THE ABSOLUTE OPTICAL INSTRUMENT*

BY

J. L. SYNGE

1. **Introduction.** An *absolute optical instrument* is a system of transparent media which gives a precise point-image of each object-point lying in a three-dimensional region, the law of light propagation being that of geometrical optics. For a long time only two absolute instruments were known, the plane mirror (or combination of plane mirrors) and the "fish's eye" of Maxwell,† which consists of a single isotropic medium of variable refractive index

$$(1.1) \qquad n = n_0 a^2/(a^2 + r^2),$$

where n_0, a are constants, and r is the distance from a fixed point. Maxwell's medium has been generalized by Lenz,‡ and Boegehold and Herzberger§ have given an infinite class of absolute instruments, consisting of homogeneous isotropic media bounded by concentric spheres.

Less attention has been paid to the question of the realization of an absolute instrument by a distribution of transparent media than to the relations between object and image which must hold if the instrument is absolute. In the simplest case, in which the initial and final media are homogeneous and isotropic, it has been shown in various ways by Maxwell, Bruns, Klein, and Liebmann‖ that if A, B are the images of object-points A', B', then

$$(1.2) \qquad n'A'B' = nAB,$$

where n', n are the (constant) refractive indices of the initial and final media. This relation may also be written

$$(1.3) \qquad [A'B'] = [AB],$$

where the brackets indicate optical lengths.

Carathéodory* has extended this result to the case of more general media, namely, to those in which the velocity of propagation of light is the same in opposite directions and this velocity, considered as a function of the direction of a ray, satisfies certain conditions of analyticity.† Further, by introducing the idea of the *field* of the instrument, he made his result applicable to actual instruments.

In the present paper the approach to the problem of the absolute instrument differs from that of Carathéodory. Attention is directed to the surface of components rather than the wave-surface, and we are thereby enabled to see that when the optical character of one of the terminal media is assigned, the optical character of the other is rather closely restricted if the instrument is absolute.

Extension to the case where the space is Riemannian and of N dimensions may easily be made; but since the problem is one of considerable optical interest, it has been thought best not to complicate the presentation by this extension. The problem may also be presented as a problem in Finsler space.

2. **Fundamental theory.** Hamilton's great achievement in geometrical optics was the reconciliation, in the form of a single mathematical treatment, of the emission theory of light and the wave theory,‡ or, in the language of pure mathematics, the calculus of variations and the theory of contact transformations. This comprehensive view places at our disposal for the discussion of any problem in geometrical optics two alternative methods—the ray-method and the wave-method.

The methods employed in the present paper are essentially those of Hamilton, although for brevity an indicial notation will be used, suffixes having the range 1, 2, 3, with the usual summation convention.§ Except for a modification of one of Hamilton's definitions to suit present purposes, there is nothing in the present paper which might not have been given by Hamilton a century ago as an immediate deduction from his theory. The dis-

---

* C. Carathéodory, loc. cit. The part of this paper containing the fundamental theorem is reproduced, with no essential modification, in M. Born, *Optik*, Berlin, 1933, pp. 61–63.

† Professor Carathéodory has informed me by letter that in establishing the result he had exclusively in mind media of the Fresnel type, or media in which similar conditions of analyticity are satisfied. Unfortunately these conditions were not stated in his paper, and the reader might come to the false conclusion that the result follows from the single assumption that the velocity of propagation is the same in opposite directions. That is not so, and it is not difficult to construct artificial examples in illustration. Carathéodory's method is discussed in §9.

‡ Cf. G. Prange, *W. R. Hamiltons Abhandlungen zur Strahlenoptik*, Leipzig, 1933, Anmerkungen, p. 104; *The Mathematical Papers of Sir W. R. Hamilton*, vol. 1, Cambridge, 1931, pp. 277, 497. The latter will be referred to as M. P. H.

§ J. L. Synge, Journal of the Optical Society of America, vol. 27 (1937), pp. 75–82.

cussion of an absolute instrument does not seem to have occurred to him; the idea arose out of Gaussian optics.

The essential parts of the theory will now be developed, care being taken to present them in a form applicable to an absolute instrument. Certain exceptional features are present in such an instrument which may render invalid statements valid in more general cases.

