Numerical Integration of an Aspheric Surface Profile

Juan L. Rayces
5890 N. Placita Alberca
Tucson AZ 85718, U.S.A.
j.rayces@juno.com

Xuemin Cheng
Graduate School at Shenzhen, Tsinghua University
Shenzhen, Guangdong, 518055, China
c.cheng_2004@yahoo.com.cn

Abstract

This paper deals with the design of aspherics on surfaces with steep curves in systems with very large numerical apertures, and small angular fields of view, with a goal to achieve rigorous axial stigmatism for specified finite axial conjugate points. In lens design, where numerical apertures are moderate, aspheric surface descriptions based on Feder's equation of a conic section with several polynomial terms give satisfactory results to a certain point. For larger numerical apertures plot of aberrations start showing ripples; increasing the number of polynomial terms is no remedy. This problem is similar to Runge's phenomenon and it is probably related to it, but yet not well understood. Alternative surface descriptions have been proposed with apparently some degree of success.

Here we propose a method for designing aspheric surfaces in systems with large numerical apertures. In this method the profile of the aspheric is presented as a table of coordinates of key points, the points of intersection of meridian key rays with the surface. Such rays are rigorously corrected for spherical aberration for specified finite conjugates. For tracing other rays through the gap between key points, interpolation with a spline of degree 3 is proposed. Correction with this method has proved satisfactory for certain high aperture examples designed with a small number of key rays. For systems with still higher apertures the number of key rays may have to be increased and the interval between key points reduced accordingly but the degree of the spline is maintained to avoid ripples similar to Runge's phenomenon.

1. INTRODUCTION

In lens design correction of axial spherical aberration may be done with aspheric surfaces depending on how it affects other aberrations. Chances are that the solution with one aspheric surface is possible if the field of view is sufficiently small even when numerical apertures are very large. Such systems include many high efficiency illumination condensers, some refractive microscope objectives, most reflective microscope objectives, and all optical disks pick up heads, etc..

D. Feder [1] proposed in 1951 a description of an aspheric surface consisting of an explicit equation

\[ x = f(y, z), \]

of a sphere plus a polynomial:

\[
x = \frac{1 - \sqrt{1 - Q^2(y^2 + z^2)}}{Q} + \sum_{n=1}^{N} a_n(y^2 + z^2)^n,
\]

This equation is given in a Cartesian frame \( \bar{X}, \bar{Y}, \bar{Z} \), with \( \bar{X} \) as of rotational symmetry axis, \( x \) denotes sag of the sphere, \( Q \) spherical curvature, and \( a \) are polynomial coefficients.

A coefficient \( K \) called "conic constant" was anonymously included into the first term of this equation to turn it into a general equation for all surfaces of revolution generated by a conic section. In addition it was modified to a form similar to the alternative formula for solution of the equation of the second degree [2] in order to avoid indetermination.
when $Q \to 0$. The constant $Q$ in this modified equation denotes curvature of the osculating sphere at the pole or paraxial curvature.

$$x = \frac{Q(y^2 + z^2)}{1 + \sqrt{1 - (\kappa + 1)Q^2(y^2 + z^2)}} + \sum_{n=2}^{N} a_n(y^2 + z^2)^n,$$

This equation became a satisfactory standard for many years and designers were happy using a $10^{th}$ degree even polynomial. Note that the polynomial is missing the second degree (term $n = 1$) because it is implicitly contained in the irrational function as a power series expansion would show.

Whether intentionally or not Feder did not specify the polynomial degree. Many designers, confusing curve fitting with series expansion, had the false hope that any steep aspheric could be designed provided that sufficient terms were included in the polynomial expression, and some of the lens design programs providers fomented this misconception by allowing the possibility of using polynomials of practically any degree.

Professor Hildebrand [3] gave good advice on the subject: "it is foolish (and, indeed, inherently dangerous) to attempt to determine a polynomial of high degree which fits the vagaries of such data exactly and hence, in all probability, is represented by a curve which oscillates violently about the curve which represents the true function"

In addition, to reinforce Hildebrand's advice, there is the famous "Runge's Phenomenon" mentioned by many authors (e.g. Lanczos [4] but sorry, nobody gives a reference to an original article). Trying to interpolate a specific simple function with a polynomial of $n$th degree Runge in 1901 discovered that outside a given interval the polynomial diverges, and the oscillation gets worse the higher the polynomial degree.

