The crossed-cylinder aberroscope: an alternative method of calculation of the aberrations

G. Smith¹, R. A. Applegate² and H. C. Howland³

¹Department of Optometry and Vision Sciences, University of Melbourne, Parkville, Victoria 3052, Australia; ²Department of Ophthalmology, University of Texas Health Science Center at San Antonio, San Antonio, Texas, USA; and ³Section of Neurobiology and Behavior, Cornell University, Ithica, New York, USA

Summary
The distorted retinal grid image in the cross-cylinder aberroscope is conventionally analysed using a method based upon orthogonal polynomials. This method restricts the amount of data that can be extracted from the grid image and requires the real grid that is placed between the cross-cylinders to be pre-distorted, with the amount of pre-distortion depending upon the vertex distance. We present an alternative method based upon the minimization of least squares that does not have these restrictions and show that it gives essentially the same results as the original orthogonal polynomial method. Furthermore, the minimization of least squares method also provides a measure of "goodness" of fit (e.g. the minimum of the sum of the squares of the deviations). Copyright © 1996 The College of Optometrists. Published by Elsevier Science Ltd

Introduction
It is 27 years since Howland (1968) published a paper on the use of crossed-cylinders to examine the aberrations of an optical system, and 19 years since Howland and Howland (1976) showed how the technique could be applied to the eye. In all published implementations of this technique (Howland and Howland, 1976, 1977; Walsh et al., 1984; Walsh and Charman, 1985), the original method of calculation (orthogonal polynomials) has been used. In this paper, we present an alternative method, based upon minimization of least squares, and show that the two methods give essentially the same results. However, we will argue that the minimization of least squares is more general.

The wave and transverse aberrations
The wave aberration \( W(X,Y) \) is the optical distance between the wavefront and the ideal spherical wavefront (the reference sphere), i.e. the optical distance \( BB' \) in Figure 1, expressed as a function of the intersection coordinates \( (X,Y) \) in the exit pupil. Due to the aberration of the wavefront, the rays which are normal to the wavefront do not meet the image plane at the ideal or paraxial image point. For any ray, the intersection distances \( \delta \xi' \) and \( \delta \eta' \) from the ideal image point \( Q' \) are the transverse aberrations. An equation given by Welford (1986, p. 98), shows that the above wave aberration and the transverse aberrations are connected. If the aberrations are small, Welford's equations reduce to the following:

\[
\delta \xi' = -\frac{1}{F} \frac{\partial W(X,Y)}{\partial X} \quad \delta \eta' = -\frac{1}{F} \frac{\partial W(X,Y)}{\partial Y}
\]

(1)

where \( F \) is the equivalent power of the lens system.

For a normally healthy eye, we can assume that the wave aberration is a smoothly and slowly varying function, and thus we can express it accurately as a power series in the variables \( X \) and \( Y \). Following Howland and Howland (1976), we can write this up to the fourth power as:

\[
W(X,Y) = w_0 + w_1X + w_2Y + w_3X^2 + w_4XY + w_5Y^2 + w_6X^3 + w_7X^2Y + w_8XY^2 + w_9Y^3 + w_{10}X^4 + w_{11}XY^3 + w_{12}Y^4 + \text{higher order terms}
\]

(2)

Received: 7 March 1995
Revised form: 25 May 1995

222
The crossed-cylinder aberroscope: G. Smith et al. 223

**Figure 1.** The wavefront, transverse aberrations and the wave aberration.

We have not used the A, B, C, etc. notation of Howland and Howland because it is less suitable for use in computer programs, as we will see in a moment.

We can combine Equations (1) and (2) and write the final result in the form:

\[ \delta \xi_i^j F_i = - \sum_{j=1}^{N} w_j f_j (X, Y) \]  \hspace{1cm} (3a)

and

\[ \delta \eta_i^j F_i = - \sum_{j=1}^{N} w_j g_j (X, Y) \]  \hspace{1cm} (3b)

where:

- the \( i \)-th subscript refers to the \( i \)-th ray in a bundle of rays;
- \( N \) refers to the number of terms in the wave aberration function. In Equation (2) above, \( N \) would be 14; and \( f_j (X, Y) \) and \( g_j (X, Y) \) are the differentials \( \partial / \partial X \) and \( \partial / \partial Y \), respectively, of the \( j \)-th 'xy' product term in the polynomial given by Equation (2), and expressions for each are given in Table 1.