The instrument is composed of any number of media. Each medium has a positive *medium-function* $v(x_r, \alpha_r)$, a function of rectangular cartesian coordinates x_r and of direction cosines α_r; v is homogeneous of degree unity in the direction cosines. It is in general multiple-valued. If the units of space and time are so chosen that the velocity of light in vacuo is unity, then $1/v$ is the *ray velocity* in the medium for a ray having direction cosines α_r at the point x_r. (For an isotropic medium, v is equal to the refractive index.)

By Fermat's principle (accepted as a basic hypothesis) rays satisfy the variational principle $\delta \int v ds = 0$, ds being an element of arc.

The *components of normal slowness* corresponding to a ray through x_r with direction cosines α_r are defined as

$$(2.1) \qquad \sigma_r = \partial v / \partial \alpha_r.$$

They were so called by Hamilton* because the vector σ_r stands normal to the wave and its magnitude is equal to the reciprocal of the wave velocity.

The right-hand side of (2.1) being homogeneous of degree zero in the direction cosines, we can eliminate their ratios and obtain the *medium-equation*

$$(2.2) \qquad \Omega(x_r, \sigma_r) = 0.$$

If, from an assigned point x_r, we measure off along each straight line with direction cosines α_r a length $1/v$, obtaining a point with relative coordinates α_r/v, the surface so obtained is the *wave-surface* corresponding to x_r. On the other hand if we draw from x_r the totality of vectors σ_r satisfying (2.2), we get the *surface of components* corresponding to x_r. The wave-surface and the surface of components are reciprocal surfaces with respect to the unit sphere with centre x_r.†

---

* M.P.H., p. 278; M. Herzberger, *Strahlenoptik*, Berlin, 1931, p. 9, calls this vector the "normal vector." "Slowness vector" might be a more descriptive abbreviation for the full title namely, "the vector representing the normal slowness of wave-propagation." These components are of course the optical analogues of the generalized momenta (p_r) of Hamilton's dynamical theory.

† It seems a pity to reject Hamilton's very suggestive terminology, *wave-surface* and *surface of components*, in favor of the names *indicatrix* and *figuratrix*, which carry no intrinsic meaning. Unless the extension to n dimensions appears an important advance, it does not seem historically correct to assign priority in the consideration of these surfaces to Minkowski and Hadamard (cf. C. Carathéodory, *Variationsrechnung*, Leipzig und Berlin, 1935, p. 247). Hamilton had a priority of seventy years, and even he assigned priority to Cauchy.

The *optical length* of any curve C joining points P, Q is

$$(2.3) \qquad [PQ] = \int_C v\,ds,$$

the directional arguments of v being the direction cosines of the tangent to C, in the sense P to Q, and ds being a positive element of arc. Giving a weak variation with, in general, displacements of the end points P, Q, we know from the calculus of variations that

$$(2.4) \qquad \delta[PQ] = \left(\frac{\partial v}{\partial \alpha_r}\delta x_r\right)_Q - \left(\frac{\partial v}{\partial \alpha_r}\delta x_r\right)_P - \int_C\left(\frac{d}{ds}\frac{\partial v}{\partial \alpha_r} - \frac{\partial v}{\partial x_r}\right)\delta x_r\,ds.$$

Suppose now that C is a ray; then, by the Euler equations, the integral vanishes and we have, in the notation of (2.1),

$$(2.5) \qquad \delta[PQ] = (\sigma_r\delta x_r)_Q - (\sigma_r\delta x_r)_P.$$

The *characteristic function* of an instrument is a function V of the coordinates x_r' of a point A' in the initial medium M' and the coordinates x_r of a point A in the final medium M, such that V is equal to the optical length of a ray joining A' and A. The function V may not be defined for all values of the six variables x_r', x_r corresponding to points A', A in the initial and final media respectively, because there may exist no ray joining A' and A. In general V is multiple-valued, since v is multiple-valued.

We shall use accents to denote quantities pertaining to the initial medium M' of the instrument; quantities pertaining to the final medium will be left unaccented. Different rectangular axes for x_r' and x_r may be used for M' and M respectively. Inspection of the method by which (2.5) has been established shows that the axes to which quantities at P are referred need not be the same as the axes to which quantities at Q are referred. From (2.5) we see at once that if we pass from a ray joining points A', A to an adjacent ray joining points B', B, the variation in V is*

$$(2.6) \qquad \delta V = \sigma_r\delta x_r - \sigma_r'\delta x_r'.$$

---

* This is the fundamental equation of Hamilton's theory, called by him the *equation of the characteristic function* (M.P.H., p. 168). It is of course the optical analogue of the fundamental equation defining contact transformations in dynamics. Hamilton did not frame his theory to take into account the exceptional case presented by an absolute instrument; he assumed that six arbitrary variations $\delta x_r'$, δx_r were permissible in (2.6), so that

$$\partial V/\partial x_r = \sigma_r, \qquad \partial V/\partial x_r' = -\sigma_r'.$$

This cannot be done in the case of an absolute instrument if the points x_r' and x_r correspond to object and image respectively. The argument of the present paper, which makes no such illegitimate assumption, has already been given by Herzberger (op. cit., p. 12) in a different notation.