Indeed neither Hildebrand's advice nor Runge's experiment apply exactly to aspheric lens design since this problem is neither curve fitting nor polynomial interpolation but both are healthy reminders of the danger of trusting too much on the power of polynomials.

Greynolds [5] discussed several alternatives to the standard aspheric description: "more general optical conicoid", "superconic surface", "subconic surface", "rational conic Bezier". This author leaves the choice up to the user stating that all of them have limitations. However, if the image is formed at infinity there is an exact solution, as proposed by Professor Emil Wolf [6]

Chase [7] discussed the use of "Non-Uniform Rational B-Splines" (NURBS) as aspherics. In the Conclusions Section he states: "Optical design with NURBS is fraught with peril. The cost of using NURBS is a dramatic increase in the complexity of the design process".

Wassermann and Wolf [5] introduced a method of numerical integration of simultaneous differential equations for the design of pairs of aspheric profiles, specifically to correct spherical aberration for a pair of conjugate foci and, at the same time, to satisfy the sine condition, that is, to achieve exact aplanatism.

In the Wassermann-Wolf solution rays are traced in opposite directions from conjugates axial points in object and image space forming angles that satisfy the sine condition. These rays, possibly traced through several spherical surfaces, reach the region of two consecutive aspheric surfaces where the integration of the differential equations takes place.

The output of this algorithm is a table of coordinates $x, y$ of a large number of points on the aspheric profiles. Since this method is meant for a surface of revolution the extension to the 3-D case is trivial.

Wassermann and Wolf state in the referenced paper: "Our differential equations permit a rigorous solution of the problem and can be solved to any desired accuracy by standard numerical methods". This statement obviously applies only to the set of key points selected for the integration of the differential equations. As long as there is a finite gap between these key points some sort of interpolation will be necessary for tracing rays to verify performance for conjugate points other than the pair of axial points for which the system is made aplanatic. It is here that the question of accuracy comes up. Wassermann and Wolf paper was published 58 years ago at the beginning of the electronic computer age. Though lacking a method of interpolation between key points to be of a practical application, it was a novel, interesting and elegant investigation.
Only Snell's law, Fermat's principle and the Newton-Raphson iterative process are needed for a point-by-point design of the aspheric profile on a steep curve, as it will be shown here.

2. A CASE STUDY

A method of aspheric design such as proposed here will be used as a way to improve performance of either a preliminary design or of an already existing design that turned out to be short of expectations. For an example, we use throughout this paper a two-element condenser, corrected for axial spherical aberration with paraxial magnification $m_o = 5.7x$. As customary it is designed from long conjugate to short conjugate. Object space, object point, etc. are names of entities associated with the long conjugate. Image space, image point, etc. are names of entities associated with the short conjugate. Specifications require numerical aperture $NA = 0.875$ in image space, therefore $\sin U = NA/m_o$ in object space.

![Figure 1. NA=0.875 condenser: in this example 16 rays are used to compute coordinates of key points on aspheric profile.](image)

Figure 1 above shows a scale drawing of the condenser system, the aspheric surface aperture diameter is 6.36cm and the working distance is 1.75cm. The diameter of an aspheric is important to note because residual aberrations vary linearly with dimension. An optical disk lens with size 1/20 of this diameter should be a breeze to design.

Construction data of NA=0.875 condenser is given in Table I below

<table>
<thead>
<tr>
<th>Table I. NA=0.875 Aspheric Condenser, construction data, $\lambda = 0.00005461cm$</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1=11.383039cm</td>
</tr>
<tr>
<td>R2=-11.383039cm</td>
</tr>
<tr>
<td>R3= 2.479980cm</td>
</tr>
<tr>
<td>R4=10.291983cm</td>
</tr>
</tbody>
</table>

3. POINT-BY-POINT GENERATION OF THE ASPHERIC PROFILE

Basically this method of design consists in finding by trial and error a sequence of points in the optical space of the aspheric where, pairs of rays coming from $F$ in object space and from $F'$ in image space (see figure 1) will mutually intersect and become one ray throughout. At each tentative point of intersection the Optical Path Difference is computed to check if Fermat's principle is satisfied: that is a point on the aspheric profile. The specific purpose of this section is to give details of this trial and error method.

