STUDIES IN OPTICS

I. General Coordinates for Optical Systems with Central or Axial Symmetry*

BY

M. HERZBERGER

Communication No. 960 from the Kodak Research Laboratories

In previous papers[1,2] the author has proposed an approach to geometrical optics different from that developed by Hamilton and his successors.

The purpose of the present paper is to generalize the formulas in these papers, and to find the most general treatment of systems with central (point) symmetry and with axial symmetry. By leaving the coordinates general, subject only to the symmetry conditions of the problem, we retain the symmetry in the formulas up to the point where we desire to draw conclusions for a special problem. We can then introduce special coordinates adapted to the problem in question, and find the particular answers.

The fundamental invariants of geometrical optics show no preference for either object or image side, nor for point or angle coordinates as variables. The different approaches suggested by Hamilton, as well as the direct approach just mentioned, are special cases of the methods developed here corresponding to special choices of coordinates. Several different choices of coordinates will be given as examples.

The fundamental formulas (A, B, B′, C below) are based only on symmetry conditions and on the validity of the Lagrange invariant (A). They are therefore not restricted to optical problems,[3] but are also valid for problems in mechanics, hydrodynamics, and electron optics.

1. Ray tracing formulas, the Lagrangian invariant. Let us assume a ray traversing a number of optical media with refractive indexes $n, n_{12}, n_{23}, \cdots, n'$. Let $\mathbf{a}(x, y, z)$, $\mathbf{a}'(x', y', z')$ be a vector from an arbitrary origin to a point on the object and image rays, respectively. Let $\mathbf{a}_k(x_k, y_k, z_k)$ be the vector from the same origin to the intersection point of the ray with the kth surface. Let $\mathbf{s}_{k,k+1}(\xi_{k,k+1}, \eta_{k,k+1}, \zeta_{k,k+1})$ be a vector along the ray in the medium between kth and $(k+1)$th surface, a vector of length equal to the refractive index $n_{k,k+1}$ of the medium.

Let $\mathbf{o}_k$ be a vector perpendicular to the kth surface at the intersection point. Its length may remain arbitrary, for the moment. The refraction law then reads:

$$\mathbf{s}_{k,k+1} \times \mathbf{o}_k = \mathbf{s}_{k-1,k} \times \mathbf{o}_k, \tag{1}$$

where the multiplication sign indicates vector multiplication. Equation (1) shows that $\mathbf{s}_{k,k+1} - \mathbf{s}_{k-1,k}$ has the direction of the surface normal $\mathbf{o}_k$, or

* Received Jan. 24, 1944.

[1] M. Herzberger, *Direct methods in geometrical optics*, Trans. Am. Math. Soc., **53**, 218–229 (1943).

[2] M. Herzberger, *A direct image error theory*, Quarterly of Applied Mathematics, **1**, 69–77 (1943).

[3] For the connection of the Lagrangian invariant with different branches of mathematics and physics, see M. Herzberger, *Theory of transversal curves and the connections between the calculus of variations and the theory of partial differential equations*, Proc. Nat. Acad. Sci. U.S.A., **24**, 466–473 (1938).

$$\mathbf{s}_{k,k+1} - \mathbf{s}_{k-1,k} = \phi_k \mathbf{o}_k. \tag{2}$$

We now can describe the path of the ray through the system by means of the vector equations

$$\begin{gathered} \mathbf{a}_1 = \mathbf{a} + \lambda \mathbf{s}, \qquad \mathbf{s}_{12} = \mathbf{s} + \phi_1 \mathbf{o}_1, \qquad \mathbf{a}_2 = \mathbf{a}_1 + \lambda_{12}\mathbf{s}_{12}, \\ \cdots\cdots\cdots\cdots, \\ \mathbf{s}' = \mathbf{s}_{\nu-1,\nu} + \phi_\nu \mathbf{o}_\nu, \qquad \mathbf{a}' = \mathbf{a}_\nu + \lambda'\mathbf{s}'. \end{gathered} \tag{3}$$

The geometrical significance of λ and ϕ can be seen by multiplying (3) scalarly by $\mathbf{s}_{k,k+1}$ and $\mathbf{o}_k$, respectively, keeping in mind that, by definition,

$$\mathbf{s}_{k,k+1}^2 = n_{k,k+1}^2. \tag{3a}$$

We then find that

$$\lambda_{k-1,k} = (\mathbf{a}_k - \mathbf{a}_{k-1}) \cdot \mathbf{s}_{k-1,k}/n_{k-1,k}, \tag{3b}$$

$$\phi_k = (\mathbf{s}_{k,k+1} - \mathbf{s}_{k-1,k}) \cdot \mathbf{o}_k/\mathbf{o}_k^2, \tag{3c}$$

i.e., the λ's are proportional to the distance between the two surfaces along the ray, and the ϕ's are proportional to what might be called the *power* of the surface for the individual ray.