Following Carathéodory, we shall say that a ray passing through a point A' of M' with direction cosines α_r' *lies in the field of the instrument* if it passes through the instrument to the final medium M. At a given point A' of M' the directions of the rays lying in the field of the instrument will be bounded by a cone of unidirectional straight lines drawn out from A'.

3. **Necessary conditions for an absolute instrument.** Consider the rays emanating in the field of the instrument from an initial point A'. As the medium-functions of the media in the instrument may be multiple-valued, these rays may pass through the final medium in a number of congruences, which we distinguish from one another by calling them congruences of different types.

If all the rays in the final congruence of type T pass through a single point A, we may say that the instrument is *absolute (type T) with respect to the source A'*, A being the image (type T) of A'. If the instrument is absolute as above with respect to all initial points and directions in the field of the instrument, we may say that the instrument is *absolute of type T*. If it is absolute for all types, we may say that it is *completely absolute*.

In the present paper we shall be concerned with absolute character for final congruences of a definite type. For simplicity, we shall refer to an instrument possessing this absolute character as *absolute*, without further explicit qualification.

In an absolute instrument there exists a one-to-one correspondence

$$(3.1) \qquad\qquad x_r' = x_r'(x_s), \qquad x_r = x_r(x_s'),$$

between the points of the initial and final media, A with coordinates x_r being the image of A' with coordinates x_r'.

If, given A', we seek points A such that a ray joins A' to A, we find our choice of A restricted to that part of the final medium traversed by the congruence of rays from A'. Thus, given A' with coordinates x_r', the characteristic $V(x_r', x_r)$ is defined only for a restricted range of values of x_r. This range includes of course the image of A'; the coordinates of that image are, by (3.1), functions of x_r'. When A is chosen at the image of A', there is an infinity of rays joining the two points, but by Fermat's principle their optical lengths are all the same. Hence *in an absolute instrument there exists an absolute characteristic, namely, a function of the three initial coordinates, or of the three final coordinates, whose value is equal to the optical length of a ray joining object to image*. This absolute characteristic will be denoted by F.

Let us now take two adjacent image-points x_r, $x_r + \delta x_r$, and compare the optical lengths of the rays drawn to these points from their respective ob-

jects. The difference between these optical lengths, by (2.6) and (3.1), is given by

(3.2)
$$\delta F = \sigma_r \delta x_r - \sigma_r' \delta x_r'$$
$$= \sigma_r \delta x_r - \sigma_s' \frac{\partial x_s'}{\partial x_r} \delta x_r,$$

where σ_r' are the components of any ray (lying in the field) through the object x_r' and σ_r the components of the corresponding ray through the image x_r. Since δx_r are arbitrary, (3.2) gives the three equations

(3.3)
$$\frac{\partial F}{\partial x_r} = \sigma_r - \sigma_s' \frac{\partial x_s'}{\partial x_r},$$

which are the fundamental equations of the present paper.

The implications of (3.3) may be explored in various ways. Thus we may suppose the absolute characteristic F and the transformation (3.1) given, and inquire into the conditions imposed by (3.3) on the optical characters of the initial and final media. Or we may suppose that the initial and final media are of assigned optical characters, given by medium-functions $v'(x_r',\alpha_r')$, $v(x_r, \alpha_r)$ or by surfaces of components

(3.4)
$$\Omega'(x_r', \sigma_r') = 0, \qquad \Omega(x_r, \sigma_r) = 0;$$

we may then inquire into the conditions imposed by (3.3) on the absolute characteristic F and the transformation (3.1). The latter type of approach has been the usual one.

