off axis. The systems differ in their image-space angular distribution of rays from an axial object point. When these emergent ray angles satisfy the sine condition there is freedom from coma. More precisely, there is freedom from linear coma, involving a zero sum of aberration terms $({}_1W_{31}\tau r^3\cos\phi + {}_1W_{51}\tau r^5\cos\phi + {}_1W_{71}\tau r^7\cos\phi + \ldots)$. The sine condition is satisfied when

$$\frac{\sin U'}{u'} = \frac{\sin U}{u} \tag{13.2}$$

where U, U' are finite aperture ray angles (Figure 13.8) and u, u' are paraxial angle variables associated with the paraxial marginal ray.

The form (13.2) is the conventional statement of the sine condition. However, it can easily be restated as

$$y' = y \tag{13.3}$$

where y, y' are the relative pupil coordinates of the ray in question. This form has the advantage that it is also valid when the system contains spherical aberration - it remains the condition that the system is free from all orders of linear coma.

The sine condition can also be extended to an object point considerably off axis. It then takes the form

$$\left.\begin{aligned} x'_S &= x_S \\ y'_T &= y_T \end{aligned}\right\} \tag{13.4}$$

where (x_S, y_T), (x'_S, y'_T) are suitably normalised pupil coordinates, and become the condition for stationarity of aberration with respect to image height (isoplanatism).

The original importance of the sine condition lays in the assessment of the amount of linear coma in a system by a calculation based on the offence against the sine condition. Nowadays it is often the relevance of isoplanatism in the theory of the optical transfer function (OTF) that is its primary significance.

5. Primary Aberrations for a Single Surface

The primary aberrations associated with a single surface can in fact be calculated from data obtained during the tracing of two paraxial rays. We obtain the result that the increment in wave aberration due to a single surface is of the form

$$W = \tfrac{1}{8}S_I r^4 + \tfrac{1}{2}S_{II}\tau r^3\cos\phi + \tfrac{1}{2}S_{III}\tau^2 r^2\cos^2\phi$$
$$+ \tfrac{1}{4}(S_{III} + S_{IV})\tau^2 r^2 + \tfrac{1}{2}S_V\tau^3 r\cos\phi + \text{higher order terms} \qquad (13.5)$$

where S_I, S_{II}, S_{III}, S_{IV}, S_V are known as the Seidel coefficients and are given by

$$\left.\begin{aligned} S_I &= A^2 h\Delta\left(\frac{u}{n}\right) & S_{IV} &= -H^2 c\Delta\left(\frac{1}{n}\right) \\ S_{II} &= ABh\Delta\left(\frac{u}{n}\right) & S_V &= \frac{B}{A}\left(S_{III} + S_{IV}\right) \\ S_{III} &= B^2 h\Delta\left(\frac{u}{n}\right) \end{aligned}\right\} \qquad (13.6)$$

In these expressions h, u are the usual height and angle variables associated with the parxial marginal ray, $\Delta(\ )$ indicates a change on refraction (thus $\Delta(x) = x' - x$), and A, B are the refraction invariants for the paraxial marginal and pupil rays.

$$\left.\begin{aligned} A &= ni = n'i' \\ &= n(g - u) = n'(g - u') \\ &= n(hc - u) = n'(hc - u') \end{aligned}\right.$$

and

$$\left.\begin{aligned} B &= n\bar{i} = n'\bar{i}' \\ &= n(\bar{g} - \bar{u}) = n'(\bar{g} - \bar{u}') \\ &= n(\bar{h}c - \bar{u}) = n'(\bar{h}c - \bar{u}') \end{aligned}\right\} \qquad (13.7)$$

These Seidel terms are such that the aberrations for a system of several surfaces are simply the sums of the surface-by-surface contributions.

6. The Aplanatic Surface

A spherical surface whose position relative to an object point (either real or virtual) is such that rays are refracted through the surface free from all orders of spherical aberration and linear coma is called an aplanatic surface. Any surface is aplanatic with respect to one particular object point. This point and its image conjugate are known as the aplanatic points of the surface.

Looking at equations (13.6) reveals that primary spherical aberration and coma can be made zero at a surface if either $A = 0$, $h = 0$, or $\Delta\left(\frac{u}{n}\right) = 0$. Any of these three cases can arise in practice: the first implies that the object point lies at the centre of curvature of the surface, the second implies that the object is at the surface itself (as in the case of a pure "field lens"), and the third, which has a less obvious meaning, is called the aplanatic condition. This is because it is only this condition that produces freedom from all orders of spherical aberration and linear coma.

