V

Geometrical theory of aberrations

In §4.9 it was mentioned that within the domain of geometrical optics the departure of the path of light from the predictions of the Gaussian theory may be studied either with the help of ray-tracing or by means of algebraic analysis. In the latter treatment, which forms the subject matter of this chapter, terms which involve off-axis distances in powers higher than the second in the expansion of the characteristic functions are retained. These terms represent *geometrical aberrations*.

The discovery of photography in 1839 by Daguerre (1789–1851) was chiefly responsible for early attempts to extend the Gaussian theory. Practical optics, which until then was mainly concerned with the construction of telescope objectives, was confronted with the new task of producing objectives with large apertures and large fields. J. Petzval, a Hungarian mathematician, attacked with considerable success the related problem of supplementing the Gaussian formulae by terms involving higher powers of the angles of inclination of rays with the axis. Unfortunately, Petzval's extensive manuscript on the subject was destroyed by thieves; what is known about this work comes chiefly from semipopular reports.* Petzval demonstrated the practical value of his calculations by constructing in about 1840 his well-known portrait lens [shown in Fig. 6.3(b)] which proved greatly superior to any then in existence. The earliest systematic treatment of geometrical aberrations which was published in full is due to Seidel,† who took into account all the terms of the third order in a general centred system of spherical surfaces. Since then, his analysis has been extended and simplified by many writers.

As the wave-fronts are the orthogonal trajectories of the pencils of rays it follows that the lack of homocentricity of the image-forming pencil is accompanied by departures of the associated wave-fronts from the spherical form. Knowledge of the shape of the wave-fronts is of particular importance in more refined treatments of aberrations, based on diffraction theory (see Chapter IX). For this reason, and also because of the intimate connection between the nonhomocentricities of the ray pencils and the asphericities of the associated wave-fronts, we shall consider these two defects side by side. Our analysis will in part be based on important investigations of Schwarzschild‡ adapted somewhat to this purpose.

* J. Petzval, *Bericht über die Ergebnisse einiger dioptrischer Untersuchungen* (Pesth, 1843); also Bericht über optische Untersuchungen, *Ber. Kais. Akad. Wien, Math. naturwiss. Kl.,* **24** (1857), 50, 92, 129.

† L. Seidel, *Astr. Nachr.*, **43** (1856), No. 1027, 289, No. 1028, 305, No. 1029, 321.

‡ K. Schwarzschild, *Abh. Königl. Ges. Wis. Göttingen. Math-phys. Kl.*, **4** (1905–1906), Nos. 1, 2, 3.
Footnote continued on page 229

5.1 Wave and ray aberrations; the aberration function

Consider a rotationally symmetrical optical system. Let P_0', P_1' and P_1 be the points in which a ray from an object point P_0 intersects the plane of the entrance pupil, the exit pupil, and the Gaussian image plane respectively. If $P_1^\star$ is the Gaussian image of P_0, the vector $\boldsymbol{\delta}_1 = \overrightarrow{P_1^\star P_1}$ will be called *the aberration of the ray*, or simply *the ray aberration* (see Fig. 5.1).

Let W be the wave-front through the centre O_1' of the exit pupil, associated with the image-forming pencil which reaches the image space from P_0. In the absence of aberrations, W coincides with a sphere S which is centred on the Gaussian image point $P^\star$ and which passes through O_1'; S will be called the *Gaussian reference sphere* (see Fig. 5.2).

Let Q and $\overline{Q}$ be the points of intersection of the ray $P_1'P_1$ with the Gaussian reference sphere and with the wave-front W respectively. The optical path length $\Phi = [\overline{Q}Q]$ may be called *the aberration of the wave element at* Q, or simply the *wave aberration*, and will be regarded as positive if $\overline{Q}$ and P_1 are on the opposite sides of Q. In ordinary instruments, the wave aberrations may be as much as forty or fifty wavelengths, but in instruments used for more precise work (such as astronomical telescopes or microscopes) they must be reduced to a much smaller value, only a fraction of a wavelength.

One can easily derive expressions for the wave aberration in terms of Hamilton's

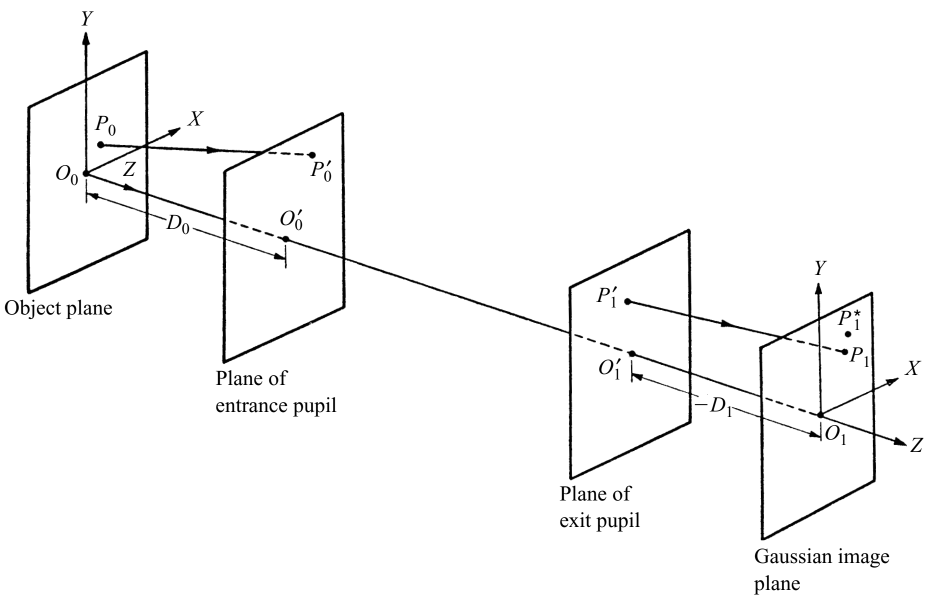

Fig. 5.1 The object plane, the image plane and the pupil planes.

Reprinted in *Astr. Mitt. Königl. Sternwarte Göttingen* (1905). An extension of Schwarzschild's analysis to systems without rotational symmetry was made by G. D. Rabinovich, *Akad. Nauk. SSSR, Zh. Eksp. Teor. Fiz.*, **16** (1946), 161.

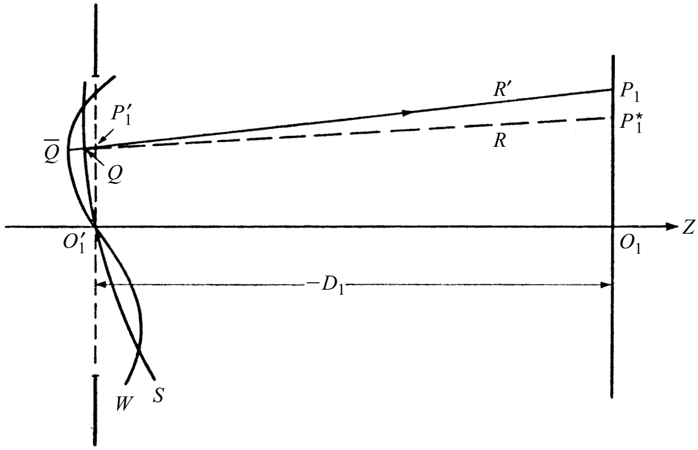

Fig. 5.2 Wave aberration and ray aberration. (The ray $\overline{Q}QP_1$ does not necessarily lie in the meridional plane $O_1'O_1P_1{}^{\star}$.)

point characteristic function of the system. If brackets [. . .] denote as before the optical path length, we have

$$\begin{aligned}\Phi &= [\overline{Q}Q] \\ &= [P_0Q] - [P_0\overline{Q}] \\ &= [P_0Q] - [P_0O_1']. \end{aligned} \tag{1}$$

Here use was made of the fact that $\overline{Q}$ and O_1' are on the same wave-front, so that $[P_0\overline{Q}] = [P_0O_1']$.

Let us introduce two sets of mutually parallel Cartesian rectangular axes, with origins at the axial points O_0 and O_1 of the object and image planes, and with the Z directions along the axis of the system. Points in the object space will be referred to the axes at O_0, and points in the image space to the axes at O_1. The Z coordinates of the pupil planes will be denoted by D_0 and D_1 (D_1 is negative in Fig. 5.1).

In terms of the point characteristic V, the wave aberration is, according to (1), given by

$$\Phi = V(X_0, Y_0, 0; X, Y, Z) - V(X_0, Y_0, 0; 0, 0, \mathrm{D}_1), \tag{2}$$

(X_0, Y_0) being the coordinates of P_0 and (X, Y, Z) those of Q. Now the coordinates (X, Y, Z) are not independent; they are connected by a relation which expresses the fact that Q lies on the Gaussian reference sphere:

$$(X - X_1^{\star})^2 + (Y - Y_1^{\star})^2 + Z^2 = R^2. \tag{3}$$

Here

$$X_1^{\star} = MX_0, \qquad Y_1^{\star} = MY_0, \tag{4}$$

are the coordinates of the Gaussian image point $P_1^{\star}$, M is the Gaussian lateral

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

magnification between the object and image planes, and R is the radius of the Gaussian reference sphere

$$R = \sqrt{X_1^{\star 2} + Y_1^{\star 2} + D_1^2}. \tag{5}$$

Using (3), Z may be eliminated from (2), so that Φ may be regarded as a function of X_0, Y_0, X and Y only, i.e.*

$$\Phi = \Phi(X_0,\ Y_0;\ X,\ Y).$$

In terms of the aberration function $\Phi(X_0, Y_0; X, Y)$, simple expressions for the ray aberrations may be obtained. We have, from (2),

$$\frac{\partial \Phi}{\partial X} = \frac{\partial V}{\partial X} + \frac{\partial V}{\partial Z}\frac{\partial Z}{\partial X}. \tag{6}$$

If α_1, β_1 and γ_1 are the angles which the ray QP_1 makes with the axes, and (X, Y, Z) and $(X_1, Y_1, 0)$ are the coordinates of the points Q and P_1, respectively, we have, according to §4.1 (7), and from Fig. 5.2,

$$\frac{\partial V}{\partial X} = n_1 \cos \alpha_1 = n_1 \frac{X_1 - X}{R'}, \qquad \frac{\partial V}{\partial Z} = n_1 \cos \gamma_1 = -n_1 \frac{Z}{R'}, \tag{7}$$

where

$$R' = \sqrt{(X_1 - X)^2 + (Y_1 - Y)^2 + Z^2} \tag{8}$$

is the distance from Q to P_1 and n_1 is the refractive index of the image space. Further, from (3),

$$\frac{\partial Z}{\partial X} = -\frac{X - X_1^\star}{Z}. \tag{9}$$

Substitution from (7) and (9) into (6) gives the following expressions for the ray aberration components:

$$\left.\begin{aligned} X_1 - X_1^\star &= \frac{R'}{n_1}\frac{\partial \Phi}{\partial X}, \\ \text{and similarly} \qquad Y_1 - Y_1^\star &= \frac{R'}{n_1}\frac{\partial \Phi}{\partial Y}. \end{aligned}\right\} \tag{10}$$

The relations (10) are exact, but they involve, on the right-hand side, the distance R', which itself depends on the coordinates of P_1, i.e. on the ray aberrations.† However, for most practical purposes R' may be replaced by the radius R of the Gaussian reference sphere or by another approximation (see (15) below).

It is easily seen that, on account of symmetry, Φ depends on the four variables only through the three combinations $X_0^2 + Y_0^2$, $X^2 + Y^2$ and $X_0 X + Y_0 Y$. For if polar coordinates are introduced in the X, Y-planes, i.e. if we set

* A more general aberration function, suitable for discussing the imaging of extended objects, was introduced by E. Wolf, *J. Opt. Soc. Amer.*, **42** (1952), 547.

† A somewhat different pair of exact equations which connect the ray aberrations and the wave aberrations was derived by J. L. Rayces, *Optica Acta*, **11** (1964), 85.

$$\left.\begin{aligned} X_0 &= r_0\cos\theta_0, \qquad & X &= r\cos\theta, \\ Y_0 &= r_0\sin\theta_0, \qquad & Y &= r\sin\theta, \end{aligned}\right\} \tag{11}$$

Φ becomes a function of r_0, r, θ_0 and θ; or, what amounts to the same thing, a function of r_0, r, $\theta_0 - \theta$ and θ. Suppose now that the X- and Y-axes at O_0 and O_1 are rotated through the same angle and in the same sense about the axis of the system. This leaves r_0, r and $\theta_0 - \theta$ unchanged, whilst θ increases by the angle of rotation. Since Φ is invariant with respect to such rotations, it must be independent of the last variable, i.e. it depends on r_0, r and $\theta_0 - \theta$ only. Hence the aberration function Φ may be expressed as a function of the three scalar products

$$\mathbf{r}_0{}^2 = X_0^2 + Y_0^2, \quad \mathbf{r}^2 = X^2 + Y^2 \quad \text{and} \quad \mathbf{r}_0\cdot\mathbf{r} = X_0X + Y_0Y, \tag{12}$$

of the two vectors $\mathbf{r}_0(X_0, Y)$ and $\mathbf{r}(X, Y)$.