4. **The optical character of one terminal medium in an absolute instrument deduced from that of the other.** Let us write (3.3) as

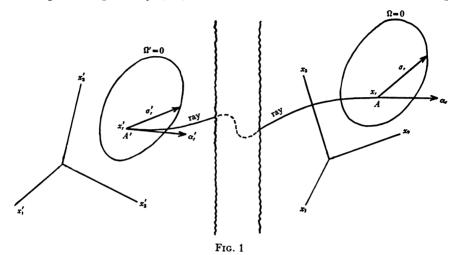(4.1)
$$\sigma_r = a_{rs}\sigma_s' + b_r,$$

where

(4.2)
$$a_{rs} = \frac{\partial x_s'}{\partial x_r}, \qquad b_r = \frac{\partial F}{\partial x_r}.$$

Let us consider (4.1) as a transformation expressing final components σ_r in terms of initial components σ_r', the corresponding object and image-points (A' and A) and the absolute characteristic being given. Thus for present purposes a_{rs} and b_r are to be treated as constants; (4.1) is a linear transformation.

Consider now the surface of components S' corresponding to A' (with equation $\Omega'=0$) and the surface of components S corresponding to A (with

equation $\Omega = 0$). Consider any ray through A' in the field of the instrument. Let its components be σ_r'. Then the vector drawn from A' with components σ_r' parallel to the axes of x_r' has its end on S'. The corresponding final ray has components given by (4.1). But the vector drawn from A with compo-

FIG. 1

nents σ_r parallel to the axes of x_r must have its end on S (Fig. 1). Hence we have the following result:

THEOREM I. *In an absolute instrument the portions lying in the field of the instrument of the surfaces of components corresponding to an object-point A' and its image A are transformable into one another by a linear transformation, whose coefficients are in general functions of the coordinates of A'.*

Thus, insofar as the condition (3.3) is concerned, we may choose arbitrarily the following:

(i) the absolute characteristic F,

(ii) the correspondence $x_r = x_r(x_s')$,

(iii) the surfaces of components (that is, the equation $\Omega'(x_r', \sigma_r') = 0$) throughout the initial medium.

Then the surfaces of components for the final medium (and hence its medium-function v) can be found to satisfy (3.3). We have in fact by (3.3)

$$(4.3) \qquad \sigma_r' = \sigma_s \frac{\partial x_s}{\partial x_r'} - \frac{\partial F}{\partial x_r'},$$

so that $\Omega'(x_r', \sigma_r') = 0$ implies

$$(4.4) \qquad \Omega'\left(x_r'(x_s), \sigma_s \frac{\partial x_s}{\partial x_r'} - \frac{\partial F}{\partial x_r'} \right) = 0;$$

these are the required surfaces of components in the final medium, (4.4) being immediately expressible in the form $\Omega(x_r, \sigma_r) = 0$.

In particular we may observe that if there is an absolute instrument in which the surfaces of components in one of the terminal media are ellipsoids (the wave-surfaces being consequently, by the reciprocal relation, ellipsoids also), then the surfaces of components in the other terminal medium are also ellipsoids (and so are the wave-surfaces).

More generally, if in an absolute instrument the surfaces of components throughout one of the terminal media are undegenerate algebraic surfaces of the nth degree, then the surfaces of components throughout the other terminal medium are also algebraic surfaces of the nth degree.

The assumption of algebraic character carries with it an important implication (as would also a suitable assumption regarding analyticity), namely, that a linear transformation which carries the *portion* of one of the surfaces of components corresponding to rays lying in the field of the instrument into a portion of the other surface of components, also carries the *whole* of one surface of components into the whole of the other. It is this fact that enables us to make statements about complete surfaces of components, although our data deal only with the portions corresponding to rays lying in the field of the instrument.

5. **The case where the surfaces of components in the terminal media are general algebraic surfaces.** Let us now suppose that there is an instrument with given terminal media in which the surfaces of components are undegenerate algebraic surfaces of the nth degree. We may write

(5.1) $\Omega' \equiv 1 + c_r' \sigma_r' + c_{rs}' \sigma_r' \sigma_s' + \cdots$ (to $n + 1$ terms) $= 0,$ $(c_{rs}' = c_{sr}'),$

(5.2) $\Omega \equiv 1 + c_r \sigma_r + c_{rs} \sigma_r \sigma_s + \cdots$ (to $n + 1$ terms) $= 0,$ $(c_{rs} = c_{sr}),$

where c_r', c_{rs}', $\cdots$ are assigned functions of x_r', and c_r, c_{rs}, $\cdots$ are assigned functions of x_r.

We wish to investigate the possibility of constructing an absolute instrument with these terminal media, taking into consideration the condition (3.3).