Since Eqs. (3) are valid for every ray, we now consider a two-dimensional manifold of rays, i.e., we assume the **a**'s, **o**'s, and **s**'s to be vector functions of two variables t_1 and t_2. From the definition of $\mathbf{s}_{k,k+1}$ and $\mathbf{o}_k$, we find

$$\mathbf{s}_{k,k+1} \cdot (\partial \mathbf{s}_{k,k+1}/\partial t_\mu) = 0, \qquad \mathbf{o}_k \cdot (\partial \mathbf{a}_k/\partial t_\mu) = 0, \qquad (\mu = 1, 2). \tag{4}$$

We now differentiate (3) with respect to t_1 and multiply scalarly by $\partial \mathbf{s}_{k,k+1}/\partial t_2$ and $\partial \mathbf{a}_k/\partial t_2$, respectively. Then we differentiate with respect to t_2 and multiply scalarly by $\partial s_{k,k+1}/\partial t_1$ and $\partial \mathbf{a}_k/\partial t_1$, respectively. Subtraction of the two sets of equations yields the "Lagrangian invariant":

$$\begin{vmatrix} \dfrac{\partial \mathbf{a}}{\partial t_1} & \dfrac{\partial \mathbf{s}}{\partial t_1} \\ \dfrac{\partial \mathbf{a}}{\partial t_2} & \dfrac{\partial \mathbf{s}}{\partial t_2} \end{vmatrix} = \begin{vmatrix} \dfrac{\partial \mathbf{a}_1}{\partial t_1} & \dfrac{\partial \mathbf{s}_{1,2}}{\partial t_1} \\ \dfrac{\partial \mathbf{a}_1}{\partial t_2} & \dfrac{\partial \mathbf{s}_{1,2}}{\partial t_2} \end{vmatrix} = \cdots = \begin{vmatrix} \dfrac{\partial \mathbf{a}'}{\partial t_1} & \dfrac{\partial \mathbf{s}'}{\partial t_1} \\ \dfrac{\partial \mathbf{a}'}{\partial t_2} & \dfrac{\partial \mathbf{s}'}{\partial t_2} \end{vmatrix}. \tag{A}$$

This formula was introduced by Lagrange in his astronomical investigations. It is known by the name of the Lagrangian bracket in the theory of partial differential equations. Herzberger[3] used it in his theory of transversal curves, as the starting point. Let us now see what conclusions can be drawn if the system in question fulfills certain conditions of symmetry.

2. Centrally symmetric systems. In this case all refracting surfaces are concentric spheres with radii $r_1, \cdots, r_n$. It is therefore appropriate to consider the common center as the coordinate origin, and to choose concentric spheres as the object surface **a** and the image surface **a′**. All the surface normals pass through the common center. We shall give them the length r_k from center to surface, with a positive sign if the surface is convex towards the incident light. In other words, we make $\mathbf{o}_k = \mathbf{a}_k$. Under these conditions, Eqs. (3) become

$$\mathbf{a}_1 = \mathbf{a} + \lambda \mathbf{s}, \qquad \mathbf{s}_{12} = \mathbf{s} + \phi_1 \mathbf{a}_1,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots, \tag{5}$$
$$\mathbf{s}' = \mathbf{s}_{n-1,n} + \phi_n \mathbf{a}_n, \qquad \mathbf{a}' = \mathbf{a}_n + \lambda' \mathbf{s}'.$$

From Eqs. (3) we can find an invariant vector, namely,

$$\mathbf{a} \times \mathbf{s} = \mathbf{a}_1 \times \mathbf{s} = \mathbf{a}_1 \times \mathbf{s}_{12} = \cdots = \mathbf{a}' \times \mathbf{s}' = \mathbf{p}. \tag{6}$$

Therefore, in a concentric system, both object and image rays lie in a plane through the center, and the optical length p of the perpendicular from the center (the length of the invariant vector) remains constant.