From this result it follows that, if Φ is expanded into a power series with respect to the four coordinates, the expansion will contain only terms of even degree. There will be no terms of zero degree, since $\Phi(0, 0; 0, 0) = 0$. Moreover, there will be no terms of the second degree, except possibly for a term proportional to $(X_0^2 + Y_0^2)$; for according to (10) such terms would give rise to ray aberrations depending linearly on the coordinates, and this contradicts the fact that $P_1{}^\star$ is the Gaussian image of P_0. Hence the expansion is of the form

$$\Phi = c(X_0^2 + Y_0^2) + \Phi^{(4)} + \Phi^{(6)} + \cdots, \tag{13}$$

where c is a constant and $\Phi^{(2k)}$ is a polynomial of degree $2k$ in the coordinates, and contains these coordinates in powers of the three scalar invariants (12) only. A term of a particular degree $2k$ is said to represent a *wave aberration of order* $2k$. The aberrations of the lowest order ($2k = 4$) are usually called *primary* or *Seidel aberrations** and will be studied in detail in §5.3.

To show the order of magnitude of certain expressions and the degree of approximation involved in some of our calculations, it is convenient to introduce a parameter μ. This may be any quantity of the first order, e.g. the angular aperture of the system. Then all the rays passing through the system may be assumed to make angles $O(\mu)$ with the optical axis, where the symbol $O(\mu)$ means 'not exceeding a moderate multiple' of μ.

Consider the error involved on replacing R' in the basic relations (10) by quantities independent of X_1 and Y_1. From (3) and (5),

$$Z^2 = \mathrm{D}_1^2 - (X^2 + Y^2) + 2(XX_1^\star + YY_1^\star) \tag{14}$$

and (8) then gives

$$\begin{aligned} R' &= -\mathrm{D}_1\left[1 + \frac{X_1^2 + Y_1^2 - 2X(X_1 - X_1^\star) - 2Y(Y_1 - Y_1^\star)}{\mathrm{D}_1^2}\right]^{1/2} \\ &= -\mathrm{D}_1 - \frac{X_1^{\star 2} + Y_1^{\star 2}}{2\mathrm{D}_1} + O(\mathrm{D}_1\mu^4). \end{aligned} \tag{15}$$

* Since the *ray aberrations* associated with wave aberrations of this order are of the third degree in the coordinates, they are also sometimes called the *third-order aberrations*.

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

The relations (10) for the ray aberration components then become

$$X_1 - X_1^\star = -\frac{1}{n_1}\left(\mathrm{D}_1 + \frac{X_1^{\star 2} + Y_1^{\star 2}}{2\mathrm{D}_1}\right)\frac{\partial \Phi}{\partial X} + O(\mathrm{D}_1\mu^7) \tag{16a}$$

$$= -\frac{\mathrm{D}_1}{n_1}\frac{\partial \Phi}{\partial X} + O(\mathrm{D}_1\mu^5), \tag{16b}$$

$$Y_1 - Y_1^\star = -\frac{1}{n_1}\left(\mathrm{D}_1 + \frac{X_1^{\star 2} + Y_1^{\star 2}}{2\mathrm{D}_1}\right)\frac{\partial \Phi}{\partial Y} + O(\mathrm{D}_1\mu^7) \tag{17a}$$

$$= -\frac{\mathrm{D}_1}{n_1}\frac{\partial \Phi}{\partial Y} + O(\mathrm{D}_1\mu^5). \tag{17b}$$

5.2 The perturbation eikonal of Schwarzschild

In his researches on geometrical aberrations, Schwarzschild employed a method similar to that used in the calculations of orbital elements in celestial mechanics. In such calculations variables are introduced which remain constant in the unperturbed motion, and the small changes which these quantities undergo in the actual motion are then determined with the help of a perturbation function. By analogy with this procedure, Schwarzschild introduced, in the papers already referred to, certain variables which, within the accuracy of Gaussian optics, have constant values along each ray passing through the optical system. With the help of a certain perturbation function which he introduced and called the *Seidel eikonal*, he then studied the changes which these variables undergo when the fourth-order terms in the expansion of the characteristic function are taken into account.* Schwarzschild called these special variables the *Seidel variables*, as they are related to those employed previously by Seidel.

Within the accuracy of the Seidel theory, the aberration function Φ defined in the preceding section is very closely related to the perturbation eikonal of Schwarzschild; and by following closely Schwarzschild's procedure one can derive the expressions for the fourth-order coefficients in the expansion of the aberration function for any centred system. This will be done in detail in §5.5. Here the Seidel variables will be defined and the connection between our aberration function and the perturbation function of Schwarzschild will be examined.

Let us introduce new units of length l_0 in the object plane and l_1 in the Gaussian image plane, such that

$$\frac{l_1}{l_0} = M \tag{1}$$

is the lateral Gaussian magnification between the two planes. Points in the object plane

* Schwarzschild also considered the general types of fifth-order ray aberrations. Expressions for the associated sixth-order coefficients in the expansion of the characteristic function (Schwarzschild's perturbation eikonal) were first given by his pupil A. Kohlschütter in his dissertation (Göttingen, 1908).

Fifth-order aberrations were also investigated by M. Herzberger, *J. Opt. Soc. Amer.*, **29** (1939), 395, and also his *Modern Geometrical Optics* (New York, Interscience Publishers, 1958), H. A. Buchdahl, *Optical Aberration Coefficients* (Oxford, Oxford University Press, 1954) and by other workers. A review of higher order aberration theory was published by J. Focke in *Progress in Optics*, Vol. 4, ed. E. Wolf (Amsterdam, North Holland Publishing Company and New York, J. Wiley and Sons, 1965), p. 1.

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

will be specified by the coordinates x_0, y_0, and points in the image space by the coordinates x_1, y_1, such that

$$\left.\begin{aligned} x_0 &= \mathrm{C}\frac{X_0}{l_0}, \qquad x_1 = \mathrm{C}\frac{X_1}{l_1}, \\ y_0 &= \mathrm{C}\frac{Y_0}{l_0}, \qquad y_1 = \mathrm{C}\frac{Y_1}{l_1}, \end{aligned}\right\} \tag{2}$$

where (X_0, Y_0), (X_1, Y_1) are the ordinary coordinates of P_0 and P_1 (see Fig. 5.1), and C is a constant to be specified later. Within the accuracy of Gaussian optics, $x_1 = x_0$ and $y_1 = y_0$.

The coordinates (X_0', Y_0') of the points where the ray from (X_0, Y_0) meets the entrance pupil are connected with the ray components by the relations

$$\frac{X_0' - X_0}{\mathrm{D}_0} = \frac{p_0}{\sqrt{n_0^2 - p_0^2 - q_0^2}}, \qquad \frac{Y_0' - Y_0}{\mathrm{D}_0} = \frac{q_0}{\sqrt{n_0^2 - p_0^2 - q_0^2}}, \tag{3}$$

the expressions for the intersection points of the ray with the exit pupil being strictly similar. Within the accuracy of Gaussian optics, the square roots in the denominators may be replaced by n_0 and n_1, respectively, and the following linear relations between the coordinates are then obtained:

$$\left.\begin{aligned} \frac{X_0' - X_0}{\mathrm{D}_0} &= \frac{p_0}{n_0}, \qquad \frac{X_1' - X_1}{\mathrm{D}_1} = \frac{p_1}{n_1}, \\ \frac{Y_0' - Y_0}{\mathrm{D}_0} &= \frac{q_0}{n_0}, \qquad \frac{Y_1' - Y_1}{\mathrm{D}_1} = \frac{q_1}{n_1}. \end{aligned}\right\} \tag{4}$$

Next we introduce new units of length λ_0 and λ_1 in the plane of the entrance and the exit pupil such that

$$\frac{\lambda_1}{\lambda_0} = M' \tag{5}$$

is the lateral magnification between the two planes. In place of X_0', Y_0', X_1', Y_1', the following variables will be used:

$$\left.\begin{aligned} \xi_0 &= \frac{X_0'}{\lambda_0} = \frac{X_0}{\lambda_0} + \frac{\mathrm{D}_0 p_0}{\lambda_0 n_0}, \qquad \xi_1 = \frac{X_1'}{\lambda_1} = \frac{X_1}{\lambda_1} + \frac{\mathrm{D}_1 p_1}{\lambda_1 n_1}, \\ \eta_0 &= \frac{Y_0'}{\lambda_0} = \frac{Y_0}{\lambda_0} + \frac{\mathrm{D}_0 q_0}{\lambda_0 n_0}, \qquad \eta_1 = \frac{Y_1'}{\lambda_1} = \frac{Y_1}{\lambda_1} + \frac{\mathrm{D}_1 q_1}{\lambda_1 n_1}. \end{aligned}\right\} \tag{6}$$

Within the accuracy of Gaussian optics, $\xi_1 = \xi_0$, $\eta_1 = \eta_0$.

In order to simplify later calculations, it will be convenient to choose C as*

* Another choice of C, which is often useful in the study of images on the basis of diffraction theory, is

$$\mathrm{C} = kn_0 l_0 \sin\theta_0 = -kn_1 l_1 \sin\theta_1.$$

Here k is the vacuum wave number of the light and $2\theta_0$ and $2\theta_1$ are the angular apertures of the system, i.e. the angles which the entrance pupil and the exit pupil subtend at the axial object and image points respectively. With this choice of C,

$$\begin{aligned} x_0 &= kn_0 X_0 \sin\theta_0, \qquad x_1 = -kn_1 X_1 \sin\theta_1, \\ y_0 &= kn_0 Y_0 \sin\theta_0, \qquad y_1 = -kn_1 Y_1 \sin\theta_1. \end{aligned}$$

$$\mathrm{C} = \frac{n_0 l_0 \lambda_0}{\mathrm{D}_0} = \frac{n_1 l_1 \lambda_1}{\mathrm{D}_1}. \tag{7}$$

The equality of the two terms on the right-hand side follows from the Smith–Helmholtz formula [§4.4 (48)].

The quantities defined by (2) and (6) are the *Seidel variables*. The inverse relations expressing the old variables in terms of the Seidel variables will also be needed. We have, on solving (2) and (6):

$$\left.\begin{aligned} X_0 &= \frac{\mathrm{D}_0}{n_0\lambda_0} x_0, & X_1 &= \frac{\mathrm{D}_1}{n_1\lambda_1} x_1, \\ Y_0 &= \frac{\mathrm{D}_0}{n_0\lambda_0} y_0, & Y_1 &= \frac{\mathrm{D}_1}{n_1\lambda_1} y_1, \end{aligned}\right\} \tag{8}$$

$$\left.\begin{aligned} p_0 &= \frac{n_0\lambda_0}{\mathrm{D}_0}\xi_0 - \frac{1}{\lambda_0}x_0, & p_1 &= \frac{n_1\lambda_1}{\mathrm{D}_1}\xi_1 - \frac{1}{\lambda_1}x_1, \\ q_0 &= \frac{n_0\lambda_0}{\mathrm{D}_0}\eta_0 - \frac{1}{\lambda_0}y_0, & q_1 &= \frac{n_1\lambda_1}{\mathrm{D}_1}\eta_1 - \frac{1}{\lambda_1}y_1. \end{aligned}\right\} \tag{9}$$

Next the aberration function will be expressed in terms of the Seidel variables. We note first of all that the arguments X and Y may be replaced by X_1' and Y_1' in Φ without changing the error term in §5.1 (16b) and §5.1 (17b). Let us denote by ϕ the aberration function, when regarded as a function of the Seidel variables,

$$\Phi(X_0, Y_0; X_1', Y_1') = \phi(x_0, y_0; \xi_1, \eta_1). \tag{10}$$

Then

$$\frac{\partial\Phi}{\partial X_1'} = \frac{\partial\phi}{\partial\xi_1}\frac{\partial\xi_1}{\partial X_1'} = \frac{1}{\lambda_1}\frac{\partial\phi}{\partial\xi_1}, \tag{11a}$$

and from (2), (1) and (7) and §5.1 (4)

$$X_1 - X_1^{\star} = \frac{\mathrm{D}_1}{n_1\lambda_1}(x_1 - x_0). \tag{11b}$$

With the help of Eqs. (11), the formulae (16b) and (17b) of §5.1 give

$$\left.\begin{aligned} x_1 - x_0 &= -\frac{\partial\phi}{\partial\xi_1} + O(\mathrm{D}_1\mu^5), \\ y_1 - y_0 &= -\frac{\partial\phi}{\partial\eta_1} + O(\mathrm{D}_1\mu^5). \end{aligned}\right\} \tag{12}$$

It was mentioned earlier that within the accuracy of the Seidel theory, ϕ is closely related to a perturbation function introduced by Schwarzschild and called by him the *Seidel eikonal*. This perturbation function is defined by the expression

$$\psi = T + \frac{\mathrm{D}_0}{2n_0\lambda_0^2}(x_0^2 + y_0^2) - \frac{\mathrm{D}_1}{2n_1\lambda_1^2}(x_1^2 + y_1^2) + x_0(\xi_1 - \xi_0) + y_0(\eta_1 - \eta_0), \tag{13}$$

where $T = T(p_0, q_0; p_1, q_1)$ is the angle characteristic, referred to origins at O_0 and

