Using matrices to understand geometrical optics

José B. Almeida

Universidade do Minho, Departamento de Física

4710-057 Braga, Portugal

ABSTRACT

The optimization of an optical system benefits greatly from a study of its aberrations and an identification of each of its elements' contribution to the overall aberration figures. The matrix formalism developed by one of the authors was the object of a previous paper and allows the expression of image-space coordinates as high-order polynomials of object-space coordinates. In this paper we approach the question of aberrations, both through the evaluation of the wavefront evolution along the system and its departure from the ideal spherical shape and the use of ray density plots. Using seventh-order matrix modeling, we can calculate the optical path between any two points of a ray as it travels along the optical system and we define the wavefront as the locus of the points with any given optical path; the results are presented on the form of traces of the wavefront on the tangential plane, although the formalism would also permit sagital plane plots. Ray density plots are obtained by actual derivation of the seventh-order polynomials.

1. INTRODUCTION

In previous papers[1,2] it was shown that it is possible to determine coefficients for matrix modeling of optical systems up to any desired order, computing power being the only limiting factor. Ref.[2] lists the calculated seventh-order coefficients for systems comprising only spherical surfaces.

The optical path length (henceforth designated *opl*) of any ray is the sum of the path length multiplied by the medium refractive index, for all the media that compose the optical system. The matrix modeling of the optical system is based on translations between reference planes and orientation changes at the surfaces separating two different media. In the following paragraphs we will show that it is possible to evaluate the optical path for all the translations incurred by any ray and add them up to get an overall *opl* between any two points on any ray path.

If a known wavefront is used as origin for the evaluation of all *opls*, then all subsequent wavefronts are loci of points equidistant from the first wavefront in *opl* terms. It is then a question of preference the choice of method to display the wavefront shape. The traces on the tangential and sagital planes lead to simplified calculations and we will show examples of the former. Every departure from a spherical wavefront is a manifestation of aberrations; the common choice for reference sphere is one that is centered on the paraxial image point and contains the center of the exit pupil.[3]

Ray-density plots are also useful diagnosis tools because they are similar to the actual images that the system will produce. We will use the analytical expressions of image-space coordinates to produce those plots.

2. OPTICAL SYSTEM MODEL

If complex coordinates are used, an axis symmetric optical system is modeled in the seventh-order by a product of 40×40 square matrices with real elements, each describing a particular ray transformation. The elementary transformations can be classified in four different classes:

- Translation: A straight ray path.

- Surface refraction: Change in ray orientation governed by Snell's law.

- Forward offset: Ray path between the surface vertex plane and the surface.

- Reverse offset: Ray path from the surface back to the vertex plane, along the refracted ray direction.

The ray itself is described by a 40-element vector comprising the monomials of the complex position and orientation coordinates that have non-zero coefficients. The product of all the elementary transformation matrices yields the system matrix which must be right-multiplied by the incident ray vector to result in the exit ray vector.

The construction of elementary transformation matrices is facilitated by the method described previously.[2,4] It can then be assumed that for any system comprising only spherical surfaces all the necessary coefficients are known and the system is perfectly described up to the seventh-order. All the equations presented in the following paragraphs, relating complex ray coordinates in the form of polynomials, were evaluated by matrix multiplication using the software Mathematica. The size of the matrices and the complexity of the expressions imposes some care on the choice of elements to display; we will usually show just the matrix element or the expression relevant for the explanation under way.

In an aberration free optical system the wavefronts should have a spherical shape throughout, or could eventually be flat in a limiting case.[3,5,6] The departure from a spherical wavefront shape is the manifestation of aberrations. In well designed systems a wavefront may have become become aspherical to be partially corrected further along the system. The study of the distortions introduced on the wavefront by each of the elementary ray transformations can greatly elucidate about the performance of a particular system and provide clues for an optimization procedure. Walther[7–9] has performed such optimizations using eikonals and computer algebra; in this paper we use matrix formulation for the determination of wavefront shape at any point along a complex system.

The method consists on evaluating the *opl* of the rays as they are subjected to the successive transformations and adding them up until any desired position along the system is reached; the result is the charateristic function $V(X, S, z, z')$, $X = x + iy$ being the complex position coordinate, $S = s + it$ the complex orientation coordinate and z and z' the positions of reference planes on object and image space, respectively[6,10]; s and t are the direction cosines relative to axes x and y, respectively.