Equations (4) can be combined into

$$\mathbf{a}' = a\mathbf{a} + b\mathbf{s}, \qquad \mathbf{s}' = c\mathbf{a} + d\mathbf{s}, \quad \text{where} \quad ad - bc = 1. \tag{B}$$

The invariant relation, $ad - bc = 1$, is found by substituting (5) in (6). It is possible to calculate a, b, c, and d as functions of the λ's and ϕ's, if we use "Gaussian brackets."[4] We find

$$\begin{aligned} a &= [\phi_1, \lambda_{12}, \cdots, \lambda'], & b &= [\lambda, \phi_1, \cdots, \lambda'], \\ c &= [\phi_1, \lambda_{12}, \cdots, \phi_n], & d &= [\lambda, \phi_1, \cdots, \phi_n]. \end{aligned} \tag{7}$$

In the case of central symmetry, the ϕ's and λ's, and therefore a, b, c, and d, can be considered as functions of a single variable, and p can be taken as this variable, as shown in (9) and (10). Now

$$(\mathbf{a}_k \cdot \mathbf{s}_{k,k+1})^2 + (\mathbf{a}_k \times \mathbf{s}_{k,k+1})^2 = \mathbf{a}_k^2 \mathbf{s}_{k,k+1}^2, \tag{8}$$

or

$$\mathbf{a}_k \cdot \mathbf{s}_{k,k+1} = \sqrt{r_k^2 n_{k,k+1}^2 - p^2}. \tag{8a}$$

Thus we find from (3) that

$$\lambda_{k,k+1} = \frac{1}{n_{k,k+1}^2}\left[\sqrt{r_{k+1}^2 n_{k,k+1}^2 - p^2} - \sqrt{r_k^2 n_{k,k+1}^2 - p^2}\,\right]$$
$$= \frac{1}{n_{k,k+1}}\left[r_{k+1}\sqrt{1 - \left(\frac{p}{r_{k+1} n_{k,k+1}}\right)^2} - r_k\sqrt{1 - \left(\frac{p}{r_k n_{k,k+1}}\right)^2}\,\right], \tag{9}$$

and that

$$\phi_k = \frac{1}{r_k^2}\left[\sqrt{r_k^2 n_{k,k+1}^2 - p^2} - \sqrt{r_k^2 n_{k-1,k}^2 - p^2}\right]. \tag{10}$$

Equations (B) correspond to the direct equations of our theory. A more general representation is given by choosing two arbitrary vector functions, $\mathbf{l}$ and $\mathbf{m}$, in terms of which the object and image vectors can be expressed. Equations (B) may then be written,

$$\mathbf{a} = a\mathbf{l} + b\mathbf{m}, \qquad \mathbf{a}' = a'\mathbf{l} + b'\mathbf{m},$$
$$\mathbf{s} = c\mathbf{l} + d\mathbf{m}, \qquad \mathbf{s}' = c'\mathbf{l} + d'\mathbf{m}; \tag{B'}$$
$$ad - bc = a'd' - b'c'.$$

---

[4] Invented by L. Euler in 1776. See M. Herzberger, *Gaussian optics and Gaussian brackets*, J. Opt. Soc. Amer. **33**, 651–655 (1943).

a, b, c, d, and a', b', c', d' are still functions of a single variable, which can be taken as

$$\pi = (\mathbf{l} \times \mathbf{m})^2 = \mathbf{l}^2\mathbf{m}^2 - (\mathbf{l}\cdot\mathbf{m})^2. \tag{11}$$

It is obvious that under these conditions the image formation described by (11) fulfills all the conditions mentioned, and that the last condition in (B) is equivalent to the validity of the invariant (6).

Formulas (B) are a special case of Eqs. (B′) if we choose $\mathbf{l}=\mathbf{a}$, $\mathbf{m}=\mathbf{s}$, which corresponds to $a=d=1$, $b=c=0$.