In principle, if we can measure the transverse aberrations for a particular set of rays, with each ray defined by its intersection \( (X, Y) \) in the pupil of the system, then we can determine the wave aberration coefficients using Equations (3). One procedure is as follows.

**Least squares solution**

If we use \( M \) rays, we have \( 2M \) simultaneous equations of the form given by Equations (3). However, there are only \( N \) unknowns. Therefore in general, it is not possible to solve these equations to find the values of the \( w \) coefficients in Equation (2) directly, unless \( 2M = N \). Alternatively, we can solve by least squares fitting.

Let us set up the following sum of squares merit function

\[ S = \sum_{i=1}^{M} \left[ \delta \xi_i^j F_i - \sum_{j=1}^{N} w_j f_j (X, Y) \right]^2 + \left[ \delta \eta_i^j F_i - \sum_{j=1}^{N} w_j g_j (X, Y) \right]^2 \]

where \( \omega_i \) is the weighting factor for the \( i \)-th ray.

This is a measure of the weighted sum of squares of the differences between the actual transverse aberrations and those predicted from the polynomial. This merit function is minimum for certain values of the \( w \) coefficients. To find this minimum, we take the differentials:

\[ \frac{\partial S}{\partial w_k} = 0 \]  \hspace{1cm} (5)

for all values of \( k \), for \( k = 1 \) to \( N \). From Equation (4):

\[ \frac{\partial S}{\partial w_k} = \sum_{i=1}^{M} \omega_i \left[ - \delta \xi_i^j F_i - \sum_{j=1}^{N} w_j f_j (X, Y) \right] f_i (X, Y) + \left[ - \delta \eta_i^j F_i - \sum_{j=1}^{N} w_j g_j (X, Y) \right] g_i (X, Y) \]

\[ \frac{\partial S}{\partial w_k} = - \sum_{i=1}^{M} \omega_i \left\{ \delta \xi_i^j F_i f_i (X, Y) + \delta \eta_i^j F_i g_i (X, Y) \right\} \]

This can be re-expressed as:

\[ \sum_{j=1}^{N} w_j \sum_{i=1}^{M} \omega_i \left\{ f_i (X, Y) f_j (X, Y) + g_i (X, Y) g_j (X, Y) \right\} \]

or

**Table 1.** The terms in the wave aberration polynomial \( W(X, Y) \) and the \( X \) and \( Y \) partial derivatives

<table>
<thead>
<tr>
<th>( i )</th>
<th>Term</th>
<th>( f_j (X, Y) )</th>
<th>( g_k (X, Y) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>A</td>
<td>( w_0 ) 0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>B</td>
<td>( w_1 ) ( X )</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>C</td>
<td>( w_2 ) ( Y )</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>D</td>
<td>( w_3 ) ( X^2 )</td>
<td>2X 0</td>
</tr>
<tr>
<td>4</td>
<td>E</td>
<td>( w_4 ) ( XY )</td>
<td>( Y ) X</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>( w_5 ) ( Y^2 )</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>G</td>
<td>( w_6 ) ( X^3 )</td>
<td>3X^2 0</td>
</tr>
<tr>
<td>7</td>
<td>H</td>
<td>( w_7 ) ( X^2 Y )</td>
<td>2XY X^2</td>
</tr>
<tr>
<td>8</td>
<td>I</td>
<td>( w_8 ) ( XY^2 )</td>
<td>( Y^2 ) 2XY</td>
</tr>
<tr>
<td>9</td>
<td>J</td>
<td>( w_9 ) ( Y^3 )</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>K</td>
<td>( w_{10} ) ( X^4 )</td>
<td>4X^3 0</td>
</tr>
<tr>
<td>11</td>
<td>L</td>
<td>( w_{11} ) ( X^3 Y )</td>
<td>3X^2Y X^3</td>
</tr>
<tr>
<td>12</td>
<td>M</td>
<td>( w_{12} ) ( X^2 Y^2 )</td>
<td>2XY^2 2X^2Y</td>
</tr>
<tr>
<td>13</td>
<td>N</td>
<td>( w_{13} ) ( XY^3 )</td>
<td>( Y^3 ) 3XY^2</td>
</tr>
<tr>
<td>14</td>
<td>O</td>
<td>( w_{14} ) ( Y^4 )</td>
<td>0</td>
</tr>
</tbody>
</table>