Point objects are defined by a set of fixed coordinates (x, y, z) and so the total *opl* for rays originating on a point object depends only on the ray orientation and image plane position, $V(S, z')$. The locus of points with any given value of the *opl*, expressed by the equation $V(S, z') = constant$, constitutes a wavefront[6] whose shape can be plotted or compared to a reference sphere. Before we start considering each of the elementary transformations in turn we have to establish that in cases where the incident beam is parallel we will evaluate the *opl* from an incident plane wavefront and find the locus of points with constant *opl* difference.

We will start by defining a generalized ray of complex coordinates (X, S); this ray is described by the 40-element monomials vector $\mathbf{X}\&$, built according to the rules explained by Kondo[4] and Almeida.[2] If the ray is subjected to a transformation described by matrix $\mathbf{M}$, then the output ray has coordinates (X', S') and is represented by the monomials vector $\mathbf{X}'\&$, such that:

$$\mathbf{X}'\& = \mathbf{M}\mathbf{X}\&.$$ (1)

In the case of a translation the orientation coordinate does not change and the *opl* for that transformation is obviously given by:

$$l = \frac{nd}{(1 - SS^*)^{1/2}} = \frac{nd}{(1 - S'S'^*)^{1/2}} \,,$$ (2)

with n being the refractive index of the optical medium, d the distance traveled along the optical axis and the asterisk is used to represent conjugate. The product of one complex number by its conjugate is obviously one means of finding the square of its modulus.

A surface refraction introduces an orientation change but no path length is involved and so it offers no contribution to the total *opl*. One optical surface contributes to the *opl* through both the forward and reverse offsets, which are not conceptually different from the translation; both are translations between the vertex plane and the surface, respectively in the forward and the reverse directions, as represented in Fig. (1). It is legitimate to use Eq. (2) to evaluate the path length contributions of these transformations, as long as d is not given a fixed value but is evaluated for each incidence position; note, though, that there is a refractive index change from the forward to the reverse offset, besides the change in the ray orientation. In the following section we will detail this procedure.

Plane waves with oblique orientation must be dealt with separately. As the ray coordinates are referenced to planes normal to the optical axis and there are phase differences between the plane wave rays that intercept the reference plane at various points, those phase differences must be accounted for by an *opl* given by:
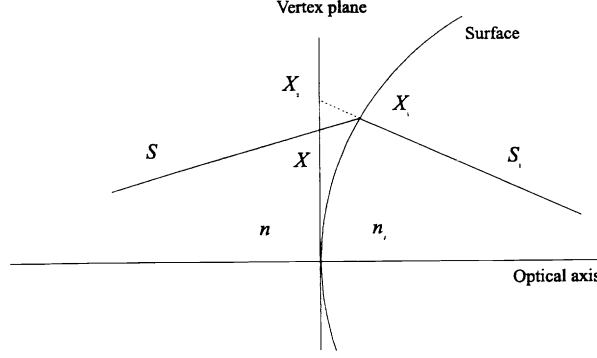
$$l_0 = n(XX^*SS^*)^{1/2} = n|X||S| \,.$$ (3)

Figure 1. The ray intersects the surface at a point X_1 which is different both from the point of intersection of the incident ray with the plane of the vertex, X, and the point of intersection of the refracted ray with the same plane, X_2. The surface is responsible for three successive transformations: 1 – an offset from X to X_1, 2 – the refraction and 3 – the offset from X_1 to X_2.

There is an implied assumption that the *opl* is zero for the ray that crosses the reference plane on the optical axis.

3. SINGLE REFRACTIVE SURFACE

We first consider the case of a single surface with parallel incidence. According to the previous argument, the first *opl* that has to be considered is l_0 given by Eq. (3), which accounts for the phase differences of the incident beam when it crosses the surface vertex plane; this will obviously vanish if the rays are parallel to the optical axis, which can always be verified by a single surface, if the axis is chosen appropriately.

For the position coordinate of the ray after the forward offset we refer to Fig. (1) and use the coefficients given by Almeida[2]:

$$
\begin{aligned}
X_1 &= X + \frac{S\left(8 + 4SS^* + 3S^2S^{*2}\right)XX^*}{16r} + \frac{S\left(2 + SS^*\right)X^2X^{*2}}{16r^3} + \frac{SX^3X^{*3}}{16r^5} \\
&= X + \frac{S}{16}\left[\left(8 + 4|S|^2 + 3|S|^4\right)\frac{|X|^2}{r} + \left(2 + |S|^2\right)\frac{|X|^4}{r^3} + \frac{|X|^6}{r^5}\right],
\end{aligned}
\tag{4}
$$

where r represents the surface curvature radius.