The connection between a, b, c, d and a', b', c', d' becomes clearer if we introduce some auxiliary angles. Let us write

$$\begin{gathered} \sphericalangle(\mathbf{a}, \mathbf{m}) = \alpha, \quad \sphericalangle(\mathbf{l}, \mathbf{a}) = \beta, \quad \sphericalangle(\mathbf{s}, \mathbf{m}) = \gamma, \quad \sphericalangle(\mathbf{l}, \mathbf{s}) = \delta, \\ \sphericalangle(\mathbf{l}, \mathbf{m}) = \psi, \quad \sphericalangle(\mathbf{a}, \mathbf{s}) = \sigma, \quad \sphericalangle(\mathbf{a}', \mathbf{s}') = \sigma'. \end{gathered} \tag{12}$$

From (B), if we write r for the absolute value of vector $\mathbf{a}$, we find that

$$\begin{aligned} a &= (rm/\pi)\sin\alpha, & b &= (rl/\pi)\sin\beta, \\ c &= (nm/\pi)\sin\gamma, & d &= (nl/\pi)\sin\delta, \end{aligned} \tag{13}$$

with analogous expressions for the primed quantities. Moreover, we find from (12) that

$$\begin{gathered} \alpha + \beta = \gamma + \delta = \psi = \alpha' + \beta' = \gamma' + \delta', \\ \alpha - \gamma = \delta - \beta = \sigma, \qquad \alpha' - \gamma' = \delta' - \beta' = \sigma', \end{gathered} \tag{14}$$

where σ and σ', according to (9), are connected by

$$p = nr\sin\sigma = n'r'\sin\sigma', \qquad \pi = lm\sin\phi. \tag{15}$$

If we write

$$ad - bc = a'd' - b'c' = D, \tag{16}$$

we have finally

$$p = D\pi.$$

Thus, for a given system of coordinates l, m, and ϕ, and given object and image spheres (radii r and r'), we can calculate all the functions $a, b, c, d, a', b', c', d'$, if only one of them is given on each side. For instance, let us assume γ and γ' to be given. We then find σ and σ' from (15), and obtain

$$\begin{aligned} \alpha &= \sigma + \gamma, & \beta &= \psi - \sigma - \gamma, & \gamma &= \gamma, & \delta &= \psi - \gamma, \\ \alpha' &= \sigma' + \gamma', & \beta' &= \psi - \sigma' - \gamma', & \gamma' &= \gamma', & \delta' &= \psi - \gamma'. \end{aligned} \tag{17}$$

Thus, according to (13), we determine a, b, c, d, a', b', c', and d'.

Let us now consider some special cases.

a) *The direct method.* We choose $\mathbf{l}=\mathbf{a}$, $\mathbf{m}=\mathbf{s}$. This means that

$$\begin{gathered} \psi = \sigma, \qquad \beta = \gamma = 0, \qquad \alpha = \delta = \sigma, \\ a = d = 1, \qquad b = c = 0, \\ \alpha' = \sigma' + \gamma', \qquad \beta' = \sigma - \sigma' - \gamma', \qquad \gamma' = \gamma', \qquad \delta' = \sigma - \gamma', \end{gathered} \tag{18}$$

where $nr\sin\sigma = n'r'\sin\sigma'$.

b) Hamilton's point coordinates. We choose $\mathbf{l}=\mathbf{a}$, $\mathbf{m}=\mathbf{a}'$, or $a=1$, $b=0$, $a'=0$, $b'=1$. This means that

$$\begin{aligned} \alpha &= \psi, & \beta &= 0, & \gamma &= \psi - \sigma, & \delta &= \sigma, \\ \alpha' &= 0, & \beta' &= \psi, & \gamma' &= -\sigma', & \delta' &= \psi + \sigma', \end{aligned} \tag{16'}$$

or, since $\mathbf{l}=\mathbf{r}$, $\mathbf{m}=\mathbf{r}'$,

$$\begin{aligned} a &= 1, & b &= 0, & c &= [nr' \sin(\psi - \sigma)]/\pi, & d &= p/\pi, \\ a' &= 0, & b' &= 1, & c' &= p/\pi, & d' &= [n'r \sin(\psi + \sigma')]/\pi, \end{aligned} \tag{17'}$$

where

$$nr \sin \sigma = n'r' \sin \sigma' = p, \qquad rr' \sin \psi = \pi. \tag{18'}$$

We note especially that $c'-d=0$.