(4)
\[ \sum_{j=1}^{N} w_i A[j,k] = B[k] \quad (8) \]

where

\[ A[j,k] = \sum_{i=1}^{M} \omega_i \{ f_i(X_i,Y_i)f_k(X,Y) + g_i(X,Y)g_k(X,Y) \} \quad (9a) \]

and

\[ B[k] = -\sum_{i=1}^{M} \omega_i \{ (\delta\xi',F_i)f_k(X,Y) + (\eta',F_i)g_k(X,Y) \} \quad (9b) \]

There are \( k = 1 \) to \( N \) equations of the form given by Equation (8) and these make up a set of \( N \times N \) simultaneous equations which can be expressed in the matrix form:

\[ \mathbf{A} \mathbf{W} = \mathbf{B} \quad (10) \]

where \( \mathbf{A} \) is the two-dimensional matrix of size \( N \times N \) specified by Equation (9a); \( \mathbf{W} \) is a one-dimensional matrix of aberration coefficients \( \{ w_1, w_2, \ldots \} \); and \( \mathbf{B} \) is also a one-dimensional matrix defined by Equation (9b). If we use only the wave aberration expansion up to and including the fourth power terms as described by Equation (2), then \( N \) is 14 and hence 14 simultaneous equations in 14 unknowns have to be solved.

The matrix Equation (10) can be solved to give the \( w_i \) coefficients by standard techniques, but depending upon the nature of the matrix \( \mathbf{A} \), not all standard techniques will give a meaningful solution, e.g., if the matrix is ill-conditioned. However, we have found that the L-U method described in *Numerical Recipes in Pascal* (Press et al., 1989) gave satisfactory solutions.

It should be noted that according to the above theory, the transverse aberrations do not depend upon the \( w_i \) coefficient. Therefore the solution to Equation (10) will not provide a value for \( w_i \). This is in contrast to the ’orthogonal polynomial’ method of Howland and Howland (1976, 1977), which does give a value for this coefficient.

**Scaling the image**

Whether we use the orthogonal polynomials or minimization of least squares, both methods require: (i) the scaling of the observed distorted grid; and (ii) a choice of grid or pupil centre. Here we will show how to incorporate the scaling into the least squares method.

Since the retinal grid is distorted, we have to somehow extract an accurate estimate of the undistorted grid size. The most central grid should be the least distorted and therefore may give the best estimate. However, whatever estimate is used, let \( \Phi \) be the length of an element (e.g., mean length of the four sides) of the observed (and usually magnified) grid image. Now let the grid element length at the crossed-cylinders be \( D \). The corresponding length \( \phi \) on the retina can be deduced from Equation (30a) or (30b), i.e.

\[ \phi = D \frac{F_i}{\Phi} \quad (11) \]

Therefore the retinal grid image being examined has been magnified by a factor \( M \) given by the equation:

\[ M = \Phi \frac{\Phi F_i}{DF_i} \quad (12) \]

Thus the true length \( t \) on the actual retina grid image of a length \( T \) on the magnified retinal grid image being examined is given by the equation:

\[ t = DF_i \frac{DF_i}{\Phi} T \quad (13) \]

It will be convenient to express this in the form:

\[ tF_i = (DF_i/\Phi) T \quad (14a) \]

or

\[ tF_i = S_i T \quad (14b) \]

where

\[ S_i = (DF_i/\Phi) \quad (15) \]

In practice, the distance \( t \) may be either the transverse aberration \( \delta r' \) or \( \delta \eta' \) and so the quantity \( (tF_i) \) may be either of the quantities \( (\delta r'F_i) \) or \( (\eta'F_i) \) shown in Equation (9b).