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

O_1. Consider now the effect of small variations in the coordinates. We have, according to §4.1 (27), with $Z_0 = Z_1 = 0$, that the corresponding change in T is given by

$$\delta T = X_0\,\mathrm{d}p_0 + Y_0\,\mathrm{d}q_0 - X_1\,\mathrm{d}p_1 - Y_1\,\mathrm{d}q_1, \tag{14}$$

or, in terms of the Seidel variables,

$$\begin{aligned}\delta T = {} & x_0\left(\mathrm{d}\xi_0 - \frac{\mathrm{D}_0}{n_0\lambda_0^2}\,\mathrm{d}x_0\right) + y_0\left(\mathrm{d}\eta_0 - \frac{\mathrm{D}_0}{n_0\lambda_0^2}\,\mathrm{d}y_0\right)\\ & - x_1\left(\mathrm{d}\xi_1 - \frac{\mathrm{D}_1}{n_1\lambda_1^2}\,\mathrm{d}x_1\right) - y_1\left(\mathrm{d}\eta_1 - \frac{\mathrm{D}_1}{n_1\lambda_1^2}\,\mathrm{d}y_1\right).\end{aligned} \tag{15}$$

Using (15), we deduce from (13) that a small variation in the variables leads to a change in ψ given by

$$\mathrm{d}\psi = (\xi_1 - \xi_0)\,\mathrm{d}x_0 + (\eta_1 - \eta_0)\,\mathrm{d}y_0 + (x_0 - x_1)\,\mathrm{d}\xi_1 + (y_0 - y_1)\,\mathrm{d}\eta_1, \tag{16}$$

so that ψ may be expressed as function of x_0, y_0, ξ_1, η_1 and we then have rigorously

$$\left.\begin{aligned}\xi_1 - \xi_0 = \frac{\partial\psi}{\partial x_0}, \quad & x_1 - x_0 = -\frac{\partial\psi}{\partial\xi_1},\\ \eta_1 - \eta_0 = \frac{\partial\psi}{\partial y_0}, \quad & y_1 - y_0 = -\frac{\partial\psi}{\partial\eta_1}.\end{aligned}\right\} \tag{17}$$

Hence from knowledge of ψ the ray aberrations can be determined, both in the image plane and in the plane of the exit pupil by simple differentiation.

Comparison of (17) and (12) shows that within the accuracy of the Seidel theory $\phi - \psi$ must be independent of ξ_1 and η_1, i.e.

$$\phi(x_0, y_0; \xi_1, \eta_1) = \psi(x_0, y_0; \xi_1, \eta_1) + \chi(x_0, y_0) + O(\mathrm{D}_1\mu^6), \tag{18}$$

where $\chi(x_0, y_0)$ is some function of x_0 and y_0. Now from the definition of ϕ, $\phi(x_0, y_0; 0, 0) = 0$. Hence $\chi(x_0, y_0) = -\psi(x_0, y_0; 0, 0)$ and consequently

$$\phi(x_0, y_0; \xi_1, \eta_1) = \psi(x_0, y_0; \xi_1, \eta_1) - \psi(x_0, y_0; 0, 0) + O(\mathrm{D}_1\mu^6). \tag{19}$$

Within the range of validity of the Seidel theory, the error terms in (12) may be neglected. If, however, terms of order higher than fourth are taken into account, the expressions for the ray aberration components in terms of the aberration function ϕ are more complicated. On the other hand, the simple relations (17) for the ray aberration components in terms of the perturbation eikonal are exact; the perturbation eikonal appears, however, to have no simple physical meaning.

The determination of terms of order higher than fourth is very laborious in all but the simplest cases. For this reason, algebraic calculations are usually restricted to the domain of the Seidel theory, supplemented where necessary by ray tracing.

5.3 The primary (Seidel) aberrations

By means of considerations strictly similar to those of §5.1 relating to the aberration function, it follows on account of symmetry that the power series expansion of the perturbation eikonal of Schwarzschild is of the form

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

$$\psi = \psi^{(0)} + \psi^{(4)} + \psi^{(6)} + \psi^{(8)} + \cdots, \tag{1}$$

where $\psi^{(2k)}$ is a polynomial of degree $2k$ in the four variables; and, moreover, that the four variables enter only in the three combinations

$$\mathrm{r}^2 = x_0^2 + y_0^2, \qquad \rho^2 = \xi_1^2 + \eta_1^2, \qquad \kappa^2 = x_0\xi_1 + y_0\eta_1. \tag{2}$$

In (1) the term of the second degree is absent, for the presence of such a term would, according to §5.2 (17), contradict the fact that within the accuracy of Gaussian optics, $x_1 = x_0$, $y_1 = y_0$, $\xi_1 = \xi_0$, and $\eta_1 = \eta_0$.

Since the variables enter only in the combinations (2), it follows that $\psi^{(4)}$ must be of the form

$$\psi^{(4)} = -\tfrac{1}{4}A\mathrm{r}^4 - \tfrac{1}{4}B\rho^4 - C\kappa^4 - \tfrac{1}{2}D\mathrm{r}^2\rho^2 + E\mathrm{r}^2\kappa^2 + F\rho^2\kappa^2, \tag{3}$$

where A, B, ... are constants. The signs and the numerical factors in (3) have been chosen in agreement with the usual practice; they lead to convenient expressions for the ray aberrations. The evaluation of the aberration constants for any centred system will be discussed in §5.5.

Clearly the power series expansion of ϕ is of the same form as (1), but it contains no term of zero order ($\phi^{(0)} = 0$), and the leading term $\phi^{(4)}$ differs from $\psi^{(4)}$ by the absence of the term $-\tfrac{1}{4}A\mathrm{r}^4$, as is immediately obvious from §5.2 (19). Hence the general expression for the wave aberration of the lowest order (fourth) is

$$\phi^{(4)} = -\tfrac{1}{4}B\rho^4 - C\kappa^4 - \tfrac{1}{2}D\mathrm{r}^2\rho^2 + E\mathrm{r}^2\kappa^2 + F\rho^2\kappa^2, \tag{4}$$

B, C, ... being the same coefficients as in (3).

Substitution from (4) into §5.2 (12) gives the general expression for the ray aberration components of the lowest (third) order:

$$\left.\begin{aligned}
\Delta^{(3)}x &= x_1 - x_0 = \frac{n_1\lambda_1}{\mathrm{D}_1}(X_1 - X_1^\star)\\
&= x_0(2C\kappa^2 - E\mathrm{r}^2 - F\rho^2) + \xi_1(B\rho^2 + D\mathrm{r}^2 - 2F\kappa^2),\\
\Delta^{(3)}y &= y_1 - y_0 = \frac{n_1\lambda_1}{\mathrm{D}_1}(Y_1 - Y_1^\star)\\
&= y_0(2C\kappa^2 - E\mathrm{r}^2 - F\rho^2) + \eta_1(B\rho^2 + D\mathrm{r}^2 - 2F\kappa^2).
\end{aligned}\right\} \tag{5}$$

The coefficient A does not enter (4) and (5), so that there are altogether five types of aberrations of the lowest order characterized by the five coefficients B, C, D, E and F. As already mentioned on p. 233 these aberrations are known as *primary* or *Seidel aberrations*.

To discuss the effect of the Seidel aberrations it will be convenient to choose the axes so that the y, z-plane passes through the object point; then $x_0 = 0$. If further polar coordinates are introduced, so that

$$\xi_1 = \rho\sin\theta, \qquad \eta_1 = \rho\cos\theta, \tag{6}$$

(4) becomes

$$\phi^{(4)} = -\tfrac{1}{4}B\rho^4 - Cy_0^2\rho^2\cos^2\theta - \tfrac{1}{2}D{y_0}^2\rho^2 + Ey_0^3\rho\cos\theta + Fy_0\rho^3\cos\theta, \tag{7}$$

and (5) takes the form

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

$$\left.\begin{aligned}\Delta^{(3)}x &= B\rho^3\sin\theta - 2Fy_0\rho^2\sin\theta\cos\theta + Dy_0^2\rho\sin\theta,\\ \Delta^{(3)}y &= B\rho^3\cos\theta - Fy_0\rho^2(1+2\cos^2\theta) + (2C+D)y_0^2\rho\cos\theta - Ey_0{}^3.\end{aligned}\right\} \tag{8}$$

In the special case when all the coefficients in (7) have zero values, the wave-front in the exit pupil coincides (within the present degree of accuracy) with the Gaussian reference sphere (see Fig. 5.2). In general, the coefficients will have finite values. Each term then represents a particular type of departure of the wave-front from the ideal spherical form; the five different types are illustrated in Fig. 5.3.

The significance of the ray aberrations associated with a given object point may be illustrated graphically by means of so-called *aberration* (or *characteristic*) *curves*. These are the curves traced out in the image plane by the intersection points of all the rays emerging from a fixed zone $\rho =$ constant of the exit pupil. The area covered by the aberration curves which correspond to all the possible values of ρ then represent the imperfect image.

We shall consider in turn each of the Seidel aberrations.*

(a) Spherical aberration ($B \neq 0$)

When all the coefficients except B are zero, (8) reduces to

$$\left.\begin{aligned}\Delta^{(3)}x &= B\rho^3\sin\theta,\\ \Delta^{(3)}y &= B\rho^3\cos\theta.\end{aligned}\right\} \tag{9}$$

The aberration curves are therefore concentric circles whose centres are at the Gaussian image point and whose radii increase with the third power of the zonal radius ρ, but are independent of the position (y_0) of the object in the field of view. This defect of the image is known as the *spherical aberration*.

The spherical aberration, being independent of y_0, affects the off-axis as well as the axial image. The rays from an axial object point which make an appreciable angle with the axis will intersect the axis in points which lie in front of or behind the Gaussian focus (Fig. 5.4). The point at which the rays from the edge of the aperture intersect the axis is called the *marginal focus*. If a screen is placed in the image region at right angles to the axis, there is a position for which the circular image spot appearing on the screen is a minimum; this minimal 'image' is called the *circle of least confusion*.

(b) Coma ($F \neq 0$)

The aberration characterized by the coefficient F is known as *coma*. According to (8) the ray aberration components are in this case

$$\left.\begin{aligned}\Delta^{(3)}x &= -2F\rho^2 y_0\sin\theta\cos\theta = -Fy_0\rho^2\sin 2\theta,\\ \Delta^{(3)}y &= -F\rho^2 y_0(1+2\cos^2\theta) = -Fy_0\rho^2(2+\cos 2\theta).\end{aligned}\right\} \tag{10}$$

It is seen that, if y_0 is fixed and the zonal radius ρ is kept constant, the point P_1 in

* Aberration curves associated with aberrations of higher orders are discussed by G. C. Steward, *Trans. Camb. Phil. Soc.*, **23** (1926), 235 and by N. Chako, *Trans. Chalmers University of Technology* (Gothenburg), Nr. 191 (1957).

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

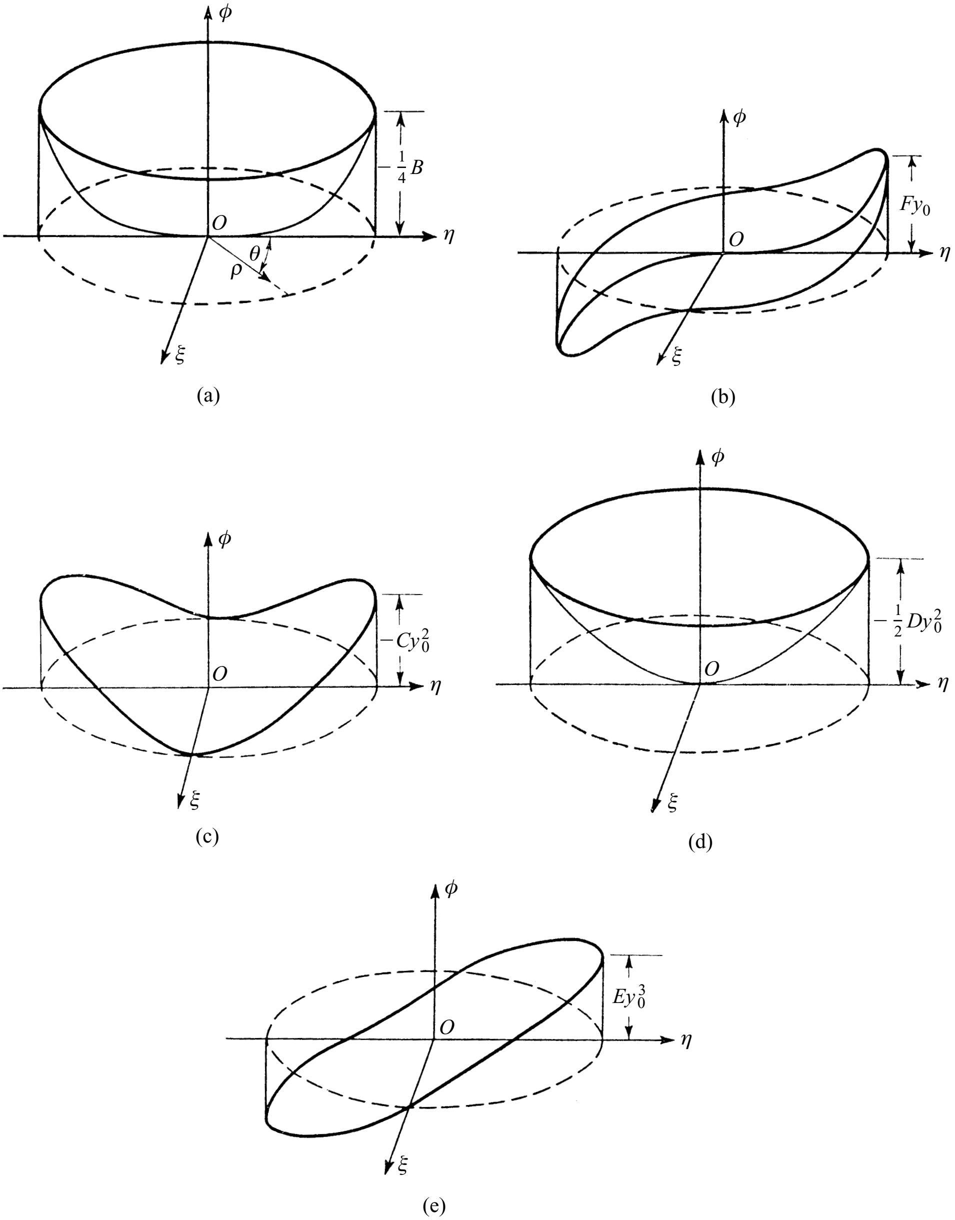

Fig. 5.3 The primary wave aberrations: (a) spherical aberration $\phi = -\frac{1}{4}B\rho^4$; (b) coma $\phi = Fy_0\rho^3 \cos\theta$; (c) astigmatism $\phi = -Cy_0^2\rho^2 \cos^2\theta$; (d) curvature of field $\phi = -\frac{1}{2}Dy_0^2\rho^2$; (e) distortion $\phi = Ey_0^3\rho\cos\theta$.