c) Hamilton's angle characteristic. We choose $\mathbf{l}=\mathbf{s}$, $\mathbf{m}=\mathbf{s}'$, or $c=1$, $d=0$, $c'=0$, $d=1$. We find from (B′) that

$$ad - bc = a'd' - b'c' \tag{19}$$

or $-b=a'=D=p/\pi$. For the auxiliary angles, we obtain

$$\begin{aligned} \alpha &= \sigma + \psi, & \beta &= -\sigma, & \gamma &= \psi, & \delta &= \sigma, \\ \alpha' &= \sigma', & \beta' &= \psi - \sigma', & \gamma' &= 0, & \delta' &= \psi, \end{aligned} \tag{20}$$

or

$$a = (n'r/\pi) \sin(\psi + \sigma), \quad b = -a' = -p/\pi, \quad b' = (n^2/\pi) \sin(\psi - \sigma'), \tag{21}$$

where

$$\pi = nn' \sin \psi, \qquad p = nr \sin \sigma = n'r' \sin \sigma'.$$

d) We take as coordinates the intersection points of the ray with two spheres in the object space: the object sphere (radius r), and a second sphere (radius R, and $\mathbf{A}$, the vector to the intersection point on this sphere). We define

$$\mathbf{l} = \mathbf{a}, \qquad \mathbf{m} = \mathbf{A},$$

and

$$\begin{aligned} \mathbf{a} &= \mathbf{a}, & \mathbf{a}' &= a'\mathbf{a} + b'\mathbf{A}, \\ \mathbf{s} &= -c\mathbf{a} + d\mathbf{A}, & \mathbf{s}' &= c'\mathbf{a} + d'\mathbf{A}, \end{aligned} \tag{22}$$

where $a'd'-b'c'=c$ and

$$c = [nR \sin(\psi - \sigma)]/\pi = -[nr \sin \sigma]/\pi = -p/\pi. \tag{23}$$

We find that

$$\sin \sigma' = [nr \sin \sigma]/n'r', \qquad \sin(\psi - \sigma) = [n \sin \sigma]/nR,$$

and, finally that

$$\begin{aligned} a' &= [r'R \sin(\sigma' + \gamma')]/\pi, & c' &= [n'R \sin \gamma']/\pi, \\ b' &= [rr' \sin(\psi - \sigma' - \gamma')]/\pi, & d' &= [n'r \sin(\psi - \gamma')]/\pi. \end{aligned} \tag{24}$$

These are the equations for the image formation.

3. Systems with rotational symmetry. Let us now assume that the system has symmetry only around an axis, the unit vector along which we shall designate by $\mathbf{k}$. In this case, all the surface normals intersect the axis, and we shall give to the vector $\mathbf{o}_k$ in (3) the length r_k, which is the distance along the normal between its intersection points with the axis and the surface.

We now project all the vectors on a plane perpendicular to the axis, and define the projected vectors $\mathbf{b}_k$ and $\mathbf{t}_{k,k+1}$ by the equations

$$\mathbf{a}_k = \mathbf{b}_k + z_k\mathbf{k}, \qquad \mathbf{s}_{k,k+1} = \mathbf{t}_{k,k+1} + \zeta_{k,k+1}\mathbf{k}, \qquad \mathbf{o}_k = \mathbf{b}_k + (z_k - z_{Nk})\mathbf{k}, \tag{25}$$

where z_N is the quantity known in geometry as the subnormal, and

$$\zeta_{k,k+1} = \sqrt{n^2_{k,k+1} - \xi^2_{k,k+1} - \eta^2_{k,k+1}};$$

$\zeta_{k,k+1}$ is the (optical) cosine of the angle between the ray and the axis.