Standard notation is used here for quantities associated with a single ray: $U$ is angle of a typical ray with the axis positive counterclockwise, $L$ is distance of intersection of the ray with the axis positive from the vertex of a surface to the right, quantities to the right (image space side) are primed and they are plain otherwise, $X$ and $Y$ are coordinates of point of intersection of ray and surface with $x$ coincident with the optical axis. Additional notation will be clarified as necessary in the course of this explanation.
Starting from left (object space) surfaces of the system will have subscripts $1, 2, \ldots, j, \ldots, fJ$ to denote the aspheric surface and $\omega$ to denote the last surface.

A fan of key rays is traced, one at a time, starting at object point $F$, and forward through all optical surfaces to an $X, Y$, coordinate system with origin at the pole of the aspheric surface, the $X$-axis coincident with the optical axis and the $Y$-axis normal to it. These key rays in the fan originated at point $F$ will have, if necessary a second subscript $1, 2, 3, \ldots n_{\text{max}}$

Consider the $mth$ key ray, defined in object space by:

$$\sin U_{1,n} = n \cdot \Delta \sin U.$$  \hspace{1cm} (3)

where $\Delta \sin U$ is an increment such that

$$\Delta \sin U = \frac{NA}{m_0}.$$ \hspace{1cm} (4)

Standard equations are used for ray tracing. At each surface the optical path difference contribution is computed with Hopkins’ [9] optical path difference formula:

$$\Omega_j = \left( X_j + \frac{Y_j \cdot \sin U_j'}{1 + \cos U_j'} \right) N_j' - \left( X_j + \frac{Y_j \cdot \sin U_j}{1 + \cos U_j} \right) N_j$$ \hspace{1cm} (5)

Comparing equation (5) with Hopkins’ original formula this one is simpler because (1) the ray is on the meridian plane and (2) the ray corresponds to axial conjugate points and its difference is taken to a ray along the axis: that does not appear explicitly in this formula.

The sum of optical path differences of all surfaces from the first to the one immediately before the aspheric surface is saved.

For each key ray traced forward, three trial rays, (denoted with a bar over the symbol) and labeled with superscripts $(\alpha), (\beta), (\gamma)$, will be traced backwards from axial image point $F'$ starting with:

$$\sin \bar{U}_{\alpha \omega} = m(\alpha) \sin U_1, \quad m(\alpha) = m_{n-1},$$

$$\sin \bar{U}_{\beta \omega} = m(\beta) \sin U_1, \quad m(\beta) = (1 + 0.001) \cdot m(\alpha),$$

$$\sin \bar{U}_{\gamma \omega} = m(\gamma) \sin U_1, \quad m(\gamma) = (1 - 0.001) \cdot m(\alpha).$$ \hspace{1cm} (6)

where $m_{n-1} = \left[ \sin U_{\alpha \omega}/\sin U_1 \right]_{n-1}$ is the computed magnification of the previous ray. For the first three trial rays the paraxial magnification $m_{n-1} = m_0$ will be used.

In equations (6) above and (7), (8) below symbols with a bar on top denote quantities associated with trial rays, while plain symbols denote quantities associated with the key ray.

At the intersections of each of the three trial rays with each of the surfaces going backwards from $\omega$, to $\mu + 1$, (the one immediately following the aspheric), optical path differences $\Omega_j$ are computed with Hopkins’ formula.

At the end of this ray tracing the summation of these will be available, as well as data for rays in the region of the coordinate system of the aspheric:

$$\sum_{j=\mu+1}^{\omega} \Omega_j^{(\alpha)}, \quad \bar{H}_{\mu}^{(\alpha)}, \quad \sin \bar{U}_{\mu}^{(\alpha)}, \quad \cos \bar{U}_{\mu}^{(\alpha)};$$

$$\sum_{j=\mu+1}^{\omega} \Omega_j^{(\beta)}, \quad \bar{H}_{\mu}^{(\beta)}, \quad \sin \bar{U}_{\mu}^{(\beta)}, \quad \cos \bar{U}_{\mu}^{(\beta)};$$

$$\sum_{j=\mu+1}^{\omega} \Omega_j^{(\gamma)}, \quad \bar{H}_{\mu}^{(\gamma)}, \quad \sin \bar{U}_{\mu}^{(\gamma)}, \quad \cos \bar{U}_{\mu}^{(\gamma)}.$$ \hspace{1cm} (7)