the image plane describes a circle twice over, as θ runs through the range $0 \leqslant \theta < 2\pi$. The circle is of radius $|Fy_0\rho^2|$ and its centre is at a distance $-2F\rho^2 y_0$ from the Gaussian focus, in the y-direction. The circle therefore touches the two straight lines which pass through the Gaussian image and which are inclined to the y-axis at 30°. As

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

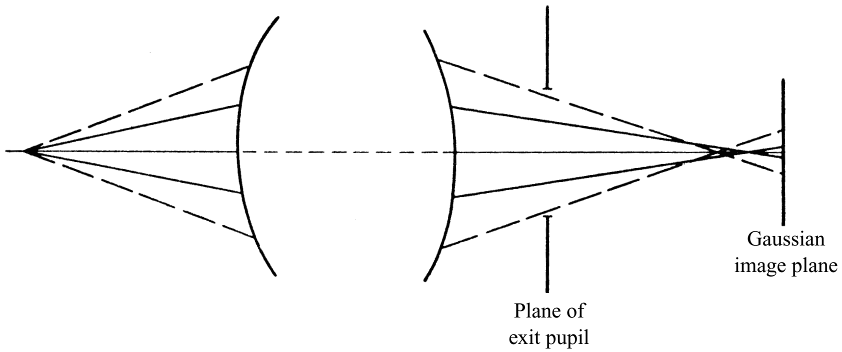

Fig. 5.4 Spherical aberration.

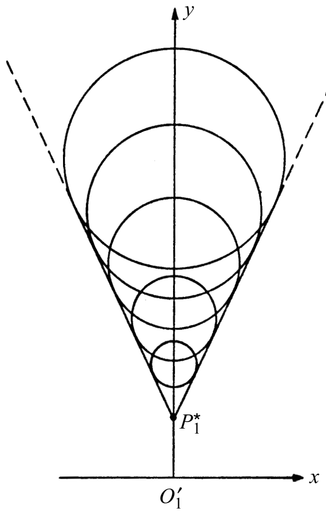

Fig. 5.5 Coma.

ρ takes on all possible values, the circles cover a region bounded by segments of the two straight lines and by an arc of the largest aberration circle (see Fig. 5.5). The overall size of this pattern increases linearly with the off-axis distance of the object point. Now in §4.5 it was shown that, if the Abbe sine condition is satisfied, an element of the object plane in the immediate neighbourhood of the axis will be imaged sharply by the system. In such a case the expansion of the aberration function cannot therefore contain terms which depend linearly on y_0. Hence if the sine condition is satisfied, the primary coma, in particular, will be absent.

(c) Astigmatism $(C \neq 0)$ *and curvature of field* $(D \neq 0)$

The effects of the aberrations which are characterized by the coefficients C and D can best be studied together. We have from (8), if all the other coefficients have zero values,

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

$$\left.\begin{aligned}\Delta^{(3)}x &= D\rho y_0^2 \sin\theta,\\ \Delta^{(3)}y &= (2C+D)\rho y_0^2 \cos\theta.\end{aligned}\right\} \tag{11}$$

To see the significance of these aberrations, assume to begin with that the image-forming pencil is very narrow. According to §4.6, the rays of such a pencil intersect two short lines, of which one (the tangential focal line) is at right angles to the meridional plane, and the other (the sagittal focal line) lies in this plane. Consider now light from all the points in a finite region of the object plane. The focal lines in the image space give rise to two surfaces, the *tangential* and the *sagittal focal surface*. To a first approximation these surfaces may be considered to be spheres. Let R_t and R_s be their radii, regarded as positive if the corresponding centre of curvature lies on the side of the image plane from which the light is propagated ($R_t > 0$, $R_s < 0$ in Fig. 5.6).

The radii of curvature may be expressed in terms of the coefficients C and D. To show this it will be convenient to calculate the ray aberrations in the presence of curvature, using first the ordinary rather than the Seidel coordinates. We have (see Fig. 5.7)

$$\frac{\Delta^{(3)}Y_1}{Y_1'} = \frac{u}{\mathrm{D}_1 + u}, \tag{12}$$

where u is the small distance between the sagittal focal line and the image plane. If v denotes the distance of this focal line from the axis,

$$R_t^2 = v^2 + (R_t - u)^2,$$

i.e.

$$v^2 = 2R_t u - u^2.$$

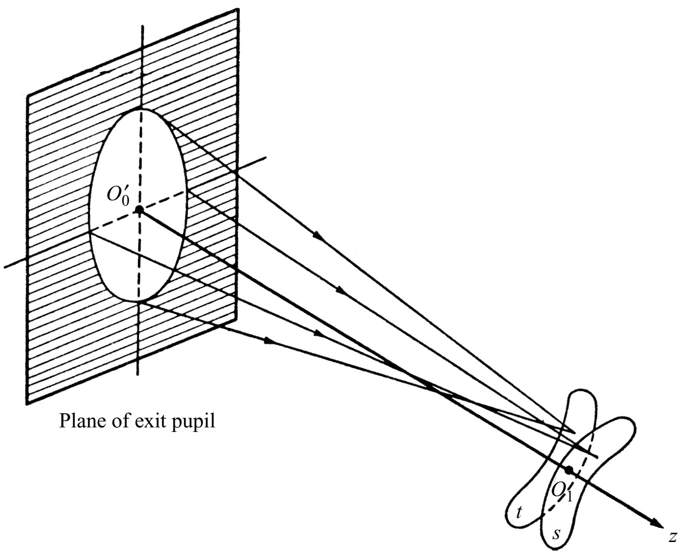

Fig. 5.6 The tangential and sagittal focal surfaces.

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

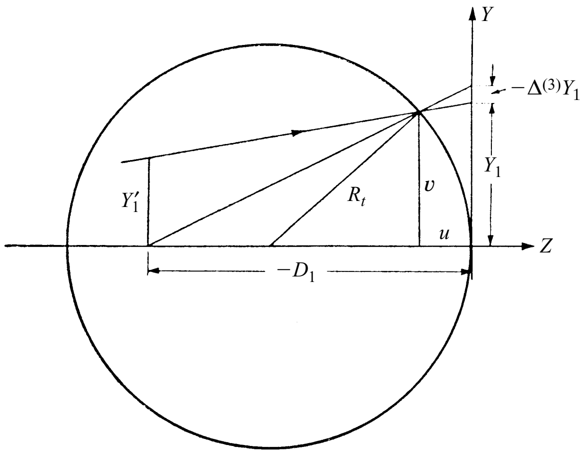

Fig. 5.7 Astigmatism and curvature of field.

If u is regarded as a small quantity of the first order, v may be replaced by Y_1, and u^2 may be neglected in the last equation, so that

$$u \sim \frac{Y_1^2}{2R_t}, \tag{13}$$

and (12) gives, if we also neglect u in comparison with D_1,

$$\Delta^{(3)} Y_1 = \frac{Y_1^2}{2R_t} \frac{Y_1'}{\mathrm{D}_1}. \tag{14}$$

Similarly

$$\Delta^{(3)} X_1 = \frac{Y_1^2}{2R_s} \frac{X_1'}{\mathrm{D}_1}. \tag{15}$$

Next these relations will be expressed in terms of the Seidel variables. Substitution from §5.2 (6) and §5.2 (8) leads to

$$\Delta^{(3)} y = \frac{n_1\lambda_1}{\mathrm{D}_1} \Delta^{(3)} Y_1 = \frac{n_1\lambda_1}{\mathrm{D}_1} \frac{Y_1^2}{2R_t} \frac{Y_1'}{\mathrm{D}_1} = \frac{n_1\lambda_1}{2\mathrm{D}_1^2 R_t} \frac{\mathrm{D}_1^2}{n_1{}^2\lambda_1^2} y_1^2 \lambda_1 \eta_1,$$

i.e.

$$\Delta^{(3)} y = \frac{y_1^2 \eta_1}{2n_1 R_t}, \tag{16}$$

and similarly

$$\Delta^{(3)} x = \frac{y_1^2 \xi_1}{2n_1 R_s}. \tag{17}$$

In (16) and (17) y_1 may be replaced by y_0, and we finally obtain, on comparison with (11) and on using (6),

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

$$\frac{1}{R_t} = 2n_1(2C + D), \qquad \frac{1}{R_s} = 2n_1 D. \tag{18}$$

The quantity $2C + D$ is usually called the *tangential field curvature* and D the *sagittal field curvature*, and the quantity

$$\frac{1}{R} = \frac{1}{2}\left(\frac{1}{R_t} + \frac{1}{R_s}\right) = 2n_1(C + D), \tag{19}$$

which is proportional to their arithmetic mean, is called, simply, *the field curvature*.

From (13) and (18) it is seen that, at height Y_1 from the axis, the separation between the two focal surfaces (i.e. the astigmatic distance of the image-forming pencil) is

$$\frac{Y_1^2}{2}\left(\frac{1}{R_t} - \frac{1}{R_s}\right) = 2n_1 C Y_1^2. \tag{20}$$

The semidifference

$$\frac{1}{2}\left(\frac{1}{R_t} - \frac{1}{R_s}\right) = 2n_1 C \tag{21}$$

is accordingly called *astigmatism*. In the absence of astigmatism ($C = 0$), $R_t = R_s = R$. It will be shown in §5.5.3 that the radius R of the common focal surface may then be calculated from a simple formula which only involves the radii of curvature of the individual surfaces of the system and the refractive indices of all the media.

(d) Distortion ($E \neq 0$)

If the E coefficient alone differs from zero, we have according to (8)

$$\left.\begin{aligned} \Delta^{(3)}x &= 0, \\ \Delta^{(3)}y &= -Ey_0^3. \end{aligned}\right\} \tag{22}$$

Since these expressions are independent of ρ and θ, the imaging will be stigmatic and independent of the radius of the exit pupil; the off-axis distance of the image will, however, not be proportional to that of the object. This aberration is therefore called *distortion*.

If distortion is present, the image of any straight line in the object plane which meets the axis will itself be a straight line, but the image of any other straight line will be curved. This effect is shown in Fig. 5.8(a), where the object has the form of a mesh formed by equidistant lines parallel to the x- and y-axes. Fig. 5.8(b) illustrates so-called *barrel distortion* ($E > 0$), and Fig. 5.8(c) illustrates *pincushion distortion* ($E < 0$).

It has been seen that, of the five Seidel aberrations, three (namely spherical aberration, coma and astigmatism) are responsible for lack of sharpness of the image. The other two (namely curvature of field and distortion) are related to the position and the form of the image. In general it is impossible to design a system which is free from all the primary as well as the higher-order aberrations, and a suitable compromise as to their relative magnitudes has, therefore, to be made. In some cases the effects of the Seidel aberrations are reduced by balancing them against aberrations of higher orders. In others one has to eliminate certain aberrations completely even at the price of

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

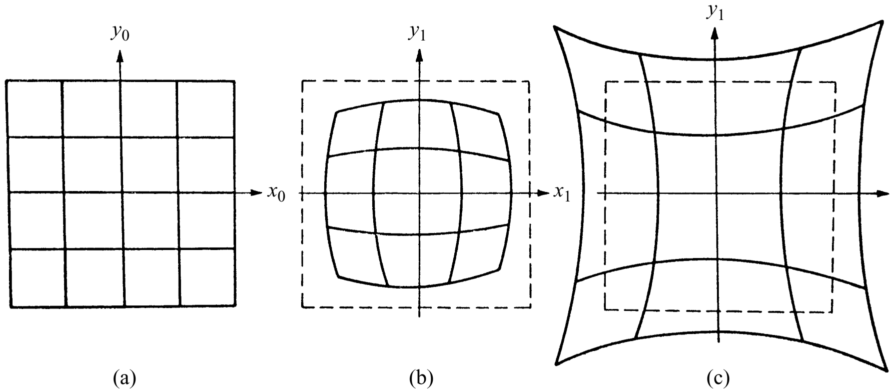

Fig. 5.8 (a) Object. (b) Image in the presence of barrel distortion ($E > 0$). (c) Image in the presence of pincushion distortion ($E < 0$).

introducing aberrations of other types. For example, because of the asymmetric appearance of an image in the presence of coma, this aberration must always be suppressed in telescopes, as it would make precise positional measurements impracticable. On the other hand a certain amount of curvature of the field and of distortion is then relatively harmless as these can be eliminated by calculation.