If we rescale the retinal grid image by the factor \( S_i \), instead of the factor \( 1/M \), then the scaling is independent of the power \( F_i \), but, if we do this, the rescaled values of \( T \) become the values of the quantities \( (\delta r'F_i) \) or \( (\eta'F_i) \) in Equation (9b). That equation now would become:

\[ B[k] = -\sum_{i=1}^{M} \omega_i \{ (DF_i T_i/\Phi)f_k(X,Y) + (DF_i T_i/\Phi)g_k(X,Y) \} \quad (16) \]

where \( T_i \) and \( T_i \) are the \( X \) and \( Y \) transverse aberrations, respectively.

The procedure that is used to determine the grid element size \( \Phi \) can also be used to estimate the pupil centre. While it is not central to this paper, we can note that an error or change in this chosen centre changes the numerical values of all but the highest terms in the wave aberration polynomial specified by Equation (2). This effect has been observed by Atchison et al. (1995) and can readily be proved by replacing the \( X \) and \( Y \) quantities in this equation.
by shifted quantities \((X - X_o)\) and \((Y - Y_o)\), respectively, and then expanding.

We can now conclude that the values of the wave aberration coefficients \(w_i\) are independent of the power \(F\) of the optical system under test. This follows from the fact that neither the coefficients of the \(A\) or \(B\) matrices now involve this power.

We can also conclude that the above equations do not depend upon the particular physical construction of the optical system under test. This follows because as stated in the preceding paragraph, the coefficients of the \(A\) and \(B\) matrices do not involve any details of the system construction. However, there is one exception to this conclusion. In the derivations up to and including Equation (10), the ray coordinates \(X\) and \(Y\) are assumed to be in the pupil. However, in the derivation of Equations (11) to (15), the ray heights are assumed to be measured at the principal planes. Therefore the final equations assume that the entrance pupil and front principal plane coincide. Therefore it follows that the equations are equally applicable to an emmetropic eye as to an ametropic eye with a correction provided that, in both cases, the pupils of the total system approximately coincide with its principal planes. We should note that in Gullstrand’s schematic eye, the separation of the entrance pupil and the front principal plane is only 1.7 mm.

**Pre-distortion of the crossed-cylinder grid**

The ‘orthogonal polynomials’ method of Howland (1968) requires a square grid in the pupil plane. If the crossed-cylinder grid is not placed in this plane, the actual grid must be appropriately distorted (Atchison et al., 1995) to ensure that the ‘virtual’ or ‘equivalent pupil’ grids are square.

The minimization of the least squares method, described here, does not require a square grid and hence does not require any pre-distortion of the actual grid as done by Atchison et al. (1995). Therefore we can use a square grid at any vertex distance provided we calculate the effective grid point coordinates in the pupil plane of the test lens. This can be done as follows.

Using Equation (25) given in Appendix 2, a ray arising from a point \((X_o, Y_o)\) in the original grid, will meet the eye pupil at the point \((X_p, Y_p)\), where:

\[
X_p = X_o + hKF, Y_o
\]

\[
Y_p = Y_o + hKF, Y_o
\]

where:

\(h\) = the distance between the grid and the eye pupil;
\(K = 2 \sin(\psi) \cos(\psi)\);
\(\psi\) = the orientation of the positive cylinder axis;
\(F\) = power of the positive component of the crossed-cylinder.

The coordinates \((X_p, Y_p)\) would then be those used in Equations (9a) and (9b).

**Comparison of the two methods**

We will now show that the two methods (least squares and orthogonal polynomials) give satisfactory comparable results. We could compare the results in two ways: (i) compare the values of the \(w\) coefficients; and (ii) compare the goodness of fit. To carry out a comparison, a theoretical lens was created and the shape of the distorted grid and subsequent transverse aberrations were calculated by standard ray-tracing techniques. The lens was a plano-convex lens (anterior radius of curvature) 16.6666 mm, refractive index of 2.0 and thickness 8.0 mm). These values were chosen to give an equivalent power of 60 D (i.e. close to that of the eye). The lens was placed 25 mm from the 5 D crossed-cylinder and 5 x 5 grid, which had a grid element length of 1.25 mm. The rays were traced from infinity at 5° off-axis to provide both on- and off-axis aberrations. The resulting transverse aberrations in the image plane are shown in Table 2 for no pre-distortion of the actual grid. This table also contains the original grid points at the distance of 25 mm and the corresponding intersection points with the pupil.