Let us now assume that object and image origins lie on two planes perpendicular to the axis. We can then replace all the vectors in (3) and (A) by their projections in these planes, and find, instead of (A), for a two-dimensional manifold of rays (parameters t_1, t_2),

$$\begin{vmatrix} \dfrac{\partial \mathbf{b}}{\partial t_1} & \dfrac{\partial \mathbf{t}}{\partial t_1} \\ \dfrac{\partial \mathbf{b}}{\partial t_2} & \dfrac{\partial \mathbf{t}}{\partial t_2} \end{vmatrix} = \begin{vmatrix} \dfrac{\partial \mathbf{b}'}{\partial t_1} & \dfrac{\partial \mathbf{t}'}{\partial t_1} \\ \dfrac{\partial \mathbf{b}'}{\partial t_2} & \dfrac{\partial \mathbf{t}'}{\partial t_2} \end{vmatrix}, \tag{26}$$

and instead of (B),

$$\mathbf{b}' = a'\mathbf{b} + b'\mathbf{t}, \qquad \mathbf{t}' = c'\mathbf{b} + d'\mathbf{t}, \tag{27}$$

where $\mathbf{b}'\times\mathbf{t}' = \mathbf{b}\times\mathbf{t}$ and therefore $a'd' - b'c' = 1$. The functions a', b', c', d' are given by formula (7), where ϕ and λ have the same meaning as before.

Moreover, a', b', c', d' are no longer functions of a single variable, but are functions of the three symmetric functions and $\mathbf{b}$ and $\mathbf{t}$, namely,

$$u = \mathbf{b}^2, \qquad v = \mathbf{b}\cdot\mathbf{t}, \qquad w = \mathbf{t}^2. \tag{28}$$

Equation (27) corresponds to the formulas of the direct image error theory. The most general choice of coordinates might be described as follows ($\mathbf{t}$ and $\mathbf{m}$, as well as the other vectors, lie in a plane perpendicular to the axis): let

$$\begin{aligned} \mathbf{b} &= a\mathbf{l} + b\mathbf{m}, & \mathbf{b}' &= a'\mathbf{l} + b'\mathbf{m}, \\ \mathbf{t} &= c\mathbf{l} + d\mathbf{m}, & \mathbf{t}' &= c'\mathbf{l} + d'\mathbf{m}, \end{aligned} \tag{29}$$

where

$$\mathbf{b}\times\mathbf{t} = \mathbf{b}'\times\mathbf{t}',$$

or

$$ad - bc = a'd' - b'c'. \tag{30}$$

Let us assume that a, b, c, and d are functions of the symmetric functions of $\mathbf{l}$ and $\mathbf{m}$; that is,

$$u = \mathbf{l}^2, \qquad v = \mathbf{l}\cdot\mathbf{m}, \qquad w = \mathbf{m}^2. \tag{31}$$

$\mathbf{b}$, $\mathbf{t}$, $\mathbf{b}'$, $\mathbf{t}'$ must fulfill Eq. (26), if we set t_1 and t_2 alternatively equal to u, v, and w. Thus, we find the following equations for a, b, c, and d. Let us write

$$A = au + bv, \qquad C = cu + dv,$$
$$B = av + bw, \qquad D = cv + dw, \tag{32}$$

and introduce the abbreviation Δ for the difference of an expression in the object and image spaces. An easy computation then gives

$$\Delta\left\{\begin{vmatrix} \frac{\partial A}{\partial t_1} & \frac{\partial A}{\partial t_2} \\ \frac{\partial c}{\partial t_1} & \frac{\partial c}{\partial t_2} \end{vmatrix} + \begin{vmatrix} \frac{\partial B}{\partial t_1} & \frac{\partial B}{\partial t_2} \\ \frac{\partial d}{\partial t_1} & \frac{\partial d}{\partial t_2} \end{vmatrix} + \begin{vmatrix} \frac{\partial a}{\partial t_1} & \frac{\partial a}{\partial t_2} \\ \frac{\partial C}{\partial t_1} & \frac{\partial C}{\partial t_2} \end{vmatrix} + \begin{vmatrix} \frac{\partial b}{\partial t_1} & \frac{\partial b}{\partial t_2} \\ \frac{\partial D}{\partial t_1} & \frac{\partial D}{\partial t_2} \end{vmatrix}\right\} = 0, \tag{C}$$
$$\Delta(ad - bc) = 0.$$

These are the necessary and sufficient conditions that (29) describe an axially symmetric image formation. We repeat that Eqs. (29) and (C) demand only the validity of Lagrange's invariant (26), and axial symmetry. Their application is therefore not restricted to optical problems.