So far we have studied the aberration effects on the basis of geometrical optics only. If, however, the aberrations are very small (wave aberrations of the order of a wavelength or less), diffraction plays an important part. The geometrical theory must then be supplemented by more refined considerations. This will be done in Chapter IX.

5.4 Addition theorem for the primary aberrations

Having considered the significance of the primary aberrations, we must turn our attention to the more difficult task of calculating the primary aberration coefficients for a general centred system. As was seen, this is equivalent to determining the fourth-order terms in the power series expansion of the perturbation eikonal of Schwarzschild. In order not to interrupt the main calculations, it will be convenient to consider first the manner in which the perturbation eikonal of an optical system depends on the perturbation eikonals associated with the imaging by each surface of the system.

Consider a centred system which consists of two surfaces of revolution, and let O_0 be the axial object point and O_1 and O_2 its Gaussian images by the first surface and by both the surfaces, respectively. Further let

$$T_1 = T_1(p_0, q_0;\ p_1, q_1) \tag{1}$$

be the angle characteristic for refraction at the first surface, and

$$T_2 = T_2(p_1, q_1;\ p_2, q_2) \tag{2}$$

the angle characteristic for refraction at the second surface, the former being referred to axes at O_0 and O_1 and the latter to parallel axes at O_1 and O_2; the Z-axes are taken along the axis of the system. Since the media are assumed to be homogeneous, the

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

angle characteristic represents the optical path between the feet of the perpendiculars dropped on to the initial and final portions of the ray from the two origins (see Fig. 4.3). Hence the angle characteristic T of the system referred to axes at O_0 and O_2 is obtained from T_1 and T_2 by addition:

$$T = T_1 + T_2. \tag{3}$$

In this expression, the variables p_1, q_1 referring to the ray in the intermediate space must be eliminated with the help of formulae describing the imaging by each surface. It will be convenient to carry out the elimination explicitly in the corresponding relation for the perturbation eikonal.

According to §5.2 (13), the perturbation eikonals ψ_1 and ψ_2 are given by

$$\psi_1 = T_1 + \frac{\mathrm{D}_0}{2n_0\lambda_0^2}(x_0^2 + y_0^2) - \frac{\mathrm{D}_1}{2n_1\lambda_1^2}(x_1^2 + y_1^2) + x_0(\xi_1 - \xi_0) + y_0(\eta_1 - \eta_0),$$

and

$$\psi_2 = T_2 + \frac{\mathrm{D}_1}{2n_1\lambda_1^2}(x_1^2 + y_1^2) - \frac{\mathrm{D}_2}{2n_2\lambda_2^2}(x_2^2 + y_2^2) + x_1(\xi_2 - \xi_1) + y_1(\eta_2 - \eta_1);$$

and the perturbation eikonal ψ of the combination is

$$\psi = T + \frac{\mathrm{D}_0}{2n_0{\lambda_0}^2}({x_0}^2 + {y_0}^2) - \frac{\mathrm{D}_2}{2n_2\lambda_2^2}(x_2^2 + y_2^2) + x_0(\xi_2 - \xi_0) + y_0(\eta_2 - \eta_0).$$

Hence, using (3),

$$\psi = \psi_1 + \psi_2 + (x_0 - x_1)(\xi_2 - \xi_1) + (y_0 - y_1)(\eta_2 - \eta_1),$$

or, using §5.2 (17)

$$\psi = \psi_1 + \psi_2 + \frac{\partial\psi_1}{\partial\xi_1}\frac{\partial\psi_2}{\partial x_1} + \frac{\partial\psi_1}{\partial\eta_1}\frac{\partial\psi_2}{\partial y_1}. \tag{4}$$

If in (4) the power series expansions for ψ_1 and ψ_2 are substituted, it follows (since according to §5.3 (1) there are no terms of second order) that

$$\psi = \psi_1^{(0)} + \psi_2^{(0)} + \psi_1^{(4)} + \psi_2^{(4)} + \cdots, \tag{5}$$

the terms not shown being of sixth and higher orders. Now in $\psi_1^{(4)}$ and $\psi_2^{(4)}$ we may obviously replace the arguments by their Gaussian values. Hence the variables relating to the intermediate space may be eliminated by means of the following rule: In $\psi_1^{(4)}(x_0, y_0; \xi_1, \eta_1)$ replace ξ_1, η_1 by ξ_2, η_2 and in $\psi_2^{(4)}(x_1, y_1; \xi_2, \eta_2)$ replace x_1, y_1 by x_0, y_0, and add the resulting expressions. This gives $\psi^{(4)}$ as a function of the four variables x_0, y_0, and ξ_2, η_2 as required. This elimination will now be carried out more explicitly.

According to §5.3 (3), $\psi_1^{(4)}$ and $\psi_2^{(4)}$ are of the form

$$\left.\begin{aligned} &\psi_1^{(4)} = -\tfrac{1}{4}A_1\mathrm{r}_0^4 - \tfrac{1}{4}B_1\rho_1^4 - C_1\kappa_{01}^4 - \tfrac{1}{2}D_1\mathrm{r}_0^2\rho_1^2 + E_1\mathrm{r}_0^2\kappa_{01}^2 + F_1\rho_1^2\kappa_{01}^2, \\ &\text{with} \\ &\mathrm{r}_0^2 = x_0^2 + y_0^2, \qquad \rho_1^2 = \xi_1^2 + \eta_1^2, \qquad \kappa_{01}^2 = x_0\xi_1 + y_0\eta_1; \end{aligned}\right\} \tag{6}$$

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

$$\left.\begin{aligned} \psi_2^{(4)} &= -\tfrac{1}{4}A_2\mathrm{r}_1^4 - \tfrac{1}{4}B_2\rho_2^4 - C_2\kappa_{12}^4 - \tfrac{1}{2}D_2\mathrm{r}_1^2\rho_2^2 + E\mathrm{r}_1^2\kappa_{12}^2 + F_2\rho_2^2\kappa_{12}^2, \\ &\text{with} \\ \mathrm{r}_1^2 &= x_1^2 + {y_1}^2, \qquad \rho_2^2 = \xi_2^2 + \eta_2^2, \qquad \kappa_{12}^2 = x_1\xi_2 + y_1\eta_2. \end{aligned}\right\} \tag{7}$$

Replacing ξ_1, η_1 by ξ_2, η_2 and x_1, y_1 by x_0, y_0, i.e. ρ_1 by ρ_2, r_1 by r_0, κ_{12}^2 and κ_{01}^2 by

$$\kappa_{02}^2 = x_0\xi_2 + y_0\eta_2, \tag{8}$$

we obtain on addition

$$\begin{aligned} \psi^{(4)} = &-\tfrac{1}{4}(A_1 + A_2)\mathrm{r}_0^4 - \tfrac{1}{4}(B_1 + B_2)\rho_2^4 - (C_1 + C_2)\kappa_{02}^4 - \tfrac{1}{2}(D_1 + D_2)\mathrm{r}_0^2\rho_2^2 \\ &+ (E_1 + E_2)\mathrm{r}_0^2\kappa_{02}^2 + (F_1 + F_2)\rho_2^2\kappa_{02}^2. \end{aligned} \tag{9}$$

A similar result obviously holds for a centred system consisting of any number of surfaces. We have thus proved the following theorem:

Each primary aberration coefficient of a centred system is the sum of the corresponding aberration coefficients associated with the individual surfaces of the system.

It is at this point of our analysis that the advantage of the Seidel variables is particularly evident; for this simple and important result depends basically upon the use of the Seidel variables and has no analogy when ordinary variables are used.

5.5 The primary aberration coefficients of a general centred lens system

The theorem established in the preceding section reduces the problem of determining the primary aberration coefficients of a general centred system to that of calculating the corresponding coefficients for each of its surfaces. This calculation will now be carried out.

5.5.1 The Seidel formulae in terms of two paraxial rays

It may be recalled that the Seidel aberration coefficients may (apart from simple numerical factors) be identified with the coefficients of the fourth-order terms in the power series expansion of the perturbation eikonal ψ of Schwarzschild. According to §5.2 (13), this function is obtained by adding to the angle characteristic T certain quadratic terms, and by expressing the resulting expression in terms of the Seidel variables. Since the relations between the Seidel variables and the ray components are linear, the order of the terms does not change by transition from the one set of variables to the other. Hence

$$\psi^{(4)}(x_0, y_0; \xi_1, \eta_1) = T^{(4)}(p_0, q_0; p_1, q_1). \tag{1}$$

The expansion of the angle characteristic up to the fourth order for a refracting surface of revolution was derived in §4.1. The fourth-order contribution may, according to §4.1 (42), be written in the form*

* As in §4.1, r denotes the radius of the surface. This symbol should not be confused with the symbol r which denotes the rotational invariant $\sqrt{x_0^2 + y_0^2}$ introduced in §5.3 (2).

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

$$T^{(4)}(p_0, q_0;\ p_1, q_1) = -\frac{r}{4(n_1 - n_0)^2}[(p_0 - p_1)^2 + (q_0 - q_1)^2]$$
$$\times\left(\frac{p_0^2 + q_0^2}{n_0} - \frac{p_1^2 + q_1^2}{n_1}\right)$$
$$-\frac{(1+b)r}{8(n_1 - n_0)^3}[(p_0 - p_1)^2 + (q_0 - q_1)^2]^2$$
$$+\frac{a_0}{8n_0^3}(p_0^2 + q_0^2)^2 - \frac{a_1}{8n_1^3}(p_1^2 + q_1^2)^2. \qquad (2)$$

It will be convenient to choose the axial points $z = a_0$, $z = a_1$ as the axial object point and its Gaussian image, and to set (see Fig. 5.9)

$$s = a_0, \qquad s' = a_1, \qquad t = a_0 + \mathrm{D}_0, \qquad t' = a_1 + \mathrm{D}_1. \qquad (3)$$

The associated Abbe invariants (§4.4 (7)) will be denoted by K and L respectively:

$$n_0\left(\frac{1}{r} - \frac{1}{s}\right) = n_1\left(\frac{1}{r} - \frac{1}{s'}\right) = K, \qquad (4)$$

$$n_0\left(\frac{1}{r} - \frac{1}{t}\right) = n_1\left(\frac{1}{r} - \frac{1}{t'}\right) = L. \qquad (5)$$

Before substituting into (1) the expressions for the ray components in terms of the Seidel variables, it will be useful to re-write (1) in a slightly different form. Because of (4) we have

$$\frac{r}{8(n_1 - n_0)^3}[(p_0 - p_1)^2 + (q_0 - q_1)^2]^2 =$$
$$\frac{r^2}{8(n_1 - n_0)^4}\left(\frac{n_1}{s'} - \frac{n_0}{s}\right)[(p_0 - p_1)^2 + (q_0 - q_1)^2]^2.$$

Using this relation, (2) may be written as

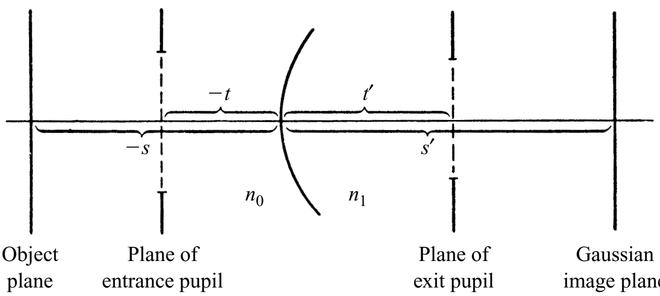

Fig. 5.9 Notation used in the calculation of the primary aberration coefficients.

