Journal of
Synchrotron Radiation

ISSN 0909-0495

Received 1 May 2007
Accepted 2 February 2008

Aberration theory of plane-symmetric grating systems

Li-Jun Lu

Department of Fine Mechanical Engineering, Shanghai University, Shanghai 200072, People's Republic of China. E-mail: lulijun@shu.edu.cn

Aberration theory of plane-symmetric optical systems of mirror and grating has been developed based on the wavefront aberration method. A toroidal reference wavefront surface is used to define the wavefront aberration. Based on the ray geometry, the coordinate mapping relationships of the ray between the optical element and the incident and aberrated wavefronts are derived using a polynomial-fit method; this enables the resultant coefficients of the wavefront and the transverse aberration to be kept to the fourth-order accuracy of the aperture-ray coordinates. By setting up the transfer equations of the field and aperture rays, the contribution to wavefront aberrations from each mirror and grating can be added to make the aberration calculation of multi-element systems feasible. The theory is validated by the analytic formulae of the spot diagram.

Keywords: aberration; synchrotron optics; mirror and grating; aspheric optics.

1. Introduction

Recent developments in synchrotron radiation instrumentation are toward higher spectral and/or spatial resolution X-ray and ultraviolet (XUV) optics. They usually consist of a series of mirrors and gratings arranged with a plane of symmetry. The initial determination and optimization of optical parameters of these instruments need aberration analysis of the optical systems although ray tracing can provide a powerful tool for evaluating XUV optics. The grating theory developed by Beutler and Namioka has often been used to analyze these optical systems, but it is essentially suitable for a single grating, and its high-order aberration coefficients can correctly describe the aberration of the grating only in the case where the angle of diffraction is zero and the meridional and sagittal focal curves cross on the grating normal (Beutler, 1945; Noda *et al.*, 1974). A number of attempts have been made on the aberration analysis of multi-element systems. The analytic formulae of the spot diagram can reproduce the ray-traced spot diagram with a high degree of accuracy (Namioka *et al.*, 1994); its aberration expressions on the image plane are correct; and recently the mathematical tool of Lie optics has been applied to the grating imaging systems (Goto & Kurosaki, 1993). The aberration coefficients for double-element systems from the two methods have been studied (Masui & Namioka, 1999*a*,*b*; Namioka *et al.*, 2001; Palmer *et al.*, 1998*a*,*b*), but they are very complex and difficult to be used to optimize the optical parameters.

An important advance in the wavefront aberration (WFA) theory for multi-element systems was made by Chrisp (1983*a*,*b*). He defines the wavefront aberration with reference to an astigmatic wavefront surface and an aperture stop displaced from the grating center to make possible the aberration calculation of the off-plane object. However, if we apply the theory to calculate the image of mirror–grating systems, we find that it can give a good agreement to ray-tracing results only with a small acceptance angle of the ray pencil; besides, the calculation of the transverse aberration is limited to the case when the image plane is just on the focus position and the wavefront of the exit ray pencil is almost stigmatic. These limitations result from the following. (i) In finding the transverse aberration with the wavefront aberration, the deviation angle of the aberrated ray from the reference wavefront surface normal in the meridional and sagittal directions, θ_x', θ_y', are determined by

$$\theta_x' = \frac{\mathrm{d}W}{\mathrm{d}x_0'}, \qquad \theta_y' = \frac{\mathrm{d}W}{\mathrm{d}y_0'}, \tag{1}$$

where x_0', y_0' are the coordinates of the ray position on the aberrated wavefront at the exit pupil, whereas Chrisp (1983*a*,*b*) used those on the reference wavefront, and it will lead to a significant error in the image when a large wavefront aberration exists. This may be due to the difficulty in the wavefront aberration method that the intermediate variables x_0', y_0' depend on the ultimate variables, the image coordinates x', y'. (ii) In finding the wavefront aberration coefficients, a few significant errors exist; for example, in finding the focal line of the reference wavefront, the position of the aperture-stop center and in dealing with wavefront aberration. Namioka *et al.* (1994) have indicated that the aberrations from Chrisp (1983*a*,*b*) significantly deviate from those of the analytic formulae of the spot diagram or ray tracing calculations.

In this study we have developed an accurate third-order aberration theory of plane-symmetric mirror or grating systems based on Chrisp's work. The theory is applicable to the aberration analysis of the optical figure of a plane, cylindrical, spherical, toroidal, ellipsoidal and paraboloidal. It circumvents the above limitations of Chrisp (1983*a*,*b*) by two measures: (i) a toroidal reference wavefront surface is used to define the wavefront aberration in §2.2; (ii) the mapping relationships of the coordinates of the ray between the optical element and the wavefront surfaces are obtained by solving the complex ray equations using the polynomial-fit method. Consequently, the wavefront and transverse aberration coefficients are derived in §2.3 to §2.4 and §3. In §4, where the transfer equations of the field and aperture rays between optical elements are set up, the wavefront aberration of multi-element systems is investigated. Finally the theory is validated with respect to the analytic formulae of the spot diagram (Namioka *et al.*, 1994).

2. Wavefront aberrations

2.1. Definition of reference coordinate systems and rays

For convenience in the following discussions it is necessary to make clear the reference coordinate systems and rays used to describe the optical system. Fig. 1 shows a plane-symmetry optical system with an off-plane object point S_0. The coordinate systems, xyz, $x_0y_0z_0$ and $x_0'y_0'z_0'$ corresponding to the aperture stop, the entrance and exit pupils are stipulated to have a common origin, $\bar{P}$, the center of the aperture stop; the axes of z, z_0 and z_0' are along the grating surface normal at the vertex, the incident and exit principal rays, respectively; the axes of x, x_0, x_0' are all on the horizontal plane. The three coordinate systems are used to describe the position of the ray on the optical surface, the wavefronts on the entrance and exit pupils, respectively. The ray $S_0\bar{P}S_0'$ is referred to as the principal ray; and the base ray $O'OO''$, lying in the symmetry plane, comes from the central field point O' and is diffracted at the grating vertex O. In the following, the wavefront aberration, the positions of the object and image plane and the parameters of the optical parameters are specified with respect to this ray. The field ray S_0OS_0' is displaced from the symmetry plane and specified with a field angle of u and u' in the object and image space.

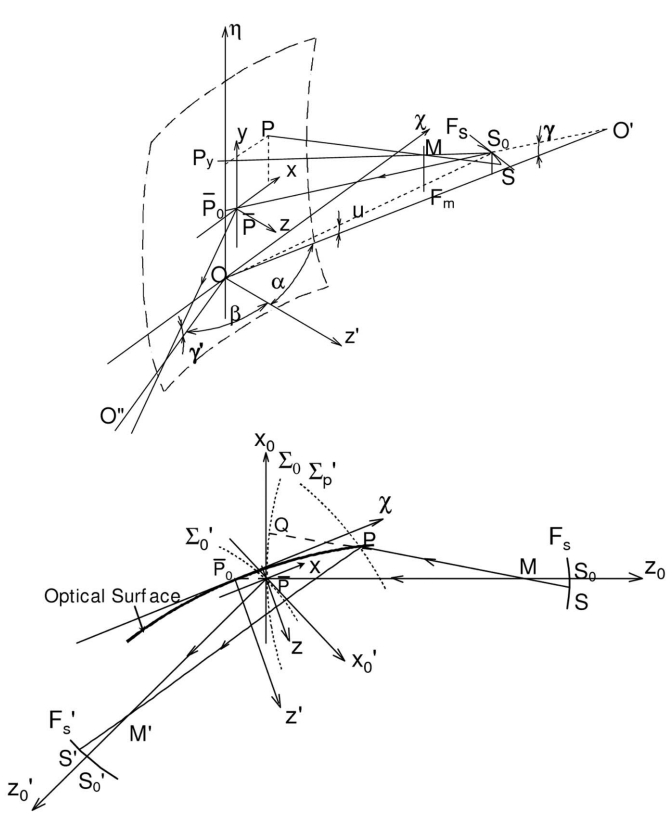

Figure 1
Optical scheme of a plane-symmetry optical system and its reference coordinate systems, reference rays and meridional and sagittal foci.

Since there is no optical axis on a plane-symmetric optical system as in the rotationally symmetric one, the field coordinate is defined with respect to the base ray, and the aperture coordinate with respect to the principal ray. The basal theory denotes the aberration terms obtained by using field coordinates. The first- and second-order wavefront aberrations determine the principal ray and the focal positions on it (*i.e.* F_m and F_s as in Fig. 1), respectively. The parabasal theory below discusses the aberrations based on the definite principal ray and the foci.

2.2. Reference wavefront surface

The wavefront aberration is defined as the displacement of the wavefront from a specified reference surface. For rotationally lens systems a spherical surface is used, and the position of its center is determined from the paraxial theory (Gaussian optics). For plane-symmetric optical systems, especially grazing-incidence optics, very large astigmatism exists; the wavefront may be cylindrical. Chrisp (1983*a*,*b*) adopts an astigmatic reference surface to define the wavefront aberration, but it does not specify the figure of this astigmatic reference surface. A specific figure is required in order to obtain the exact information of its focal lines. A toroidal figure is chosen here because it is relatively simple with different meridional and sagittal foci.