Adopting the notation of (4.1) and substituting in (5.2), we get

(5.3) $\Omega \equiv 1 + c_r(a_{rs}\sigma_s' + b_r) + c_{rs}(a_{rt}\sigma_t' + b_r)(a_{su}\sigma_u' + b_s) + \cdots = 0;$

this must be the same as the surface (5.1). Thus we require

$$1 + c_r b_r + c_{rs} b_r b_s + \cdots = \theta,$$

(5.4)

$$c_s a_{sr} + 2c_{st} a_{sr} b_t + \cdots = \theta c_r',$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad ,$$

where θ is a factor of proportionality. Here we have a set of equations for five unknown functions of x_r, namely, F, x_r', θ. Omitting from consideration for obvious reasons the case $n = 1$, let us pass to the case $n = 2$, for which the surfaces are quadrics. There are then ten coefficients in Ω and hence ten equations in (5.4) for five unknowns. If $n > 2$, there are of course more equations. *It is therefore in general impossible to construct an absolute instrument with arbitrarily assigned terminal media in which the surfaces of components (or the wave-surfaces) are undegenerate algebraic surfaces of the second degree or of higher degree.*

6. **The case where the surfaces of components in the terminal media are central algebraic surfaces.** By a *central surface* we shall understand a surface which has a centre at the origin; for a surface of components or wave-surface, the origin is the point in the medium to which the surface corresponds.

A linear transformation (4.1) which carries a central surface into a central surface must lack the absolute term. Hence we have this result (of importance in connection with Carathéodory's theorem and implicit in his proof, unless replaced by a condition of analyticity):

THEOREM II. *In an absolute optical instrument in which the surfaces of components (or, equivalently, the wave-surfaces) are undegenerate central algebraic surfaces, the absolute characteristic is a constant. This result also holds for degenerate central surfaces, provided that each of the separate surfaces given by degeneration is central.*

Under these circumstances the transformation of components is simply

$$(6.1) \qquad \sigma_r = \sigma_s' \, \frac{\partial x_s'}{\partial x_r}.$$

7. **The case of isotropic terminal media.** When the terminal media are isotropic, the surfaces of components are spheres with the equations

$$(7.1) \qquad \Omega' \equiv \sigma_r' \sigma_r' - n'^2 = 0,$$

$$(7.2) \qquad \Omega \equiv \sigma_r \sigma_r - n^2 = 0,$$

where n', n are refractive indices, functions of position. Applying (6.1), we see that if the instrument is absolute, the correspondence of object and image-points must satisfy

$$(7.3) \qquad \frac{\partial x_s'}{\partial x_r} \frac{\partial x_t'}{\partial x_r} = \frac{n^2}{n'^2} \delta_{st},$$

where δ_{st} is the Kronecker delta. These partial differential equations, six in

number, are not in general soluble for the three unknowns x_r'. We deduce from them

$$(7.4) \qquad \frac{\partial x_s'}{\partial x_r} = \frac{n^2}{n'^2} \frac{\partial x_r}{\partial x_s'},$$

and hence

$$
\begin{aligned}
(7.5) \qquad n^2 ds^2 = n^2 dx_r dx_r &= n^2 \frac{\partial x_r}{\partial x_s'} \frac{\partial x_r}{\partial x_t'} dx_s' dx_t' \\
&= \frac{n'^4}{n^2} \frac{\partial x_s'}{\partial x_r} \frac{\partial x_t'}{\partial x_r} dx_s' dx_t' \\
&= n'^2 dx_s' dx_s' \\
&= n'^2 ds'^2.
\end{aligned}
$$

This establishes (by a method different from, and perhaps more direct than, the method of Carathéodory) the equality of optical lengths for object and image elements in an absolute instrument for which the terminal media are isotropic and in general heterogeneous. A necessary condition for the absolute character of such an instrument is the applicability of the terminal media, considered as Riemannian spaces with metrics nds, $n'ds'$.