In order to use Eq. (2) we must first find d in terms of the incidence point X_1; this is done by the following equation:

$$
d_1 = r - \left(r^2 - X_1X_1^*\right)^{1/2} = r - \left(r^2 - |X_1|^2\right)^{1/2}.
\tag{5}
$$

Now we can substitute Eq. (5) in Eq. (2) to obtain the forward offset path length l_1.

After refraction the ray's orientation coordinate is changed according to Snell's law; in the seventh-order approximation the new coordinate is given by:

$$
\begin{aligned}
S_1 &= \nu S + \frac{(-1+\nu)X}{r} + \frac{\left(-\nu+\nu^2\right)XSS^*}{2r} + \frac{\left(-\nu+\nu^2\right)X^2S^*}{2r^2} \\
&\quad + \frac{\left(-\nu+\nu^4\right)XS^2S^{*2}}{8r} + \frac{\left(-\nu^2+\nu^4\right)X^2SS^{*2}}{4r^2} + \frac{\left(-\nu^2+\nu^4\right)X^3S^{*2}}{8r^3} \\
&\quad + \frac{\nu\left(-1+\nu^5\right)XS^3S^{*3}}{16r} + \frac{\nu^2\left(-1-2\nu^2+3\nu^4\right)X^2S^2S^{*3}}{16r^2} \\
&\quad + \frac{3\nu^4\left(-1+\nu^2\right)X^3SS^{*3}}{16r^3} + \frac{\left(-\nu^4+\nu^6\right)X^4S^{*3}}{16r^4} + \frac{S\left(-\nu+\nu^2\right)XX^*}{2r^2} \\
&\quad + \frac{\left(-\nu+\nu^2\right)X^2X^*}{2r^3} + \frac{\left(-\nu^2+\nu^4\right)XX^*S^2S^*}{4r^2}
\end{aligned}
$$

$$+\frac{\nu\left(1-3\nu+2\nu^3\right)X^2X^*SS^*}{4r^3}+\frac{\left(-\nu^2+\nu^4\right)X^3X^*S^*}{4r^4}$$

$$+\frac{\nu^2\left(-1-2\nu^2+3\nu^4\right)XX^*S^3S^{*2}}{16r^2}+\frac{\nu\left(1-10\nu^3+9\nu^5\right)X^2X^*S^2S^{*2}}{16r^3}$$

$$+\frac{\nu^2\left(2-11\nu^2+9\nu^4\right)X^3X^*SS^{*2}}{16r^4}+\frac{3\nu^4\left(-1+\nu^2\right)X^4X^*S^{*2}}{16r^5}$$

$$+\frac{\left(-\nu^2+\nu^4\right)XX^{*2}S^2}{8r^3}+\frac{\left(-\nu^2+\nu^4\right)X^2X^{*2}S}{4r^4}+\frac{\left(-\nu+\nu^4\right)X^3X^{*2}}{8r^5}$$

$$+\frac{3\nu^4\left(-1+\nu^2\right)XX^{*2}S^3S^*}{16r^3}+\frac{\nu^2\left(2-11\nu^2+9\nu^4\right)X^2X^{*2}S^2S^*}{16r^4}$$

$$+\frac{\nu\left(1-10\nu^3+9\nu^5\right)X^3X^{*2}SS^*}{16r^5}+\frac{\nu^2\left(-1-2\nu^2+3\nu^4\right)X^4X^{*2}S^*}{16r^6}$$

$$+\frac{\nu^4\left(-1+\nu^2\right)XX^{*3}S^3}{16r^4}+\frac{3\nu^4\left(-1+\nu^2\right)X^2X^{*3}S^2}{16r^5}$$

$$+\frac{\nu^2\left(-1-2\nu^2+3\nu^4\right)X^3X^{*3}S}{16r^6}+\frac{\left(-\nu+\nu^6\right)X^4X^{*3}}{16r^7}\ , \tag{6}$$

where ν represents the refractive index ratio from the first medium to the second.