Fig. 2 shows a toroidal reference surface with its vertex on the center of the entrance pupil, $\bar{P}$; $Q(x_0, y_0, z_0)$ is any point on the surface; the z_0 coordinate of Q is given by the toroidal equation,

$$z_0 = r_m - r_m\left\{1 - \frac{(x_0^2 + y_0^2)}{r_m^2} + \frac{2r_s}{r_m}\left(\frac{r_s}{r_m} - 1\right)\left[1 - \left(1 - \frac{y_0^2}{r_s^2}\right)^{1/2}\right]\right\}^{1/2}, \qquad (2)$$

where r_m and r_s are the major and minor radii of the toroidal surface, which here mean the distances of the meridional and sagittal focal lines from the optical element. The direction of the ray is defined as the normal of the wavefront; thus the ray

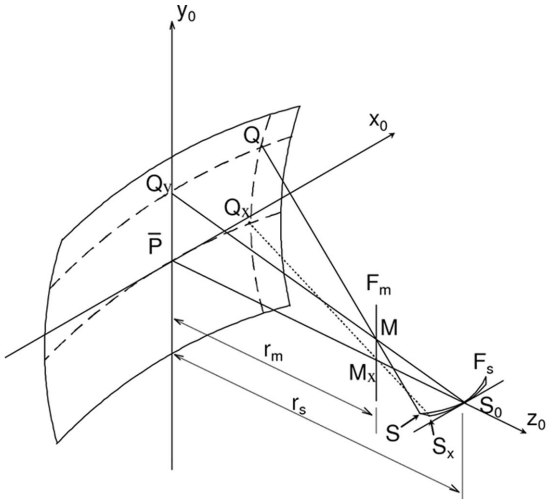

Figure 2
Toroidal reference wavefront surface and its foci on the meridional and sagittal plane.

equation at point Q, from equation (2), is formulated as follows,

$$\frac{x-x_0}{\frac{x_0}{r_m-z_0}}=\frac{y-y_0}{\frac{y_0}{r_m-z_0}\left[1+\frac{r_m-r_s}{\left(r_s^2-y_0^2\right)^{1/2}}\right]}=\frac{z-z_0}{-1}. \tag{3}$$

Its projection trajectory F_m on the plane of $x = 0$ is then

$$y_m=\frac{(r_s-r_m)y_0}{\left(r_s^2-y_0^2\right)^{1/2}},\qquad z_m=r_m, \tag{4}$$

and that of F_s on the plane of $y = 0$ is

$$x_s=x_0-\frac{(z-z_0)x_0}{r_m-z_0},\qquad z_s=z_0+\frac{r_m-z_0}{1+\frac{r_m-r_s}{\left(r_s^2-y_0^2\right)^{1/2}}}. \tag{5}$$

F_m and F_s are referred to as the meridional and sagittal foci, and $M(0, y_m, z_m)$ and $S(x_s, 0, z_s)$ represent the corresponding projection points on them to point Q. Expanding y_m, x_s, z_s in equations (4) and (5) as the power series of x_0 and y_0 to the fourth order yields

$$\begin{aligned}
y_m&=\frac{(r_s-r_m)}{r_s}y_0+\frac{(r_s-r_m)}{2r_s^3}y_0^3,\\
x_s&=\frac{(r_m-r_s)}{r_m}x_0+\frac{(r_m-r_s)}{2r_m^2r_s}x_0y_0^2,\\
z_s&=r_s+\frac{(r_m-r_s)}{2r_m^2}x_0^2+\frac{(r_m-r_s)}{8r_m^4}x_0^4+\frac{(r_m-r_s)}{2r_m^3r_s}x_0^2y_0^2.
\end{aligned} \tag{6}$$

The above equations show that the meridional focus, F_m, is related to only y_0 and it is a straight line perpendicular to the z_0 axis, while the sagittal focus, F_s, is related to both x_0 and y_0 and looks like a curved strip as shown in Fig. 2. The first term on the right-hand side of the above equations just describes the meridional and sagittal focal lines given by Chrisp (1983*a*,*b*). However, the remaining terms are found to have contributions to high-order wavefront aberration coefficients in the later discussions.

The point Q_y on the reference surface is defined to be on the y_0–z_0 plane with the same y_0 coordinate as Q. Setting $x_0 = 0$ in (6) indicates that the ray passing through Q_y will intersect the z_0 axis at point $S_0(0, 0, r_s)$ and the line QS at point $M(0, y_m, r_m)$; consequently, the lengths of QM, Q_yM, Q_yS_0, and $\bar{P}S_0$ can be found up to the fourth-order accuracy of x_0, y_0,

$$\begin{aligned}
QM&=Q_yM=r_m-\frac{(r_s-r_m)(4r_s^2+3y_0^2)y_0^2}{8r_s^4},\\
Q_yS_0&=\bar{P}S_0=r_s.
\end{aligned} \tag{7}$$

2.3. Wavefront aberration

Chrisp (1983*a*,*b*) has defined the wavefront aberration of a plane-symmetric optical system as given in the following. Fig. 3 shows ray paths through a plane-symmetry optical system; $\bar{P}^+$, $\bar{P}$ are the centers of the entrance and exit pupils; the z_0 axis is along the principal ray $S_0\bar{P}^+\bar{P}S_0'$; the x_0, x_0' axes are on the horizontal plane passing through $\bar{P}^+$, $\bar{P}$; the y_0, y_0' axes are perpendicular to the x_0, x_0' and the principal ray. Parabasal rays in the x_0–z_0 plane pass through the meridional focus F_m', and the y_0–z_0 plane contains the corresponding sagittal focus F_s'.

A perfect toroidal wavefront Σ^+ enters the system and emerges as the aberrated wavefront Σ. The notation (+) is used to distinguish the corresponding points, $\bar{P}^+$, $Q_y{}^+$, Q^+ on the toroidal reference surface of the object pencil. The aberrations of the image pencil cause its rays to be displaced from the parabasal focal lines, F_m' and F_s', which correspond to an exit toroidal reference surface Σ_r. The aberrated wavefront Σ will be compared with Σ_r to define the wavefront aberration.

The wavefront aberration W of a ray passing through a point P in the system is defined as the distance along the ray between the aberrated wavefront and the reference surface. So for ray SQ^+PQS', the wavefront aberration is

$$W=\langle Q_rQ\rangle=\langle Q^+Q_r\rangle-\langle Q^+Q\rangle. \tag{8}$$

Using $\langle\bar{P}^+\bar{P}\rangle = \langle Q^+Q_r\rangle$, since they are the ray distances between the same wavefronts, then the wavefront aberration is equivalent to

$$W=\langle\bar{P}^+\bar{P}\rangle-\langle Q^+Q\rangle. \tag{9}$$

To obtain the wavefront aberration, an auxiliary ray $S_0Q_y{}^+Q_yS_0'$ is constructed. This ray lies in the sagittal plane and intersects the four focal lines. It passes $Q_y{}^+$, Q_y as defined

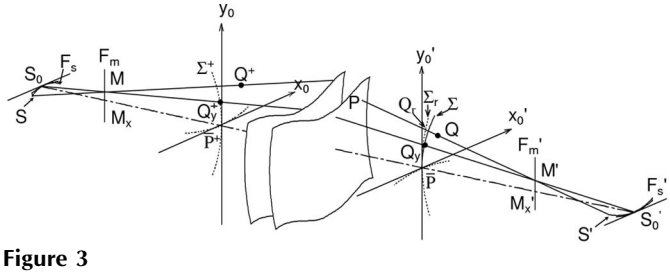

Figure 3
Ray paths through the plane-symmetry optical system, the wavefronts on the entrance and exit pupil and their corresponding foci.

in the last section. Splitting the aberration into the path difference between the principal and auxiliary ray, and the path difference between the auxiliary and the general ray, (9) can be written as

$$W = \langle \bar{P}^+ \bar{P} \rangle - \langle Q_y^+ Q_y \rangle + \langle Q_y^+ Q_y \rangle - \langle Q^+ Q \rangle. \tag{10}$$

The additional ray distances to the focal lines as indicated in (7) can now be added; thus the first two terms in the above equation are then converted into

$$\langle \bar{P}^+ \bar{P} \rangle - \langle Q_y^+ Q_y \rangle = \langle S_0 \bar{P} S_0' \rangle - \langle S_0 Q_y S_0' \rangle, \tag{11}$$

and the latter two terms into

$$\langle Q_y^+ Q_y \rangle - \langle Q^+ Q \rangle = \langle M Q_y M' \rangle - \langle M Q M' \rangle. \tag{12}$$

Substitution of (11) and (12) into (10) gives the final result of the wavefront aberration of the ray through P,

$$W = \langle S_0 \bar{P} S_0' \rangle - \langle S_0 Q_y S_0' \rangle + \langle M Q_y M' \rangle - \langle M Q M' \rangle. \tag{13}$$

Now let us consider the wavefront aberration of a grating. Referring to Figs. 1 and 3, if P, P_y are the intersection points of rays MQ^+QM' and $S_0{Q_y}^+Q_yS_0{}'$ with the grating surface, the replacement of Q, Q_y in (13) by P, P_y gives the wavefront aberration of the grating,

$$W = \langle S_0 \bar{P} S_0' \rangle - \langle S_0 P_y S_0' \rangle + \langle M P_y M' \rangle - \langle M P M' \rangle. \tag{14}$$

To convert from optical path distances to real distances through the grating, the additional optical path difference produced by the groove of the grating must be taken into account. The optical path length of each term equals the geometrical distance. If the groove number between P and $\bar{P}$ is n, equation (14) for the mth diffraction order of the grating is then

$$W = S_0 \bar{P} S_0' - S_0 P_y S_0' + M P_y M' - M P M' - nm\lambda. \tag{15}$$

The general form of a plane-symmetric surface can be expressed, in its vertex coordinate system, by the equation (Noda *et al.*, 1974; Peatman, 1997)

$$z = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} c_{i,j} \chi^i \eta^j,$$
$$c_{0,0} = c_{1,0} = 0, \quad j = \text{even}.$$

For the third-order aberration theory, the power series needs to be kept to the fourth order; thus the figure equation is then given by

$$z = c_{2,0}\chi^2 + c_{0,2}\eta^2 + c_{3,0}\chi^3 + c_{1,2}\chi\eta^2 + c_{4,0}\chi^4 + c_{0,4}\eta^4 + c_{2,2}\chi^2\eta^2. \tag{16}$$

Peatman (1997) has given the coefficients $c_{i,j}$ for a toroid, ellipsoid and paraboloid. For a toroidal surface, $c_{i,j}$ is as follows,

$$c_{2,0} = \frac{1}{2R}, \quad c_{0,2} = \frac{1}{2\rho}, \quad c_{3,0} = 0, \quad c_{1,2} = 0,$$
$$c_{4,0} = \frac{1}{8R^3} \quad c_{0,4} = \frac{1}{8\rho^3} \quad c_{2,2} = \frac{1}{4R^2\rho}, \tag{17}$$

where R and ρ are the major and minor curvature radii of the toroid. If $R = \rho$, (16) is just a spherical equation, and if R or ρ tend to infinity it is then a cylindrical equation.