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

$$\begin{aligned} T^{(4)} = {} & \frac{1}{8n_0 s}\left\{\frac{n_0 r}{(n_1-n_0)^2}[(p_0-p_1)^2+(q_0-q_1)^2]-\frac{s}{n_0}(p_0^2+q_0^2)\right\}^2 \\ & -\frac{1}{8n_1 s'}\left\{\frac{n_1 r}{(n_1-n_0)^2}[(p_0-p_1)^2+(q_0-q_1)^2]-\frac{s'}{n_1}(p_1^2+q_1^2)\right\}^2 \\ & -\frac{br}{8(n_1-n_0)^3}\{(p_0-p_1)^2+(q_0-q_1)^2\}^2. \end{aligned} \tag{6}$$

In (6), the arguments may be replaced by their Gaussian approximations; in particular, the Seidel variables referring to points on the incident and the refracted ray may be interchanged. In order to obtain $\psi^{(4)}$ as a function of x_0, y_0, ξ_1 and η_1, we may then use in place of §5.2 (9) the relations

$$\left.\begin{aligned} p_0 &= \frac{n_0\lambda_0}{D_0}\xi_1-\frac{1}{\lambda_0}x_0, \quad & p_1 &= \frac{n_1\lambda_1}{D_1}\xi_1-\frac{1}{\lambda_1}x_0, \\ q_0 &= \frac{n_0\lambda_0}{D_0}\eta_1-\frac{1}{\lambda_0}y_0, \quad & q_1 &= \frac{n_1\lambda_1}{D_1}\eta_1-\frac{1}{\lambda_1}y_0. \end{aligned}\right\} \tag{7}$$

It will be useful to make one further modification. The Gaussian lateral magnification between the object and the image plane (l_1/l_0) and between the planes of the entrance and the exit pupil (λ_1/λ_0) may be obtained from §4.4 (14) and §4.4 (10), or more simply by noting that imaging by a spherical surface is a projection from the centre of the sphere. Hence, if (4) is also used,

$$\frac{l_1}{l_0}=\frac{r-s'}{r-s}=\frac{n_0 s'}{n_1 s}, \qquad \frac{\lambda_1}{\lambda_0}=\frac{r-t'}{r-t}=\frac{n_0 t'}{n_1 t}. \tag{8}$$

Introducing the abbreviations

$$h=\frac{\lambda_0 s}{D_0}=\frac{\lambda_1 s'}{D_1}, \qquad H=\frac{t}{\lambda_0 n_0}=\frac{t'}{\lambda_1 n_1}, \tag{9}$$

where (8) and §5.2 (7) were used, (7) becomes

$$\left.\begin{aligned} p_0 &= n_0\left(\frac{h}{s}\xi_1-\frac{H}{t}x_0\right), \quad & p_1 &= n_1\left(\frac{h}{s'}\xi_1-\frac{H}{t'}x_0\right), \\ q_0 &= n_0\left(\frac{h}{s}\eta_1-\frac{H}{t}y_0\right), \quad & q_1 &= n_1\left(\frac{h}{s'}\eta_1-\frac{H}{t'}y_0\right). \end{aligned}\right\} \tag{10}$$

If as before, r^2, ρ^2 and κ^2 denote the three rotational invariants

$$\mathrm{r}^2=x_0^2+y_0^2, \qquad \rho^2=\xi_1^2+\eta_1^2, \qquad \kappa^2=x_0\xi_1+y_0\eta_1, \tag{11}$$

the terms in the curly brackets of (6) become

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

$$
\begin{aligned}
(p_0 - p_1)^2 + (q_0 - q_1)^2 &= \left(\frac{n_1 - n_0}{\mathrm{r}}\right)^2 (H^2\mathrm{r}^2 + h^2\rho^2 - 2hH\kappa^2),\\
\frac{n_0 r}{(n_1 - n_0)^2}[(p_0 - p_1)^2 + (q_0 - q_1)^2] &- \frac{s}{n_0}(p_0^2 + q_0^2)\\
&= H^2\mathrm{r}^2\left[L - (K - L)\frac{s}{t}\right] + h^2\rho^2 K - 2hH\kappa^2 L,\\
\frac{n_1 r}{(n_1 - n_0)^2}[(p_0 - p_1)^2 + (q_0 - q_1)^2] &- \frac{s'}{n_1}(p_1^2 + q_1^2)\\
&= H^2\mathrm{r}^2\left[L - (K - L)\frac{s'}{t'}\right] + h^2\rho^2 K - 2hH\kappa^2 L.
\end{aligned}
$$

Substituting from these relations into (6) and recalling (1), we finally obtain the required expression for $\psi^{(4)}$:

$$
\begin{aligned}
\psi^{(4)} = \frac{1}{8}\mathrm{r}^4 H^4 &\left[\frac{b}{r^3}(n_0 - n_1) + L^2\left(\frac{1}{n_0 s} - \frac{1}{n_1 s'}\right) - 2L(K - L)\left(\frac{1}{n_0 t} - \frac{1}{n_1 t'}\right)\right.\\
&\left. + (K - L)^2\left(\frac{s}{n_0 t^2} - \frac{s'}{n_1 t'^2}\right)\right]\\
&+ \frac{1}{8}\rho^4 h^4\left[\frac{b}{r^3}(n_0 - n_1) + K^2\left(\frac{1}{n_0 s} - \frac{1}{n_1 s'}\right)\right]\\
&+ \frac{1}{2}\kappa^4 H^2 h^2\left[\frac{b}{r^3}(n_0 - n_1) + L^2\left(\frac{1}{n_0 s} - \frac{1}{n_1 s'}\right)\right]\\
&+ \frac{1}{4}\mathrm{r}^2\rho^2 H^2 h^2\left[\frac{b}{r^3}(n_0 - n_1) + KL\left(\frac{1}{n_0 s} - \frac{1}{n_1 s'}\right) - K(K - L)\left(\frac{1}{n_0 t} - \frac{1}{n_1 t'}\right)\right]\\
&- \frac{1}{2}\mathrm{r}^2\kappa^2 H^3 h\left[\frac{b}{r^3}(n_0 - n_1) + L^2\left(\frac{1}{n_0 s} - \frac{1}{n_1 s'}\right) - L(K - L)\left(\frac{1}{n_0 t} - \frac{1}{n_1 t'}\right)\right]\\
&- \frac{1}{2}\rho^2\kappa^2 H h^3\left[\frac{b}{r^3}(n_0 - n_1) + KL\left(\frac{1}{n_0 s} - \frac{1}{n_1 s'}\right)\right]. \qquad (12)
\end{aligned}
$$

This formula gives, on comparison with the general expression §5.3 (3), the fourth-order coefficients A, B, $\ldots$, F of the perturbation eikonal of a refracting surface of revolution.

Generalization to a centred system consisting of any number of refracting surfaces is now straightforward. Let us denote by the suffix i quantities referring to the ith surface, and let n_i be the refractive index of the medium which follows the ith surface. Then, using the addition theorem of §5.4 it follows from (12) on comparison with §5.3 (3) that

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

$$
\left.
\begin{aligned}
B &= \frac{1}{2}\sum_i h_i^4\left[\frac{b_i}{r_i^3}(n_i - n_{i-1}) + K_i^2\left(\frac{1}{n_i s_i'} - \frac{1}{n_{i-1}s_i}\right)\right],\\
C &= \frac{1}{2}\sum_i H_i^2 h_i^2\left[\frac{b_i}{r_i^3}(n_i - n_{i-1}) + L_i^2\left(\frac{1}{n_i s_i'} - \frac{1}{n_{i-1}s_i}\right)\right],\\
D &= \frac{1}{2}\sum_i H_i^2 h_i^2\left[\frac{b_i}{r_i^3}(n_i - n_{i-1}) + K_i L_i\left(\frac{1}{n_i s_i'} - \frac{1}{n_{i-1}s_i}\right)\right.\\
&\qquad\qquad\left. -K_i(K_i - L_i)\left(\frac{1}{n_i t_i'} - \frac{1}{n_{i-1}t_i}\right)\right],\\
E &= \frac{1}{2}\sum_i H_i^3 h_i\left[\frac{b_i}{r_i^3}(n_i - n_{i-1}) + L_i^2\left(\frac{1}{n_i s_i'} - \frac{1}{n_{i-1}s_i}\right)\right.\\
&\qquad\qquad\left. -L_i(K_i - L_i)\left(\frac{1}{n_i t_i'} - \frac{1}{n_{i-1}t_i}\right)\right],\\
F &= \frac{1}{2}\sum_i H_i h_i^3\left[\frac{b_i}{r_i^3}(n_i - n_{i-1}) + K_i L_i\left(\frac{1}{n_i s_i'} - \frac{1}{n_{i-1}s_i}\right)\right].
\end{aligned}
\right\}
\tag{13}
$$

These are the *Seidel formulae* for the primary aberration coefficients of a general centred system of refracting surfaces.*

Eqs. (13) express the primary aberration coefficients in terms of data specifying the passage of two paraxial rays through the system, namely a ray from the axial object point and a ray from the centre of the entrance pupil. It will be useful to summarize the relevant Gaussian formulae. Let d_i be the distance between the poles of the ith and the $(i+1)$th surface. Since the Gaussian image formed by the first i surfaces of the system is the object for the $(i+1)$th surface, we have the transfer formulae

$$
s_{i+1} = s_i' - d_i, \qquad t_{i+1} = t_i' - d_i. \tag{14}
$$

Given the distances s_1 and t_1 of the object plane and the plane of the entrance pupil from the pole of the first surface, the distances s_1', t_1', s_2, t_2, s_2', t_2', ... and the corresponding values K_1, L_1, K_2, L_2, ... may then be calculated successively from the Abbe relations

$$
\left.
\begin{aligned}
n_{i-1}\left(\frac{1}{r_i} - \frac{1}{s_i}\right) &= n_i\left(\frac{1}{r_i} - \frac{1}{s_i'}\right) = K_i,\\
n_{i-1}\left(\frac{1}{r_i} - \frac{1}{t_i}\right) &= n_i\left(\frac{1}{r_i} - \frac{1}{t_i'}\right) = L_i
\end{aligned}
\right\}
\tag{15}
$$

and from (14). We also have to determine the quantities h_i and H_i. If for simplicity the arbitrary length λ_0 in the plane of the entrance pupil is taken equal to *unity*, and the

* Seidel actually considered centred systems consisting of spherical surfaces only. The effects of asphericities (characterized by the constants b_i) were taken into account by later writers (see M. von Rohr, *Geometrical Investigation of Formation of Images in Optical Systems*, translated from German by R. Kanthack (London, HM Stationery Office, 1920), p. 344).

relations $D_i = t'_i - s'_i = t_{i+1} - s_{i+1}$ are used, it is seen from (9) that h_i and H_i may be calculated in succession from the relations*

$$\left.\begin{aligned} H_1 &= \frac{t_1}{n_0}, & h_1 &= \frac{s_1}{t_1 - s_1}, \\ \frac{H_{i+1}}{H_i} &= \frac{t_{i+1}}{t'_i}, & \frac{h_{i+1}}{h_i} &= \frac{s_{i+1}}{s'_i}. \end{aligned}\right\} \tag{16}$$

From (9) and from the Abbe relations (4) and (5) we obtain the following relation, which may be used as a check on the calculations and which will be needed later:

$$h_i H_i = \frac{s_i t_i}{n_{i-1}(t_i - s_i)} = \frac{s'_i t'_i}{n_i(t'_i - s'_i)} = \frac{1}{L_i - K_i}. \tag{17}$$

5.5.2 *The Seidel formulae in terms of one paraxial ray*

It is often desirable to express the primary aberration coefficients in a form which shows as clearly as possible the dependence of the coefficients on parameters which specify the optical system, and for this purpose (13) are not the most convenient set of formulae. To evaluate them, two rays must be traced through the system in accordance with the laws of Gaussian optics, namely a ray from the axial object point and a ray from the centre of the entrance pupil. It was, however, shown by Seidel that we may eliminate the data relating to the second ray and thus express the primary aberration coefficients by means of quantities which relate to the first ray only. Naturally one quantity depending on the position of the entrance pupil remains in the formulae, since alteration of the position of the stop will obviously influence the aberrations.