8. **The case of Fresnel media.** In a biaxial crystal, or homogeneous Fresnel medium, the surface of components referred to principal axes has the equation[*]

$$(8.1) \qquad \frac{\sigma_1^2}{1 - c_1^2 q^2} + \frac{\sigma_2^2}{1 - c_2^2 q^2} + \frac{\sigma_3^2}{1 - c_3^2 q^2} = 0, \qquad q^2 \equiv \sigma_r \sigma_r,$$

where c_r are constants (the principal velocities). Cleared of fractions, this equation reads

$$(8.2) \qquad q^2 A_{mn}^0 \sigma_m \sigma_n - B_{mn}^0 \sigma_m \sigma_n + 1 = 0, \qquad q^2 \equiv \sigma_r \sigma_r,$$

where

$$(8.3) \qquad \begin{cases} A_{11}^0 = c_2^2 c_3^2, & A_{22}^0 = c_3^2 c_1^2, & A_{33}^0 = c_1^2 c_2^2, \\ B_{11}^0 = c_2^2 + c_3^2, & B_{22}^0 = c_3^2 + c_1^2, & B_{33}^0 = c_1^2 + c_2^2, \end{cases}$$

and $A_{mn}^0 = B_{mn}^0 = 0$ if $m \neq n$.

If we change to an arbitrary set of rectangular axes (not principal) the surface of components has an equation of the form

$$(8.4) \qquad q^2 A_{mn} \sigma_m \sigma_n - B_{mn} \sigma_m \sigma_n + 1 = 0, \qquad q^2 \equiv \sigma_r \sigma_r,$$

---

[*] Cf. M.P.H., p. 280; Born, op. cit., p. 224.

where $A_{mn} = A_{nm}$, $B_{mn} = B_{nm}$, these quantities being found from A^0_{mn}, B^0_{mn} respectively by application of the rules of tensor transformation. Given an equation (8.4), necessary and sufficient conditions that it should define the surface of components in a homogeneous Fresnel medium are as follows:

(i) The roots of the determinantal equations

$$(8.5) \qquad \left| A_{mn} - \lambda \delta_{mn} \right| = 0, \qquad \left| B_{mn} - \mu \delta_{mn} \right| = 0,$$

are of the form

$$(8.6) \qquad \begin{cases} \lambda_1 = c_2^2 c_3^2, & \lambda_2 = c_3^2 c_1^2, & \lambda_3 = c_1^2 c_2^2, \\ \mu_1 = c_2^2 + c_3^2, & \mu_2 = c_3^2 + c_1^2, & \mu_3 = c_1^2 + c_2^2, \end{cases}$$

where the c_r are constants.

(ii) Each of the three perpendicular directions defined by

$$(8.7) \qquad A_{mn}\sigma_n - \lambda_r \sigma_m = 0$$

for fixed r, coincides with the direction defined by

$$(8.8) \qquad B_{mn}\sigma_n - \mu_r \sigma_m = 0$$

for the same value of r.

A heterogeneous Fresnel medium may be defined as a medium in which at each point there exists a set of axes such that the surface of components corresponding to that point has the form (8.1). Equivalently, we may say that, for a single set of rectangular cartesian coordinates throughout the medium, a heterogeneous Fresnel medium is one in which the surface of components at any point has the form (8.4), where A_{mn}, B_{mn} are functions of the coordinates x_r satisfying the conditions (i) and (ii) above, in which c_r are no longer constants but functions of x_r. We shall confine our attention to the undegenerate case in which c_1, c_2, c_3 are distinct.

Let us now consider an absolute instrument in which the terminal media are Fresnel media, in general heterogeneous. Their surfaces of components have the equations

$$(8.9) \qquad q'^2 A'_{mn}\sigma'_m \sigma'_n - B'_{mn}\sigma'_m \sigma'_n + 1 = 0, \qquad q'^2 \equiv \sigma'_r \sigma'_r,$$

$$(8.10) \qquad q^2 A_{mn}\sigma_m \sigma_n - B_{mn}\sigma_m \sigma_n + 1 = 0, \qquad q^2 \equiv \sigma_r \sigma_r.$$

Since these are undegenerate central algebraic surfaces, the formula for transformation of components is homogeneous of the form (6.1). Hence we must have

$$(8.11) \qquad B_{mn}\sigma_m \sigma_n \equiv B'_{mn}\sigma'_m \sigma'_n,$$

$$(8.12) \qquad q^2 A_{mn}\sigma_m \sigma_n \equiv q'^2 A'_{mn}\sigma'_m \sigma'_n.$$

Two cases must now be distinguished, according to the way in which we break up (8.12):

Case I:

(8.13) $q^2 \equiv \theta^2 q'^2,$ $A_{mn}\sigma_m\sigma_n \equiv \theta^{-2}A'_{mn}\sigma'_m\sigma'_n \; ;$

Case II:

(8.14) $q^2 \equiv \phi^2 A_{mn}\sigma'_m\sigma'_n,$ $A_{mn}\sigma_m\sigma_n = \phi^{-2}q'^2;$

here θ, ϕ are undetermined functions of x_r.