The last section has given the positions of $M(0, y_m, r_m)$ and $S_0(0, 0, r_s)$ in the entrance-pupil coordinate system, and y_m is determined by (6). The optical path lengths in (15) can thus be determined for the given points P and P_y on a specific optical surface. However, the optical elements discussed here are usually a mirror or grating with a ray pencil of oblique or even extremely grazing incidence, so that the principal ray is at an oblique angle to the optical surface normal. The position of P or P_y can be described conveniently by the figure equation of the optical surface on the coordinate system xyz or $\chi\eta z'$ (called the vertex coordinate system) as shown in Fig. 1. Therefore we need to find the mapping relationships of the ray coordinates from the optical element to the reference wavefront surface to express y_m by x and y.

Figs. 1 and 2 show an object ray pencil incident on a grating, the related coordinate systems and foci. Since u or γ is usually very small and r_m and r_s are large, the following approximations can be made,

$$\overline{OM_x} \simeq \overline{\bar{P}_0 M_x} \simeq r_m, \qquad \overline{OS_0} \simeq \overline{P_0 S_0} \simeq r_s.$$

α and β are the angles of incidence and diffraction at the grating. The distance of $\bar{P}_0$ from O is

$$\bar{y} = s\gamma = s'\gamma' \equiv ul, \tag{18}$$

where $s = \overline{OO'}$ and γ is the intercept and angle of the principle ray with respect to the base ray, and l is a non-physical parameter so that the effect on aberrations due to $\bar{y}$ can be correlated to u.

Fig. 4 shows the optical scheme of the field and aperture rays passing through a system of two elements on the sagittal plane. The ray geometry gives the proportion relationship

$$\frac{r_s u}{\bar{y}} = \frac{s - r_s}{s}. \tag{19}$$

Equations (18) and (19) result in

$$\bar{y} = \frac{r_s s}{s - r_s} u = lu = s\gamma. \tag{20}$$

Hence the relationships between s, γ and l, u are derived as follows,

$$\gamma = \left(\frac{l}{r_s} - 1\right) u \equiv -\Lambda_l u, \quad l = \frac{r_s s}{s - r_s},$$
$$\gamma' = \left(\frac{l'}{r_s'} - 1\right) u' \equiv \Lambda_l' u, \quad s' = \frac{r_s' l}{l + r_s'}, \tag{21}$$

where $\Lambda_l = 1 - l/r_s$ and $\Lambda_l' = 1 + l/r_s'$.

Actually, the intersection point of the principal ray on the grating surface, $\bar{P}$, is not exactly on the η axis like $\bar{P}_0$, but also displaced in the axes of χ and z owing to the oblique incidence of the ray pencil and the sagittal curvature of the grating. The displacement of $\bar{P}$ from O is

$$\bar{x} = u^2 l^2 c_{0,2} \tan\alpha \equiv x_c u^2, \quad \bar{y} = ul, \quad \bar{z} = u^2 l^2 c_{0,2}. \tag{22}$$

So (16), representing the optical surface, becomes

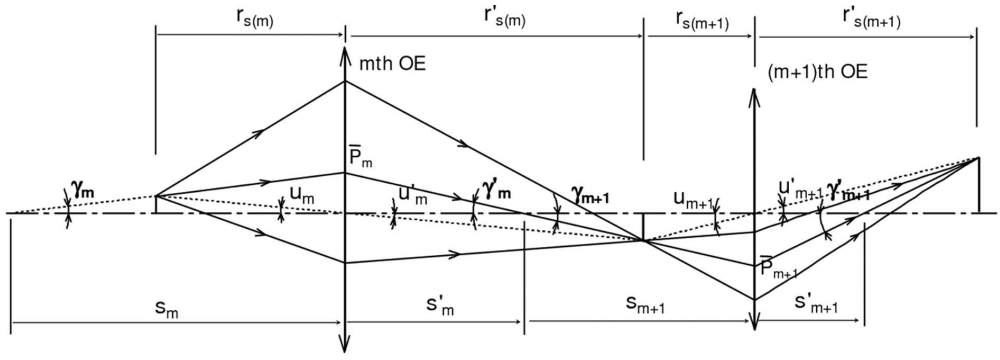

Figure 4
Optical scheme of the field and aperture rays passing through a system of two elements on the sagittal plane.

$$\begin{aligned} z = {} & c_{2,0}\left(x + x_c u^2\right)^2 + c_{0,2}(y + ul)^2 + c_{3,0}\left(x + x_c u^2\right)^3 \\ & + c_{1,2}\left(x + x_c u^2\right)(y + ul)^2 + c_{4,0}\left(x + x_c u^2\right)^4 \\ & + c_{0,4}(y + ul)^4 + c_{2,2}\left(x + x_c u^2\right)^2 (y + ul)^2 - u^2 l^2 c_{0,2}. \end{aligned} \tag{23}$$

Referring to Fig. 1, (x_0, y_0, z_0), (x_p, y_p, z_p) and $(0, y_m, r_m)$ are the coordinates of Q, P, M, and z_0 and z_p are determined by (2) and (23), respectively. The coordinates of M, from §2.1, are given by

$$\begin{aligned} & x_m = 0, \quad y_m = \frac{(r_s - r_m)}{r_s} y_0 + \frac{(r_s - r_m)}{2r_s^3} y_0^3 + r_m \Lambda_l u, \\ & z_m = r_m. \end{aligned} \tag{24}$$

To set up the equation of the line QPM, the coordinate x, y, z of P should be transformed to the entrance-pupil coordinate system,

$$x_p = x\cos\alpha - z\sin\alpha, \quad y_p = y, \quad z_p = x\sin\alpha + z\cos\alpha. \tag{25}$$

Then the line equation of QPM is

$$\frac{x_0}{x_p} = \frac{y_0 - y_m}{y_p - y_m} = \frac{z_0 - r_m}{z_p - r_m}. \tag{26}$$

Obviously this equation cannot be solved analytically. The general approach is to use an iterative method. Because all variables are not numerical but symbolic, it is difficult to judge whether the result reaches the required accuracy or not. The polynomial fit can circumvent the difficulty. Assuming that the mapping relationship of the ray coordinates from the optical element to the reference wavefront surface is a polynomial expression,

$$x_0 = \sum_{i=0}^{4}\sum_{j=0}^{4} a'_{ij} x^i y^j, \quad y_0 = \sum_{i=0}^{4}\sum_{j=0}^{4} b'_{ij} x^i y^j \quad (i + j \le 4), \tag{27}$$

the coefficients a'_{ij} and b'_{ij} can be fitted using (26). With the same process we can also obtain the reverse mapping relationship,

$$x = \sum_{i=0}^{4}\sum_{j=0}^{4} a_{ij} x_0^i y_0^j, \quad y = \sum_{i=0}^{4}\sum_{j=0}^{4} b_{ij} x_0^i y_0^j \quad (i + j \le 4). \tag{28}$$

2.4. Wavefront aberration coefficients

The terms related to the optical path length in (15) can be separated into the contributions from the object and image ray pencils,

$$\begin{aligned} W_{obj} &= S_0\bar{P} - S_0 P_y + MP_y - MP, \\ W_{ima} &= S'_0\bar{P} - S'_0 P_y + M'P_y - M'P. \end{aligned} \tag{29}$$

where $S_0\bar{P} = r_s$ and $S'_0\bar{P} = r_s'$. Hence (15) is equivalent to

$$W = W_{obj} + W_{ima} - nm\lambda. \tag{30}$$

The aberration coefficients can now be found from the pencil coefficients obtained from the expansion of each pencil contribution. The ray pencils of the image and object differ in their coordinates with respect to the substrate. So once the coefficients have been found for one of them, the coefficients of another can be obtained by substitution of the corresponding optical parameters. Note that the sign convention of α or β is that they are taken to be positive if positioned at the positive side of the x axis. In the following we first expand each term in W_{obj} into a power series of coordinates of x, y, u.

The length MP is calculated by

$$MP = \left[(x - r_m\sin\alpha)^2 + (y - \tilde{y}_m)^2 + (z - r_m\cos\alpha)^2\right]^{1/2}, \tag{31}$$

where $\tilde{y}_m$ is the coordinate of M in the y axis and, from Fig. 1, is given by

$$\tilde{y}_m = u r_m\left(1 - \frac{l}{r_s}\right) + \frac{(r_s - r_m)}{r_s} y_0 + \frac{(r_s - r_m)}{2r_s^3} y_0^3. \tag{32}$$

Substituting y_0 with (27) and z with (23) in the above equation, MP is now correlated to only the coordinates x, y, u and can be developed into the power series form

$$MP = \sum_{ijk}^{4} \tau_{ijk} x^i y^j u^k \quad (i + j + k \le 4), \tag{33}$$

where τ_{ijk} are the coefficients of a power series.