The ray could now be traced back to the vertex plane along the S_1 direction and the respective opl calculated. We prefer not to do this but rather to use an equivalent procedure which consists on evaluating the next translation opl from the point X_1 on the surface instead of from the vertex plane. The rays will now follow a straight path to the image plane at some distance z' from the surface; we calculate the corresponding opl (l_2) by means of Eq. (2) with d replaced by $d_2 = z' - d_1$ and S replaced by S_1 taken from Eq. (6).

We are now able to evaluate the total path length, l_t, in any position in the second medium, just by adding the three contributions $l_t = l_0 + l_1 + l_2$.

4. WAVEFRONT PLOTS

In the previous paragraph we established the method to evaluate the path length for any ray as it intercepts any given reference plane along the optical axis. In fact we defined a function of $l_t(X, z')$ which is no other than the characteristic function linking points on a wavefront in object space to points on a reference plane in image space. In order to define the wavefront surface we must specify a reference value for the characteristic function and find the locus of the points where that reference value holds; for convenience we take the value for the ray that intercepts the reference plane on the optical axis and call this l_r.

Fig. 2 represents a ray crossing a reference plane normal to the optical axis; the plane of the figure is not necessarily a meridional plane but it is rather the plane defined by the ray and the normal to the reference plane on the point of intersection. The ray coordinates on the point of intersection are (X', S') and the medium refractive index is n'; the optical path difference is given by the difference $\Delta = l_r - l_t$. If we were to follow along the ray the distance Δ/n' we would find a point with the same opl as the reference; this point is necessarily on the same wavefront as the reference point.

From the figure we see that the projection of the distance Δ/n' on the reference plane is given by:

$$\rho = \frac{\Delta}{n'}\cos\alpha\ . \tag{7}$$

The factor $\cos\alpha$ can be decomposed on the direction cosines relative to axes x and y, leading to two components ρ_x and ρ_y, which must be added to the position coordinates of the intersection point in order to obtain the coordinates of the wavefront point; in complex notation it is:
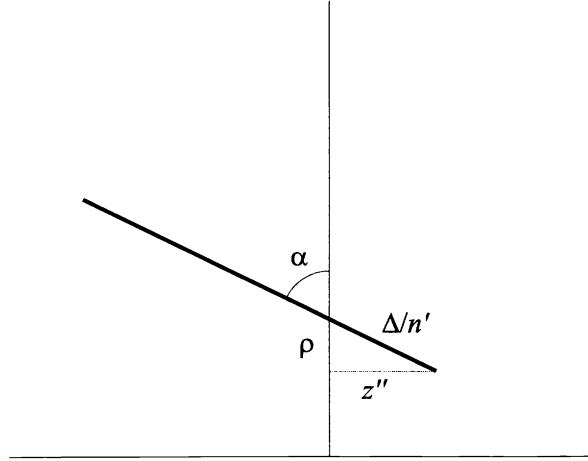
$$X'' = X' + \frac{\Delta}{n'}S'\ . \tag{8}$$

Figure 2. The figure represents a ray crossing a reference plane normal to the optical axis; the plane of the figure is not necessarily a meridional plane but it is rather the plane containing the ray, which is normal to the reference plane.

The position of the wavefront point relative to the reference plane is given by z'', according to the equation:

$$z'' = \frac{\Delta}{n'} \sin \alpha \; ; \tag{9}$$

again in complex notation this can be rewritten:

$$z'' = \frac{\Delta}{n'} \left(1 - S'S'^{*}\right)^{1/2} = \frac{\Delta}{n'} \left(1 - |S'|^2\right)^{1/2} \; . \tag{10}$$

The two equations (8) and (10) define a surface whose points have all the same optical path and so, by definition, they are the wavefront equations.

5. NUMERICAL EXAMPLE

For this example we chose a convex spherical surface of 1 m radius, which marks the boundary between air and a 1.5 refractive index optical medium, upon which impinges a bundle of parallel rays; the optical axis is chosen to be the line containing the center of curvature which is parallel to the impinging rays. This simple optical system has a paraxial focal distance of 3 m and the paraxial focus is the center of all the aberration–free wavefronts considered after refraction.