To eliminate the data referring to the ray from the centre of the entrance pupil, it is necessary first of all to express the heights H_i in terms of h_i. From (16), (14) and (17)

$$\begin{aligned} \frac{H_{i+1}}{h_{i+1}} - \frac{H_i}{h_i} &= \frac{1}{h_i h_{i+1}} H_i h_i \left(\frac{t_{i+1}}{t'_i} - \frac{s_{i+1}}{s'_i} \right) \\ &= \frac{1}{h_i h_{i+1}} H_i h_i d_i \left(\frac{1}{s'_i} - \frac{1}{t'_i} \right) \\ &= \frac{d_i}{n_i h_i h_{i+1}}. \end{aligned} \tag{18}$$

If quantities k_i are defined by

$$H_i = k_i h_i, \tag{19}$$

it follows from (18) that

$$k_{i+1} = k_1 + \sum_{j=1}^{i} \frac{d_j}{n_j h_j h_{j+1}}, \tag{20}$$

* The relation $h_{i+1}/h_i = s_{i+1}/s'_i$ implies that h_i is proportional to the height from the axis, at which a paraxial ray from the axial object point intersects the ith surface. The relation $H_{i+1}/H_i = t_{i+1}/t'_i$ gives a similar interpretation to H_i, in terms of a paraxial ray from the centre of the entrance pupil.

where, according to (16),

$$k_1 = \frac{H_1}{h_1} = \frac{t_1(t_1 - s_1)}{n_0 s_1}. \tag{21}$$

The other quantities in (13) which relate to the paraxial ray from the centre of the entrance pupil may also easily be expressed in terms of quantities referring to the other ray. From (17) and (19)

$$L_i = K_i + \frac{1}{k_i h_i^2}. \tag{22}$$

Further from (15) and (22)

$$\begin{aligned}
\frac{1}{n_i t_i'} - \frac{1}{n_{i-1} t_i} &= \frac{1}{n_i r_i} - \frac{1}{n_{i-1} r_i} - \left(\frac{1}{n_i^2} - \frac{1}{n_{i-1}^2}\right) L_i \\
&= \frac{1}{n_i}\left(\frac{K_i}{n_i} + \frac{1}{s_i'}\right) - \frac{1}{n_{i-1}}\left(\frac{K_i}{n_{i-1}} + \frac{1}{s_i}\right) - \left(\frac{1}{n_i^2} - \frac{1}{n_{i-1}^2}\right)\left(K_i + \frac{1}{k_i h_i^2}\right) \\
&= \frac{1}{n_i s_i'} - \frac{1}{n_{i-1} s_i} - \frac{1}{k_i h_i^2}\left(\frac{1}{n_i^2} - \frac{1}{n_{i-1}^2}\right).
\end{aligned} \tag{23}$$

Substitution from (19), (22) and (23) into (13) finally gives

$$\left.\begin{aligned}
B &= \frac{1}{2}\sum_i h_i^4 \frac{b_i}{r_i^3}(n_i - n_{i-1}) + h_i^4 K_i^2 \left(\frac{1}{n_i s_i'} - \frac{1}{n_{i-1} s_i}\right), \\
C &= \frac{1}{2}\sum_i h_i^4 k_i^2 \frac{b_i}{r_i^3}(n_i - n_{i-1}) + (1 + h_i^2 k_i K_i)^2 \left(\frac{1}{n_i s_i'} - \frac{1}{n_{i-1} s_i}\right), \\
D &= \frac{1}{2}\sum_i h_i^4 k_i^2 \frac{b_i}{r_i^3}(n_i - n_{i-1}) + h_i^2 k_i K_i (2 + h_i^2 k_i K_i) \left(\frac{1}{n_i s_i'} - \frac{1}{n_{i-1} s_i}\right) \\
&\quad - K_i \left(\frac{1}{n_i^2} - \frac{1}{n_{i-1}^2}\right), \\
E &= \frac{1}{2}\sum_i h_i^4 k_i^3 \frac{b_i}{r_i^3}(n_i - n_{i-1}) + k_i (1 + h_i^2 k_i K_i)(2 + h_i^2 k_i K_i) \left(\frac{1}{n_i s_i'} - \frac{1}{n_{i-1} s_i}\right) \\
&\quad - \frac{1 + h_i^2 k_i K_i}{h_i^2} \left(\frac{1}{n_i^2} - \frac{1}{n_{i-1}^2}\right), \\
F &= \frac{1}{2}\sum_i h_i^4 k_i \frac{b_i}{r_i^3}(n_i - n_{i-1}) + h_i^2 K_i (1 + k_i h_i^2 K_i) \left(\frac{1}{n_i s_i'} - \frac{1}{n_{i-1} s_i}\right).
\end{aligned}\right\} \tag{24}$$

This is the required form of the Seidel formulae. The position of the entrance pupil enters here only through the factor k_1 which is related to the distance t_1 of the plane of the entrance pupil from the pole of the first surface by (21). The quantities s_i, s_i', K_i and h_i are again computed from the appropriate relations in (14)–(16), whilst the k_i's are determined from (20).

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

5.5.3 Petzval's theorem

From the expressions for the coefficients of astigmatism and curvature we can derive an interesting relation due to Petzval. We have

$$
\begin{aligned}
C - D &= \frac{1}{2}\sum_i \left\{ [(1 + h_i^2 k_i K_i)^2 - 2h_i^2 k_i K_i - h_i^4 k_i^2 K_i^2]\left(\frac{1}{n_i s_i'} - \frac{1}{n_{i-1} s_i}\right) \right.\\
&\qquad \left. + K_i\left(\frac{1}{n_i^2} - \frac{1}{n_{i-1}^2}\right)\right\}\\
&= \frac{1}{2}\sum_i \left\{\frac{1}{n_i s_i'} - \frac{1}{n_{i-1} s_i} + K_i\left(\frac{1}{n_i^2} - \frac{1}{n_{i-1}^2}\right)\right\}\\
&= \frac{1}{2}\sum_i \frac{1}{r_i}\left(\frac{1}{n_i} - \frac{1}{n_{i-1}}\right),
\end{aligned}
\tag{25}
$$

where (15) was used.

According to §5.3, C and D determine the sagittal and the tangential field curvature. If n_α denotes the refractive index of the last medium, it follows from §5.3 (18) and §5.3 (21) that

$$
C = \frac{1}{4n_\alpha}\left(\frac{1}{R_t} - \frac{1}{R_s}\right), \qquad D = \frac{1}{2n_\alpha}\frac{1}{R_s}. \tag{26}
$$

Hence (25) may be written as

$$
\frac{1}{R_t} - \frac{3}{R_s} = 2n_\alpha \sum_i \frac{1}{r_i}\left(\frac{1}{n_i} - \frac{1}{n_{i-1}}\right). \tag{27}
$$

We thus obtain a relation between the curvatures of the two focal surfaces which contains only the radii of the refracting surfaces of the system and the corresponding refractive indices. If the system is free of spherical aberration, coma, and astigmatism, then a sharp image is formed on a surface of radius $R_s = R_t = R$; and the radius of this surface is according to (27) given by

$$
\frac{1}{R} = -n_\alpha \sum_i \frac{1}{r_i}\left(\frac{1}{n_i} - \frac{1}{n_{i-1}}\right). \tag{28}
$$

This result is known as *Petzval's theorem*.

The condition

$$
\sum_i \frac{1}{r_i}\left(\frac{1}{n_i} - \frac{1}{n_{i-1}}\right) = 0 \tag{29}
$$

is known as *Petzval's condition*. It is a necessary condition for *flatness of the field*. It should, however, be remembered that this condition belongs to the domain of the Seidel theory; it loses its significance outside this domain.

Whether or not aberrations are present, the spherical surface which touches the two focal surfaces at their common axial point and whose radius R is given by (28) is called the *Petzval surface*. According to (27) and (28) the radii of curvature of the

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

sagittal focal surface, the tangential focal surface, and the Petzval surface are related by

$$\frac{3}{R_s} - \frac{1}{R_t} = \frac{2}{R}. \tag{30}$$

5.6 Example: The primary aberrations of a thin lens

The Seidel formulae will now be used to find the primary aberration coefficients of a thin lens of refractive index n, situated in air (vacuum). In this case

$$n_0 = n_2 = 1, \qquad n_1 = n. \tag{1}$$

Since the thickness d of the lens is assumed to be negligible, we have according to §5.5 (14)

$$s_2 = s_1'; \tag{2}$$

and §5.5 (15) gives

$$\left.\begin{aligned} \frac{1}{r_1} - \frac{1}{s_1} &= n\left(\frac{1}{r_1} - \frac{1}{s_1'}\right) = K_1, \\ n\left(\frac{1}{r_2} - \frac{1}{s_2}\right) &= \left(\frac{1}{r_2} - \frac{1}{s_2'}\right) = K_2. \end{aligned}\right\} \tag{3}$$

Further from §5.5 (16), §5.5 (21) and §5.5 (20), using (2)

$$\left.\begin{aligned} &h_1 = h_2 = h, \qquad k_1 = k_2 = k, \\ &\text{where} \\ &h = \frac{s_1}{t_1 - s_1}, \qquad k = \frac{t_1(t_1 - s_1)}{s_1}. \end{aligned}\right\} \tag{4}$$

It will be convenient to characterize the lens by a number of simple parameters. Let $\mathcal{P}$ be the power of the lens (see §4.4 (35) and §4.4 (36))

$$\mathcal{P} = \frac{1}{f} = (n-1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right), \tag{5}$$

and let

$$\sigma = (n-1)\left(\frac{1}{r_1} + \frac{1}{r_2}\right). \tag{6}$$

From (3),

$$(n-1)\frac{1}{r_1} = \frac{n}{s_1'} - \frac{1}{s_1}, \qquad (n-1)\frac{1}{r_2} = \frac{n}{s_2} - \frac{1}{s_2'}, \tag{7}$$

so that, because of (2),

$$\frac{1}{s_2'} - \frac{1}{s_1} = \mathcal{P}. \tag{8}$$

This relation will be written in the form

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

$$-\frac{1}{s_1}-\frac{\mathcal{P}}{2}=-\frac{1}{s_2'}+\frac{\mathcal{P}}{2}=\mathcal{K}; \tag{9}$$

$\mathcal{K}$ plays a similar role for the lens as K does for a single surface. Accordingly, $\mathcal{K}$ will be referred to as the *Abbe invariant of the lens*.

The deformation coefficients b_1 and b_2 of the two surfaces of the lens will be seen later to enter the formulae only through the quantity

$$\beta=(n-1)\left(\frac{b_1}{r_1^3}-\frac{b_2}{r_2^3}\right); \tag{10}$$

this quantity may be called the *deformation coefficient of the lens*.

We now express the various quantities which enter the Seidel formulae §5.5 (24) in terms of the parameters $\mathcal{P}$, σ, β and $\mathcal{K}$. The first three specify the lens and the last specifies the position of the object. First, we obtain from (3)

$$K_1=\mathcal{K}+\frac{\sigma+n\mathcal{P}}{2(n-1)},\qquad K_2=\mathcal{K}+\frac{\sigma-n\mathcal{P}}{2(n-1)}. \tag{11}$$

Further

$$\left.\begin{aligned}\frac{1}{ns_1'}-\frac{1}{s_1}&=\frac{n^2-1}{n^2}\mathcal{K}+\frac{\sigma}{2n^2}+\frac{\mathcal{P}}{2},\\ \frac{1}{s_2'}-\frac{1}{ns_2}&=-\frac{n^2-1}{n^2}\mathcal{K}-\frac{\sigma}{2n^2}+\frac{\mathcal{P}}{2}.\end{aligned}\right\} \tag{12}$$

Substitution into the Seidel formulae §5.5 (24) leads to the following expressions for the primary aberration coefficients of a thin lens:

$$\left.\begin{aligned}B&=h^4U,\\ F&=h^4kU+h^2V,\\ C&=h^4k^2U+2h^2kV+\tfrac{1}{2}\mathcal{P},\\ D&=h^4k^2U+2h^2kV+\frac{n+1}{2n}\mathcal{P},\\ E&=h^4k^3U+3h^2k^2V+k\frac{3n+1}{2n}\mathcal{P},\end{aligned}\right\} \tag{13}$$

where

$$\left.\begin{aligned}U&=\frac{1}{2}\beta+\frac{n^2}{8(n-1)^2}\mathcal{P}^3-\frac{n}{2(n+2)}\mathcal{K}^2\mathcal{P}\\ &\quad+\frac{1}{2n(n+2)}\mathcal{P}\left[\frac{n+2}{2(n-1)}\sigma+2(n+1)\mathcal{K}\right]^2,\\ V&=\frac{1}{2n}\mathcal{P}\left[\frac{n+1}{2(n-1)}\sigma+(2n+1)\mathcal{K}\right].\end{aligned}\right\} \tag{14}$$

Only the case when the *entrance pupil coincides with the lens* ($t_1=0$) will be considered here. According to (4) we then have

$$h=-1,\qquad k=0,$$

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

and (13) become

$$\left.\begin{aligned} B &= U, \\ F &= V, \\ \frac{1}{2}(C+D) &= \frac{2n+1}{4n}\mathcal{P}, \\ \frac{1}{2}(C-D) &= -\frac{1}{4n}\mathcal{P}, \\ E &= 0. \end{aligned}\right\} \tag{15}$$

The quantities U and V defined by (14) are seen to represent the spherical aberration and coma of a lens working with the stop in the plane of the lens. Such a lens is free from distortion ($E = 0$) but astigmatism and field curvature are always present ($C \neq 0$, $D \neq 0$).

Let us examine whether, for a given lens which is working with the stop in the plane of the lens, there exists a pair of aplanatic points ($B = F = 0$). For coma to be absent we must have, according to (15), $V = 0$, i.e.

$$\frac{n+1}{2(n-1)}\sigma + (2n+1)\mathcal{K} = 0. \tag{16}$$

Eq. (9) then gives the object distance s_1:

$$\frac{1}{s_1} = -\frac{1}{2}\mathcal{P} + \frac{n+1}{2(n-1)(2n+1)}\sigma. \tag{17}$$

With this choice of s_1, the spherical aberration of the lens is fully determined. Hence in general there is no aplanatic point pair for a given lens working with the stop at the lens. However, if only $\mathcal{P}$ and σ are given, one may eliminate the spherical aberration by an appropriate choice of the deformation coefficient β.

Conversely, if the object distance s_1 and the focal length $f = 1/\mathcal{P}$ are given one may first choose the radii r_1 and r_2 in such a way that coma vanishes. According to (17), (5) and (6) we have in this case,

$$\left.\begin{aligned} \frac{1}{r_1} &= \frac{2n+1}{n+1}\frac{1}{s_1} + \frac{n^2}{n^2-1}\mathcal{P}, \\ \frac{1}{r_2} &= \frac{2n+1}{n+1}\frac{1}{s_1} + \frac{n^2-n-1}{n^2-1}\mathcal{P}. \end{aligned}\right\} \tag{18}$$

This gives, for example, for an object at infinity ($s_1 = \infty$) and with $n = 1.5$,

$$r_1 = \frac{5}{9}\frac{1}{\mathcal{P}} = \frac{5}{9}f, \qquad r_2 = -5\frac{1}{\mathcal{P}} = -5f. \tag{19}$$

Such a coma-free lens is shown in Fig. 5.10, where C_1 and C_2 are the centres of curvature of the two surfaces of the lens.