To calculate S_0P_y and MP_y we need to know the position of P_y. Since P_y is the intersection point of the ray $S_0Q_y{}^+Q_yS_0'$ on the optical surface, applying the mapping relationship (28) and setting $x_0 = 0$ can derive its position,

$$x_{py} = a_{01}y_0 + a_{02}y_0^2 + a_{03}y_0^3 + a_{04}y_0^4,$$
$$y_{py} = b_{01}y_0 + b_{02}y_0^2 + b_{03}y_0^3, \quad (34)$$

where a_{ij} and b_{ij} have been fitted above. Replacing y_0 in the above equation by (27) enables x_{py} and y_{py} to be expressed as the polynomial of x, y, u. Substituting them for x, y and (23) for z in (31) will give the power series of MP_y as in (33). Concerning S_0P_y, we only need to replace the related parameters of M by those of S_0 further; that is, to use r_s and $\tilde{y}_s = u(rs - l)$ instead of r_m and $\tilde{y}_m$ in (31).

Adding the corresponding τ_{ijk} of each ray pencil of the same index i, j, k, W_{obj} can be expressed as the power series

$$W_{obj} = \sum_{ijk}^{4} M_{ijk}x^i y^j u^k \quad (i + j + k \leq 4), \quad (35)$$

where M_{ijk} are the wavefront aberration coefficients of the object pencil. The resultant 16 coefficients are listed in Appendix A: equations (78)–(85) for in-plane aberration ($u = 0$) and equations (86)–(93) for off-plane aberration, depending on the parameters of α, r_m, r_s, l: $M_{ijk} = M_{ijk}(\alpha, r_m, r_s, l)$.

Similarly, the contributions from the image pencil can be derived in the same way as above, and W_{ima} is expressed by

$$W_{ima} = \sum_{ijk}^{4} M'_{ijk}x^i y^j u'^k \quad (i + j + k \leq 4). \quad (36)$$

The coefficients of the image-pencil aberration can be obtained only by substitution of β, r_m', r_s', l' of the image pencil for α, r_m, r_s, l of the object pencil in each M_{ijk}, respectively: $M_{ijk}' = M_{ijk}(\beta, r_m', r_s', l')$.

The groove function $n = n(\chi, \eta)$ in (30) for holographic and mechanically ruled gratings has been derived by Namioka *et al.* (1994),

$$n = \frac{\chi}{\sigma} + \frac{\Gamma}{\sigma}\left(\frac{n_{20}}{2}\chi^2 + \frac{n_{02}}{2}\eta^2 + \frac{n_{30}}{2}\chi^3 + \frac{n_{12}}{2}\chi\eta^2 + \frac{n_{40}}{8}\chi^4 + \frac{n_{22}}{4}\chi^2\eta^2 + \frac{n_{04}}{8}\eta^4 + \ldots\right), \quad (37)$$

where the effective grating constant σ is defined by

$$\sigma \equiv 1/(\partial n/\partial\chi)_{\chi=\eta=0}, \quad (38)$$

and Γ and n_{ij} are given by equations (20)–(22) of Namioka *et al.* (1994). With the coordinate transformation $\chi = x + x_c u^2$, $\eta = y + ul$, n can be expressed as

$$n = \sum_{ijk}^{4} (\Gamma/\sigma)N_{ijk}x^i y^j u^k, \quad (39)$$

where N_{ij} are given in Table 1.

Consequently, the wavefront aberration W of (30) in terms of the pencil coefficients can now be represented by

$$W = \sum_{ijk}^{4} w_{ijk}x^i y^j u^k \quad (i + j + k \leq 4). \quad (40)$$

w_{ijk}, referred to as the wavefront aberration coefficients, are given by

Table 1
N_{ij} as a function of n_{ij}.

In-plane		Off-plane	
$N_{100} = 1/\Gamma$	$N_{120} = n_{12}/2$	$N_{011} = n_{02}l$	$N_{102} = n_{12}l^2/2 + n_{20}x_c$
$N_{200} = n_{20}/2$	$N_{400} = n_{40}/8$	$N_{111} = n_{12}l$	$N_{202} = n_{22}l^2/4 + 3n_{30}x_c/2$
$N_{020} = n_{02}/2$	$N_{220} = n_{22}/4$	$N_{031} = n_{04}l/2$	$N_{022} = 3n_{04}l^2/4 + n_{12}x_c/2$
$N_{300} = n_{30}/2$	$N_{040} = n_{04}/8$	$N_{211} = n_{22}l/2$	$N_{013} = n_{04}l^3/2 + n_{12}x_cl$

$$w_{ijk} = M_{ijk}(\alpha, r_m, r_s, l) + (-1)^k M_{ijk}(\beta, r'_m, r'_s, l') - \Lambda N_{ijk}, \quad (41)$$

where $\Lambda = (m\lambda/\sigma)\Gamma$ and the factor $(-1)^k$ results from $u' = -u$ since the field ray is displaced from the base ray only on the sagittal plane without the effect of the groove diffraction of the grating.

The direction of the base ray after diffraction must be such that the meridional tilt of the wavefront is zero, *i.e.* $w_{100} = 0$. This directly leads to the familiar grating equation,

$$\sin\alpha + \sin\beta = m\lambda/\sigma. \quad (42)$$

To determine the direction of the principal ray, an additional condition, $w_{011} = 0$, should be satisfied,

$$l\left(2c_{0,2}\cos\alpha - \frac{1}{r_s}\right) - l'\left(2c_{0,2}\cos\beta - \frac{1}{r'_s}\right) = n_{02}\Lambda l. \quad (43)$$

The position along the base ray of the meridional focal line is given from $w_{200} = 0$, *i.e.*

$$2c_{2,0}(\cos\alpha + \cos\beta) - \left(\frac{\cos^2\alpha}{r_m} + \frac{\cos^2\beta}{r'_m}\right) = n_{20}\Lambda. \quad (44)$$

Similarly for the sagittal focal line position, $w_{020} = 0$, and

$$2c_{0,2}(\cos\alpha + \cos\beta) - \left(\frac{1}{r_s} + \frac{1}{r'_s}\right) = n_{02}\Lambda. \quad (45)$$

Equations (43) and (45) directly lead to $l' = -l$.

From the above discussions, the explicit expression of the wavefront aberration for parabasal theory will be

$$W = w_{300}x^3 + w_{120}xy^2 + w_{400}x^4 + w_{220}x^2y^2 + w_{040}y^4 + w_{102}xu^2 + w_{013}yu^3 + w_{202}x^2u^2 + w_{022}y^2u^2 + w_{111}xyu + w_{031}y^3u + w_{211}x^2yu. \quad (46)$$

The former five terms give the wavefront aberration for an in-plane object and the latter seven terms give the additional off-plane wavefront aberration for an off-plane object.

3. Transverse aberration on the image plane

The issue of determining the position of an aberrated ray on the image plane is to find the diffraction direction of the ray. If the mapping relationships of the ray from the optical surface to the aberrated wavefront can be obtained, the direction of the aberrated ray will be determined. Equation (27) is used to calculate the position of the rays on the incident wavefront for any point P. But it does not apply to the case of an aberrated ray because aberration from the optical surface takes place.

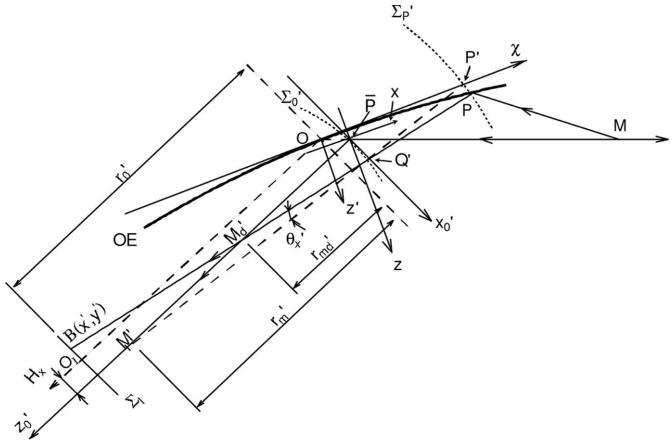

Figure 5
Ray diffracted from point P of the grating surface intersects point B on the image plane Σ_I on the meridional plane, with the incident and exit wavefront as well as the focus.

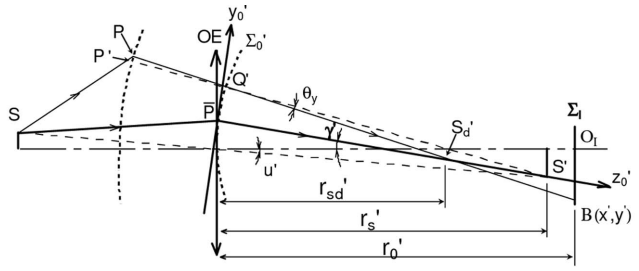

Figure 6
Ray diffracted from point P at the grating surface intersects point B on the image plane Σ_I on the sagittal plane, with the incident and exit wavefront as well as the focus.

The effect of aberration must be considered. A toroidal surface is also employed as the reference surface for the aberrated wavefront.