We want to depict the wavefront shape through its trace on the meridional plane; this allows an important simplification, as the rays' position coordinate has null imaginary component and is thus represented by the real component x; furthermore, the orientation coordinate is zero because all the impinging rays are parallel to the optical axis. As a result we have $x_1 = x$ and from Eq. (5):

$$l_1 = d_1 = r - (r^2 - x_1^2)^{1/2} \; . \tag{11}$$

The orientation coordinate after refraction, S_1, is real for all rays on the meridional plane and so it is represented in lower case: $S_1 = s_1$. This was evaluated by matrix multiplication but we could just as well have used Eq. (6) with suitable substitutions. We applied Eq. (2) to evaluate the optical path contribution of the translation from the surface vertex plane to a reference plane located 2.8 m after the surface; the refractive index was set to $n = 1.5$ and the distance was set to $d_2 = 2.8 - d_1$. Eqs. (8) and (10), with real position coordinates, were used to evaluate the curve of the wavefront trace which was then plotted as shown in Fig. 3 superimposed on the traces of meridional rays; these are naturally normal to the wavefront in every point. We notice that the ends of the wavefront are folded and show a convex curvature, indicating spherical aberration. The points on the curve with zero curvature radius are points on a caustic arising from the crossing of rays with different directions.
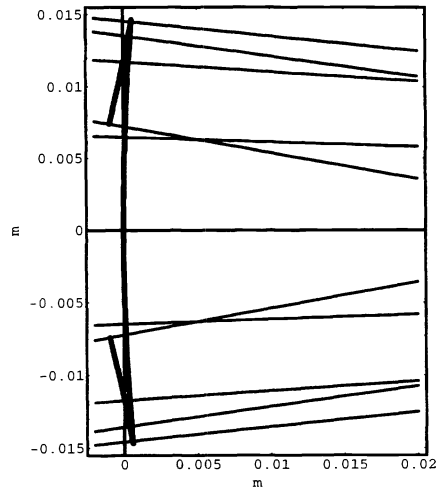
Figure 3. Meridional wavefront trace for a single refracting surface, superimposed on the traces of meridional rays. Notice that the ends of the wavefront are folded and show a convex curvature, indicating spherical aberration.

6. SINGLE LENS

We turn our attention now to a thin lens with oblique incidence. The lens is convex on the first surface and flat on the second surface, the convex surface has a curvature radius of 31.123 mm and the center thickness is 5.8 mm; the glass is BK7, defined as having a refractive index of 1.5168. This lens has a nominal focal distance of 60 mm. The rays incident upon the lens form a parallel bundle with a direction cosine $s = 0.1$.

The only added complication to the situation of the single surface results from the consideration of the second surface, which marks the transition from glass to air with no associated curvature. The optical axis is now clearly identified by the line normal to the flat surface and containing the first surface's center of curvature and cannot be aligned with the direction of incidence. Oblique incidence promotes the emergence of the various aberration terms but does not imply any new equations.

The wavefront is studied at a distance of 52 mm and its meridional trace is plotted on Fig. 4. Again we notice that the rays are normal to the wavefront and that the ends of this are folded backwards. The sharper bend of the upper end is an indication of coma. The other aberration terms are not clearly noticeable on the figure because in the case of astigmatism we would have to compare with the sagital plot and in the cases of field curvature and distortion the wavefront is still spherical but its center is shifted from the paraxial position.

7. RAY-DENSITY PLOTS

In order to study the ray-density plots we use the lens of the previous example wit the image plane moved to a position just past the tangential focus, i.e. 56 mm, because this is a natural position and also because all the rays are divergent from this position onwards. This avoids the complication of having to deal with overlapping wavefront folds, each of them contributing independently to the overall ray-density.

The overall system matrix for the lens above followed by a straight path to the image plane was evaluated with the help of Mathematica and then right-multiplied by the input ray vector $\mathbf{X}$& composed with the variable complex position coordinate, X, and fixed orientation coordinate, $s = 0.1$. The result is a 40–element vector, the first of which is a 7th–order polynomial on X, representing the dependence of the point of intersection on the image plane on the input variable; we call this X'.