The transverse aberrations given in Table 2 were analysed by the least squares method and by the conventional ‘orthogonal polynomial’ method of Howland and Howland (1976, 1977). The least squares method was performed using a PASCAL program. Since this program was written to analyse the output of video monitor image of the retinal grid image, the transverse aberration data shown in Table 2 had to be transformed into the equivalent video array of points. Similarly, this data was not in the form valid for the Howland program and therefore also had to be transformed.

A comparison of the optics of the two methods is given in Appendix 3.

The resulting wave aberration \(\{w\}\) coefficients are shown in Table 3 along with a ‘goodness of fit’ based upon the residual sum of squares. The coefficients are essentially the same, but the orthogonal polynomial method does not give the same goodness of fit. However, the goodness of fit comparison may not be a valid comparison for the following two reasons:

1. The least squares method used a double precision PASCAL program, whereas the orthogonal polynomial method used an early BASIC program of Howland and Howland (1976, 1977). The BASIC languages uses only 7–8 significant digits compared to 12 for the PASCAL program.

2. The least square method would be expected to give the minimum residual sum of least squares, since it in fact minimizes the sum of squares.
Table 2. Ray intersection or grid coordinates at different stages or planes for an undistorted grid placed 25 mm from the lens

<table>
<thead>
<tr>
<th>X, Y</th>
<th>X, Y</th>
<th>X, Y</th>
<th>X, Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.500, 2.500</td>
<td>-1.250, 2.500</td>
<td>0.000, 2.500</td>
<td>1.250, 2.500</td>
</tr>
<tr>
<td>-2.500, 1.250</td>
<td>-1.250, 1.250</td>
<td>0.000, 1.250</td>
<td>1.250, 1.250</td>
</tr>
<tr>
<td>-2.500, 0.000</td>
<td>-1.250, 0.000</td>
<td>0.000, 0.000</td>
<td>1.250, 0.000</td>
</tr>
<tr>
<td>-2.500, -1.250</td>
<td>-1.250, -1.250</td>
<td>0.000, -1.250</td>
<td>1.250, -1.250</td>
</tr>
<tr>
<td>-2.500, -2.500</td>
<td>-1.250, -2.500</td>
<td>0.000, -2.500</td>
<td>1.250, -2.500</td>
</tr>
</tbody>
</table>

Virtual grid points in pupil plane:

<table>
<thead>
<tr>
<th>X, Y</th>
<th>X, Y</th>
<th>X, Y</th>
<th>X, Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2.188, 2.188</td>
<td>-0.938, 2.344</td>
<td>0.313, 2.500</td>
<td>1.563, 2.656</td>
</tr>
<tr>
<td>-2.344, 0.938</td>
<td>-1.094, 1.094</td>
<td>0.156, 1.250</td>
<td>1.406, 1.406</td>
</tr>
<tr>
<td>-2.500, -0.313</td>
<td>-1.250, -0.156</td>
<td>0.000, 0.000</td>
<td>1.250, 0.156</td>
</tr>
<tr>
<td>-2.656, -1.563</td>
<td>-1.406, -1.406</td>
<td>-0.156, -1.250</td>
<td>1.094, -1.094</td>
</tr>
<tr>
<td>-2.813, -2.813</td>
<td>-1.563, -2.656</td>
<td>-0.313, -2.500</td>
<td>0.938, -2.344</td>
</tr>
</tbody>
</table>

Transverse aberrations of finite rays, i.e. intersection points with image plane (retina):

<table>
<thead>
<tr>
<th>X, Y</th>
<th>X, Y</th>
<th>X, Y</th>
<th>X, Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.292, 1.081</td>
<td>0.239, 1.197</td>
<td>0.198, 1.290</td>
<td>0.145, 1.356</td>
</tr>
<tr>
<td>0.161, 1.170</td>
<td>0.124, 1.285</td>
<td>0.101, 1.385</td>
<td>0.072, 1.468</td>
</tr>
<tr>
<td>0.041, 1.215</td>
<td>0.012, 1.325</td>
<td>0.000, 1.429</td>
<td>-0.015, 1.523</td>
</tr>
<tr>
<td>-0.067, 1.240</td>
<td>-0.095, 1.344</td>
<td>-0.104, 1.447</td>
<td>-0.113, 1.549</td>
</tr>
<tr>
<td>-0.162, 1.265</td>
<td>-0.196, 1.360</td>
<td>-0.208, 1.461</td>
<td>-0.217, 1.567</td>
</tr>
</tbody>
</table>