For a surface of radius of curvature r, and bounded by media of refractive indices, n,n', the aplanatic object point is located a distance from the surface given by

$$\ell = \left(\frac{n + n'}{n}\right)r$$

and the corresponding image point lies at (13.8)

$$\ell' = \left(\frac{n + n'}{n'}\right)r$$

The situation is shown in Figure 13.8. The results (13.8) mean that the aplanatic points always lie on the same side of the centre of curvature of the surface. It is therefore never possible for an aplanatic surface to focus a diverging beam into a converging one, so that we cannot build an optical system comprising only aplanatic surfaces which will form a real image of a real object. Notwithstanding this the aplanatic surface is very useful in optical design.

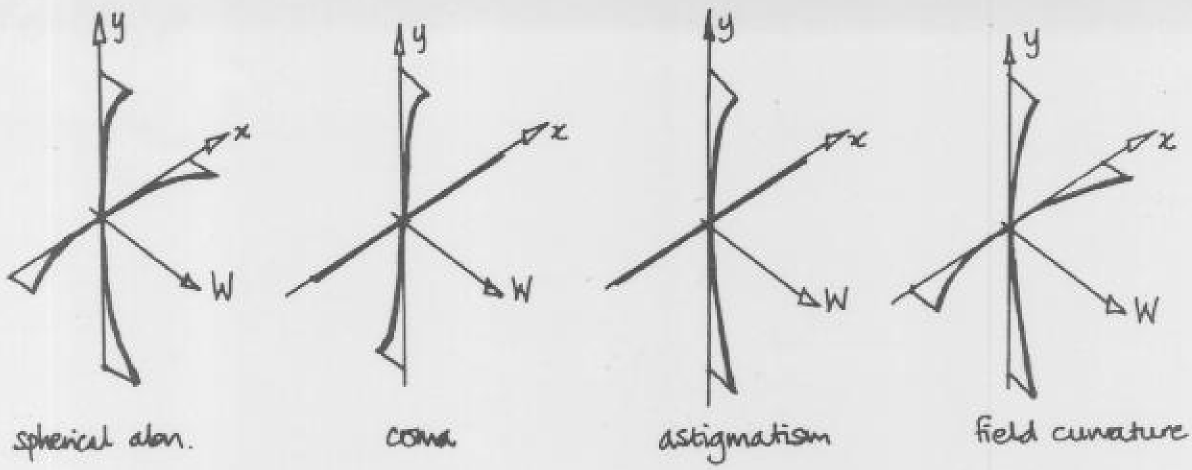


Figure 13.1 : Wave Aberration Plots

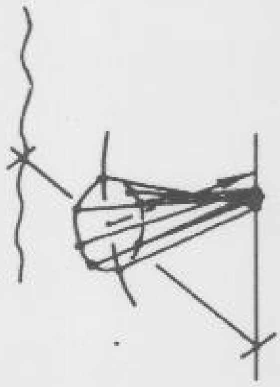
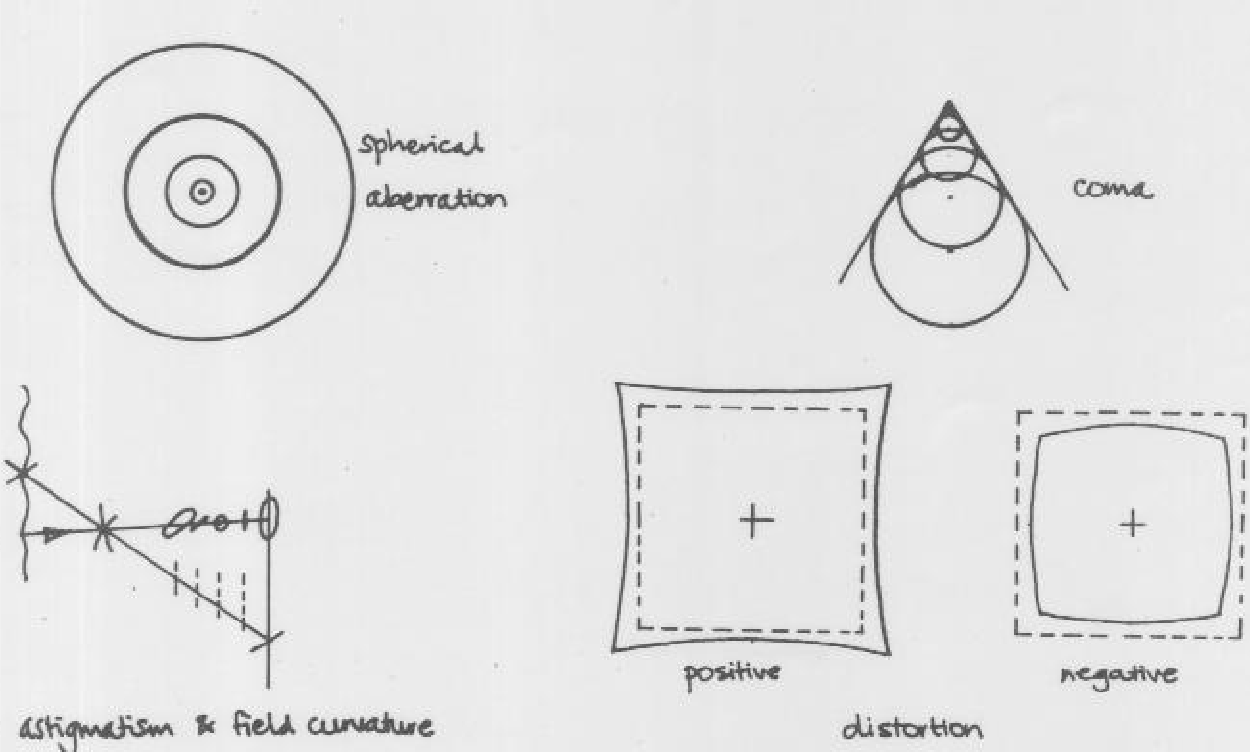