As shown in Figs. 5 and 6, Σ_0' and Σ_I denote the exit wavefront surface at the exit pupil and the image plane at a distance r_0' from the optical element. The reference surface normal at Q' intersects the principal ray $S\bar{P}S'$ at M' and S' displaced from the optical element by r_m' and r_s', meridionally and sagittally. The direction of the aberrated ray in relation to the normal is given from the slope of the wavefront with respect to the reference surface. The wavefront aberration measures the separation between these two surfaces; its differentials in the two pupil coordinates will give this slope in the meridional and sagittal direction. So the angular displacement of the aberrated ray from the reference surface normal in the meridional and sagittal directions, θ_x', θ_y', are

$$\theta_x' = \frac{\mathrm{d}W}{\mathrm{d}x_0'}, \qquad \theta_y' = \frac{\mathrm{d}W}{\mathrm{d}y_0'}, \tag{47}$$

where x_0', y_0' are the coordinates of the ray on the aberrated wavefront at the exit pupil. Consequently, the actual intersection points, M_d' and S_d', of the aberrated ray with the principal ray are at distances of r_{md}' and r_{sd}' from the optical element. This aberrated ray finally reaches point B on the image plane.

In (47), x_0', y_0' are correlated to the wavefront aberration but the relationships between them are unknown. This is one difficulty in wavefront aberration. However, the aberrated ray equation can be set up based on the ray geometry. Assuming the mapping relationship of the ray from the optical element to the aberrated wavefront surface is as follows,

$$x_0' = \sum_{i=0}^{4}\sum_{j=0}^{4} A_{ij}x^i y^j, \quad y_0' = \sum_{i=0}^{4}\sum_{j=0}^{4} B_{ij}x^i y^j \quad (i+j \le 4), \tag{48}$$

where $A_{00} = 0$ and $B_{00} = 0$, then differentiating both sides of the above equations with respect to x_0', y_0' will result in the following equations,

$$\begin{aligned}
1 &= \sum_{i=0}^{4}\sum_{j=0}^{4}\left(i\frac{\partial x}{\partial x_0'}A_{ij}x^{i-1}y^j + j\frac{\partial y}{\partial x_0'}A_{ij}x^i y^{j-1}\right),\\
0 &= \sum_{i=0}^{4}\sum_{j=0}^{4}\left(i\frac{\partial x}{\partial y_0'}A_{ij}x^{i-1}y^j + j\frac{\partial y}{\partial y_0'}A_{ij}x^i y^{j-1}\right),\\
1 &= \sum_{i=0}^{4}\sum_{j=0}^{4}\left(i\frac{\partial x}{\partial y_0'}B_{ij}x^{i-1}y^j + j\frac{\partial y}{\partial y_0'}B_{ij}x^i y^{j-1}\right),\\
0 &= \sum_{i=0}^{4}\sum_{j=0}^{4}\left(i\frac{\partial x}{\partial x_0'}B_{ij}x^{i-1}y^j + j\frac{\partial y}{\partial x_0'}B_{ij}x^i y^{j-1}\right).
\end{aligned} \tag{49}$$

From these linear equations, the four derivatives, $\partial x/\partial x_0'$, $\partial x/\partial y_0'$, $\partial y/\partial x_0'$, $\partial y/\partial y_0'$, can be solved with the coefficients A_{ij} and B_{ij} to be determined. Thus θ_x', θ_y' will be correlated to x, y, u by the following expressions,

$$\theta_x' = \frac{\partial W}{\partial x}\frac{\partial x}{\partial x_0'} + \frac{\partial W}{\partial y}\frac{\partial y}{\partial x_0'}, \quad \theta_y' = \frac{\partial W}{\partial x}\frac{\partial x}{\partial y_0'} + \frac{\partial W}{\partial y}\frac{\partial y}{\partial y_0'}. \tag{50}$$

On the other hand, as shown in Figs. 5 and 6, the auxiliary line equations of $P'Q'M'$ and $P'Q'S'$ can be written as

$$\begin{aligned}
x_0' - \frac{r_m' - z_0'}{r_m' - z_p}\left[x_p + (z_p - z_0')\theta_x'\right] &= 0,\\
y_0' - \frac{z_s' - z_0'}{z_s' - z_p}\left[y_p + \left(z_p - z_0'\right)\left(\theta_y' + \Lambda_I' u\right)\right] &= 0,
\end{aligned} \tag{51}$$

where x_0', y_0', z_0' of Q' and θ_x', θ_y' can be expressed by the coordinates of x, y, u; (x_p, y_p, z_p), the coordinate of P in the exit-pupil coordinate system, is calculated by (25) with the replacement of α by β. z_0' and z_s' are well approximated by

$$z_0' = \frac{x_0'^2}{2r_{md}'} + \frac{y_0'^2}{2r_{sd}'} + \frac{x_0'^4}{8r_{md}'^3} + \frac{y_0'^4}{8r_{sd}'^3} + \frac{x_0'^2 y_0'^2}{4r_{md}'^2 r_{sd}'}, \tag{52}$$

$$z_s' = r_s' + \frac{(r_m' - r_s')}{2r_m'^2}x_0'^2 + \frac{(r_m' - r_s')}{8r_m'^4}x_0'^4 + \frac{(r_m' - r_s')}{2r_m'^3 r_s'}x_0'^2 y_0'^2. \tag{53}$$

After substitution of the above parameters into (51), the coefficients A_{ij}, B_{ij} can be fitted, leading to the determination of θ_x', θ_y'.

For a plane-symmetric system, the position of an image plane is usually specified with the distance from the grating vertex along the base ray, so the origin of the image coordinate

x', y', $O_{\rm I}$, is set at the intersection point of the image plane with the base ray, as illustrated in Figs. 5 and 6. Second, the displacement of $\bar{P}$ from O causes the principal ray to displace the base ray both horizontally and vertically. From (22) the displacement in the x' and y' directions is determined by

$$H_{x'} = \frac{\sin(\alpha-\beta)x_c}{\sin\alpha}\,u^2, \qquad H_{y'} = ul. \tag{54}$$

Consequently, the position of image B is calculated by

$$\begin{aligned} x' &= \frac{x_p - x_0'}{z_p - z_0'}r_0' + \frac{\left(z_p x_0' - z_0' x_p\right)}{z_p - z_0'} + H_{x'},\\ y' &= \frac{y_p - y_0'}{z_p - z_0'}r_0' + \frac{\left(z_p y_0' - z_0' y_p\right)}{z_p - z_0'} + H_{y'}, \end{aligned} \tag{55}$$

Thus the expressions on the right-hand side of (55) are expanded into the power series,

$$\begin{aligned} x' &= \sum_{i=0}^{4}\sum_{j=0}^{4}\sum_{k=0}^{3} d_{ijk}x^i y^j u^k,\\ y' &= \sum_{i=0}^{4}\sum_{j=0}^{4}\sum_{k=0}^{3} h_{ijk}x^i y^j u^k \quad (i+j+k\le 4). \end{aligned} \tag{56}$$

As a result, 32 non-zero coefficients d_{ijk} and h_{ijk}, called transverse aberration coefficients, yield; 13 of them are fourth-order terms and 19 terms for first- to third-orders. $\{d_{400}, d_{040}, d_{220}, h_{310}, h_{130}\}$ of the fourth-order terms pertain to in-plane aberrations and $\{d_{211}, d_{031}, d_{013}, d_{022}, d_{202}, h_{301}, h_{121}, h_{112}\}$ to off-plane ones. The fourth-order terms are quite complex but with much smaller contributions to the aberration than those from lower-order terms. They, in principle, should include the fifth-order wavefront aberration coefficients like w_{500}, w_{320} and so on, but it is found that the results will no longer be correct if they are included. The resultant fourth-order aberration coefficients removing only the term related to the fifth-order wavefront aberration still have a positive contribution especially when a large aberration of the optical system exists.

In Appendix B, only the transverse aberration coefficients from the first to third order are listed. The imaging formulae of the third-order aberration will be

$$\begin{aligned} x' &= d_{100}x + d_{200}x^2 + d_{020}y^2 + d_{300}x^3 + d_{120}xy^2 + d_{002}u^2\\ &\quad + d_{011}yu + d_{111}xyu + d_{102}xu^2,\\ y' &= h_{010}y + h_{110}xy + h_{210}x^2y + h_{030}y^3 + h_{003}u^3 + h_{001}u\\ &\quad + h_{101}xu + h_{201}x^2u + h_{021}y^2u + h_{012}yu^2. \end{aligned} \tag{57}$$

The expressions of d_{ijk} and h_{ijk} show that the contributions to the transverse aberration consist of two parts: wavefront aberration and the effect caused by the displacement of the image plane from the focus (*i.e.* defocus). Because we have $w_{200} = 0$, $w_{020} = 0$ in parabasal theory, the effect of defocus cannot be obtained from the wavefront aberration itself. The previous aberration theories developed with the optical-path function or wavefront aberration cannot include the defocus effect, so they are limited only to dealing with imaging close to the focal position. The limitation is now removed.

Secondly, the position of the principal ray is determined by the terms of $d_{002}u^2$ and $h_{001}u + h_{003}u^3$ because they specify the diffraction ray from the center of the exit pupil.

4. Wavefront aberration of multi-element systems

4.1. Transfer equations of field and aperture rays

For plane-symmetric multi-element optical systems, all the optical elements have a common symmetry plane; the base ray passing through the center of each element lies in this plane and its direction is determined by the grating equation. The principal ray is displaced from the base ray on the sagittal plane and governed by $w_{011} = 0$ besides the grating equation. The parabasal foci are found by applying the equations $w_{200} = 0$ and $w_{020} = 0$ repeatedly throughout the system.