Table 3. A comparison of the aberration coefficients determined by the two methods [see text for details of optical system]. Note that the theory outlined in the text shows that the transverse aberrations do not depend upon the value of \( w_0 \), which is therefore not given by the method of least squares. The pre-distorted grid was for a vertex distance (h) of 25 mm and calculated using Equations (17a, b)

<table>
<thead>
<tr>
<th>j</th>
<th>Term</th>
<th>Orthogonal polynomials</th>
<th>Least squares</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Pre-distortion</td>
<td>No distortion</td>
</tr>
<tr>
<td>0</td>
<td>A ( w_0 )</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>1</td>
<td>B ( w_1 )</td>
<td>0.00019</td>
<td>-0.0455</td>
</tr>
<tr>
<td>2</td>
<td>C ( w_2 )</td>
<td>-85.765</td>
<td>-85.795</td>
</tr>
</tbody>
</table>

Cylindrical:

<table>
<thead>
<tr>
<th>j</th>
<th>Term</th>
<th>Orthogonal polynomials</th>
<th>Least squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>D ( w_1 )</td>
<td>0.55346</td>
<td>0.56072</td>
</tr>
<tr>
<td>4</td>
<td>E ( w_4 )</td>
<td>-0.02266</td>
<td>-0.02856</td>
</tr>
<tr>
<td>5</td>
<td>F ( w_5 )</td>
<td>0.95345</td>
<td>0.96444</td>
</tr>
</tbody>
</table>

Coma-like terms:

<table>
<thead>
<tr>
<th>j</th>
<th>Term</th>
<th>Orthogonal polynomials</th>
<th>Least squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>G ( w_6 )</td>
<td>0.00668</td>
<td>0.00676</td>
</tr>
<tr>
<td>7</td>
<td>H ( w_7 )</td>
<td>0.17179</td>
<td>0.17483</td>
</tr>
<tr>
<td>8</td>
<td>I ( w_8 )</td>
<td>0.02284</td>
<td>0.02058</td>
</tr>
<tr>
<td>9</td>
<td>J ( w_9 )</td>
<td>0.17315</td>
<td>0.17525</td>
</tr>
</tbody>
</table>

Spherical-like terms:

<table>
<thead>
<tr>
<th>j</th>
<th>Term</th>
<th>Orthogonal polynomials</th>
<th>Least squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>K ( w_{10} )</td>
<td>0.02542</td>
<td>0.02573</td>
</tr>
<tr>
<td>11</td>
<td>L ( w_{11} )</td>
<td>-0.00540</td>
<td>-0.00713</td>
</tr>
<tr>
<td>12</td>
<td>M ( w_{12} )</td>
<td>0.05743</td>
<td>0.05812</td>
</tr>
<tr>
<td>13</td>
<td>N ( w_{13} )</td>
<td>-0.00530</td>
<td>-0.00349</td>
</tr>
<tr>
<td>14</td>
<td>O ( w_{14} )</td>
<td>0.02964</td>
<td>0.03000</td>
</tr>
</tbody>
</table>

Residual mean sum of squares:

\[ 4.9e^{-7} \quad 6.2e^{-9} \quad 7.7e^{-9} \]

From these results, we can conclude that the two methods give essentially the same result.

Having established that the least squares method is a valid method of solving for the wave aberration coefficients, we now establish that it is superior in two ways over the orthogonal polynomial method.

(i) It does not require a square array of grid points

In the development of the least square equations, no assumptions were made concerning the values of the coordinates \((X, Y)\). Therefore a grid could be constructed with any set of coordinates. This is in contrast to the orthogonal polynomial method which requires a square grid. This result has two advantages:

1. There is no necessity of a pre-distorted grid. Therefore a square grid could be used, providing the intersection points with the pupil plane are calculated using Equations (17).