Next, the spherical aberration may be eliminated by a suitable choice of β. When the object is at infinity, we have, according to (9), $\mathcal{K} = -\mathcal{P}/2$, and hence from (16) $\sigma = (2n+1)(n-1)\mathcal{P}/(n+1)$; the condition $U = 0$ for the absence of spherical aberration then gives the following expression for β:

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

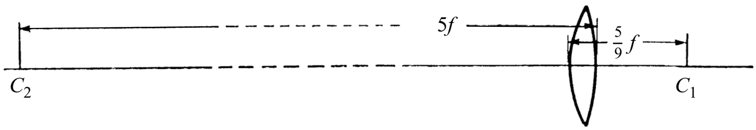

Fig. 5.10 Coma-free thin lens (object at infinity).

$$-\frac{1}{2}\beta = \mathcal{P}^3\left\{\frac{n^2}{8(n-1)^2} - \frac{n}{8(n+2)} + \frac{1}{2n(n+2)}\left[\frac{(n+2)(2n+1)}{2(n+1)} - (n+1)\right]^2\right\},$$

whence

$$\beta = -\frac{n^3}{(n^2-1)^2}\mathcal{P}^3. \tag{20}$$

For $n = 1.5$ this gives

$$\beta = -\tfrac{5}{2}\cdot\tfrac{4}{5}\mathcal{P}^3, \tag{21}$$

so that by (19) and (10)

$$729b_1 + b_2 = -540. \tag{22}$$

This may be satisfied, for example, by taking $b_1 = b_2 = -0.74$.

Instead of eliminating coma, one may first attempt to reduce the spherical aberration as much as possible. It is then found that, if the focal length is positive, the spherical aberration may be removed completely only if the object lies in a certain limited range; for $n = 1.5$ this range extends from $0.36f$ to $0.44f$ behind the lens. For any other position of the object a certain amount of spherical aberration will always be present.

In order to affect the curvature of field, the lens must be used with a stop placed in an appropriate position, which may be determined from the general formulae (13). For a lens with a positive focal length and with $n = 1.5$, the curvature may be eliminated only if the object lies within a range extending from one focal length in front of the lens to half of the focal length behind it.

For a system which consists of a number of lenses, the calculations become more complicated. The removal of the chromatic aberration is then of primary importance. The primary aberrations of an achromatic telescope objective which consists of two cemented thin lenses may still be investigated with relative ease, and it is found that all the primary aberrations except astigmatism and curvature of field may be eliminated; it may therefore be used over a narrow field only (at most up to about 3° for an $f/10$ system).

5.7 The chromatic aberration of a general centred lens system

Chromatic aberration has already been briefly discussed in §4.7. Here explicit expressions will be derived for the chromatic change in the position of the image, and in the magnification, using the general formalism of the preceding sections.* The system

* We again follow substantially the analysis of K. Schwarzschild, *Abh. Königl. Ges. Wiss. Göttingen, Math.-phys. Kl.*, **4** (1905–1906), No. 3. Reprinted in *Astr. Mitt. Königl. Sternwarte Göttingen* (1905).

will be assumed to be perfectly corrected for monochromatic light. This is a permissible idealization if effects of the leading order only are being considered; for changes in the monochromatic, aberrations associated with a small change in the wavelength may be assumed to be small in comparison with the magnitudes of the aberrations themselves; in general these changes will therefore be of the same order of magnitude as the terms neglected in the Seidel theory.

According to §5.5 (15) and §5.5 (14),

$$n_{i-1}\left(\frac{1}{r_i}-\frac{1}{s_i}\right)=n_i\left(\frac{1}{r_i}-\frac{1}{{s_i}'}\right)=K_i, \tag{1}$$

$$s_{i+1}=s_i'-d_i. \tag{2}$$

If δ denotes the change in a quantity associated with a small change $\delta\lambda$ in the wavelength, we have from (1) and (2),

$$n_i\frac{\delta s_i'}{s_i'^2}-n_{i-1}\frac{\delta s_i}{{s_i}^2}=-\left(\frac{\delta n_i}{n_i}-\frac{\delta n_{i-1}}{n_{i-1}}\right)K_i, \tag{3}$$

$$\delta s_{i+1}=\delta s_i'. \tag{4}$$

Multiplying (3) by ${h_i}^2$ and using the relation

$$\frac{h_{i+1}}{h_i}=\frac{s_{i+1}}{s_i'}, \tag{5}$$

we find that

$$n_i\left(\frac{h_{i+1}}{s_{i+1}}\right)^2\delta s_{i+1}-n_{i-1}\left(\frac{h_i}{s_i}\right)^2\delta s_i=-h_i^2K_i\left(\frac{\delta n_i}{n_i}-\frac{\delta n_{i-1}}{n_{i-1}}\right). \tag{6}$$

Let α denote the number of surfaces in the system. Then by adding all the equations of the form (6) it follows that

$$n_\alpha\left(\frac{h_{\alpha+1}}{s_{\alpha+1}}\right)^2\delta s_{\alpha+1}=n_\alpha\left(\frac{h_\alpha}{s_\alpha'}\right)^2\delta s_\alpha'=n_0\left(\frac{h_1}{s_1}\right)^2\delta s_1-\sum_{i=1}^{\alpha}h_i^2K_i\left(\frac{\delta n_i}{n_i}-\frac{\delta n_{i-1}}{n_{i-1}}\right).$$

Since the position of the object is independent of the wavelength, $\delta s_1=0$, and we obtain the following expression for the *chromatic change* $\delta s_\alpha'$ *in the position of the image:*

$$\delta s_\alpha'=-\frac{1}{n_\alpha}\left(\frac{s_\alpha'}{h_\alpha}\right)^2\sum_{i=1}^{\alpha}{h_i}^2K_i\left(\frac{\delta n_i}{n_i}-\frac{\delta n_{i-1}}{n_{i-1}}\right). \tag{7}$$

Next consider the chromatic change in the magnification M of the system. We have

$$M=\frac{l_1}{l_0}\frac{l_2}{l_1}\frac{l_3}{l_2}\cdots\frac{l_\alpha}{l_{\alpha-1}}. \tag{8}$$

Now by §5.5 (8),

$$\frac{l_i}{l_{i-1}}=\frac{n_{i-1}}{n_i}\frac{s_i'}{s_i},$$

so that

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

$$M = \frac{n_0}{n_\alpha}\left(\frac{s_1'}{s_1}\frac{s_2'}{s_2}\frac{s_3'}{s_3}\cdots\frac{s_\alpha'}{s_\alpha}\right). \tag{9}$$

By logarithmic differentiation,

$$\begin{aligned}\frac{\delta M}{M} &= \frac{\delta n_0}{n_0} - \frac{\delta n_\alpha}{n_\alpha} + \sum_{i=1}^{\alpha}\left(\frac{\delta s_i'}{s_i'} - \frac{\delta s_i}{s_i}\right)\\ &= \left(\frac{\delta n_0}{n_0} - \frac{\delta n_\alpha}{n_\alpha}\right) + \frac{\delta s_\alpha'}{s_\alpha'} - \frac{\delta s_1}{s_1} - \sum_{i=2}^{\alpha}\delta s_i\left(\frac{1}{s_i} - \frac{1}{s_{i-1}'}\right).\end{aligned} \tag{10}$$

The sum entering this expression may be rewritten as follows

$$\sum_{i=2}^{\alpha}\delta s_i\left(\frac{1}{s_i} - \frac{1}{s_{i-1}'}\right) = \sum_{i=2}^{\alpha}\frac{d_{i-1}\delta s_i}{s_i s_{i-1}'} = \sum_{i=2}^{\alpha}\frac{d_{i-1}}{s_i^2}\frac{h_i}{h_{i-1}}\delta s_i.$$

But by §5.5 (18),

$$\frac{d_{i-1}}{h_{i-1}} = (k_i - k_{i-1})h_i n_{i-1},$$

so that

$$\sum_{i=2}^{\alpha}\delta s_i\left(\frac{1}{s_i} - \frac{1}{s_{i-1}'}\right) = \sum_{i=2}^{\alpha}h_i^2 n_{i-1}\left(\frac{k_i - k_{i-1}}{s_i^2}\right)\delta s_i,$$

or, re-arranging according to the k_i's,

$$\begin{aligned}\sum_{i=2}^{\alpha}\delta s_i\left(\frac{1}{s_i} - \frac{1}{s_{i-1}'}\right) = &- n_0\left(\frac{h_1}{s_1}\right)^2 k_1\delta s_1 + n_\alpha\left(\frac{h_{\alpha+1}}{s_{\alpha+1}}\right)^2 k_\alpha\delta s_{\alpha+1}\\ &- \sum_{i=1}^{\alpha}k_i\left[n_i\left(\frac{h_{i+1}}{s_{i+1}}\right)^2\delta s_{i+1} - n_{i-1}\left(\frac{h_i}{s_i}\right)^2\delta s_i\right].\end{aligned}$$

Re-writing the second term on the right with the help of (4) and (5), and the last term with the help of (6), we obtain

$$\begin{aligned}\sum_{i=2}^{\alpha}\delta s_i\left(\frac{1}{s_i} - \frac{1}{s_{i-1}'}\right) = &- n_0\left(\frac{h_1}{s_1}\right)^2 k_1\delta s_1 + n_\alpha\left(\frac{h_\alpha}{s_\alpha'}\right)^2 k_\alpha\delta s_\alpha'\\ &+ \sum_{i=1}^{\alpha}h_i^2 k_i K_i\left(\frac{\delta n_i}{n_i} - \frac{\delta n_{i-1}}{n_{i-1}}\right).\end{aligned}$$

On substituting this expression into (10) and using the fact that $\delta s_1 = 0$, we obtain

$$\frac{\delta M}{M} = \left(\frac{\delta n_0}{n_0} - \frac{\delta n_\alpha}{n_\alpha}\right) + \frac{\delta s_\alpha'}{s_\alpha'}\left(1 - \frac{n_\alpha}{s_\alpha'}{h_\alpha}^2 k_\alpha\right) - \sum_{i=1}^{\alpha}h_i^2 k_i K_i\left(\frac{\delta n_i}{n_i} - \frac{\delta n_{i-1}}{n_{i-1}}\right).$$

But, by §5.5 (22) and §5.5 (15),

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press

$$\frac{\delta s'_\alpha}{s'_\alpha}\left[1-\frac{n_\alpha}{s'_\alpha}h_\alpha^2 k_\alpha\right]=\frac{\delta s'_\alpha}{s'_\alpha}\left[1-\frac{n_\alpha}{s'_\alpha(L_\alpha-K_\alpha)}\right]$$
$$=\frac{\delta s'_\alpha}{s'_\alpha-t'_\alpha},$$

so that we finally obtain the following expression for the *chromatic change δM in the magnification*

$$\frac{\delta M}{M}=\left(\frac{\delta n_0}{n_0}-\frac{\delta n_\alpha}{n_\alpha}\right)+\frac{\delta s'_\alpha}{s'_\alpha-t'_\alpha}-\sum_{i=1}^{\alpha}h_i^2 k_i K_i\left(\frac{\delta n_i}{n_i}-\frac{\delta n_{i-1}}{n_{i-1}}\right). \tag{11}$$

If the refractive indices of the object and the image space are equal, as is usually the case, the first two terms on the right-hand side cancel out. If the system is achromatized with respect to the position of the image, $\delta s'_\alpha$ [given by (7)] also vanishes, and we then obtain the following two conditions for the achromatization of the position of the image and of the magnification

$$\sum_{i=1}^{\alpha}h_i^2 K_i\left(\frac{\delta n_i}{n_i}-\frac{\delta n_{i-1}}{n_{i-1}}\right)=0, \tag{12}$$

and

$$\sum_{i=1}^{\alpha}h_i^2 k_i K_i\left(\frac{\delta n_i}{n_i}-\frac{\delta n_{i-1}}{n_{i-1}}\right)=0. \tag{13}$$

For a *thin lens in air*, we obtain, on using the appropriate formulae from §5.6, the following expressions for the two sums:

$$\Sigma h_i^2 K_i\left(\frac{\delta n_i}{n_i}-\frac{\delta n_{i-1}}{n_{i-1}}\right)=h^2\mathcal{P}\frac{\delta n}{n-1}, \tag{14}$$

and

$$\Sigma h_i^2 k_i K_i\left(\frac{\delta n_i}{n_i}-\frac{\delta n_{i-1}}{n_{i-1}}\right)=h^2 k\mathcal{P}\frac{\delta n}{n-1}, \tag{15}$$

where h, k, and the power $\mathcal{P}$ of the lens are given by §5.6 (4) and §5.6 (5). The corresponding expressions for the chromatic change in the position of the image and in the magnification are then obtained by substituting from (14) into (7) and from (15) into (11) respectively.

https://doi.org/10.1017/9781108769914.008 Published online by Cambridge University Press