Let us choose a pair of corresponding points (object and image) and choose our axes in the principal directions corresponding. Then (8.11) reads

(8.15) $$\sum (c_2^2 + c_3^2)\sigma_1^2 \equiv \sum (c_2'^2 + c_3'^2)\sigma_1'^2,$$

where $\sum$ indicates a sum of three terms obtained by cyclic permutation of 1, 2, 3. Therefore, renumbering the axes if necessary,

(8.16) $$\begin{cases} (c_2^2 + c_3^2)\sigma_1^2 \equiv (c_2'^2 + c_3'^2)\sigma_1'^2, \\ (c_3^2 + c_1^2)\sigma_2^2 \equiv (c_3'^2 + c_1'^2)\sigma_2'^2, \\ (c_1^2 + c_2^2)\sigma_3^2 \equiv (c_1'^2 + c_2'^2)\sigma_3'^2. \end{cases}$$

Let us now consider separately Case I and Case II.

Case I: Substitution from (8.16) in (8.13) gives

(8.17) $$\sum \frac{c_2'^2 + c_3'^2}{c_2^2 + c_3^2} \sigma_1'^2 \equiv \theta^2 \sum \sigma_1'^2,$$

(8.18) $$\sum c_2^2 c_3^2 \frac{c_2'^2 + c_3'^2}{c_2^2 + c_3^2} \sigma_1'^2 \equiv \theta^{-2} \sum c_2'^2 c_3'^2 \sigma_1'^2.$$

From (8.17) we have

(8.19) $$\begin{cases} c_2'^2 + c_3'^2 = \theta^2(c_2^2 + c_3^2), \\ c_3'^2 + c_1'^2 = \theta^2(c_3^2 + c_1^2), \\ c_1'^2 + c_2'^2 = \theta^2(c_1^2 + c_2^2), \end{cases}$$

from which it follows that

(8.20) $c_1' = \theta c_1,$ $c_2' = \theta c_2,$ $c_3' = \theta c_3;$

then (8.18) is automatically satisfied.

Case II: Substitution from (8.16) in (8.14) gives

(8.21) $$\sum \frac{c_2'^2 + c_3'^2}{c_2^2 + c_3^2} \sigma_1'^2 \equiv \phi^2 \sum c_2'^2 c_3'^2 \sigma_1'^2,$$

$$(8.22) \qquad \sum c_2^2 c_3^2 \frac{c_2'^2 + c_3'^2}{c_2^2 + c_3^2} \sigma_1'^2 \equiv \phi^{-2} \sum \sigma_1'^2 .$$

From (8.21) we have

$$(8.23) \qquad \begin{cases} c_2'^{-2} + c_3'^{-2} = \phi^2(c_2^2 + c_3^2), \\ c_3'^{-2} + c_1'^{-2} = \phi^2(c_3^2 + c_1^2), \\ c_1'^{-2} + c_2'^{-2} = \phi^2(c_1^2 + c_2^2), \end{cases}$$

and hence

$$(8.24) \qquad c_1'^{-1} = \phi c_1, \qquad c_2'^{-1} = \phi c_2, \qquad c_3'^{-1} = \phi c_3;$$

then (8.22) is automatically satisfied. Hence we have this result:

THEOREM III. *In an absolute instrument in which the terminal media are undegenerate Fresnel media (in general heterogeneous), the principal velocities at a point in one medium are either directly or inversely proportional to the principal velocities at the corresponding point in the other medium.*

9. **Carathéodory's method.** By (3.2), valid for an absolute instrument, we have

$$(9.1) \qquad \delta F = \sigma_r \delta x_r - \sigma_r' \delta x_r' .$$

Now it is immediately obvious that if we take adjacent points A', B' on a ray in the initial medium in the field of an absolute instrument, their images A, B lie on the final portion of the ray in question. Taking in (9.1) $\delta x_r'$ to be the displacement from A' to B' and δx_r to be the displacement from A to B, and denoting the direction cosines of initial and final rays as usual by α_r', α_r, we may substitute in (9.1)

$$(9.2) \qquad \delta x_r = \alpha_r ds, \qquad \delta x_r' = \alpha_r' ds',$$

where ds, ds' are respectively equal to AB, $A'B'$. Thus we obtain, on division by ds,

$$(9.3) \qquad \frac{\partial F}{\partial x_r} \alpha_r = \sigma_r \alpha_r - \sigma_r' \alpha_r' \frac{ds'}{ds},$$

or, by (2.1) and the homogeneity of v,

$$(9.4) \qquad \frac{\partial F}{\partial x_r} \alpha_r = v(x_r, \alpha_r) - v'(x_r', \alpha_r') \frac{ds'}{ds} .$$