2. Experience with the crossed-cylinder technique shows that not every grid point image is clearly defined and it is sometimes difficult to identify a suitable square grid to measure. Because the least squares method does not require a square grid, any grid crossing point can be used providing that its relative position is known. The choice of good and bad grid points can be discriminated through the use of the weighting values \( \omega_i \) in Equations (9).

(ii) It provides a goodness of fit

The current procedure based upon the orthogonal polynomials and associated program written in BASIC does not
provide a measure of goodness of fit. While this procedure could be modified, the least squares method is based upon a goodness of fit and the appropriate value for this quantity would be the mean of the residual sum of squares.

(iii) It readily allows much larger grid arrays

While both methods can use a grid array of any size, the grid size in the orthogonal polynomial method is limited by the availability of sets of orthogonal polynomials. The least squares method has no such restriction.

Acknowledgements

This work was supported in part by a grant from the Australian Research Council and an International Scholars Travel Award from the Research to Prevent Blindness, New York, New York, USA to G.S., Grant EY-02994 from the National Eye Institute, NIH to H.C.H.

References


Appendix 1: Paraxial refraction and transfer equations

For definitions of symbols, refer to Symbol list at the end of the paper.

Spherical surface and systems

Refraction equation:

\[ n'v' - nv = -YF \]  \hspace{1cm} (18)

Transfer equation:

\[ Y' = Y + v'd \]  \hspace{1cm} (19)

Cylindrical surfaces or lenses

Refractive equations:

\[ n'u' - nu = -F, [X\sin(\psi) - Y\sin(\psi)\cos(\psi)] \]  \hspace{1cm} (20a)
\[ n'v' - nv = -F, [-X\sin(\psi)\cos(\psi) + Y\cos^2(\psi)] \]  \hspace{1cm} (20b)

Transfer equations:

\[ X' = X + du' \]
\[ Y' = Y + dv' \]  \hspace{1cm} (21)

Thin crossed-cylinders in contact

Suppose we have two equal and opposite power cylinders, with the positive lens of power \(F,\) with its axis inclined at an angle \(\psi,\) to the \(X\)-axis and the negative lens inclined at an angle \(-\psi,\)

Refraction equation (object at infinity): \(u' = YF/K\)  \hspace{1cm} (22a)

and \(v' = XF/K\)  \hspace{1cm} (22b)

Ray coordinates at a distance \(d\) after the crossed-cylinders: Combining Equations (21) and (22) we have:

\[ X' = X + dF, KY \]
\[ Y' = Y + dF, KX \]  \hspace{1cm} (23)

Appendix 2: Formation of the retinal grid

We can use the paraxial ray equations given in Appendix 1 to examine the formation of the grid of rays used to measure the transverse aberrations of the equivalent real rays. Let us begin by tracing a paraxial ray from infinity through the grid and find where this ray crosses the image plane which is also the back focal plane of the test lens.

Figure 2 shows a ray travelling parallel to the optical axis and meeting the crossed-cylinder at a point \((X, Y),\) Using Equation (23) in Appendix 1, after refraction at the crossed-cylinder we have:

\[ u' = KF, Y \]
\[ v' = KF, X \]  \hspace{1cm} (24)

where the symbols are defined in the Symbol list at the end of the paper.

We now transfer this ray to the test lens which is a
distance \( h \) away. The X-coordinate \( (X_p) \) at the front principal plane of the test lens is given by the transfer equation:

\[
X_p = X_n + hKF_nY_n
\]  \hspace{1cm} (25)

After refraction at this lens, we have:

\[
n''u'' - u' = -X_pF_i
\]  \hspace{1cm} (26)

Combining equations (24), (25) and (26) gives:

\[
n''u'' = -X_iF_i + KF_i(1 - hF_i)Y_i
\]  \hspace{1cm} (27)

This ray meets the back focal plane (which is at a distance \( n''/F_i \) from the back principal plane of the test lens) at the X-coordinate \( X_i \), which is given by the transfer equation:

\[
X_i = X_n + (n''/F_i)u''
\]  \hspace{1cm} (28)