Let us now again investigate what forms the fundamental formulas assume for special choices of coordinates.

a) Hamilton's (Bruns') point characteristic. Hamilton (Bruns) chose as variables the coordinates of a point in the object and image spaces. That corresponds to taking $\mathbf{l}=\mathbf{b}$, $\mathbf{m}=\mathbf{b}'$. Equations (29) become

$$\mathbf{b} = \mathbf{b}, \qquad \mathbf{b}' = \mathbf{b}',$$
$$\mathbf{t} = c\mathbf{b} + d\mathbf{b}', \qquad \mathbf{t}' = c'\mathbf{b} + d'\mathbf{b}', \tag{33}$$

or

$$a = b' = 1, \qquad b = a' = 0.$$

These are conditions for the coefficients. Equations (32) now become

$$A = u, \quad C = cu + dv, \quad A' = v, \quad C' = c'u + d'v,$$
$$B = v, \quad D = cv + dw, \quad B' = w, \quad D' = c'v + d'w, \tag{34}$$

and we find instead of (C) that

$$d = -c', \tag{35a}$$

$$\begin{vmatrix} \frac{\partial u}{\partial t_1} & \frac{\partial u}{\partial t_2} \\ \frac{\partial c}{\partial t_1} & \frac{\partial c}{\partial t_2} \end{vmatrix} + \begin{vmatrix} \frac{\partial v}{\partial t_1} & \frac{\partial v}{\partial t_2} \\ \frac{\partial d}{\partial t_1} & \frac{\partial d}{\partial t_2} \end{vmatrix} + \begin{vmatrix} \frac{\partial c'}{\partial t_1} & \frac{\partial c'}{\partial t_2} \\ \frac{\partial v}{\partial t_1} & \frac{\partial v}{\partial t_2} \end{vmatrix} + \begin{vmatrix} \frac{\partial d'}{\partial t_1} & \frac{\partial d'}{\partial t_2} \\ \frac{\partial w}{\partial t_1} & \frac{\partial w}{\partial t_2} \end{vmatrix} = 0. \tag{35b}$$

Equation (35b) stands for three equations, which we can obtain by replacing t_1 and t_2 in (35) by u and v, u and w, v and w, respectively. This yields

$$\frac{\partial c}{\partial v} - \frac{\partial d}{\partial u} = -\frac{\partial c'}{\partial u}, \qquad \frac{\partial c}{\partial w} = -\frac{\partial d'}{\partial u}, \qquad \frac{\partial d}{\partial w} = \frac{\partial c'}{\partial w} - \frac{\partial d'}{\partial v}. \tag{35c}$$

Equations (35a) and (C), when integrated, lead to a function $V(u, v, w)$ such that

$$c = 2\frac{\partial V}{\partial u}, \qquad c' = -\frac{\partial V}{\partial v}, \qquad d = \frac{\partial V}{\partial v}, \qquad d' = -2\frac{\partial V}{\partial w}. \tag{36}$$

V is the characteristic function of Hamilton, the "eiconal" of Bruns. Formulas (36) agree with Hamilton's formulas, except that he used $\mathbf{l}^2/2$ and $\mathbf{m}^2/2$ as variables. Our choice of variables simplifies the form of the general formulas (C).

b) The angle characteristic. Hamilton chose as coordinates the direction cosines of the object and image rays. This means that $\mathbf{l}=\mathbf{t}$, $\mathbf{m}=\mathbf{t}$, or

$$\mathbf{b} = a\mathbf{t} + b\mathbf{t}, \qquad \mathbf{b}' = a'\mathbf{t} + b'\mathbf{t}', \qquad \mathbf{t} = \mathbf{t}, \qquad \mathbf{t}' = \mathbf{t}', \tag{37}$$

or

$$c = 1, \qquad d = 0, \qquad c' = 0, \qquad d' = 1. \tag{38}$$

Equations (29) now give

$$C = u, \qquad C' = v, \qquad D = v, \qquad D' = w, \tag{39}$$

and (C) becomes

$$b + a' = 0,$$
$$\frac{\partial b}{\partial u} - \frac{\partial a}{\partial v} = \frac{\partial a'}{\partial u}, \qquad -\frac{\partial a}{\partial w} = \frac{\partial b'}{\partial u}, \qquad -\frac{\partial b}{\partial w} = -\frac{\partial a'}{\partial w} + \frac{\partial b'}{\partial v}\cdot \tag{40}$$