The coordinates of the three intersections of trial rays with key ray can now be computed with these equations:
\[
X^{(\alpha)}_\mu = \frac{\bar{H}^{(\alpha)}_\mu - H_\mu}{\tan \bar{U}^{(\alpha)}_\mu - \tan U_\mu}, \quad y^{(\alpha)}_\mu = H_\mu + X^{(\alpha)}_\mu \tan U_\mu,
\]
\[
X^{(\beta)}_\mu = \frac{\bar{H}^{(\beta)}_\mu - H_\mu}{\tan \bar{U}^{(\beta)}_\mu - \tan U_\mu}, \quad y^{(\beta)}_\mu = H_\mu + X^{(\beta)}_\mu \tan U_\mu,
\]
\[
X^{(\gamma)}_\mu = \frac{\bar{H}^{(\gamma)}_\mu - H_\mu}{\tan \bar{U}^{(\gamma)}_\mu - \tan U_\mu}, \quad y^{(\gamma)}_\mu = H_\mu + X^{(\gamma)}_\mu \tan U_\mu.
\] (8)

At this point all that is necessary to compute optical path differences at intersection points is available, as well as the corresponding three values of the magnification used to compute them:
\[
\Omega^{(\alpha)}_\mu, \quad m^{(\alpha)}_\mu,
\]
\[
\Omega^{(\beta)}_\mu, \quad m^{(\beta)}_\mu,
\]
\[
\Omega^{(\gamma)}_\mu, \quad m^{(\gamma)}_\mu.
\] (9)

The optical path difference \( \Omega \) is now a function \( f(m) \). We find the value of \( m^* \) that will give a value of \( \Omega^* \) that will balance the summation of contributions of optical path differences before and after the aspheric and \( \Delta \Omega = 0 \):
\[
\Delta \Omega = \Omega^*_\mu - \left( \sum_{j=1}^{\mu-1} \Omega_j + \sum_{j=\mu+1}^{\sigma} \Omega_j \right)
\] (10)

We use the Newton-Raphson process to solve this problem
\[
\Delta \Omega_{i+1} = \Delta \Omega_i - \frac{f(m_i)}{f'(m_i)}.
\] (11)

Figure 2.- Numerical integration of the aspheric profile.

We choose the intersection of the key ray with trial ray \((\alpha)\) as a first approximation, that is: \( \Omega_1 = \Omega^{(\alpha)} \). We use rays \((\beta), (\gamma)\) to compute a sufficiently good approximation to the derivative:
\[
f'(m) \approx \frac{\Omega^{(\beta)} - \Omega^{(\gamma)}}{m^{(\beta)} - m^{(\gamma)}},
\] (12)

At each \( n \)th key point and key ray Newton-Raphson method is applied iteratively until a sufficiently small residual \( \Delta \Omega \) is achieved, small meaning a fraction of the wavelength.
Each time the goal is reached direction cosines of the normal (angle $\sigma$) are computed with these formulas derived from the graphic construction, see figure 3.

$$
\cos \sigma_\mu = \frac{N'_\mu \cos U'_\mu - N_\mu \cos U_\mu}{\sqrt{(N'_\mu \cos U'_\mu - N_\mu \cos U_\mu)^2 + (N'_\mu \sin U'_\mu - N_\mu \sin U_\mu)^2}}
$$

$$
\sin \sigma_\mu = \frac{N'_\mu \sin U'_\mu - N_\mu \sin U_\mu}{\sqrt{(N'_\mu \cos U'_\mu - N_\mu \cos U_\mu)^2 + (N'_\mu \sin U'_\mu - N_\mu \sin U_\mu)^2}}
$$

(13)

Fig. 3. Equations (13) can be derived from a graphical construction of Snell’s law.

The whole process is repeated for each one of the key points from $n = 1$ to $n = n_{\text{max}}$. The computer program output would be a lookup table similar to Table II below.