If we substitute for \( X_n \) using Equation (25) and \( (n''/F_i) \) using Equation (26), Equation (28) reduces to:

\[
X_i = +KF_iY_i/F_i
\]  \hspace{1cm} (29a)

Similarly,

\[
Y_i = +KF_iX_i/F_i
\]  \hspace{1cm} (29b)

If we set the axis \( \psi \) of the positive power cylinder at \( +45^\circ \), then \( K = 1 \) and we would then have:

\[
X_i = + (F_i/F_i)Y_i
\]  \hspace{1cm} (30a)

Similarly,

\[
Y_i = + (F_i/F_i)X_i
\]  \hspace{1cm} (30b)

From these last two equations, it follows that the grid is reduced by a factor \( (F_i/F_i) \) and this is independent of the vertex distance \( h \).

Appendix 3: Subjective aberroscope and video camera set-ups

Originally, the aberroscope data was taken from subjective drawings of the apparent grid shape. With the advent of sensitive video cameras, it is quicker, more convenient and more accurate to photograph the retinal grid. The grid points are then taken from the video image. These two processes are optically different and the video image is not that seen by the subject. The two processes are shown in Figure 3. The diagram shows a square ABCD representing the grid in the pupil. If we ignore aberrations, this square is imaged as a square. For the positive power cylindrical lens orientation as shown (i.e. at \( +45^\circ \)), rays through the points A, B, C and D meet the retina at the points as shown.

In the subjective method, the observed orientation of these four points would not be as shown on the retina in Figure 3. We can determine the difference as follows. Imagine that we observe the grid image from behind the retina. The observer will see the grid as shown in the top right-hand side of the diagram. Now, since the brain 'erects' retinal images, it in fact rotates the retinal image through \( 180^\circ \). Therefore, the final subjective view of the retinal image is as shown at the bottom right of the diagram.

If we now use a video camera to view the image, we probably will use the geometry as shown in Figure 3, i.e., view the grid directly and use a beam splitter at \( 45^\circ \) to photograph the retina. Since the optics of the eye give an erect image of the retina and video camera images are also erect, the video image would have the same orientation as the retinal image, but for the presence of the \( 45^\circ \) mirror. This mirror flips the image about a vertical axis and so the final video image has the orientation as shown in the lower middle of the diagram.

If we now compare the two final images, we note that the video image is the same as the subjective view, but rotated through \( 180^\circ \). Therefore if we use the original Howland program with the video image, the grid pattern must be rotated through \( 180^\circ \). This rotation can be done by either:
(i) rotating the grid after it has been digitized; or (ii) by simply rotating the camera through 180°.

**Definition of symbols**

- $X, Y$: ray heights at a lens or surface, with the $Z$-axis being the optical axis
- $X', Y'$: ray heights at next lens or surface
- $W(X, Y)$: wave aberration function
- $\delta \xi', \delta \eta'$: transverse aberrations in $X$ and $Y$ axes direction, respectively.
- $X_p, Y_p$: corresponding coordinates at the principal planes of the eye (very close to the coordinates at the pupils)
- $X_r, Y_r$: the coordinates in the back focal plane (the retina in the case of the eye)
- $F_c$: power of a cylinder or the positive component of the crossed-cylinder
- $\psi$: axis of the cylinder or of the above positive cylinder
- $K$: $2 \sin(\psi) \cos(\psi)$
- $F$: power of a surface or equivalent power of a lens
- $F_e$: power of the eye
- $h$: vertex distance of the crossed-cylinder to the entrance pupil
- $n, n'$: refractive indices in the incident and refraction media, respectively
- $n''$: refractive index in image space
- $u''$: angle of the paraxial ray in image space
- $v, v'$: paraxial angles the ray subtends to the axis before and after refraction, respectively, in the $Y-Z$ section
- $u, u'$: paraxial angles the ray subtends to the axis before and after refraction, respectively, in the $X-Z$ section

For angles, the standard trigonometric sign convention applies.

- $D$: grid dimension (of actual grid)
- $\phi$: corresponding grid dimension in the retinal image
- $\Phi$: corresponding distance on magnified or recorded grid image
- $t$: any distance on actual retinal grid image
- $T$: corresponding distance on magnified or recorded grid image
- $S$: scaling factor