To calculate the wavefront and transverse aberration coefficients of a multi-element system, information on the field and aperture rays of each element are required; thus their transfer equations between successive optical elements are needed.

Fig. 4 illustrates the transfer scheme of the field and aperture rays between the mth and $(m+1)$th elements on the sagittal plane. A prime ($'$) is used to distinguish $F_{\rm s}'$, $r_{\rm s}'$, s', γ', u' of the exit ray pencil from the corresponding parameters, $F_{\rm s}$, $r_{\rm s}$, s, γ, u of the incident ray pencil.

Obviously the transfer equations of u are

$$u'_{(m)} = -u_{(m)}, \quad u_{(m+1)} = \frac{r'_{{\rm s}(m)}}{r_{{\rm s}(m+1)}}u'_{(m)} = -\frac{r'_{{\rm s}(m)}}{r_{{\rm s}(m+1)}}u_{(m)}. \tag{58}$$

From (21), $l_{(m)}$ is determined by

$$l_{(m)} = \frac{r_{{\rm s}(m)}s_{(m)}}{s_{(m)} - r_{{\rm s}(m)}}, \quad l'_{(m)} = -l_{(m)}. \tag{59}$$

Calculation of $l_{(m)}$ requires the value of $s_{(m)}$. The sagittal focusing equation (45) leads to

$$\begin{aligned} \frac{1}{r_{{\rm s}(m)}} + \frac{1}{r'_{{\rm s}(m)}} = \frac{1}{s_{(m)}} + \frac{1}{s'_{(m)}} &= 2c_{0,2(m)}\left[\cos\alpha_{(m)} + \cos\beta_{(m)}\right]\\ &\quad - n_{02}\Lambda, \end{aligned} \tag{60}$$

$s_{(m)}{}'$ is thus calculated by

$$\frac{1}{s'_{(m)}} = \frac{1}{r_{{\rm s}(m)}} + \frac{1}{r'_{{\rm s}(m)}} - \frac{1}{s_{(m)}}, \tag{61}$$

and its transfer equation is simply given by

$$s_{(m+1)} = r'_{{\rm s}(m)} + r_{s(m+1)} - s'_{(m)}. \tag{62}$$

As discussed in §3, $d_{002(m)}u_{(m)}{}^2$ and $h_{001(m)}u_{(m)}$ [$h_{003(m)}u_{(m)}{}^3$ is neglected here] give the position of the principal ray on the image plane. If the image distance equals $r_{{\rm s}(m)}{}' + r_{{\rm s}(m\,+\,1)}$, the position of the aperture-stop center on the $(m+1)$th element, $\bar{P}_{(m+1)}$, will be given by

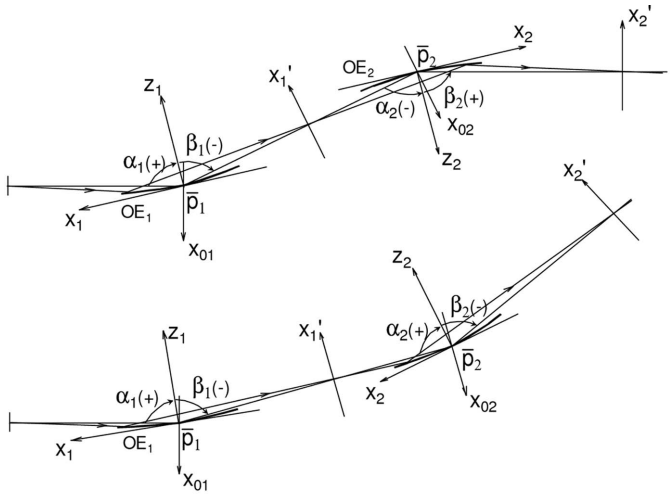

Figure 7
The coordinate system of two-element optics: the upper one arranged with the figure of 'Z' and the lower one arranged with the figure of 'U'.

$$\begin{aligned}\bar{x}_{(m+1)} &= l^2_{(m+1)}c_{0,2(m+1)}\tan\alpha_{(m+1)}u^2_{(m+1)} - \frac{d_{002(m)}u^2_{(m)}}{\cos\alpha_{(m+1)}}\\ &\equiv x_{c(m+1)}u^2_{(m+1)},\\ \bar{y}_{(m+1)} &= h_{001(m)}u_{(m)} = -\left[\frac{r_{s(m+1)}l_{(m)}}{r'_{s(m)}} + r'_{s(m)} + r_{s(m+1)}\right]u_{(m)}.\end{aligned} \tag{63}$$

The first term in $\bar{x}_{(m+1)}$ comes from $x_{c(m)}u_{(m)}{}^2$ of (22). The coordinate system for two-element optics is shown in Fig. 7.

With equations (58)–(62), $\bar{y}_{(m+1)}$ in (63) can be, as expected, simplified to $u_{(m+1)}l_{(m+1)}$; and $x_{c(m+1)}$ is now derived,

$$\begin{aligned}x_{c(m+1)} &= \frac{1}{\cos\alpha_{(m+1)}}\left\{l^2_{(m+1)}c_{0,2(m+1)}\sin\alpha_{(m+1)} - \left[\frac{r_{s(m+1)}}{r'_{s(m)}}\right]^2 d_{002(m)}\right\}\\ d_{002(m)} &= -\frac{\left[r'_{s(m)} + r_{s(m+1)}\right]w_{102(m)}}{\cos\beta_{(m)}} + \frac{x_{c(m)}\sin\left[\alpha_{(m)} - \beta_{(m)}\right]}{\sin\alpha_{(m)}},\end{aligned} \tag{64}$$

and $d_{002(0)} \equiv 0$ is assumed for the case of the first element.

Taking the first-order approximation, (57) gives the position of an aperture ray at the entrance pupil of the $(m + 1)$th element,

$$\begin{aligned}x_{0(m+1)} &= -x'_{(m)} = -d_{100(m)}x_{(m)} = \frac{r_{m(m+1)}\cos\beta_{(m)}}{r'_{m(m)}}x_{(m)},\\ y_{0(m+1)} &= y'_{(m)} = h_{010(m)}y_{(m)} = -\frac{r_{s(m+1)}}{r'_{s(m)}}y_{(m)}.\end{aligned} \tag{65}$$

Mapping it onto the aperture-stop coordinate system, the transfer equation of aperture ray will then be

$$x_{(m+1)} = \frac{r_{m(m+1)}\cos\beta_{(m)}}{r'_{m(m)}\cos\alpha_{(m+1)}}x_{(m)}, \quad y_{(m+1)} = -\frac{r_{s(m+1)}}{r'_{s(m)}}y_{(m)}. \tag{66}$$

4.2. Wavefront aberration of multi-element systems

For multi-element optical systems, the total wavefront aberration is the sum of the contribution from each optical element in the region of the third-order aberration. Thus, for a system of n gratings the wavefront aberration should be

$$W = W_{(1)} + W_{(2)} + \ldots + W_{(n)} = \sum_{m=1}^{n}\sum_{ijk}^{4} w_{ijk(m)}x^i_{(m)}y^j_{(m)}u^k_{(m)}$$
$$(i + j + k \leq 4), \tag{67}$$

where w_{ijk} is given by (41), and $x_{(m)}$, $y_{(m)}$, $u_{(m)}$ are the coordinates of the ray and the field angle of the mth element. To derive the wavefront aberration coefficients of a multi-element system, we need the corresponding relationships between $x_{(m)}$, $y_{(m)}$, $u_{(m)}$ $(m = 1, 2 \ldots, n - 1)$ and $x_{(n)}$, $y_{(n)}$, $u_{(n)}$.

From (66) the relationship between $x_{(m)}$, $y_{(m)}$ and $x_{(n)}$, $y_{(n)}$ can be obtained,

$$x_{(m)} = A_{(m)}x_{(n)}, \quad y_{(m)} = (-1)^{n-m}B_{(m)}y_{(n)}. \tag{68a}$$

$$\begin{aligned}A_{(m)} &\equiv \frac{r'_{m(m)}r'_{m(m+1)}\ldots r'_{m(n-1)}}{r_{m(m+1)}r_{m(m+2)}\ldots r_{m(n)}}\frac{\cos\alpha_{(m+1)}\cos\alpha_{(m+2)}\ldots\cos\alpha_{(n)}}{\cos\beta_{(m)}\cos\beta_{(m+1)}\ldots\cos\beta_{(n-1)}},\\ B_{(m)} &\equiv \frac{r'_{s(m)}r'_{s(m+1)}\ldots r'_{s(n-1)}}{r_{s(m+1)}r_{s(m+2)}\ldots r_{s(n)}}.\end{aligned} \tag{68b}$$

Equation (58) gives the relationship of the field angle u between two adjacent elements, and thus leads to

$$\begin{aligned}u_{(m)} &= (-1)^{n-m}\ \frac{r_{s(m+1)}r_{s(m+2)}\ldots r_{s(n)}}{r'_{s(m)}r'_{s(m+1)}\ldots r'_{s(n-1)}}u_{(n)}\\ &= (-1)^{n-m}\ \frac{u_{(n)}}{B_{(m)}}.\end{aligned} \tag{69}$$

Replacing $x_{(m)}$, $y_{(m)}$, $u_{(m)}$ in (67) by (68*a*) and (69) gives

$$\begin{aligned}W &= \sum_{m=1}^{n-1}\sum_{ijk}^{4} w_{ijk(m)}A^i_{(m)}B^{j-k}_{(m)}x^i_{(n)}y^j_{(n)}u^k_{(n)} + \sum_{ijk}^{4} w_{ijk(n)}x^i_{(n)}y^j_{(n)}u^k_{(n)}\\ &\equiv \sum_{ijk}^{4} W_{ijk}x^i_{(n)}y^j_{(n)}u^k_{(n)},\end{aligned} \tag{70}$$

where W_{ijk}, called the wavefront aberration coefficient of the multi-element system, is then given by

$$W_{ijk} = \sum_{m=1}^{n-1} w_{ijk(m)}A^i_{(m)}B^{j-k}_{(m)} + w_{ijk(n)}. \tag{71}$$

Equation (57) applies likewise to calculations of the transverse aberration of a multi-element system; only in calculations of d_{ijk} and h_{ijk}, all the related parameters are those of the final optical element, and W_{ijk} should be used instead of w_{ijk}.