Equation (4) is solved if we introduce the angle characteristic $T(u, v, w)$, and set

$$a = \frac{1}{2}\frac{\partial T}{\partial u}, \qquad a' = -\frac{\partial T}{\partial v}, \qquad b = \frac{\partial T}{\partial v}, \qquad b' = -\frac{1}{2}\frac{\partial T}{\partial w}\cdot \tag{41}$$

We see that this also agrees with Hamilton's theory.

c) The direct method. In the papers mentioned,[1,2] we took as variables the object point and the direction of the object ray, i.e., we chose $\mathbf{l}=\mathbf{b}$ and $\mathbf{m}=\mathbf{t}$. This gives

$$\mathbf{b} = \mathbf{b}, \qquad \mathbf{b}' = a'\mathbf{b} + b'\mathbf{t}, \qquad \mathbf{t} = \mathbf{t}, \qquad \mathbf{t}' = c'\mathbf{b} + d'\mathbf{t}. \tag{42}$$

That is, we put

$$a = d = 1, \qquad b = c = 0. \tag{43}$$

Equations (27) then give

$$A = u, \qquad C = v, \qquad B = v, \qquad D = w, \tag{44}$$

and Eqs. (C) give

$$a'd' - b'c' = 1,$$
$$\begin{vmatrix} \frac{\partial A'}{\partial t_1} & \frac{\partial A'}{\partial t_2} \\ \frac{\partial c'}{\partial t_1} & \frac{\partial c'}{\partial t_2} \end{vmatrix} + \begin{vmatrix} \frac{\partial B'}{\partial t_1} & \frac{\partial B'}{\partial t_2} \\ \frac{\partial d'}{\partial t_1} & \frac{\partial d'}{\partial t_2} \end{vmatrix} + \begin{vmatrix} \frac{\partial a'}{\partial t_1} & \frac{\partial a'}{\partial t_2} \\ \frac{\partial C'}{\partial t_1} & \frac{\partial C'}{\partial t_2} \end{vmatrix} + \begin{vmatrix} \frac{\partial b'}{\partial t_1} & \frac{\partial b'}{\partial t_2} \\ \frac{\partial D'}{\partial t_1} & \frac{\partial D'}{\partial t_2} \end{vmatrix} = 0. \tag{45}$$

If we denote the sum of the four determinants in (45) by I' when $t_1=u$, $t_2=v$, by II' when $t_1=u$, $t_2=w$, and by III' when $t_1=v$, $t_2=w$, we can write (45) in the form

$$a'd' - b'c' = 1, \qquad I' = II' = III' = 0. \tag{46}$$

Again, if we disregard the fact that the variables u and w differ by a factor of two from those used previously[1,2] we find Eqs. (46) to be identical with those given before.

d) Object and stop coordinates. To analyze image errors, we investigate the manner in which the image ray changes with the position of the object point and the position of the intersection with the diaphragm plane, for which we frequently substitute the entrance pupil of the system. If we choose these as the coordinates of the ray, assuming that the distance between object and entrance pupils is equal to k, we find that

$$\begin{aligned} \mathbf{b} &= \mathbf{b}, & \mathbf{b}' &= a'\mathbf{b} + b'\mathbf{b}_p, \\ \mathbf{t} &= \gamma(\mathbf{b} - \mathbf{b}_p), & \mathbf{t}' &= c'\mathbf{b} + d'\mathbf{b}_p, \end{aligned} \tag{47}$$

where

$$\gamma = \frac{n}{\sqrt{k^2 + (\mathbf{b} - \mathbf{b}_p)^2}} = \frac{n}{\sqrt{k^2 + u - 2v + w}}\,. \tag{48}$$

From (48) we conclude that

$$\gamma_u = -\frac{1}{2}\gamma_v = \gamma_w = \frac{1}{2}\frac{n}{\sqrt{(k^2 + u - 2v + w)^3}}\,. \tag{49}$$

Thus Eq. (29) gives

$$A = u, \qquad B = v, \qquad C = \gamma(u - v), \qquad D = \gamma(v - w). \tag{50}$$

The fundamental equations (C) become

$$\begin{aligned} a'd' - b'c' &= \frac{n}{\sqrt{k^2 + u - 2v + w}}, \\ I' = -II' = III' &= \frac{1}{2}\frac{n}{(k^2 + u - 2v + w)^{3/2}}\,. \end{aligned} \tag{51}$$