| Table II. Key point coordinates, and components of the unit vector in the direction of the normal $\sigma$ of the aspheric profile of the $\text{NA}=0.875$ condenser. |
|---|---|---|---|
| 0 | $X$ (cm) | $Y$ (cm) | $\cos \sigma$ | $\sin \sigma$ |
| 1 | 0.014011 | 0.264327 | 0.994396 | -0.105717 |
| 2 | 0.056200 | 0.527680 | 0.977746 | -0.209793 |
| 3 | 0.126821 | 0.789052 | 0.950514 | -0.310682 |
| 4 | 0.226383 | 1.047362 | 0.913418 | -0.407022 |
| 5 | 0.355585 | 1.301415 | 0.867359 | -0.497683 |
| 6 | 0.515321 | 1.549859 | 0.813336 | -0.581795 |
| 7 | 0.706681 | 1.791126 | 0.752376 | -0.658733 |
| 8 | 0.930923 | 2.023364 | 0.685477 | -0.728094 |
| 9 | 1.189423 | 2.244364 | 0.613570 | -0.789640 |
| 10 | 1.483563 | 2.451461 | 0.537519 | -0.843251 |
| 11 | 1.814496 | 2.641434 | 0.458150 | -0.888875 |
| 12 | 2.182666 | 2.810416 | 0.376326 | -0.926487 |
| 13 | 2.586820 | 2.953875 | 0.293100 | -0.956082 |
| 14 | 3.021966 | 3.066834 | 0.209994 | -0.977703 |
| 15 | 3.475162 | 3.144747 | 0.129547 | -0.991573 |
| 16 | 3.917565 | 3.185884 | 0.056382 | -0.998409 |
4. APPROXIMATE INTERSECTION OF A RAY.

Let all key points $P_n$, rotate about the $x$-axis: they will generate rings of radii $R_n$ and will cut the aspheric lens into circular slices where the rim surfaces are generated by arcs $\overline{P_nP_{n+1}}$ of the aspheric profile.

On the meridian plane the pair of consecutive points will have coordinates $X_n, Y_n$, and $X_{n+1}, Y_{n+1}$. Draw a segment of a line through those points. The equation of the line is

$$a \cdot x + b \cdot y - 1 = 0$$

with coefficients

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} X_n & Y_n \\ X_{n+1} & Y_{n+1} \end{bmatrix}^{-1} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$  \hfill (15)

Substitution of the radial variable $r = \sqrt{y^2 + z^2}$, for the variable $y$ converts the equation of the line into the equation of a straight cone. This equation is rationalized to take the form:

$$(1 - \alpha x)^2 - b^2(y^2 + z^2) = 0$$

The slices of aspheric lens are now slices of a cone.

---

Fig.4 – First approximation of the aspheric surface slice by a cone slice.

Let

$$x = H_x + t \cdot \cos \alpha,$$

$$y = H_y + t \cdot \cos \beta,$$

$$z = H_z + t \cdot \cos \gamma.$$ \hfill (17)

be the parametric equations of a ray. When these expressions for $x, y, z$, equation are substituted in (17) above WE obtain an equation of the second degree in $t$:

$$\overline{A} \cdot t^2 + \overline{B} \cdot t + \overline{C} = 0$$

with coefficients:

$$\overline{A} = a^2 \cdot \cos^2 \alpha - b^2 \cdot \left(\cos^2 \beta + \cos^2 \gamma\right),$$

$$\overline{B} = 2 \cdot a \cdot \cos \alpha \cdot (a \cdot H_x - 1) - 2 \cdot b^2 \cdot \left(H_y \cdot \cos \beta + H_z \cdot \cos \gamma\right),$$

$$\overline{C} = \left(1 - a \cdot H_x\right)^2 - b^2 \cdot \left(H_y^2 + H_z^2\right).$$

(18)

(19)
and solution:

\[ t = \frac{2 \cdot C}{B + \sqrt{B^2 - 4 \cdot A \cdot C}} \]

The intersection coordinates \( \hat{x}, \hat{y}, \hat{z} \) of the ray on the rim surface of cone slice are computed with \( \hat{t} \) and equations (17).

We repeat the procedure on all cone slices starting from the edge down until we find

\[ R_n^2 \leq (\hat{y}^2 + \hat{z}^2) \leq R_{n+1}^2 \]

That will tell us that the ray intersects the rim of the \( n \)th cone slice. \( \hat{t} \) will be a good first approximation for the Newton-Raphson method.

4. INTERPOLATING METHOD

Interpolation is an important part of numerical analysis. There are different methods of interpolation, some are general, other meet specific requirements.

![Diagram](image)

Fig. 5. Aspheric profile and binomial interpolating curve drawn to scale. Compare distance between pair of points to distance between maximum and minimum of spline of degree 3.