5. Comparison of the theory with analytic formulae of the spot diagram

It is of practical interest at this point to examine the validity of the theory. One way is to use ray-tracing calculations, but this cannot check single terms of the aberration. Namioka *et al.*

(1994) have derived the analytic formulae of the spot diagram following exact ray-tracing formalism, and given the following equations for image calculations,

$$\begin{aligned} x' &= r'_0 \sec(\beta+\phi)(1-\tan\beta\tan\phi) \\ &\quad\times\left[xf_{100}+x^2f_{200}+y^2f_{020}+yzf_{011}+z^2f_{002}\right. \\ &\quad\left.+x^3f_{300}+xy^2f_{120}+xyzf_{111}+xz^2f_{102}+O\left(x^4/R^3\right)\right], \\ y' &= r'_0\left[zg_{001}+yg_{010}+xyg_{110}+xzg_{101}+x^2yg_{210}\right. \\ &\quad\left.+x^2zg_{201}+y^3g_{030}+y^2zg_{021}+yz^2g_{012}+z^3g_{003}\right. \\ &\quad\left.+O\left(x^4/R^3\right)\right], \end{aligned} \tag{72}$$

where z is the field parameter ($u = z/r$, if r is the object distance); the aberration coefficients f_{ijk}, g_{ijk} are determined by equations (25N)–(28N) [note: N after the equation number means the equations from Namioka *et al.* (1994)]. The above equations are similar to (57) and comparable between them.

Namioka *et al.* (1994) discuss the imaging for an ellipsoidal grating with varied spacing and curved grooves; its figure equation (17N) has a different form from (16). They will be identical if we assume, in (17N) and (19N), that

$$\begin{aligned} &R=\frac{1}{2c_{2,0}}, \quad \rho=\frac{1}{2c_{0,2}}, \quad \frac{1}{A}=0, \\ &\varepsilon_{30}=c_{3,0}, \quad \varepsilon_{12}=c_{1,2}, \quad \varepsilon_{40}=c_{4,0}, \\ &\varepsilon_{22}=c_{2,2}, \quad \varepsilon_{04}=c_{0,4}, \end{aligned} \tag{73}$$

Second, it is reasonable to assume that $r_\mathrm{s} = r_\mathrm{m} = r$. From (44) and (45), equations (27*a*N) and (27*b*N) will become

$$\begin{aligned} F_{200} &= T_A + T_B + n_{20}\Lambda \\ &= \frac{\cos^2\beta}{r'_0} - \frac{\cos^2\beta}{r'_\mathrm{m}} = \frac{\Lambda_\mathrm{m}\cos^2\beta}{r'_0}, \end{aligned} \tag{74a}$$

$$F_{020} = \bar{S}_A + \bar{S}_B + n_{02}\Lambda = \frac{1}{r'_0} - \frac{1}{r'_\mathrm{s}} = \frac{\Lambda_\mathrm{s}}{r'_0}, \tag{74b}$$

where Λ_m and Λ_s are defined by equation (94) of Appendix *B*.

Third, in the study of Namioka *et al.* (1994), the aperture-stop center is coincident with the grating vertex, so we set $l = 0$ in the calculations of d_{ijk} and h_{ijk}.

Now we can compare their 19 transverse aberration coefficients from the first to third order according to (57) and (72) by examining the differences for the case of $\phi = 0$,

$$\Delta d_{ijk} = f_{ijk}\frac{r'_0}{\cos\beta}r^k - d_{ijk}, \tag{75a}$$

$$\Delta h_{ijk} = g_{ijk}r'_0r^k - h_{ijk}. \tag{75b}$$

Substituting the expressions of f_{ijk}, g_{ijk}, d_{ijk} and h_{ijk} into the above equations, the results show that only two terms are non-zero,

$$\Delta h_{201} = -\frac{r'_0\cos^2\beta}{2r'^2_\mathrm{m}}, \qquad \Delta h_{021} = -\frac{r'_0}{2r'^2_\mathrm{s}}. \tag{76}$$

The two terms do not contain the parameters of the object ray pencil, *i.e.* r, α and the figure coefficients $c_{i,j}$, so the differences are not caused by wavefront aberration but by the approximation of some geometry relationships in finding the transverse aberration in §3.

The error of the image coordinate y' owing to (76) is approximated to

$$\Delta y' = \Delta h_{201}x^2u + \Delta h_{021}y^2u \simeq -\frac{\left(\theta'^2_v+\theta'^2_h\right)s'_y}{2}, \tag{77}$$

where $s_y' = r_0'u$ means the half sagittal image size without aberration, and θ_v' and θ_h' are the exit converging angles of the ray, which are $\sim$0.01 rad for a marginal ray in synchrotron radiation optics; thus $\Delta y'$ is about $0.0001s_y'$ and negligible.

6. Summary

The third-order aberration theory developed here is applicable to aberration analysis of the usual plane-symmetric multi-element mirror or grating systems. The surface figure of the element can be plane, spherical, cylindrical, toroidal, elliptical, parabolic and so on. The wavefront and transverse aberration coefficients are derived as listed in Appendices *A* and *B*.

The previous aberration theories based on the optical path function or wavefront aberration can only deal with the aberration analysis when the image plane is at the foci on both the meridional and the sagittal planes. Now the theory removes this limitation. This is especially important for synchrotron radiation instrumentation because the meridional and sagittal foci of each optical element are often not coincident, leading to the image plane, at least, displaced from either the meridional or sagittal focus.

Equation (28) is used to calculate the ray distribution on the optical surface. The accuracy of the first-order approximation depends on the acceptance angle of the light beam, the angle of incidence and the curvature radius. In synchrotron radiation optics, the source size ($\sim$1 mm) and the acceptance angle ($\sim$several mrad) are usually small, so we can even use the first-order approximation with $u = 0$ to calculate the aperture-ray coordinates whereas the image can be obtained with a high degree of accuracy. Anyway, the aberration theory is mainly used for the initial choice and optimization of the parameters of optical systems not for replacement of ray tracing; thus the simplicity of aberration expressions is emphasized. The theory with the first-order approximation of the aperture ray has simple aberration expressions while remaining quite high in accuracy in most cases so far; moreover, the expressions of the aberration coefficient clearly show the dependence of the wavefront aberration. At this point it is advantageous over the other methods, especially in studying multi-element systems.

APPENDIX *A*
Wavefront aberration coefficients of the object pencil

$$M_{100}(\alpha, r_\mathrm{m}, r_\mathrm{s}, l) = \sin(\alpha), \tag{78}$$

$$M_{200}(\alpha, r_\mathrm{m}, r_\mathrm{s}, l) = \frac{\cos\alpha}{2}\left(2c_{2,0} - \frac{\cos\alpha}{r_\mathrm{m}}\right), \tag{79}$$

$$M_{300}(\alpha, r_m, r_s, l) = \frac{1}{2}\left[2c_{3,0}\cos\alpha + \frac{\sin\alpha\cos\alpha}{r_m}\left(2c_{2,0} - \frac{\cos\alpha}{r_m}\right)\right], \quad (80)$$

$$M_{020}(\alpha, r_m, r_s, l) = \frac{1}{2}\left(2c_{0,2}\cos\alpha - \frac{1}{r_s}\right), \quad (81)$$

$$M_{120}(\alpha, r_m, r_s, l) = \frac{1}{2}\left[\frac{\sin\alpha}{r_m}\left(2c_{0,2}\cos\alpha - \frac{r_m}{r_s^2}\right) + 2c_{1,2}\cos\alpha\right], \quad (82)$$

$$M_{400}(\alpha, r_m, r_s, l) = \frac{\cos^2\alpha(5\cos^2\alpha - 4)}{8r_m^3} + \frac{c_{2,0}\cos\alpha(2 - 3\cos^2\alpha)}{2r_m^2} - \frac{c_{2,0}^2\sin^2\alpha}{2r_m} + \frac{c_{3,0}\sin 2\alpha}{2r_m} + c_{4,0}\cos\alpha, \quad (83)$$

$$M_{040}(\alpha, r_m, r_s, l) = \frac{1}{8r_s^3} - \frac{c_{0,2}\cos\alpha}{2r_s^2} - \frac{c_{0,2}^2\sin^2\alpha}{2r_m} + c_{0,4}\cos\alpha, \quad (84)$$

$$M_{220}(\alpha, r_m, r_s, l) = \frac{1}{2r_s^2}\left(\frac{\cos^2\alpha}{2r_m} - \frac{\sin^2\alpha}{r_s} - c_{2,0}\cos\alpha\right) + \frac{c_{0,2}\cos\alpha(2 - 3\cos^2\alpha)}{2r_m^2} - \frac{c_{2,0}c_{0,2}\sin^2\alpha}{r_m} + \frac{c_{1,2}\sin 2\alpha}{2r_m} + c_{2,2}\cos\alpha, \quad (85)$$