Figure 13.2 : Ray Intersection Pattern



Figures 13.3 to 13.6

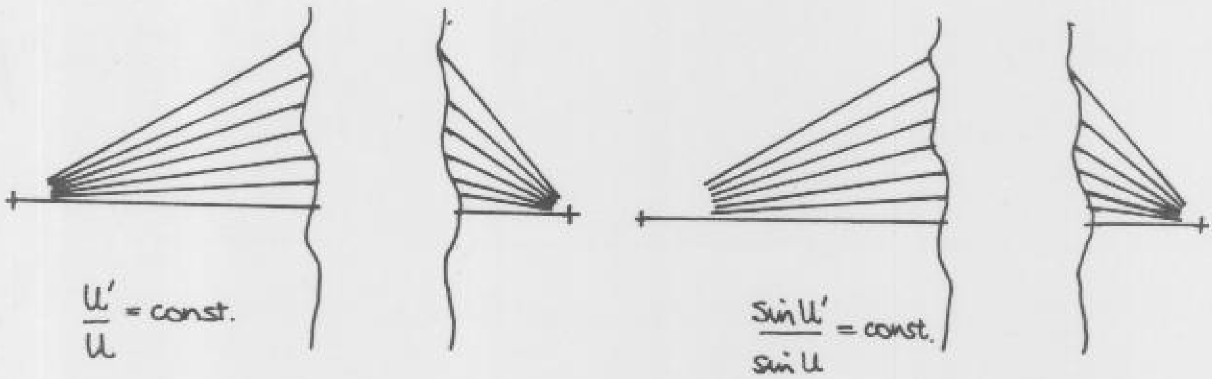


Figure 13.7 : The Sine Condition

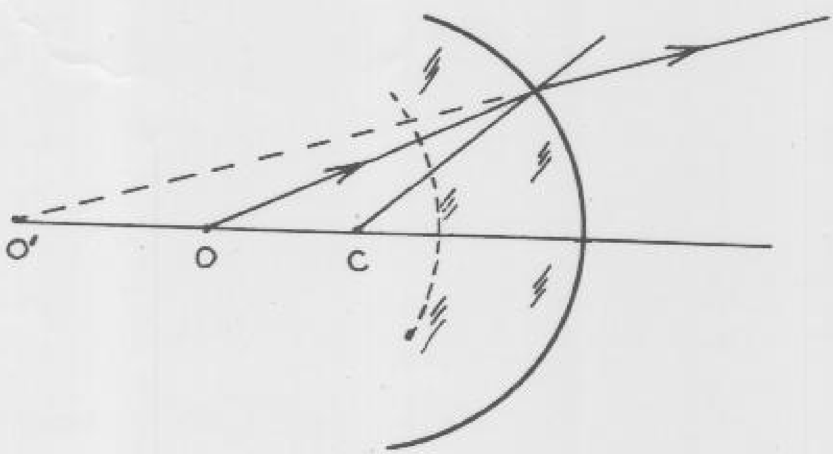


Figure 13.8 : Aplanatic Surface