Here we propose an interpolation method for consecutive points \( P_n, P_{n+1} \) on the aspheric profile. In addition to passing through these points the interpolation is required to be smooth, that is the normals match at these points.

What else is important in geometrical optics? In view of the fact that ripples seem to plague the design of aspherics it seems wise to heed the advice of Hildebrand and avoid the Runge phenomenon. For these reasons we opt to choose a spline of the form:

\[ v = A \cdot u^2 + B \cdot u^3 \]

or, in implicit form:

\[ \Phi = A \cdot u^2 + B \cdot u^3 - v, \quad \text{with} \quad \Phi = 0 \]

Let \( P_n, P_{n+1} \) be the pair of neighboring points on the aspheric profile, see figure 5. Take a system of local coordinates \( u, v \), with origin at \( P_n \). The \( v \)-axis is in the direction of the normal while the \( u \)-axis is in the direction of the tangent. Let \( u_{n+1}, v_{n+1} \) be the coordinates of \( P_{n+1} \). Let \( \sigma_{n+1} \) be the angle of the normal at \( P_{n+1} \) and \( \tau_{n+1} \) the angle the tangent at \( P_{n+1} \) makes with the \( u \)-axis. We then have \[ \tan \tau_{n+1} = \frac{dv}{du}_{P_{n+1}}. \] For simplicity we let \( \tan \tau_{n+1} = v'_{n+1}. \)

We feed into the spline equation the values of \( u_{n+1}, v_{n+1}, v'_{n+1} \). It is the clear that:
where $A_n, B_n$ are the coefficients corresponding to the $n$th arc $\overline{P_n P_{n+1}}$.

The use of a spline of degree 3 is a novel idea that promises the following. Suppose an aspheric is designed to a prescribed number of key points between the pole and the extreme aperture, then oscillation does not appear if one increases this number. This keeps the recomputed surface smooth and avoids ripples. This also avoids the annoyance of having the intersection of the marginal ray with the aspheric not satisfactory. As the number of points in increased the separation between these points decreases accordingly. For this reason one guarantees that the interpolation function described by (23) possesses no maxima and no minima on the arc $\overline{P_n P_{n+1}}$ and in fact increases monotonically on this arc.

To illustrate this point, a circular arc of unit radius and $30^\circ$ opening was replaced the spline of degree 3. The maximum error was 1 part in 1175 of the radius. When the opening was halved to $15^\circ$ the maximum error was 1 part in 24752 of the radius or about 1/20th smaller.

A lookup table like Table III below, generated at the time the key points are computed, provides the coordinates of the family of points $\{P_n\}$ as well as components of key unit vectors in the main coordinate system.

### Table III

<table>
<thead>
<tr>
<th>$P_n$</th>
<th>$X_n$</th>
<th>$Y_n$</th>
<th>$\cos \sigma_n$</th>
<th>$\sin \sigma_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{n+1}$</td>
<td>$X_{n+1}$</td>
<td>$Y_{n+1}$</td>
<td>$\cos \sigma_{n+1}$</td>
<td>$\sin \sigma_{n+1}$</td>
</tr>
</tbody>
</table>

The normal and tangential lines are related by $\tau_n = \sigma_n + \frac{\pi}{2}$

5. INTERSECTION OF A SKEW RAY WITH THE ASPHERIC APPROXIMATION

In this section we derive an expression for the intersection of a skew ray with the surface generated by a segment of the curve representing the spline revolving around the $x$-axis, i.e. the aspheric approximation. In this expression $\Phi = 0$ denotes the solution. It will be given as a function of the parameter $t$ of the ray equation. Also an expression of its derivative $\Phi'$ will be given.

The elements involved in the tracing of skew rays in an aspheric system designed as proposed in this paper are shown in Figure 6.

There is one main, 3-D, coordinate system $x, y, z$, associated with the aspheric surface of revolution with origin at the pole $P_0$ of the aspheric surface. The axis of revolution is the $x$-axis. The $y$-axis is orthogonal to the former and both define the meridian plane. The $z$-axis is orthogonal to the other two axes.