$$M_{102}(\alpha, r_m, r_s, l) = -\frac{1}{2}\Lambda_l^2\sin\alpha + c_{1,2}l^2\cos\alpha + 2c_{2,0}x_c\cos\alpha, \quad (86)$$

$$M_{202}(\alpha, r_m, r_s, l) = \frac{1}{2}\left(\frac{\cos^2\alpha}{2r_m} - \frac{\sin^2\alpha}{r_s} - c_{2,0}\cos\alpha\right)\Lambda_l^2 + \frac{c_{1,2}l^2\sin 2\alpha}{2r_m} + c_{2,2}l^2\cos\alpha + \frac{c_{2,0}x_c\sin 2\alpha}{r_m} + 3c_{3,0}x_c\cos\alpha, \quad (87)$$

$$M_{022}(\alpha, r_m, r_s, l) = \frac{3}{4r_s}\Lambda_l^2 - \frac{c_{0,2}\Lambda_l\cos\alpha}{2}\left(1 - \frac{5l}{r_s}\right) - \frac{2c_{0,2}^2l^2\sin^2\alpha}{r_m} + 6c_{0,4}l^2\cos\alpha + c_{1,2}x_c\cos\alpha, \quad (88)$$

$$M_{011}(\alpha, r_m, r_s, l) = \Lambda_l + 2c_{0,2}l\cos\alpha, \quad (89)$$

$$M_{013}(\alpha, r_m, r_s, l) = -\frac{1}{2}\left(\Lambda_l + 2c_{0,2}l\cos\alpha\right)\Lambda_l^2 + 4c_{0,4}l^3\cos\alpha + 2c_{1,2}lx_c\cos\alpha, \quad (90)$$

$$M_{111}(\alpha, r_m, r_s, l) = \frac{\Lambda_l\sin\alpha}{r_s} + \frac{c_{0,2}l\sin 2\alpha}{r_m} + 2c_{1,2}l\cos\alpha, \quad (91)$$

$$M_{211}(\alpha, r_m, r_s, l) = -\frac{1}{r_s}\left(\frac{\cos^2\alpha}{2r_m} - \frac{\sin^2\alpha}{r_s} - c_{2,0}\cos\alpha\right)\Lambda_l + \frac{c_{0,2}l\cos\alpha}{r_m^2}\left(2 - 3\cos^2\alpha\right) - \frac{2c_{2,0}c_{0,2}l\sin^2\alpha}{r_m} + \frac{c_{1,2}l\sin 2\alpha}{r_m} + 2c_{2,2}l\cos\alpha, \quad (92)$$

$$M_{031}(\alpha, r_m, r_s, l) = -\frac{\Lambda_l}{2r_s^2} + \frac{c_{0,2}\cos\alpha}{r_s}\left(1 - \frac{2l}{r_s}\right) - \frac{2c_{0,2}^2l\sin^2\alpha}{r_m} + 4c_{0,4}l\cos\alpha. \quad (93)$$

APPENDIX *B*
Transverse aberration coefficients

$$\Lambda_m = \frac{r'_m - r'_0}{r'_m}, \quad \Lambda_s = \frac{r'_s - r'_0}{r'_s}, \quad \Lambda'_l = \frac{r'_s - l'}{r'_s} = 1 + \frac{l}{r'_s}, \quad (94)$$

$$d_{100} = \Lambda_m\cos\beta, \quad (95)$$

$$d_{200} = -\frac{3r'_0w_{300}}{\cos\beta} + \Lambda_m\sin\beta\left(\frac{\cos\beta}{r'_m} - c_{2,0}\right), \quad (96)$$

$$d_{300} = -\frac{4r'_0w_{400}}{\cos\beta} + 3\tan\beta\left(1 + \frac{r'_0}{r'_m} - \frac{2r'_0c_{2,0}}{\cos\beta}\right)w_{300} + \Lambda_m\left(\frac{\cos(2\beta)c_{2,0}}{r'_m} + \frac{\cos\beta\sin^2\beta}{r'^2_m} - \sin\beta c_{3,0}\right), \quad (97)$$

$$d_{020} = -\frac{r'_0w_{120}}{\cos\beta} - \Lambda_m\sin\beta c_{0,2}, \quad (98)$$

$$d_{120} = -\frac{2r'_0w_{220}}{\cos\beta} + \tan\beta w_{120}\left(1 + \frac{r'_0}{r'_m} + \frac{2r'_0}{r'_s} - \frac{2r'_0c_{2,0}}{\cos\beta}\right) + \Lambda_m\left(\frac{\cos(2\beta)c_{0,2}}{r'_m} - \sin\beta c_{1,2}\right), \quad (99)$$

$$d_{011} = -\frac{r'_0w_{111}}{\cos\beta} - 2\Lambda_m\sin\beta c_{0,2}l, \quad (100)$$

$$d_{002} = -\frac{r'_0w_{102}}{\cos\beta} + \frac{x_c\sin(\alpha - \beta)}{\sin\alpha}, \quad (101)$$

$$d_{111} = -\frac{2r'_0w_{211}}{\cos\beta} + \tan\beta r'_0\left(\frac{1}{r'_0} + \frac{1}{r'_m} + \frac{1}{r'_s} - \frac{2c_{2,0}}{\cos\beta}\right)w_{111} + 2\tan\beta r'_0\Lambda'_lw_{120} + 2\Lambda_ml\left(\frac{\cos(2\beta)c_{0,2}}{r'_m} - \sin\beta c_{1,2}\right), \quad (102)$$

$$d_{102} = -\frac{2r'_0 w_{202}}{\cos\beta} + \tan\beta r'_0 \Lambda'_l w_{111} + \tan\beta\left(1 + \frac{r'_0}{r'_m} - \frac{2r'_0 c_{2,0}}{\cos\beta}\right) w_{102} - \Lambda_m \sin\beta\left(2c_{2,0}x_c + c_{1,2}l^2\right), \quad (103)$$

$$h_{010} = \Lambda_s, \quad (104)$$

$$h_{110} = -2r'_0 w_{120} + \frac{\Lambda_s \sin\beta}{r'_s}, \quad (105)$$

$$h_{210} = -2r'_0 w_{220} + 2\sin\beta w_{120} - 6\tan\beta r'_0 c_{0,2} w_{300} + \frac{\Lambda_s}{r'_s}\left(\cos\beta c_{2,0} + \frac{\sin^2\beta}{r'_s}\right) + \frac{(\Lambda_m - \Lambda_s)\cos^2\beta}{2r'_m r'_s}, \quad (106)$$

$$h_{001} = -r'_0 + \Lambda_s l, \quad (107)$$

$$h_{101} = -r'_0 w_{111} + \Lambda_s \Lambda'_l \sin\beta, \quad (108)$$

$$h_{201} = -r'_0 w_{211} + \sin\beta w_{111} - 6\tan\beta r'_0 c_{0,2} l w_{300} + \Lambda_s \Lambda'_l\left(\frac{\sin^2\beta}{r'_s} - \frac{\cos^2\beta}{2r'_m} + \cos\beta c_{2,0}\right), \quad (109)$$

$$h_{021} = -3r'_0 w_{031} - 2\tan\beta r'_0 c_{0,2}(l w_{120} + w_{111}) + \Lambda_s\left[\cos\beta c_{0,2}\left(1 + \frac{3l}{r'_s}\right) - \frac{\Lambda'_l}{2r'_s}\right], \quad (110)$$

$$h_{012} = -2r'_0 w_{022} - 2\tan\beta r'_0 c_{0,2}(l w_{111} + w_{102}) + 2\Lambda_s \cos\beta c_{0,2} l \Lambda'_l, \quad (111)$$

$$h_{030} = -4r'_0 w_{040} - 2\tan\beta r'_0 c_{0,2} w_{120} + \frac{\Lambda_s \cos\beta c_{0,2}}{r'_s}, \quad (112)$$

$$h_{003} = -r'_0 w_{013} - 2\tan\beta r'_0 c_{0,2} l w_{102}. \quad (113)$$

The work was supported by the National Nature Science Foundation of China (60578040).

References

Beutler, H. G. (1945). *J. Opt. Soc. Am.* **35**, 311–350.
Chrisp, M. P. (1983*a*). *Appl. Opt.* **22**, 1508–1518.
Chrisp, M. P. (1983*b*). *Appl. Opt.* **22**, 1519–1529.
Goto, K. & Kurosaki, T. (1993). *J. Opt. Soc. Am.* A**10**, 452–465.
Masui, S. & Namioka, T. (1999*a*). *J. Opt. Soc. Am.* A**16**, 2253–2268.
Masui, S. & Namioka, T. (1999*b*). *Proc. SPIE*, **3737**, 2–12.
Namioka, T., Koike, M. & Content, D. (1994). *Appl. Opt.* **33**, 7261–7274.
Namioka, T., Koike, M. & Masui, S. (2001). *Opt. Precis. Eng. (China)*, **9**, 458–466.
Noda, H., Namioka, T. & Seya, M. (1974). *J. Opt. Soc. Am.* **64**, 1031–1036.
Palmer, C., Wheeler, B. & McKinney, W. (1998*a*). *Proc. SPIE*, **3450**, 55–66.
Palmer, C., Wheeler, B. & McKinney, W. (1998*b*). *Proc. SPIE*, **3450**, 67–77.
Peatman, W. B. (1997). *Gratings, Mirrors and Slits: Beamline Design for Soft X-ray Synchrotron Radiation Sources*, pp. 71–75. Amsterdam: Gordon and Breach.