If it is established that the input beam has a uniform ray-density, then the image-plane ray-density is given by:

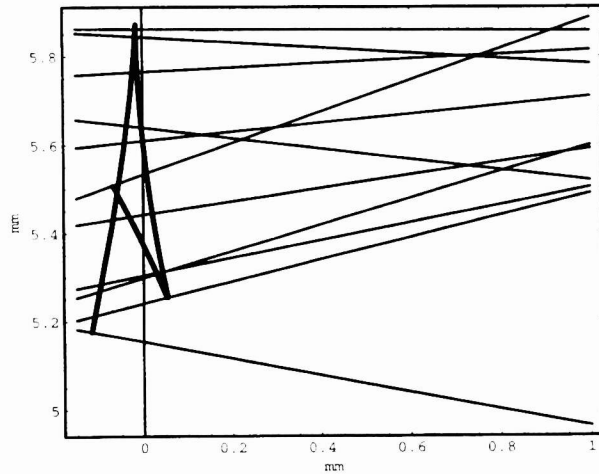$$d = 1 / \left| \frac{\mathrm{d}X'}{\mathrm{d}X} \right| . \tag{12}$$

333

Figure 4. Meridional wavefront trace for a single lens, superimposed on the traces of meridional rays. The sharper bend of the upper end is an indication of coma.

Now, X being a complex coordinate we can express it in the exponential form as $X = \chi e^{i\theta}$ and plug this into Eq. (12) to get:

$$d = 1 / \left| \frac{\mathrm{d}X'}{\mathrm{d}\chi} - \frac{i}{\chi} \frac{\mathrm{d}X'}{\mathrm{d}\theta} \right| . \tag{13}$$

The value of the ray-density given by Eq. (13) was evaluated and plotted as shades of gray on a logaritmic scale on the positions corresponding to the image coordinate X', as shown on Fig. 5. The image is just as one would expect from a lens focusing an oblique beam of light.
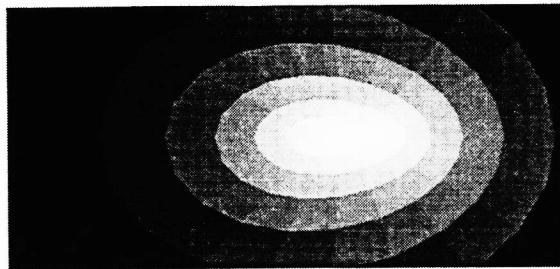

Figure 5. Ray density plot on a plane near the tangential focus.

8. CONCLUSION

Previous results had shown that optical systems could be modeled with matrices up to any desired order of approximation and the necessary coefficients for axis-symmetrical systems built with spherical surfaces had already bean reported. Those results have now been used to evaluate aberrations in non-standard ways.

An implementation of the seventh-order matrix algorithm in Mathematica allows the construction of algebraic models for very complex systems, which can be used in various ways to judge their performance and quality.

The possibility of plotting wavefront shapes at any point along a complex optical system was demonstrated with two simple examples but the same procedure could be used in more complex situations. Ray-density plots were also demonstrated, these providing a visualization of the actual image of point objects. It is expected that ray-density plots can be integrated for extended objects, thus yielding the expected aberrated images given by real optical systems.

REFERENCES

1. J. B. Almeida, "The use of matrices for third order modeling of optical systems," in *International Optical Design Conference*, K. P. Thompson and L. R. Gardner, eds., vol. 3482 of *Proc. SPIE*, pp. 917–925, 1998.

2. J. B. Almeida, "General method for the determination of matrix coefficients for high order optical system modeling," *J. Opt. Soc. Am. A* **16**, pp. 596–601, 1999.

3. W. T. Welford, *Aberrations of Optical Systems*, Adam Hilger, Bristol, 1991.

4. M. Kondo and Y. Takeuchi, "Matrix method for nonlinear transformation and its application to an optical lens system," *J. Opt. Soc. Am. A* **13**, pp. 71–89, 1996.

5. G. G. Slyusarev, *Aberration and Optical Design Theory*, Adam Hilger Ltd., Bristol, 1984.

6. M. Born and E. Wolf, *Principles of Optics*, Cambridge University Press, Cambridge, 1997.

7. A. Walther, "Eikonal theory and computer algebra," *J. Opt. Soc. Am. A* **13**, pp. 523–531, 1996.

8. A. Walther, "Eikonal theory and computer algebra II," *J. Opt. Soc. Am. A* **13**, pp. 1763–1765, 1996.

9. A. Walther, "Zoom lenses and computer algebra," *J. Opt. Soc. Am. A* **16**, pp. 198–204, January 1999.

10. D. Goodman, *Geometrical Optics*, ch. 1. Vol. 1 of Bass,[11] CD–ROM ed., 1995.

11. M. Bass, ed., *Handbook of Optics*, Mc. Graw-Hill, CD–ROM ed., 1995.