Ray tracing of a skew ray is started with the parametric equations of the ray, equations (17) as explained in section 4 to find the cone slice whose rim is intercepted by the ray. The coordinates $\hat{x}, \hat{y}, \hat{z}$ of this point of intersection of ray and cone are going to be used as the starting point in Newton-Raphson iteration method:

The local coordinates are based at $P_n$ (with $u_n = v_n = 0$; notice as well that $w_n = 0$ for all $n$). We have the following recursive formulae.

$$
\begin{bmatrix}
  u_{n+1} \\
  v_{n+1}
\end{bmatrix} =
\begin{bmatrix}
  \cos \tau_n & \sin \tau_n \\
  -\sin \tau_n & \cos \tau_n
\end{bmatrix}
\begin{bmatrix}
  X_{n+1} - X_n \\
  Y_{n+1} - Y_n
\end{bmatrix}
$$

The components of the tangent based on $P_{n+1}$ in a coordinate system at the base point given by a rotation about the $w$-axis:

$$
\begin{bmatrix}
  \cos \zeta_{n+1} \\
  \sin \zeta_{n+1}
\end{bmatrix} =
\begin{bmatrix}
  \cos \tau_n & \sin \tau_n \\
  -\sin \tau_n & \cos \tau_n
\end{bmatrix}
\begin{bmatrix}
  \cos \tau_{n+1} \\
  \sin \tau_{n+1}
\end{bmatrix}
$$
In the local coordinates system a unit vector the direction of the tangent line has a slope:

\[ v_{n+1}' = \tan \phi_{n+1} \]  

(27)

An expression of the spline in space of the main frame is needed. This is done with another coordinate transformation. The variables in our spline are replaced by

\[ u = \cos \tau_n \cdot (x - X_n) + \sin \tau_n \cdot (y - Y_n) \]  

\[ v = -\sin \tau_n \cdot (x - X_n) + \cos \tau_n \cdot (y - Y_n) \]  

(28)

The transformed equation is still that of a plane curve in the meridian plane.

To obtain our surface of revolution, all we have to do is to substitute the radial variable \( r \) for \( y \) in equation (28). Therefore one has

\[ u = \cos \tau_n \cdot (x - X_n) + \sin \tau_n \cdot (r - Y_n) \]  

\[ v = -\sin \tau_n \cdot (x - X_n) + \cos \tau_n \cdot (r - Y_n) \]  

(29)

The function \( \Phi \) defined by:

\[ \Phi = u^2 A_n + u^3 B_n - v \]  

(30)

can be thought of as a “merit function”.

Differentiating this equation gives:

\[ \frac{d\Phi}{dt} = (2 \cdot A_n + 3 \cdot u \cdot B_n) \cdot u \cdot \frac{du}{dt} - \frac{dv}{dt} \]  

(31)

The following relationships need not be computed but are needed for derivation of subsequent formulas:
from (17): \[
\frac{dx}{dt} = \cos \alpha, \quad \frac{dy}{dt} = \cos \beta, \quad \frac{dz}{dt} = \cos \gamma,
\]
and from \( r^2 = y^2 + z^2 \)
\[
\frac{dr}{dt} = \frac{y}{r} \frac{dy}{dt} + \frac{z}{r} \frac{dz}{dt} = \frac{y \cdot \cos \beta + z \cdot \cos \gamma}{r},
\]

Since:
\[
\frac{du}{dt} = \frac{du}{dx} \frac{dx}{dt} + \frac{du}{dr} \frac{dr}{dt},
\]
\[
\frac{dv}{dt} = \frac{dv}{dx} \frac{dx}{dt} + \frac{dv}{dr} \frac{dr}{dt},
\]
we have
\[
\begin{bmatrix}
\frac{du}{dt} \\
\frac{dv}{dt}
\end{bmatrix}
= \begin{bmatrix}
\cos \tau_n & \sin \tau_n \\
-\sin \tau_n & \cos \tau_n
\end{bmatrix}
\begin{bmatrix}
\frac{\cos \alpha}{r} \\
\frac{y \cdot \cos \beta + z \cdot \cos \gamma}{r}
\end{bmatrix}
\]

These are numerical values to be used in evaluating (36).

The value of \( t \) corresponding to the intersection of the ray with the aspheric surface will be found using the Newton-Raphson method:
\[
t_{m+1} = t_m - \frac{\Phi(t_m)}{\Phi'(t_m)},
\]

When the difference between two consecutive approximations is sufficiently small the computation is terminated.