But

$$(9.5) \quad v'(x_r', \alpha_r') \frac{ds'}{ds} = v'\left(x_r', \frac{dx_r'}{ds'}\right) \frac{ds'}{ds} = v'\left(x_r', \frac{dx_r'}{ds}\right) = v'\left(x_r', \frac{\partial x_r'}{\partial x_s} \alpha_s\right),$$

and (9.4) thus becomes

$$(9.6) \qquad \frac{\partial F}{\partial x_r} \alpha_r = v(x_r, \alpha_r) - v'\left(x'_r, \frac{\partial x'_r}{\partial x_s} \alpha_s\right).$$

Except for trivial changes in notation, this is the fundamental equation (14) of Carathéodory's paper (loc. cit.).

Let us now reconstruct Carathéodory's theorem regarding equality of optical lengths, filling in the part of the proof not given explicitly in his paper. We shall assume that the instrument is absolute for rays (in the field) of a certain type, corresponding to a certain one of the values (v') of the multiple-valued medium-function of the initial medium and to a certain one of the values (v) of the multiple-valued medium-function of the final medium. We assume that these single-valued functions v', v are *even analytic* functions of α'_r, α_r respectively, throughout three-dimensional spaces in which α'_r, α_r are coordinates, except perhaps for certain singular lines through the origins.

Let P be any point in the region of the α_r space which is filled with lines drawn in the directions of rays lying in the field of the instrument, and let Q be the reflection of P in the origin of the α_r space. Let C be an analytic curve joining P and Q. Then if we put

$$\phi(u) = \frac{\partial F}{\partial x_r} \alpha_r - \left\{v(x_r, \alpha_r) - v'\left(x'_r, \frac{\partial x'_r}{\partial x_s} \alpha_s\right)\right\},$$

where u is the parameter on C, $\phi(u)$ is an analytic function which vanishes for a range of values of u adjacent to P. Therefore $\phi(u) = 0$ along C and, in particular, at Q. But the first part of $\phi(u)$ is an odd function, and the second part is an even function. Hence, adding and subtracting the values of $\phi(u)$ at P and Q, we get

$$(9.7) \qquad \partial F/\partial x_r = 0, \qquad v(x_r, \alpha_r) = v'\left(x'_r, \frac{\partial x'_r}{\partial x_s} \alpha_s\right),$$

throughout the field in which P has been taken. Applying the analytic condition again, we see that the second relation holds for all values of α_r, except for the excluded singularities. Hence if $\delta x'_r$, δx_r are object and image elements with lengths $\delta s'$, δs and direction cosines α'_r, α_r respectively, we have

$$
\begin{aligned}
v(x_r, \alpha_r)\delta s &= v'\left(x'_r, \frac{\partial x'_r}{\partial x_s} \alpha_s\right)\delta s \\
(9.8) \qquad &= v'\left(x'_r, \frac{\partial x'_r}{\partial x_s} \delta x_s\right) \\
&= v'(x'_r, \alpha'_r)\delta s'.
\end{aligned}
$$

Thus Carathéodory's method establishes the equality of optical lengths for corresponding elements in an absolute instrument. The method by which the fundamental equation (9.1) is obtained in the present paper (indeed, it is deducible in a couple of lines from Hamilton's equation of the characteristic function, properly interpreted) appears more direct than the method given by Carathéodory. The latter part of the argument given above is merely the expansion of Carathéodory's proof, with a full statement of the implied conditions.

The necessary amplification of his original proof will also be found in Carathéodory, *Geometrische Optik*, Berlin, 1937, p. 70, which appeared after the present paper was written. Reference should be made to W. Blaschke, Abhandlungen aus dem Mathematischen Seminar der Hansischen Universität, vol. 11, 1936, pp. 409–412; although not based on the characteristic function, his argument is similar to that of the present paper, for he directs attention to the surface of components and the linear transformation (4.1). I owe this reference to W. C. Graustein.*

---

* The last paragraph was added in proof, May 20, 1938.

UNIVERSITY OF TORONTO,
    TORONTO, CANADA