6. DIRECTION COSINES OF THE NORMAL AT THE POINT OF INTERSECTION

To find the direction cosines of the normal, we first differentiate \( u \) and \( v \) to \( x \) and \( r \), using equations (29):
\[
\frac{du}{dx} = \cos \tau_n, \quad \frac{du}{dr} = \sin \tau_n, \quad \frac{dv}{dx} = -\sin \tau_n, \quad \frac{dv}{dr} = \cos \tau_n.
\]

Then from: \( r^2 = y^2 + z^2 \), we have:
\[
\frac{dr}{dy} = \frac{y}{r}, \quad \frac{dr}{dz} = \frac{y}{z}.
\]

We write now the implicit form of the spline equation,
\[
\Phi = u^2 \cdot A_n + u^3 \cdot B_n - v
\]

Differentiate this function to \( x \) and \( r \), to get:
\[
\frac{\partial \Phi}{\partial x} = \left( 2 \cdot u \cdot A_n + 3 \cdot u^2 \cdot B_n \right) \cdot \cos \tau_n + \sin \tau_n,
\]
\[
\frac{\partial \Phi}{\partial r} = \left( 2 \cdot u \cdot A_n + 3 \cdot u^2 \cdot B_n \right) \cdot \sin \tau_n + \cos \tau_n,
\]

The second one is split again into these two:
\[
\frac{\partial \Phi}{\partial y} = \left[ \left( 2 \cdot u \cdot A_n + 3 \cdot u^2 \cdot B_n \right) \cdot \sin \tau_n + \cos \tau_n \right] \cdot \frac{y}{r},
\]
\[
\frac{\partial \Phi}{\partial z} = \left[ \left( 2 \cdot u \cdot A_n + 3 \cdot u^2 \cdot B_n \right) \cdot \sin \tau_n + \cos \tau_n \right] \cdot \frac{z}{r}.
\]
Let

\[ K = \sqrt{\left( \frac{\partial \Phi}{\partial x} \right)^2 + \left( \frac{\partial \Phi}{\partial y} \right)^2 + \left( \frac{\partial \Phi}{\partial z} \right)^2} \]  \hspace{1cm} (41)

Then the direction cosines of the normal are:

\[
\cos \xi = \frac{1}{K} \frac{\partial \Phi}{\partial x}, \quad \cos \eta = \frac{1}{K} \frac{\partial \Phi}{\partial y}, \quad \cos \zeta = \frac{1}{K} \frac{\partial \Phi}{\partial z}.
\]  \hspace{1cm} (42)

Finding the intersection of the ray with a refracting surface and finding the direction of the normal, as we have done, are two steps that make all the difference between tracing rays through a surface of a kind and tracing rays through a surface of a different kind. The next two steps, finding the refracted ray with the laws of refraction/reflection and transferring the ray to the coordinate system of the next optical surface are common to all ray tracing methods.

7. CONCLUSIONS

It is proposed here a method of designing a rotationally symmetric aspheric surface by computing a table of coordinates of points on its profile. This method will be most useful as a final improvement on a preliminary design with aspherics described by the standard equation with a polynomial correction where oscillations (akin to Runge's phenomenon) due to many terms and high orders in the polynomial make the solution unacceptable.

It is also proposed a ray tracing scheme to suit the special problem of an aspheric defined by a table of coordinates rather than being described by an equation. This scheme consists of dividing the aspheric into slices corresponding to the computed points. The first approximation in the scheme is to replace the aspheric slice by cone slices. Further approximation involves the use of a curve represented by a spline of degree 3 to describe the shape of the rim of the true aspheric slice, and solution with Newton-Raphson method of iteration.

If this approximation is insufficient it will be necessary to add to the number of points and decrease the interval between them without the danger of increasing the amplitude of the oscillation between points. An improvement not discussed here will be to moderately expand the complexity of the spline.

ACKNOWLEDGMENTS

The authors would like to thank Prof. Duanyi Xu and Dr. Jianshe Ma at the Lab of Optical Pickup Head and Servo System, Optical Memory National Engineering Research Center in Tsinghua University, Beijing, for permitting Dr. Xuemin Cheng to participate in this project. We thank Dr. Martha Rosete-Aguilar, Universidad Nacional Autonoma de Mexico and Dr. Lotfi Hermi, Dept. of Mathematics, University of Arizona for valuable assistance in preparing the manuscript. Also thanks to Mr. James Sutter for helping us with references to "Runge's Phenomenon".

REFERENCES

2. H.B.Dwight "Tables of Integrals and Other Mathematical Data" Macmillan, 1967, eq.55.1